Internet Transparency in Developing African regions : Case of the DR Congo

by

Dieudonne Ishara MUNGANGA

Masters by Dissertation only in Computer Science

2022

Faculty of Science
University of Cape Town

Internet Transparency in Developing Africans regions : Case of DR Congo
Next line

by

Dieudonne Ishara Munganga

Supervisor: Dr. Josiah Chavula

Department: Department of Computer Science

Masters by dissertation only in Computer Science

2022

Faculty of Science
University of Cape Town

Acknowledgments

Writing this thesis has been a journey, and I would like to take this opportunity to express my gratitude to all those who have supported me throughout this process.

First and foremost, I would like to start by giving thanks to God for His grace and guidance throughout my academic journey. It is by His blessing that I have been able to successfully complete this thesis and I am grateful for His continuous presence in my life.

I would like to express my deepest appreciation to my advisor, Dr Josiah Chavula, for their unwavering support, guidance, and encouragement. Their insightful comments and suggestions have been invaluable in shaping this thesis and helping me to develop my ideas and thoughts.

I would also like to extend my thanks to the members of my research proposal defense committee, Professors Melissa Densmore and Rob Simmonds, for their constructive feedback and valuable suggestions. Their expertise and support have been instrumental in my research and in the development of this thesis.

I would like to extend my sincere gratitude to my colleagues and friends for their support and encouragement. Their camaraderie and understanding have been a source of comfort during this challenging yet rewarding journey.

I am grateful to the Computer Science department of the University of Cape Town for providing me with the resources and opportunities to pursue my studies and develop my skills.

Finally, I would like to thank my family for their unconditional love and support. Particularly my Dad Joachim MUNGANGA and my Mum Antoinette Furaha, for their unwavering support in funding my education. Their investment in my future has made this journey possible and I am grateful for their unwavering love and belief in me which have been a constant source of inspiration and motivation.

I am deeply grateful to everyone who has supported me in this endeavor.

This thesis is a reflection of all the hard work, dedication, and support that I have received, and I am proud to have had the opportunity to work with such talented and dedicated individuals.

Thank you all!

Abstract

This thesis investigates internet transparency and connectivity in the Democratic Republic of Congo (DRC) using a combination of secondary data analysis, online surveys, and primary data collected through Internet measurement on personal devices. The research aims to understand the perception of internet performance by the users, the level of interconnection among Autonomous systems and the Quality of Service (QoS) and Quality of Experience (QoE) provided by broadband networks in the DRC, and how it compares to other Central African countries. The research found that users in the DRC have a low level of satisfaction with internet performance and that the country's internet infrastructure and Autonomous systems are not well interconnected, leading to poor network performance. Additionally, the research found that the QoS and QoE provided by broadband networks in the DRC are far from optimal and lower than other Central African countries. Furthermore, the research revealed that there is a lack of transparency in the DRC's internet structure, with certain networks exerting a significant level of influence on users' internet connectivity. These findings indicate a need for increased transparency and improved internet infrastructure in the DRC to better serve the needs of its citizens and support economic and social development. The thesis concludes with a section on future work, highlighting potential avenues for further research such as enhancing the Internet measurement aspect of the study, using machine learning techniques to analyze the data, conducting a study on the impact of internet access and quality on social development, and developing a real-time monitoring system for internet connectivity in the DRC.

Declaration

I declare that the thesis here submitted is original except for the source materials explicitly acknowledged and that this thesis as a whole, or any part of this thesis has not been previously submitted for the same degree or for a different degree.

I also acknowledge that I have read and understood the Rules on Handling Student Academic Dishonesty and the Regulations of the Student Discipline of the University of Cape Town.

Table of Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

It is important to pay sufficient attention to internet infrastructure, topology, and performance in order to effectively apply our understanding of the evolution and complexity of the internet. Internet infrastructure refers to the physical components that make up the Internet, including servers, routers, and cables. The topology of the internet refers to the way that these components are arranged and connected. In Africa, the internet infrastructure and topology can vary greatly from one country to another. Some countries may have a well-developed and sophisticated internet infrastructure, while others may have a more basic and limited infrastructure.

The performance of the internet in Africa also varies widely [1]. Some areas may have fast and reliable internet connections [2], while others may experience slow speeds and frequent outages [3]. It is important to have clear insights into internet infrastructure, topology, and performance in Africa for a number of reasons.

First, understanding the state of internet infrastructure and topology in a particular country or region can help policymakers and decision-makers make informed decisions about how to allocate resources and invest in improving internet access. For example, if a country has a poorly developed internet infrastructure, policymakers may prioritize investing in infrastructure upgrades to improve internet access and connectivity.

Second, understanding the performance of the internet in Africa can help businesses and individuals make informed decisions about how to use the Internet. For example, if an area has slow internet speeds, businesses may need to adjust their online strategies or consider alternative methods of communication. Finally, understanding the Internet infrastructure, topology, and performance in Africa can help researchers and analysts track the progress and development of the Internet in the region. This information can be used to identify areas where further investment and improvement are needed and to evaluate the effectiveness of policy and investment decisions related to the Internet in Africa. Such understandings are possible through Internet measurement, which refers to the process of collecting Internet data in networks to diagnose, improve, and make them more resilient. As Kelvin et al. [4] in 1883 said: *"When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind"*. This is why Internet performance measurements have become a necessity today, especially in unexpected situations such as Covid-19.

In Africa, even though there has been an increase in the deployment of mobile broadband networks, the continent still lags behind in terms of connectivity compared to other regions in the world. To put it into perspective, as of January 2021, Eastern and Middle Africa had the lowest rates of internet penetration, with only 24% and 26%, respectively, of their populations having access to the internet [5] [6]. Furthermore, the majority of internet traffic in the region comes from mobile devices, with only around 15% of households having access to fixed broadband at home (source: [7][8]). The cost of internet access is often prohibitively high for many people, encompassing expenses related to data pricing, device affordability, frequency of access, and the challenge of limited access to basic necessities such as electricity. These factors contribute to the lack of meaningful connectivity in the region (source: [9]).

Reasons for measuring the Internet can be grouped into three categories: technical, social, and commercial reasons. From a technical perspective, measuring the Internet in Africa can provide insight into the current state of the Internet on the continent, allowing us to more accurately identify the solutions and practices that need to be implemented. For instance, Phokeer et al. [10] gathered network measurements in low-resource communities in South Africa to determine their behavior regarding Wi-Fi usage and cellular data consumption. Socially, measurements can be handy to report comprehensive details of the Internet's response during crises as elaborated by National Research Council [11] [12]. Commercial reasons for conducting measurements are related to the end-users understanding of the kind of service they are being offered by their Internet Service Providers (ISP) and to benchmark how a network behaves.

Despite these advantages, there are many challenges that researchers and business operators experience when they attempt to measure the Internet. The first is the scale of the Internet. As the Internet rapidly expands along many axes (users, businesses, devices, applications, connection technologies, and protocols), its system becomes variable. As a consequence, characterizing it accurately is complex given a large number of observations this entails [13] [14]. The second challenge is opacity. Due to the scale and complexity of the current Internet, interpreting results for Internet measurement experiments is a difficult task. Originally, the Internet simply forwarded packets from a source to a destination. This made measurement a relatively straightforward task.
However, as complexity got introduced into the forwarding of traffic in terms of proxies, firewalls, caches, replicas, NATs, ad injectors, and performance enhancers, an observation at one point might bear little resemblance to an observation of the same traffic at a different point [15][16]. Lastly, ethics comes as the third challenge to Internet measurement due to an increase in the implications of disrupting user communication and personal data. For this reason, the topic forces heightened scrutiny as opposed to past Internet experiments [17]. Other challenges include infrastructure availability and interference in terms of shutdowns and censorship [18]. In Africa, keeping vantage points up and running, and the lack of interest in Internet measurements are the common challenges

that cause the continent to be less covered research-wise despite the presence of measurement platforms [13].

Approaches taken by researchers to perform Internet measurement experiments are classified to be either passive or active measurements. The former uses real applications and their traffic to record the application experience of using the network, i.e no additional traffic is introduced across the network. Active measurement involves saturating the end-to-end connection across the network path to tell its maximum capacity. Through the use of these techniques, network performance is derived through the data collected according to specific metrics. Widely known platforms such as Ookla's SpeedTest [19], Ripe Atlas [20][21], SpeedChecker [22], PerfSONAR, CAIDA Ark and M-LAB are examples of platforms used by researchers to conduct Internet measurements.

In this study, we applied an active measurement approach to determine the structure and state of the Internet, and its performance in the Democratic Republic of Congo in terms of both the Quality of Service (QoS) of networks and the Quality of Experience (QoE) of end-users. Additionally, we intend to run an online survey that will reveal the end-users' perceptions of Internet Performance in the country. Through a combination of both qualitative and quantitative data, our research will be based on both secondary and newly generated datasets. These will be constructed by a combination of data collection tools, each used according to the accuracy, scalability, and metrics it covers concerning our research goal.

1.1 Motivation

1.2 Problem Statement and research questions

Internet performance measurements play an essential role in determining the health of networks. This, however, has been minimally adopted in Africa, especially in regions whose state of the Internet is not known to the public. The

Democratic Republic of Congo is one such region. As a consequence, end-users choice of Internet service providers is speculation-based rather than data-driven, network operators cannot accurately compare their services to their competitors, and regulators have difficulty making insightful decisions that can grow and strengthen the Internet in the country. Our research aims to investigate the current state of the Internet in the country in terms of the Quality of Service (QoS) and Quality of Experience (QoE) of mobile and fixed-line broadband network performance by the use of active measurements. The research questions are presented as follows:

1. How do users perceive Internet performance in DR Congo?
Our first aim is to collect qualitative data that will reveal the users' needs given their experience with broadband Internet services and their level of knowledge about Internet measurements. To do this, we will conduct a survey from different provinces in the country.
2. To what extent are Autonomous systems in the DRC interconnected and how does this affect network performance? Through the use of secondary Internet traffic datasets, we will analyze DRC's Internet structure by means of ASNs' rank, customer sizes, the number and structure of paths between ASNs, and Traceroute-generated active datasets . Thereafter, a correlation to the data collected in the first two questions will be performed.
3. What quality of service and quality of experience do broadband networks in the DRC provide and how does it compare to other Central African countries? We will quantify the performance of networks based on metrics such as Upload/Download speeds, latency, Page Load Time, and Domain Name System performance

Answers to the above questions will indicate the broadband subscribers' QoS and QoE in DR Congo with comparisons to its neighbors, and to what extent network operators meet their advertised speeds. Furthermore, by answering these questions, the country's Internet structure and performance will be available to the research community, a step towards better Internet research transparency in Africa. Other insights of our study include knowledge of the level of awareness and practical application of Internet measurements in the country's networks.

1.3 Research Approach

This research takes a mixed-methods approach, using both qualitative and quantitative methods to investigate internet transparency and connectivity in the Democratic Republic of Congo (DRC). The research begins with a secondary data analysis of existing datasets that reveal the country's internet structure and the level of influence certain networks have on users' internet connectivity, as well as the historical Quality of Service (QoS) of networks in the country.

To gather primary data, the researcher conducted an online survey to gain insight into participants' perceptions of internet performance in the country and their level of satisfaction towards services offered by their Internet Service Provider (ISP). The survey was designed to gather qualitative information about users' perceptions and experiences with internet access in the DRC.

To complement the survey, research participants were also recruited to perform Internet measurement on their personal devices to assess their Quality of Service (QoS) and Web Quality of Experience (QoE) with metrics such as download/upload speed, latency, and Domain Name System resolution Time (DNST), Transport Control Protocol Time (TCPT), Time To First Byte (TTFB), and Total time. TCPT underscores the reliability of data exchange through organized segmenting and retransmission, while DNST reveals the swiftness of translating user-friendly domain names into vital IP addresses. TTFB illuminates server responsiveness by gauging the promptness of the initial data byte transmission. These metrics collectively contribute to the Total Time, encompassing the complete journey from user request initiation to the receipt of the first data byte, aiding in diagnosing network and server performance, and ultimately optimizing user experiences. These measurements were quantitative in nature and aimed to provide objective data on internet access quality in the DRC.

By using both qualitative and quantitative methods, this research was able to gather a comprehensive understanding of internet transparency and connectivity in the DRC, including both subjective perceptions of users and objective measurements of internet access quality.
Additionally, this research approach allowed the researcher to triangulate the data, comparing and contrasting the results from the different methods to gain a more complete understanding of the issue at hand. For example, the survey data provided insight into users' perceptions of internet access quality, while the Internet measurement data offered a more objective view of the actual quality of service and quality of experience provided by networks in the DRC.
The use of mixed-methods also allowed for a more robust analysis of the data and helped to increase the validity and reliability of the findings. The quantitative data provided a foundation for the analysis and helped to quantify the extent of the issues identified through the qualitative data.
Overall, the mixed-methods approach taken in this research allowed the researcher to gain a comprehensive understanding of internet transparency and connectivity in the DRC, by capturing both the subjective perceptions of users and the objective measurements of internet access quality. This research approach was well-suited to the research question and provided valuable insights into the state of internet access in the DRC.

1.4 Thesis Outline

The thesis is organized into eight main sections. The introduction section 1 provides background on the problem of internet transparency and connectivity in developing African regions, specifically in the Democratic Republic of Congo (DRC). The research question and objectives are presented, along with an explanation of the significance of the study.

The literature review section 2 provides an overview of internet access and infrastructure in Africa, particularly in the DRC. It also presents current research on internet transparency and connectivity in the DRC and other African regions. Theoretical framework and conceptual model are also discussed.

The methodology section 3 describes the research design and data collection methods used in the study. The research design is a mixed-methods approach, using both qualitative and quantitative methods. Data collection methods include secondary data analysis, online surveys, and primary data collection through Internet measurement on personal devices. The section also includes information on the participants and sampling techniques, as well as the data analysis techniques used.

The results sections 4,5 and 6 present the findings from the study. The findings from secondary data analysis on internet structure and Quality of Service (QoS) in the DRC are presented. Results from online survey on users' perceptions of internet performance in the DRC are also presented. Additionally, the results from primary data collection on the Quality of Service (QoS) and Quality of Experience (QoE) of broadband networks in the DRC are presented.

The discussion and analysis section 7 provides an in-depth examination of the findings in relation to the research question and objectives. The analysis of the implications of the results for internet transparency and connectivity in the DRC is presented. This section also includes a comparison of the results to other studies on internet access in Africa.

The conclusion section 8 provides a summary of the main findings and contributions of the study. The implications for policy and practice are discussed. The section concludes with suggestions for future research.

The references section includes a list of sources cited in the thesis. The appendices include additional data and materials that provide additional information to support the analysis and findings in the main body of the thesis.

Chapter 2

BACKGROUND AND RELATED WORK

This section provides an overview of internet measurement gloably and in Africa in particular for developing countries such as DRC. The literature review presents current research on internet transparency and connectivity in the DRC and other African regions. Existing measurement tools are also discussed according to their categories. The literature review will provide the foundation for the research by identifying gaps in the existing literature and setting the stage for the study.

2.1 Internet measurements

The Internet plays a vital role in today's society. Situations such as the COVID-19 pandemic have further underscored the importance of reliable internet connectivity for everyone. As deployments of mobile broadband networks increase, there is still a significant rate of connectivity gap and digital divide around the globe, especially in low-income communities which have under-provisioned networks and lack appropriate cable Internet infrastructure or redundant interconnection systems. More importantly, low-income countries generally do not have the capacity to thoroughly audit their Internet infrastructure and, in many cases, they have not developed or adopted best practices for building and monitoring a resilient Internet infrastructure. Internet measurements are a set of practices that allow network operators, researchers, and regulators to elaborate on the above problems by expressing them in numbers so that troubleshooting and performance improvements can be considered. Reasons for measuring the Internet can be technical, commercial, and social. Technically, measurement tests are conducted on the Internet infrastructure to help interested parties better understand the state of the Internet at every geographical scope, thereby giving them a clearer picture of what solution practices need to be implemented. Business-wise, users buying Internet connectivity from an ISP may need to understand the kind of service being offered. Only by measuring the actual performance, we can verify that they are getting what they pay for and whether there is equity in access to that performance between regions are cities. Without measurements, there is no objective record or benchmark of how a network behaves. Similarly, Internet measurements show whether changes improve or degrade the network's performance, and by how much. Also, network operators can rely on measurements and analysis reports of services before their deployment in cases where they are heavily dependent on Internet connectivity. From a social point of view, measurements can be relied upon in situations of crisis (such as a natural disaster), whereby investigations are made to tell the effect of a certain event on the socio-economic growth of a region or a country.

2.2 Measurement Tools and Platforms

Internet measurement platforms are tools used to assess the performance of networks and the quality of service and experience provided to users. In this research, we will be focusing on three types of Internet measurement platforms: device-based, client-based, and Web QoE platforms. Device-based platforms, such as Iperf and PerfSONAR, measure network performance by installing software on devices connected to the network. These platforms are useful for measuring the performance of specific network segments and identifying bottlenecks in the network. Client-based platforms, such as Ookla's Speedtest, measure network performance by running tests on the client side. These platforms are useful for measuring the performance of specific devices and connections, and for gathering data from a large number of users.

Web QoE platforms, such as WebPageTest, measure the quality of experience provided to users by measuring the load time and performance of web pages. These platforms are useful for identifying issues with web pages and for measuring the performance of web-based services.

We have selected these three types of platforms because they provide a comprehensive view of network performance and the quality of service and experience provided to users. By using a combination of device-based, client-based, and Web QoE platforms, we can gain a better understanding of the existing factors that inspire Internet performance researchers and developers and identify areas for improvement.

2.2.1 Client-based

Ookla Speedtest vs MLAB's Network Diagnostic Test (NDT)

Ookla Speedtest [19] and Measurement Lab's Network Diagnostic Test (NDT) [23] are two examples of client-based network measurement tools that collect various types of performance measurements, including throughput and latency measurements. However, Ookla has extended its measured metrics to cater to situations when a network is under working conditions, the so-called responsiveness [24]. Previous research has also highlighted that these two tools do not measure the same thing: Ookla is argued to measure the performance at the last mile (from end-users to their ISP) whereas MLAB's NDT measures the performance to the Internet (networks outside the users' network) where there is a presence of NDT Server. In this research, we explored the two-speed test approaches to account for both aspects of Internet performance from users' perspectives with a focus on throughput and latency metrics.

Iperf

Iperf is a command-line tool that allows for easy measurement of network throughput, delay, jitter, and packet loss, but it can only measure between two hosts on a network and doesn't provide a comprehensive view of the network performance. It supports a variety of protocols, including TCP and UDP, and can be used on various operating systems. However, it's not as user-friendly and can require a certain level of technical knowledge to operate.

PerfSONAR

PerfSONAR is a more comprehensive tool that allows for end-to-end measurement of network performance. It allows for active measurements, such as throughput and packet loss, as well as passive measurements, such as latency and jitter. However, it requires a significant amount of setup and configuration, and it can be difficult to interpret the data it produces. Additionally, PerfSONAR also has a web interface that allows users to view and analyze network performance data, but it can be complex to navigate. It also has the ability to aggregate data from multiple measurement points, providing a more comprehensive view of network performance, but it can be challenging to set up and manage multiple measurement points. It can be used to measure performance between different networks and different domains, and it supports a variety of protocols and operating systems, but it can be resource-intensive.

MobiPerf

Mobiperf is an Android-based app previously hosted but de-commissioned later by MLAB. It supports five utilities that measure the user's Web QoE such as DNSLookup, HTTP Download, Ping, and Traceroute. Other researchers have attempted to maintain its code by updating its dependencies and implementing new features such as support for other measurement scheduling algorithms. However, the App still suffers compatibility issues for different Android versions of a user's device. While we explored this tool by customizing its measures to our needs, the aforementioned issue did not serve to our advantage, hence we looked for other alternatives as described in the methodology section.

2.2.2 Device-based

RIPE Atlas

Some platforms, notably RIPE Atlas, provide on-demand measurements, public data, and "real" network vantage points [12]. RIPE Atlas is a globally distributed network of thousands of devices that enables researchers to perform basic measurements, such as ping and traceroute. While it is one of the most successful platforms for network experimentation, it does not allow for customization and its measurement sample is based on the locations of its probes, rather than being chosen specifically for research purposes. Despite this, RIPE Atlas remains a valuable resource for researchers as it allows them to conduct custom measurements, and collect and curate data for future use.

Bismark

BISmark (Broadband Internet Service Benchmark) platform [25] is a tool that allows researchers to measure various aspects of network performance, including throughput, latency, packet loss, and jitter, from a network endpoint. It can be installed on single-board computers like the Raspberry Pi or Odroid and can be used on any network, including home networks. Recently, BISmark has been enhanced with the addition of Netrics [26], which expands the range of throughput tests available, including Ookla's Speedtest, iPerf, and the Measurement Lab's Network Diagnostic Test (NDT). While Bismart has a management suite that enables the deployment and updating of software and configurations on a distributed set of measurement nodes, it can be challenging to adopt the tool in a low financial resource research context like ours.

2.2.3 Web QoE tools

WebLar [27] is a custom tool that was used by Sheferaw et al. [28] to measure end-users web QoE and network QoS on the European mobile measurement platform [29]. The tool caters to a number of QoE metrics that include TCP connect time, and Time To First Byte (TTFB), and Page Load Time, and Above The Fold time. However, it's limited to the Chrome browser only and it's not bug-free.

Asrese et al. [28] deployed the Webget tool on SamKnowns [30] probes over a period of three and a half years. The tool can record metrics such as DNS lookup time, Time To First Byte, the download speed, the number, and the size of website components more specifically Facebook, Google, and Youtube. Nevertheless, this tool is not open source, hence not feasible to reproduce despite the comprehensiveness of data the tool collected [31]

WebPerf as explored by Sheferaw et al. [32] on the MONROE platform[29], is an application that captures detailed TCP and HTTP statistics and takes into account both static and dynamic objects of a webpage, hence its usage alongside Webget above provides accurate results. However, its implementation remains internal to the platform

Netalyzr [33], an online service that tests how open and transparent a user's connection to the Internet is. Researchers used this tool to track and quickly understand emerging connectivity problems encountered by Internet users, such as website redirect links. However, since early 2019 the tool went offline and is no longer available to users and researchers.

2.3 State of Inter-connectivity

2.3.1 BGP and ASNs

"BGP" stands for "border gateway protocol". Its first specification was first published in 1989, with the purpose of connecting the last router in a network (the gateway) to the first router in the next network (neighbor) to ex- change routing information [34]. Networks that run BGP are called autonomous systems (ASes), which are routers/networks under common administrative control. These ASes are represented by a unique 32-bit number called Autonomous System Number (ASN) with the idea that each AS presents a consistent view of itself to the outside world, and what happens inside an AS is irrelevant to other ASes, as far as BGP is concerned.

2.3.2 BGP neighbor relationships

Like all routing protocols, BGP maintains relationships with neighboring routers. They can have peer, transit and customer relationships with each other[34][35]. Because BGP does not discover these routers automatically, BGP neighbor relationships must be explicitly set up on both sides through administrative configuration:the router is told the IP addresses of its neighbors along with the remote AS number and other information that's relevant to that specific neighbor relationship [35]

Peering in Africa

Peering refers to the direct exchange of internet traffic between two networks. This exchange happens at facilities called Internet Exchange Points with the aim of keeping local Internet traffic local, hence reducing the inter-connection costs (average-per-bit-delivery) and shortening the AS-PATH for Autonomous Systems owners (i.e network operators) [36] [37] [38] [39]. The presence of an IXP within a country and the number of members connected to it can be a crucial factor to assess its Internet resilience. In Africa, the state of peering between networks is a complex issue. One challenge is the lack of infrastructure to support peering. A study by Patel et al. [40] found that many African countries have limited internet infrastructure, and the networks that do exist are often disconnected from one another. This makes it difficult for networks to establish peering agreements, as they may not have the necessary connectivity to exchange traffic directly.

Another challenge is the lack of peering policies in some African countries. A report [41] by the Internet Society found that in many African countries, peering policies are either non-existent or poorly enforced, which can make it difficult for networks to establish peering arrangements. The report also noted that the lack of peering policies can lead to a lack of transparency and fairness in the peering process.

Despite these challenges, some progress has been made in recent years to improve peering in Africa. For example, the African Peering and Interconnection Forum (AfPIF) has held annual conferences and workshops to bring together internet service providers, content providers, and other stakeholders to discuss issues related to peering and interconnection in Africa. In addition, initiatives such as the South African Peering Forum (ZAPF), the West African Peering and Interconnection Forum (WAPIF), and the East African Peering and Interconnection Forum (EAPIF) have been established to promote peering within specific regions of the continent. However, there is still much work to be done to improve peering in Africa. A study by Huitema et al. [42] found that many African networks rely on transit providers to carry their traffic, rather than establishing peering agreements with other networks. This can lead to higher costs and lower quality of service for internet users in the region.

2.3.3 Available data sets

PeeringDB is a well-known publicly available user-maintained database of networks. It provides network interconnection data to interested parties globally. Its usefulness will apply not only to the understanding of the topology and traffic flow from the Internet Exchange Point's and data centers' point of view but also to that of interconnection decisions updates. PeeringDB's API [43] is leveraged for inter-IXP comparison of advertised traffic quantities, and IP prefixes' network behavior across the peers is analyzed.

CAIDA's Autonomous Systems (AS) relationship data set [44] provides detailed information on how different networks interconnect both at a global and national level. This data is relevant to both technical and economic aspects of the Internet's inter-domain structure and stability. More crucial information can be seen from this data set, such as the geographic location of links between networks.
Ookla Open Data [45] hosted on AWS made public their dataset. Through their popular SpeedTest tool, they possess data from millions of users' tests worldwide. While the data gets updated on a quarter basis, the raw results are aggregated into tiles whose size varies depending on latitude and classified based on whether the test was taken over a mobile (cellular) connection or a fixed (including Wi-Fi) connection, making it useful for a broad range of stakeholders.
A dataset [8] was developed by the Alliance for Affordable Internet (A4AI) and the Wide World Web Foundation for Low Mobile Income Countries (LMICs) to produce the aggregate upload and download speeds by country and mobile carrier. Their 2018 and 2020 reports [8][46] and cover DRC's QoS data through M-Lab.
Other aggregated host-based measurements on broadband performance for many countries in the Sub-Saharan Africa region do exist, but with the same constraints or even worse. For instance, OOKLA's 2020 net index [19] that ranks DRC 168th, or Alexa's [47] top sites by country whose ranking [48] lacks data from the DRC. In this study, we do further investigation to understand why this is happening.

2.4 Measurement Research in Africa

There have been several related works done regarding Internet performance globally, and on the African continent in particular. Chetty et al. [21] compared the broadband performance of fixed-line and mobile providers in South Africa. The authors describe how six out of seven ISPs do not provide upstream and downstream speeds they advise for and that high latency experienced by users to popular websites and services affects broadband performance. A survey by

AFRINIC [49] [13] that involved 234 participants across 34 African countries explores several tools used to perform Internet measurements of fixed-line and mobile access networks. The authors found that these tools don't perform well in Africa and other underdeveloped regions because the servers hosting them are located in Europe. Further insights from the survey revealed that there are limited network data regarding Africa's Internet operations. Formoso et al. [50] evaluated inter-country latency among African countries based on the regions in which they are located. The authors found that regions whose transit providers route traffic through Europe and America suffer from high delays even for local connections among the countries that make a region. In the same context, Chavula et al. [14] conducted active measurements to compare Internet performance between different African countries and the latency experienced in the interconnections among them. They revealed the existence of a correlation between African regions' individual latency with the language and colonial country shared among members of each region. They also found that DRC was the country with the highest latency value on the continent without however providing data-driven details on why this was the case.

Regarding AS relationships, Gilmore et al. [51] mapped the African Internet traffic at router and Autonomous System (AS) levels. The authors came up with 3D visualizations of the African Internet and smaller regions at each level. It is crucial to do the same at a more granular level in the context of DRC's networks. Alternatively, Feamster et al. [52] focused on reducing latency in popular sites through measurements and encouraged proper proxy placement to deal with site-resolving delays. They discovered the existence of many circuitous Internet paths between regions in Africa but the study only focuses on South Africa, Kenya, and Tunisia. The study highlighted poor interconnectivity between African ISPs, who are connected to European IXPs rather than doing so with regional IXPs. The same discovery is done by Bagula et al. [53] through a characterization of the level of inter-connectivity and peering among Africa's National Research and

Education Networks (NRENs) with a focus on the latency experienced between heavily intensive applications that are run in these networks. The authors also perform active Internet measurements from five African countries(Morocco, Gambia, Senegal, South Africa, and Rwanda) serving as vantage points of an Internet measurement platform called Archipelago, owned by CAIDA [54]. Such a tool is of potential use to our research for network topology discovery.

To determine the structure of networks, Chavula et al. [55] designed a network topology visualization tool for National Research and Education Networks in Africa that relies on Traceroute. Its accuracy to reveal the logical network topology of these networks is worth leveraging. While Cuvellier et al. [56] give details on anomalies that exist in Traceroute-based measurements, they proposed Paris Traceroute to picture the flow of packets across the network path more precisely. Rahman et al. [57] explored tools necessary for a mapping of the actual network topology by grouping them into two, namely, topology generations tools and network/topology discovery tools. The authors' work was expanded by Luckie et al. [58] in terms of the number of vantage points included in Traceroute's path discovery process. The authors also compared paths inferred with UDP, ICMP, and TCP-based traceroute methods to a predefined list of IP addresses, routers, and websites. They found that ICMP-based traceroute performs better in terms of destination reachability and AS link discovery.

In terms of web performance, Zaki et al. [59] carried out a two years-long analysis of the primary bottlenecks that affect Ghana's web performance from the end-users point of view, specifically the user-perceived web page load latency. They found that slow access to websites is caused by the lack of good DNS servers and caching infrastructure. While the authors suggest simple DNS and redirection caching as possible solutions, this can be limited to the type of DNS technology that is running. Hence the need to investigate users' web experience in the current context and have a comparison to other networks.

2.5 DRC Internet Ecosystem

DRC is a vast country spread over a territory of 2,345,410 kilometers squared, surrounded by 9 countries with which it shares 9,165km of borders [60]. According to a report by the World Bank in 2018 [61], DRC has a population of around 78.7 million inhabitants with less than 40% of whom live in urban areas. In the digital sector, despite many governmental plans, the country is experiencing a digital divide that accentuates intra-urban and inter-urban disparities, between cities, the countryside as well as the provinces. The government released its National Digital Strategic Plan towards 2025 [60] whose pillars include improvements for the Infrastructures, Content, Application Uses, and Governance Regulations.

Since 1996 the country enshrined the liberalization of the audiovisual sub-sector and the telecommunication sector in 2002 as formerly public monopolies. These two sectors are experiencing strong growth with the expansion of mobile telephony (2G, 3G, and 4G), Digital Terrestrial Television (DTT), and the Internet demand is still on the rise [60]. By 2050, more than 50% of households are expected to use optical fiber, and over 90% to connect to the Internet through their mobile phones[62][60]. For instance, the country's National Strategic Plan for Development has planned to put the first Congolese satellite into orbit, the complete of 5,000 km of national fiber optic backbone and connection of 30,000,000 fixed and mobile lines to metropolitan networks.

In 2016 the country set up two functional IXPs: Kinshasa Internet Exchange Point (KINIX) and Lubumbashi Internet Exchange Point (LUBIX) through a project initiated by an association of ISPs and has managed to reach a peak capacity of 14Gbps from up 15 peering networks [7][63][60]. A report [7] by the Internet Society on Africa's networks inter-connectivity revealed that Content Delivery Networks such as Facebook have caches in DRC at the Kinshasa Exchange Point and that the content generated in the country is relatively little.

This shows the presence of a thriving technical community through local community solidarity and participation in offline and online technical meetings in Africa.

Today, the deployment of fiber optic infrastructure in the DRC is done by three operators, namely: SCPT, SNEL, and Liquid Intelligent Technologies. SCPT covers 3,600 km of national fiber optic backbone sections with an objective of 30,000 km and has two urban fiber optic rings(10 Gbps and 40 Gbps) in Kinshasa, the capital city[60]. SNEL, on the other hand, covers 2,200 km of fiber optic line, mounted on high voltage electricity transmission pylons, whereas Liquid Intelligent Technologies covers over 300 km of optical fiber from the Zambian border to Lualaba province and an urban ring in Lubumbashi(South-East) as well as in Goma (East) planned to be marketed in Open mode Access open to all operators or ISPs [64][60]. The same company also has an ongoing project to deploy an FTTx infrastructure in Kinshasa from the urban Ring of one of the mobile phone operators. Fiber to the "x" (FTTx) is a collective term used to describe a wide range of broadband network architecture options utilizing optical fiber for some or all of their last-mile connectivity. With "x" representing the fiber termination point, FTTx technology encompasses optical fiber deployments such as FTTH, FTTA, FTTB, and FTTC [65].

As for mobile telephony services, they are offered by operators: Airtel [66], Vodacom[67], Orange DRC (Tigo)[68], Africell[69], and Standard Telecom. The first three support national coverage while the fourth covers the capital city and major cities. The list covers a few districts in the capital city only. Most of these active mobile telephony operators belong to large international groups, except Standard Telecom which is the result of a foreign partnership with the public operator SCPT [60]. Internet service providers on the other hand are eight in number, which includes Microcom, Global Broadband Solutions, OrionCom, DHI, Raganet, Afrinet, PIWI, and Standard Telecom [70]. The state of interconnection and performance of these networks are covered later in this research.

When it comes to internet measurements, most of the aforementioned network operators and their users do not conduct measurements at a comprehensive level and do not host internet measurement platforms, such as Ripe Atlas probes or SpeedTest, in their networks for diagnosis or to provide internet speed awareness to users. As a result, there is a lack of data on the performance of the Democratic Republic of the Congo's (DRC) networks in public datasets like MLAB, Ookla Speedtest OpenData, and SpeedChecker. This lack of data is evident when comparing the wide deployment of Ripe Atlas probes in countries like South Africa (82 probes) and Kenya (17 probes) to the single probe at KINIX in DRC. Similarly, the presence of M-LAB servers and Ookla's vantage points in a country can also influence the ability to conduct internet research and studies on the performance of networks. This limitation on internet measurement data limits researchers' ability to study the resilience of the internet in the DRC and to improve the performance of networks.

However, the broadness of the Internet and the interconnectivity between networks still provide a way to study a country's Internet structure by looking at how it connects to the world. Platforms such as PeeringDB, Routeviews, RipeStat, Packet Clearing House, and CAIDA have capabilities to inform researchers on the overall structure, capacity, and scalability of the Internet at a country, and at more granular levels of the network. We cover more on these resources in section 2.3.3 and apply them in the DRC's context as part of our methodology in Chapter 3.

Chapter 3

METHODOLOGY

This Chapter describes the research design and data collection methods used in the study. The research design is a mixed-methods approach, using both qualitative and quantitative methods. Data collection methods include secondary data analysis, an online survey, and primary data collection through Internet measurement on Congolese users' personal devices to assess their Web QoE and the QoS of networks they are subscribed to. The section also includes information on the participants and sampling techniques, as well as the data analysis techniques used. Overall, this section provides the details of how the research was conducted, including the methods and techniques used to collect and analyze data.

3.1 Measurements workflow

This section presents the measurement workflow of our study which covers the process involved in collecting data for the research. A measurement workflow is a crucial component of any research study, as it provides a structured approach to collecting and analyzing data to ensure consistency and accuracy, reproducibility of the study by other researchers, efficient data processing, and the validation of results. Figure 3.1 shows a summary of the workflow involved in the data collection stage of this study. The lower left and right sides of the figure show iNethiPerf and Ookla Speedtest app instances triggered by users who primarily install the App on their devices. Furthermore, the online services seen on top of the figure serve as secondary datasets for this study. These three data sources including the online survey serve as input to the data pipeline upon which data

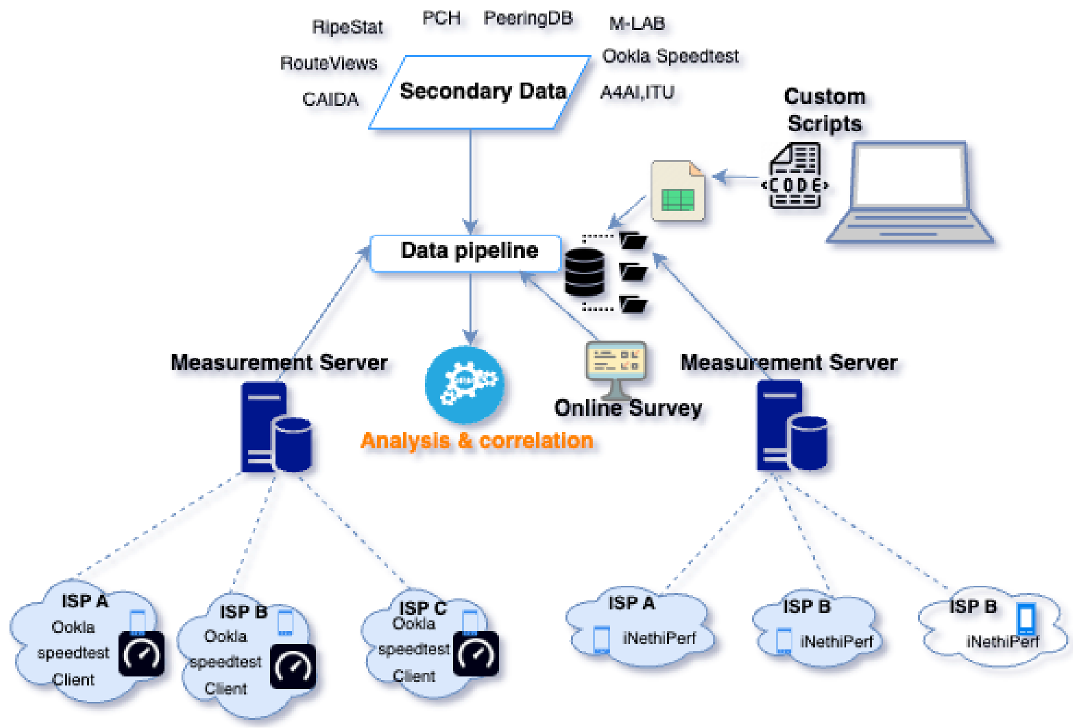

Figure 3.1: Data Collection Workflow of this study

cleaning, visualization, analysis, and correlation are done in order to present the results of our study.

3.2 User perception survey

Our first research question seeks to determine the Congolese users' perception and level of satisfaction with their experienced internet performance. To answer this question, we conducted an online survey that runs across DRC's three major network operators (Vodacom, Airtel, and Orange). The selection of networks was based on a network's customer population [70][71] where each network holds a market share value greater than 20%.

3.2.1 Survey questionnaire

Questions we asked in the survey include but are not limited to: the name and type of provider users are subscribed to (fixed or mobile), users' subscribed and experienced bandwidth, how often they face Internet disruption and delays, mobile application usage, and how each performs, their awareness of Internet measurements and their interest in Internet Speed measurements. Data collected from this survey provided qualitative and quantitative data on the users' perceived QoE in the different networks. Thereafter, the analysis of the data gathered served as input to answer the second question of our research.

3.2.2 Sampling and Participant Recruitment

Survey respondents are part of key contributors to the success of this research. Hence, a well-defined plan for recruitment was crucial. The total population size was composed of subscribers from three cellular networks (Vodacom DRC, Airtel, and Orange) who reside in cities where there is an IXP presence; namely Kinshasa, Lubumbashi, and Goma. These cities are considered clusters. From each cluster, we randomly selected 3 districts and from each district, we randomly selected 33 participants subscribed to at least one of the three mobile networks to run the measurements.

We relied on social media posts, email recruitment, and most importantly our local contacts composed of both technical and non-technical individuals. More precisely, our local contacts included 200 individuals distributed around the 11 provinces, with whom we have engaged in previous projects (for instance, Internet society DRC Chapter members) some directly and others indirectly.

3.3 Passive datasets

Passive datasets are an essential source of internet measurement data that provides insights into the behavior, business relationships, and interconnection of Autonomous Systems (ASNs) on the internet. In this research, passive datasets were used to filter and analyze the internet measurement data of ASNs in the Democratic Republic of Congo (DRC). The sources of the passive datasets include CAIDA ASNrank API, PeeringDB, RipeStats, BGPview, ROuteviews, RIS, and more. These resources provided valuable information on the network configurations, peering relationships, and routing policies of ASNs, which are crucial factors in determining the structure of the internet in the DRC. Using these passive datasets, a more comprehensive and accurate analysis of the internet measurement data collected for this research was possible.

3.3.1 Datasets and APIs

In the quest to answer the second question of our research, *"To what extent are Autonomous systems in the DRC interconnected, and how does this affect internet performance?"*[1]. The selection of the dataset we used was based on its popularity in previous research. In situations where a dataset was not available for download, we leveraged REST and GraphQL API calls to the data in case such a feature was made available[5].

3.3.2 Data Extraction

The dataset we rely upon to get the answers to our second question is produced by the Pipe-and-Filter pattern shown in Figure 3.2, based on the datasets and APIs mentioned in Section 3.3.1 above. The Pipe-and-Filter architecture is a pattern that has independent entities called filters (components) that perform transformations on data and process the input they receive, and pipes, which serve as connectors for the stream of data being transformed, each connected to

[1]We explored different BGP data-based datasets available online. This data is collected by a combination of various organizations or projects and tools such as Routeviews [72][2], RIPE NCC [73][3], CAIDA Archipelago tool [74][4], and others.

[5]These APIs include RIPEstat[6], CAIDA ASNRank v2.1[7], BGPview[8], PeeringDB[9], ARPTC[10], and Apnic Stats[11]. RIPEstat's API provides the necessary information to identify Congolese routed ASNs as seen by the Internet Routing System (RIS), whereas CAIDA ASNRank v2.1 told us valuable information such as the customer cone of ASNs (a set of IP addresses that are reachable through a specific ASN and are considered to be under its control), their rank, the number of announced prefixes, neighbors, and more details regarding the organizations to which they belong. BGPView API provided existing business relationships (upstream, downstream, and transit) information among the queried ASNs. PeeringDB, on the other hand, provides peering information on the individual networks and the IXPs they each are connected to. ARPTC and APNIC both revealed the customer population and Internet market share of Congolese networks, the basis on which we selected what networks to focus on in the analysis in subsequent chapters.

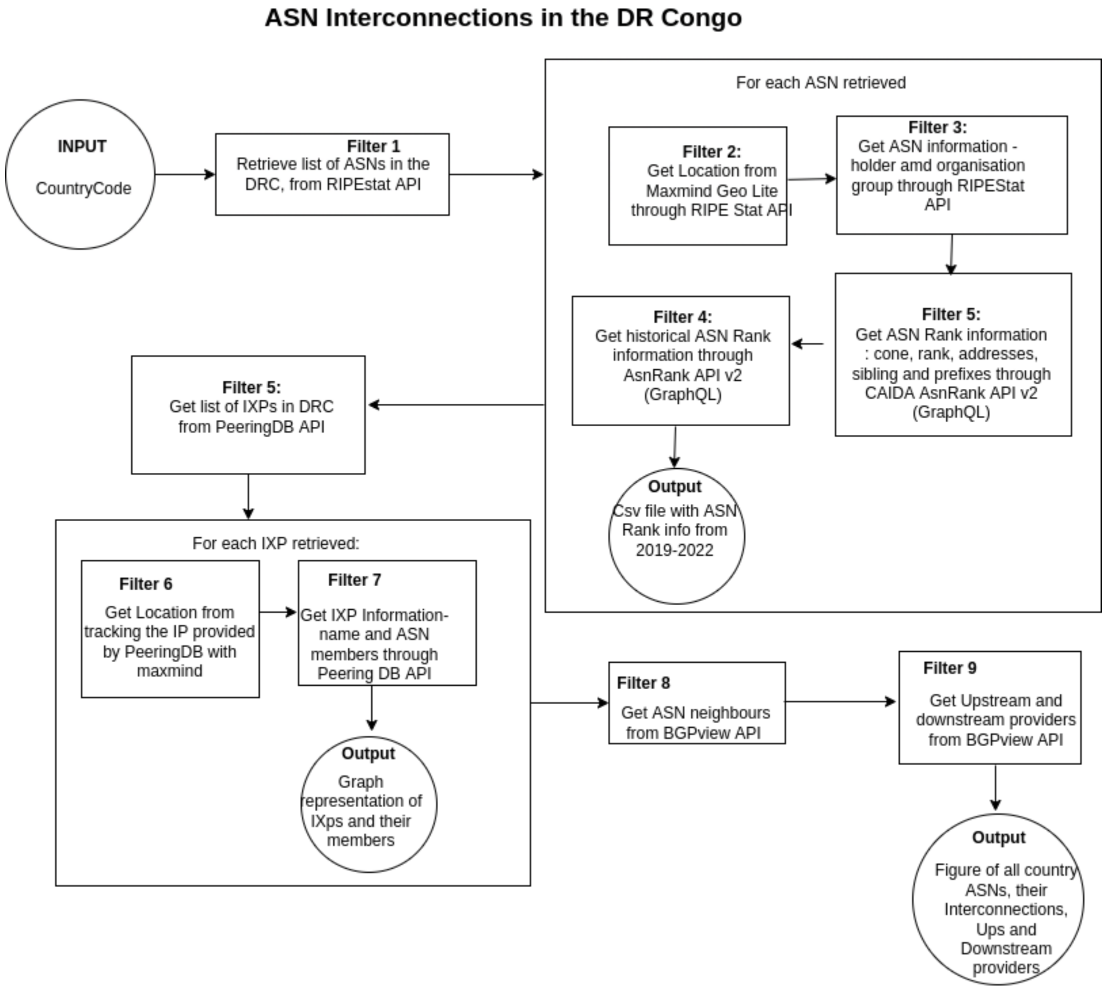

Figure 3.2: Data extraction architecture.

the next component in the pipeline [75].

Therefore, unlike a traditional Object oriented-based Pipe and Filter architecture, the filters shown in Figure 3.2 are not separate classes or components but rather separate functions in our Python script. Functions are executed after one another with the results from previous functions in the sequence used in subsequent ones to add to and transform the data stored. Python was used to write the data processing script for speed, simplicity, and visualizations libraries the language supports. While limited input data is passed at processing time, through various API calls, and data transformations we produced separate data objects on which visualizations are made according to the targeted insight we want to highlight.

List of ASes and their Organisation

We constructed minor datasets that helped in drawing some insights into the DRC's network's interconnection. From RipeStat API in Figure 3.2's filter 1, we identified 44 Congolese ASNs after filtering the data by country code ('CD' in this case). For each of the ASNs found, we performed four different filters: Firstly, we collected geographical locations for each using the GeoLite Maxmind database, through the same RipeStat API. While we did not use all this information in our subsequent results, it can be crucial for the extension of this research in future work, especially in terms of the categorization of networks per region or city, as a way to present the Internet market concentration of networks at a granular level. Secondly, we extracted organization information for the returned list of ASes and their organization holders. In knowing this, we could elaborate networks with potential sibling-to-sibling relationships which in turn would highlight the direction of the routing traffic within and outside the country.

CAIDA ASNRank

Using CAIDA's AsnRank V 2.1 API, we constructed a CSV file containing information on all DRC's ASN ranks, their number of announced prefixes and addresses, their customer cone, neighbors, and extra information regarding the organizations they belong to. The data constructed during this step was then analyzed as a way to make comparisons between the top 20 ASNs in terms of their difference in the number of announced prefixes they have, the customer cone, and also the number of IPv4 addresses they possess.

PeeringDB API

In filter 2 we generated a JSON file containing information on the presence of IXPs in DRC and identifying the number of members they have connected. Also, we put a focus on the IXPs members that are eyeball networks, and what the state of their interconnectivity is with each other. Eyeball networks refer to networks that deliver Internet access to users' homes and businesses, and are designed to handle high volumes of Internet traffic. To achieve that, we queried the PeeringDB API to identify that information, which resulted in the identification of ASNs that do not leverage the IXP Internet infrastructure available to them.

BGPView, CAIDA AS-relationship, RIPEStat ASN neighbors

Having known the state of the interconnections of DRC ASes, we looked at their relationships, and neighbors within the BGP Path hence inferring the routing structure of the country's Internet. First, we present who is the Upstream and Downstream ASes, or else which pair/pairs of ASNs have peering and those that have transit relationships with each other. Second, we look at both the AS peers/neighbors of each AS as seen by the Routing Information System [73] using RIPEstat API, BGPView API, and CAIDA's AS-relationship data sets. The BGPView API returns the list of an ASN's peers in terms of upstream and downstream. For each API call's response, a dictionary object is updated to categorize the data queried by its particular ASN. This new object is therefore passed to the second class that leverages the power of the Networkx library [76] to generate a network graph between ASNs that happen to have a relationship with each other.

The BGPView API had a limitation in terms of the number of requests we were allowed to make. Hence we considered a pre-constructed dataset that we could download locally and do the analysis accordingly. The CAIDA AS-relationship dataset catered for this. We specifically chose to use the Serial-2 directory of this dataset because it combines both the data from the 'Serial-1' category with that from the multi-lateral peering methodology [77]. This directory is composed of multiple bz2-compressed files updated on a monthly basis. Our analysis is based on data collected in February 2022. Given the list of DRC ASNs returned by our Pipe-and-Filter architecture, we filtered out ASNs that are not part of it so that our results are representative of DRC networks. Next, we associated each ASN to its corresponding AS name using the AS-Org dataset, CAIDA's data containing further details of all ASes, including the organization ID, name, and location. From this same dataset, we were able to elaborate not only on what are the AS links available but also on statistical measures of the AS relationships for DRC's networks in terms of the number of customers, providers, or transits each has, and their distribution across the entire dataset.

Another resource we explored is the ASN neighbors' information provided through RIPEStat API. ASN Neighbours offers similar functionality to ASInUse which is part of the Routing Information Service (RIS). However, while ASInUse includes information that is only found by route collecting services (such as RIS), RIPE-stat's ASN Neighbours provides a view on which ASNs are adjacent to a queried ASN in the routing tables, then identifies them as either left or right neighbors depending on their position in the BGP path. This means that an ASN which is right of the queried ASN is not necessarily the true originator of the prefix, only that it is closer to the origin than the queried ASN. An ASN with only left neighbors can be seen as an originating ASN.

3.3.3 AS Hegemony

We rely on the local AS hegemony to present the most influential networks in DRC combined with their connections to local and foreign ASes. We obtain AS hegemony scores for every Congolese AS using Internet Health Report (IHR) API [78]. AS hegemony [79] is a metric that quantifies the likelihood of an AS to lie on paths toward certain destination IP prefixes. There exist two variants of this metric, global and local AS hegemony. The global AS hegemony is computed with paths to all IP prefixes globally reported by the BGP viewpoints. ASes with a large value stand for large transit networks that are commonly used to reach any host on the Internet. For instance, tier-1 ASes, like Level 3 (AS3356), have the highest scores, and stub ASes have the lowest scores. The local AS hegemony is computed with paths from all BGP viewpoints towards only one origin AS. In this case, high values stand for ASes that are commonly used to reach the given origin AS. For instance, computing the local AS hegemony for Orange DRC (AS37447) reveals that the highest score is attributed to CELTEL DRC (AS37020), which is a main upstream provider to many networks in DRC as we shall see in the results section later on.

3.4 Active Measurements

Our third research question seeks to measure the Internet performance that DR Congo Internet users experience in both mobile and desktop environments. We run active measurements on end-user devices to assess their performance in terms of quality of service, quality of experience, and network traffic destination. These metrics are to be compared to the perceptions that participants previously stated early in the survey.

3.4.1 Metrics

Every measurement study has a set of metrics considered to produce a dataset based on which some assessment can be conducted. In our case, we have two categories of metrics. The first takes into account networks' QoS whereas the second measures users' Web QoE when accessing a predefined list of trending websites. Semrush methodology [80] helped us identify website trends in the DRC, from which we selected the top 20 to conduct our Web QoE measurements. Semrush provides a dataset composed of country-level based statistics of the most accessed websites with their respective number of visits per month, filterable by industry and containing multiple fields such as the main traffic source, desktop and mobile share in percentage, month-over-month and Year-Over-Year change in traffic as a percentage of the previous month/year's traffic stats.

Quality of service metrics include Download and Upload throughput and Latency to Ookla and MLAB servers. The Web QoE metrics measured are extracted through the execution of a Python script that makes API calls to Curl's lib curl handle [81] for each website in our selected list of the most accessed sites by Congolese users (using Semrush methodology mentioned above). These metrics include Domain Name Start Time (DNST), Time to First Byte (TTFB), Transport Control Protocol Time (TCPT), Total time, and Redirect count. DNST provides the total time in seconds from the start until the name resolution is completed. TTFB is the time, in seconds, it took from the start until the web request is just about to begin. As explained on Libcurl's options list [82] TCPT is the total time in seconds from the start until the connection to the remote host (or proxy) is completed. Total Time is the total time in seconds for the previous transfer, including name resolving, and TCP connections. Finally, the redirect count is the total number of redirections that were actually followed before the webpage could be resolved. Furthermore, we perform an analysis of network traffic paths to infer how independent and interdependent Congolese networks are by running a traceroute to a predefined list of IP destinations from the top 10 Congolese ASNs as listed by statistics by APNIC [70]. Table 3.1 summarises the metrics, their description, rationale, and the tools that have them implemented. We take into account application scalability in the choice of tools for active measurements, so that representative vantage points are explored with ease of access to the results of the conducted measurements.

3.4.2 Sampling and Participant Recruitment

Similar to Section 3.2, we relied on participants to run active measurements on their mobile devices (phone and laptop). These participants are those who were asked in the survey about their interest in participating in the subsequent experiment of the study. The geographical scope of our participants is such that they should reside in Congolese cities where an Internet Exchange Point is present. This is because IXP facilities in DRC are strategically distributed across three of its biggest cities, namely Kinshasa in the West, Lubumbashi in the South East,

QoS - QoE Metrics and tools			
Metric	Description	Rationale	Data Collection Tool
TCP (Upload/Download throughput, latency)	QoS metrics	widely used (Mobile and desktop) open-source and ease result sharing	Ookla speedtest
TCP (Upload/Download throughput, latency)	connection speeds to foreign networks	crowd-sourced data, open-source and widely used	MLAB's NDT
TTFB, DNST, TCPT and Total Time time	Web QoE metrics	Web page load times and redirects	Python script with PyCURL library
Traceroute, RTT	Network traffic destination, diagnosis	trace packets to multiple destinations	multiTraceroute

Table 3.1: QoS and QoE data collection tools used

and Goma in the East. However, we recorded participants from other cities too. We recruited a minimum of 33 participants from each mobile subscriber located in but not limited to the aforementioned cities. These cities are considered clusters. From each cluster, we randomly selected 33 participants subscribed to at least one of the three mobile networks to run the measurements.

3.4.3 Period of Measurements

The active measurements comprised four groups that we ran consecutively depending on the availability of participants: Ookla Speedtest, MLAB's NDT, Web QoE, and traffic destination custom scripts. The first group runs for two months, from July 1st to September 1st, 2022. The second group is composed of crowd-sourced NDT speed tests and covers the period from 1st September to 15 th October 2022. However, because MLAB allows researchers to query historical Internet performance data as well, we included throughput and latency data from 1st January to 1st October 2022. The third and fourth groups which required participants with a technical background (Python, Basic Linux commands, and fundamentals of computer networks) were conducted simultaneously from 15th

October to 15th November 2022.

Chapter 4

USER SURVEY RESULTS

The online survey [83] was conducted to gain insight into users' perceptions of Internet performance in the Democratic Republic of Congo (DRC) and their level of satisfaction with the services provided by their Internet Service Provider (ISP). The survey was designed to gather information on a range of topics, including participants' demographics, Internet usage habits, and their perceptions of network performance and reliability.

In this chapter, we present the results of the online survey, including a detailed analysis of the responses received and the key findings that emerged from the data. The chapter is organized into several sections, each of which covers a different aspect of the survey results. Section 4.2 provides an overview of the survey participants, including their demographics and Internet usage habits. We then present the results of the survey in relation to the key research questions, including participants' perceptions of Internet performance in the DRC and their level of satisfaction with the services provided by their ISP. Finally, in Section 4.3 we discuss the implications of the survey results for the broader research on Internet performance and connectivity in the DRC and other developing Central African countries.

4.1 Dataset

In order to understand Congolese users' perception of the Internet performance they experience, we constructed a dataset from the responses received from the web-based survey over a period of 2 months, namely from 1st September to

30th October 2022 with a target on adult Internet users residing in the DRC. The respondents were required to complete the 20 questions for their responses to be considered in our analysis. The majority of the questions were multiple choice, with a minority of short answer and Likert scale categories. From the total of 20 questions, we subdivided them into four categories. The first covered the demographics of the respondents with details such as age, gender, town, occupation, and the province from which they are participating. The second category covered the device specifications and the applications they commonly use to access the Internet. The third category was composed of a set of questions on the connection details whereas the fourth and last category covered the level of awareness of Internet measurement and its practices. The complete questionnaire can be found in the Appendix Section 8.3. As a result, we ended up with 90 fully responded.

4.2 User perceptions

4.2.1 Characteristics of Respondents

As mentioned in Section 3.2, the 90 respondents were randomly identified from all provinces in DRC. An attempt was made to include respondents from all occupation categories. In Figures 4.1 and 4.2 it can be seen that:

• Age: The majority of the respondents surveyed (40.2% of the 90 total) belonged to the age bracket of 26-35 years. Out of the remaining, 39% belonged to the age group 18-25, approximately 11% belonged to the group 36-45 and 7% belonged to the group 45-55 whereas 3% were of 56 years or above.

• Gender: in terms of gender, most of the respondents are gendered men (76.7%), and the remaining 22.2% of women. Regarding the occupation of the respondents in Figure 4.3, students and employed categories represented the majority (31.2% each), followed by self-employed (27.1%). Unemployed respondents on the other end were the minority (10.4%).

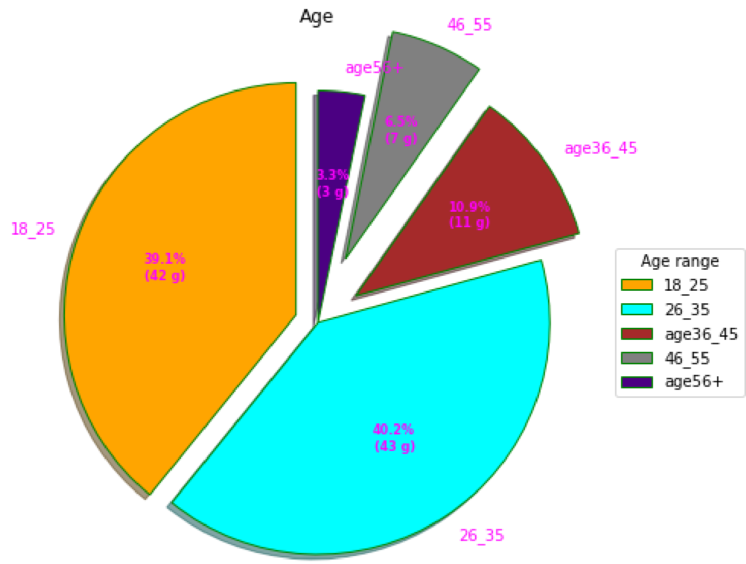

Figure 4.1: Characteristics of respondents: age

4.2.2 Service Providers and users satisfaction

Once the demographics were captured, the respondents were asked which service provider they used for mobile internet services. Multiple service providers provide mobile internet connectivity to participants. Table III provides a snapshot of the data obtained from the respondents

Analysis of respondents

1)The service provider: It was seen that the majority of respondents (75%) subscribed to Airtel for their mobile internet services. The next popular mobile internet service provider was Orange (50%) followed by Vodacom (32%) and last but not least Africell (0.6%). Given that our questionnaire had provisioned for survey respondents to specify multiple networks they are subscribed to by a multiple select option, it is pertinent to mention that we noticed 37 participants (42%) who said they are subscribed to two networks at the same time, hence the noticeable overlap in the distributions of respondents per network. We also notice

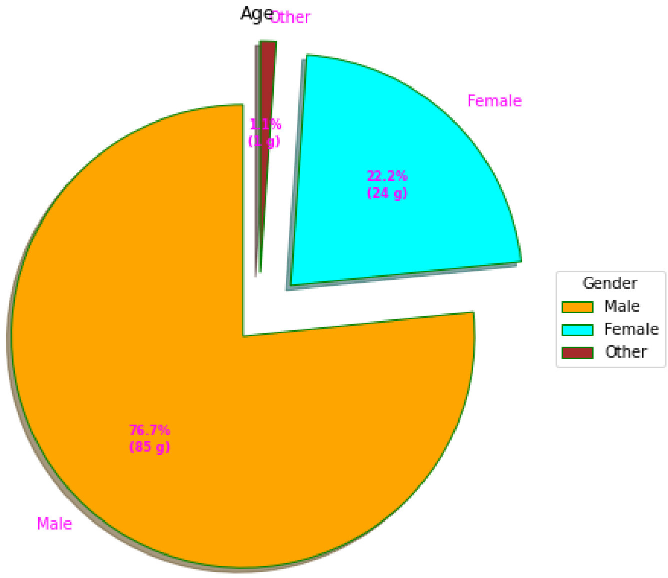

Figure 4.2: Characteristics of respondents: gender

Table 2: THE SERVICE PROVIDER THAT USED FOR MOBILE INTERNET	
Parameter	Respondents
The service provider used for mobile Internet	
Airtel	75%
Vodacom	32%
Orange	50%
Africell	0.6%
Level of Satisfaction with QoS	
Very satisfied	11.5%
Satisfied	30.6%
Neutral	25%
Dissatisfied	32.9%
Level of Satisfaction with Internet Performance	
Very satisfied	3.6%
Satisfied	20.2%
Neutral	18.5%
Dissatisfied	24.8%
Very Dissatisfied	32.5%

Table 4.1: Analysis involved in the experiments of this study

that Africell shows the lowest number of participants. This may be accredited to the fact that the network's coverage is limited to a single city in DRC, which is Kinshasa. Airtel, Orange, and Vodacom being leaders in DRC's cellular industry,

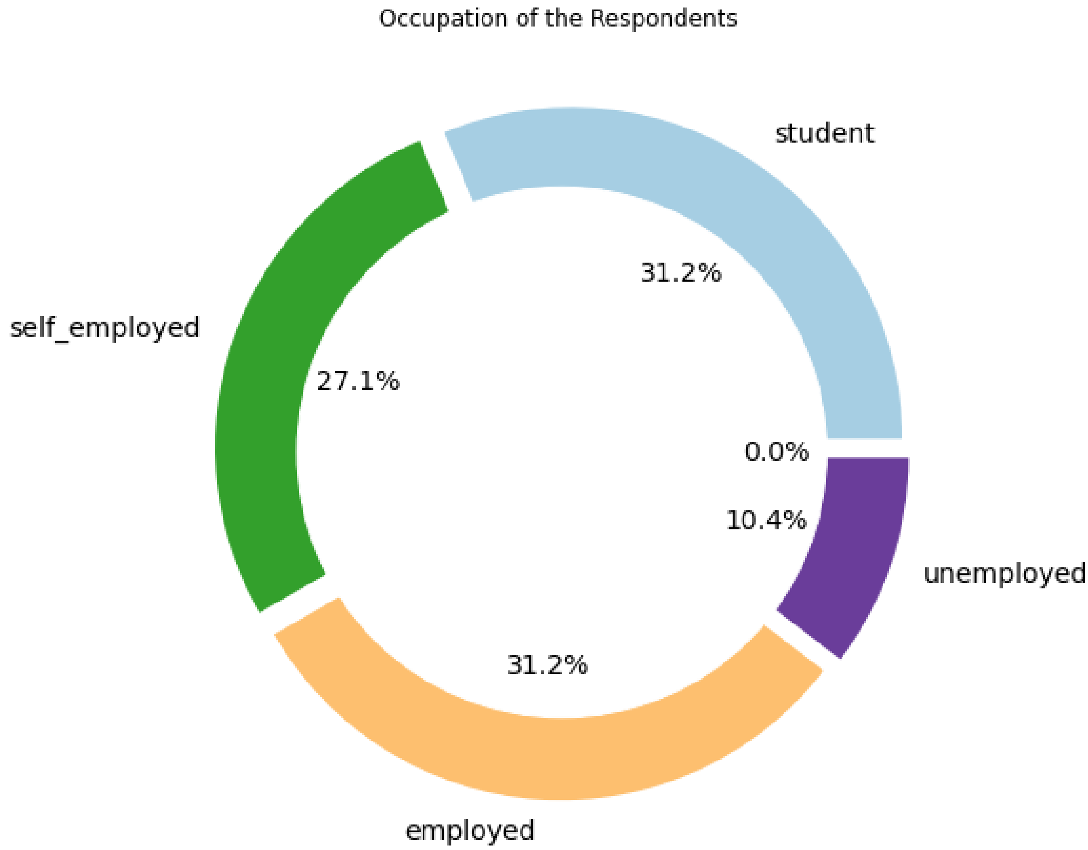

Figure 4.3: Characteristics of respondents: Occupation

provide more coverage across the region as compared to others, which contributes to the higher participation rate in our survey.

4.2.3 Awareness of bandwidth: Received vs ISP-promised

Perceptions on users awareness of subscribed mobile internet plans in terms of use are presented in Table 4.2.

Analysis of respondents

Most of the respondents (Very concerned: 64.7%, concerned: 26.4%) reported that they were concerned about being aware of the bandwidth their network

Table 2 : INTERNET SPEED AWARENESS	
Parameter	Respondents
Concern about the awareness of provided bandwidth	
Very concerned	64.7%
Concerned	26.1%
Not concerned	8.8%
I don't care	0.02%
Promised bandwidth	
1 to 3 Mbps	25%
5 to 10 Mbps	28%
Less than 25Mbps	14.7%
More than 1Gbps	0.09%
I don't know	35.2%
Perceived bandwidth	
1 to 3 Mbps	35.2%
5 to 10 Mbps	23.8%
Less than 25Mbps	0.07%
1Gbps or more	0.03%
I don't know	38.6%

Table 4.2: Internet Speed Awareness

provider serves them. The features that were inquired about were the knowledge of the promised internet speed and the actual internet speed they perceive as their day-to-day experience. For the promised speeds, common responses were; the majority (35%) said they didn't know whereas 25% and 28% said they were offered 1 to 3 Megabits per (Mbps) second and 5 to 10 Mbps respectively. Also, a small portion (0.09%) of the respondents said they were promised 1Gbps. Regarding the inquiry about the users' perceived bandwidth speed they actually experience, it was found once again that most of the respondents (38.6%) said they did not know, (35.2%) and (23.8%) perceived a 1 to 3 Mbps, and 5 to 10 Mbps bandwidth respectively. The remaining respondents said they perceived speeds less than 25Mbps and 1Gbps or more. This Internet speed perception information did not include the data incurred from accessing the internet through other services such as Wi-Fi or Fiber. It may, therefore, be concluded that most users are unaware of Internet speeds promised by their providers and do not know the speeds they experience.

4.2.4 Perceptions of Internet Performance: A Comparative Analysis Across Airtel, Orange, and Vodacom Networks

For comparison reasons, in this section we provide a comprehensive insight into the perceptions of internet users regarding their experienced internet performance across three major networks: Airtel, Orange, and Vodacom. Out of the 88 total survey responses gathered in this study, enabling a robust analysis of user satisfaction levels in terms of percentages. The findings reveal that user satisfaction varies significantly among these networks.

When examining the data on user satisfaction, it is evident that Airtel has the highest percentage of satisfied users, with 22.73% of respondents expressing satisfaction. Orange follows closely with 15.91%, while Vodacom registers at 11.36%. Notably, the category labeled "Others" had the lowest satisfaction level, with only 3.41% of respondents expressing satisfaction. When considering the "very satisfied" category, Airtel once again outperforms the others, with 28.41% of respondents expressing very high satisfaction. This highlights a notable preference for Airtel's internet services among the survey participants.

However, it is crucial to acknowledge that user dissatisfaction exists within these networks. A total of 6.82% of respondents each expressed dissatisfaction with Airtel and Orange, while Vodacom had 2.27% of users dissatisfied. This underscores the presence of areas where improvements may be needed across all three networks. Furthermore, in the "very dissatisfied" category, 6.82% of Airtel users, 4.55% of Orange users, and 4.55% of Vodacom users had extremely negative experiences, highlighting areas of concern that warrant immediate attention. In conclusion, these findings underscore the importance of continued efforts by these network providers to enhance user satisfaction and address areas of dissatisfaction, ultimately striving for improved internet performance.

4.2.5 Relevance of QoE Metrics factors

We can users' opinions on which metrics have the most influence on their perceived quality of experience on the Internet in 4.3. On the aspect of Web QoE,

Table 2 : IMPORTANCE OF QoE METRICS	
Parameter	Respondents
Page Load Time	
Extremely important	47.73%
Important	46.5%%
Not important	5.68%
Number of freezes during video calls	
Extremely important	36%%
Important	51.14%%
Not important	12.5%
Number of voice call dropouts	
Extremely important	45.45%
Important	50%%
Not important	4.55%
Text message Delays (Send/Receive)	
Extremely important	53.41%
Important	39.77%%
Not important	6.82%

Table 4.3: Importance of some QoE metrics

the majority (47,7%) responded that it is extremely important for them to access a web page fast enough, whereas 46.5% responded it is important. The remaining 12.5% claimed that web page load time's influence on their QoE is not important. Regarding video calls, the number of freezes had an extremely important influence on 36.36% of the respondents, preceded by 51.4% of respondents who considered this metric's influence just important. Only 12.5% of the respondents considered video call freezes not an important factor to the QoE. Voice calls, on the other hand, 45.45% of the respondents found the number of dropouts extremely influential to their QoE, 50% found it influential whereas 4.5% didn't consider it an important factor. Furthermore, the delays during text messaging were perceived to influence 53.4% of the respondents at an extreme level, 37.7% at a normal level, and 6.82 said they aren't influenced by this metric at all.

4.3 Conclusions

One of the key findings from our survey is the discrepancy between users' awareness of the internet speeds promised by their ISP and the actual performance they experienced after conducting measurements. Many users reported being unaware of the specific internet speeds promised by their ISP, or they had difficulty understanding the technical jargon used in the fine print of their service contract. As a result, they were unable to compare their measured internet speeds to the promised speed accurately. Second, the lack of awareness can also make it difficult for users to accurately compare the performance of different ISPs. If users are not aware of the promised speeds for each ISP, they may not be able to accurately compare the performance of one ISP to another. This can make it difficult for users to make informed decisions about which ISP to choose, and it can also make it difficult for ISPs to compete on the basis of performance.

Third, the lack of awareness also negatively impacts users' overall perceived quality of experience with the internet. Often, users who are unaware of the promised speeds do not realize when their internet service is under-performing. As a result, they do not take action to troubleshoot or resolve the issue, and they continue to experience poor performance without realizing that it is not normal.

This lack of awareness can have several negative consequences. First, it can make it difficult for users to assess whether they are getting the value for money they expect from their internet service. If they are unaware of the promised speeds, they may not be able to tell if their internet service is delivering the performance they are paying for. This can lead to dissatisfaction with the service, and potentially even to users switching to a different ISP. Hence more awareness campaign about internet measurement is needed to cover this gap.

Chapter 5

INTERCONNECTION ANALYSIS

In this chapter, we present the results of the secondary dataset analysis, including a detailed analysis of the data and the key findings that emerged. The chapter is organized into several sections, each of which covers a different aspect of the analysis. Section 5.1.1 gives an overview of the Internet structure of the DRC, including the market structure and the major players in the market. Sections 5.1.2, 5.1.4 and 5.1.3 present the results of the analysis in relation to the key research questions, including the interconnections of networks in the DRC and the influence of different networks on users' Internet connectivity. Finally, Section 5.1.5 discusses the implications of the analysis for the broader research on Internet performance and connectivity in the DRC.

5.1 Results & Discussion

5.1.1 AS Market structure

In our initial steps, we identified 44 DRC ASNs visible on the global routing system. Their distribution by region or city is unevenly distributed, which was expected due to the divide in digital access and population size disparities between big cities versus small ones. We found the majority of ASNs have their services in the capital city Kinshasa, while the rest of the cities across the country have a lower number, half and sometimes a third of those in Kinshasa. CAIDA's ASNRank API v2.1 [84] provides metrics that make the comparison of these ASNs less difficult. The metrics include AS rank (based on the number of its direct and indirect customers), the number of customers, providers, siblings, announced

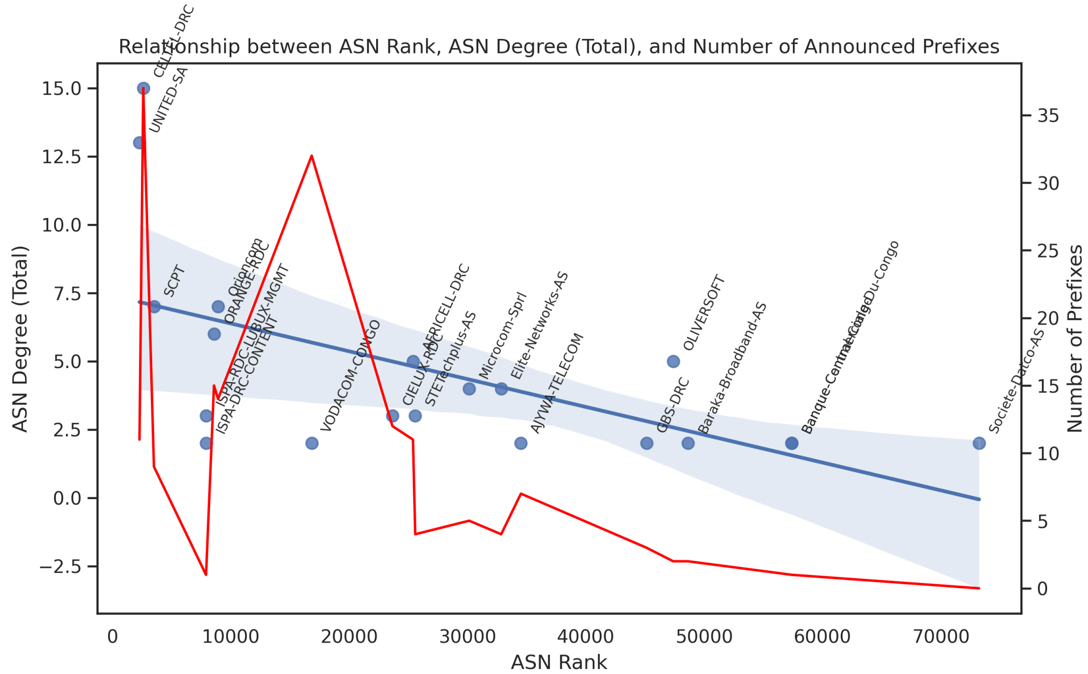

Figure 5.1: Regression analysis of the relationship between AS rank, AS degree, and number of announcing prefixes of the Top 20 DRC CAIDA AS ranking

prefixes, and more. CAIDA ASrank (Autonomous System Rank) is a method for ranking the relative importance of Autonomous Systems (ASes) on the Internet based on their interconnections and traffic exchange. It is a measure of the centrality of an AS in the Internet topology and can be used to understand the roles and relationships of different ASes on the Internet. AS Degree on another end is a metric calculated as part of CAIDA's AS Ranking project [44] to represent the number of unique prefixes an Autonomous System announces to the Internet. This metric is a way of quantifying the interconnectivity of an AS and is used to understand the relative importance and size of different ASes in the Internet.

Figure 5.1 is a regression plot that shows the relationship between CAIDA's AS rank, AS degree, and the number of announcing prefixes of the top 20 DRC ASNs. The scatter plot reveals a strong positive relationship between the AS rank and the AS degree, suggesting that higher-ranked ASNs also have a higher degree. The regression line indicates that there is a linear relationship between the AS rank and the AS degree. Additionally, the plot showed a positive relationship between the number of announcing prefixes and the AS degree, indicating that

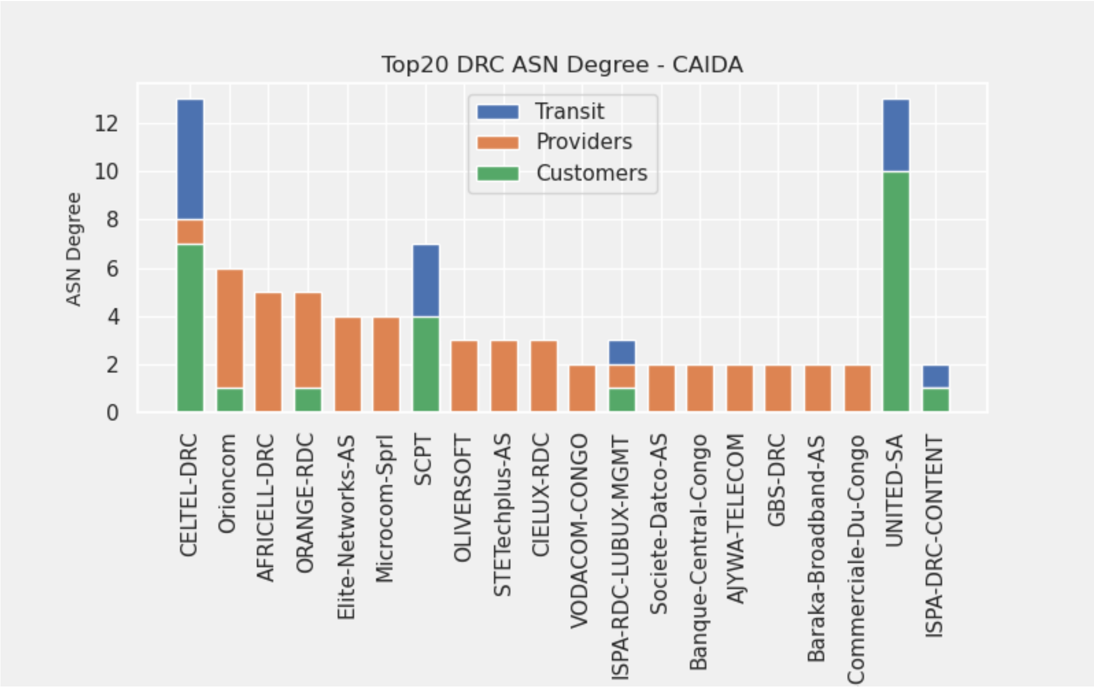

Figure 5.2: Market structure of DRC's top 20 ASes in three categories: Customer, Provider, and Transit Autonomous Systems.

as the number of announcing prefixes increases, so does the AS degree. The line plot in red for the number of announcing prefixes also indicates major differences between relatively large ASNs versus smaller ones in terms of the number of addresses being announced to the Internet.

The same API allows us to look at other metrics that are missing in Figure 5.1, especially for the remaining ASes that happen to have no customer ASes. In Figure 5.2 we show the number of transit networks an ASN routes its traffic through, its number of providers, and customers respectively. Transit ASes are those that provide connectivity to other ASes and are responsible for transmitting data between different networks. They connect multiple ASes and serve as a bridge between different parts of the Internet. These ASes typically do not host any end-users and instead focus on providing transit services to other networks. Provider ASes, on the other hand, provide Internet connectivity to end-users. They offer Internet services such as web hosting, email, and cloud computing. They may have one or more customers, who are other ASes that use the provider's network to offer Internet services to their own end-users.

We observe that the ranking is correlated to the number of customer networks an AS provides services to. We find that UNITED-SA is the highest-ranked AS in the DRC with 10 customer networks, followed by Celtel-DRC having 7 and the government's commercial company of post and telecommunications (SCPT) with 4 customers while the remaining have only 1 customer. Those cases with 0 customer value simply mean that they are ISPs providing connectivity on the last mile only, close to the end-users.

Among these top 6 ranking ASNs, two of them are identified to be IXPs: ISPA-DRC-CONTENT and ISPA-DRC-LUBUX-MGMT. ISPA-DRC is a non-governmental organization under Congolese law. Its members represent more than 85% of the total number of Internet connections in the country [63]. This explains why the AS names (ISPA-DRC-CONTENT and ISPA-DRC-LUBUX-MGMT) appear amongst the highest ranks given that they are most likely having ISPs connected to them either through direct or remote peering to facilitate the exchange of local traffic. Subsequent paragraphs will shed more light on this.

We cannot ignore the presence of eyeball networks such as CELTEL-DRC and ORANGE-DRC among the highest ranks. The two combined, represent an estimated customer population size of 44,97%, with 24.99%, and 19.98% respectively [70][71]. We also notice another eyeball network having a 21,44% customer population size, namely VODACOM-CONGO occupies a lower rank (9th from the bottom) in Figure 5.1. Why this is the case? CAIDA ASRank's methodology is based on the concept of graph centrality, which is a measure of the importance or influence of a node (in this case, an AS) in a network. Its Degree centrality metric measures the number of connections that an AS has to other ASes. An AS with a high degree of centrality is considered to be more important or influential because it has more connections to other ASes. While Vodacom has a higher customer population than Orange-DRC, it does not have other ASes as its customers, as shown in Figure 5.2. Hence the reason why it occupies a lower rank than its eyeball network counterparts.

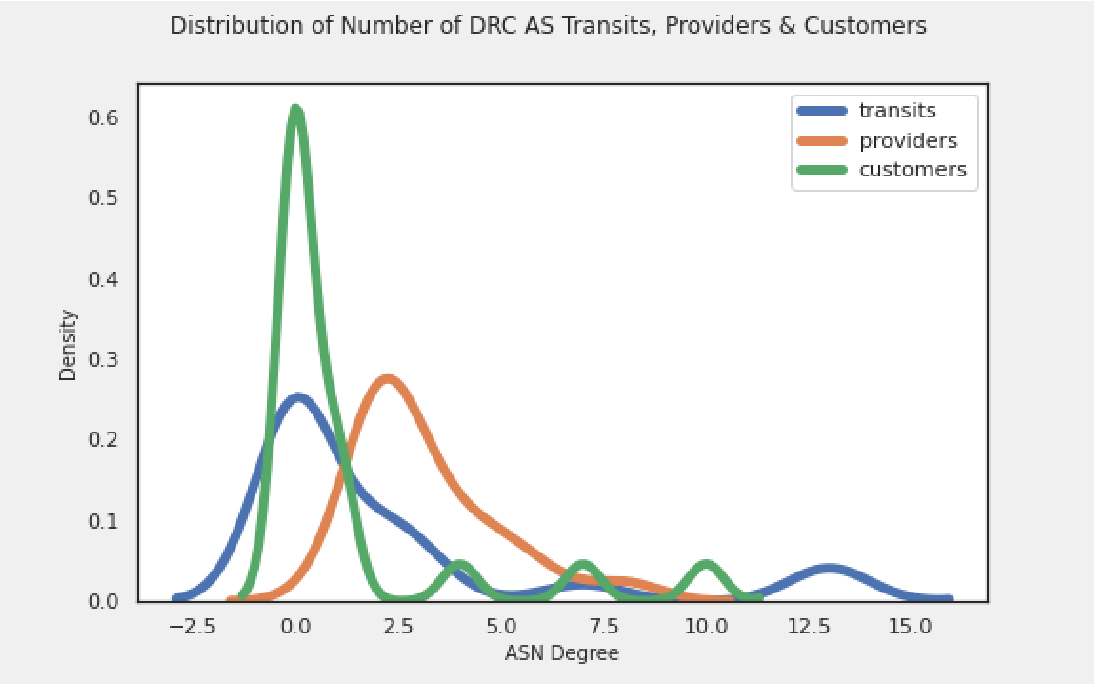

Figure 5.3: Distribution per DRC ASN's customers, providers, and transits.

Statistically, we can also see how distributed each of the AS relationship metrics presented in Figure 5.3 are, with all DRC 44 ASNs considered so that it serves as a guide to the discussion of the results we obtain from subsequent experiments of our study. Figure 5.3 shows these distributions. Considering 44 total ASNs in DRC, and counting the number of transit providers and direct customers each represented ASN in Figure 5.3, We can see that 60% of the selected ASNs in DRC are providers that have at least one customer network, and 25% of them do not connect through transit providers. 30% of these ASNs appear to have at least 2 providers whereas less than 10% of them have more than 5 transits, customers, or providers.

5.1.2 Network Interconnections

Network Interconnections at an IXP and Their Impact on End-User Experience

Network interconnections at an IXP serve as a crucial nexus where Internet service providers (ISPs), content providers, and other network entities converge to exchange traffic. For end-users, these interconnections have a direct and profound impact on the quality of their online experience. By interconnecting at an IXP, ISPs can access a broader array of content and services more efficiently, often with reduced latency. This means that when an end-user requests data, such as loading a web page or streaming a video, the content is delivered with lower delays, resulting in faster load times and smoother playback. Moreover, multiple interconnections at an IXP enhance network redundancy. If one network path experiences issues or outages, traffic can be seamlessly rerouted through alternative paths, minimizing disruptions and ensuring continuous service availability. In essence, network interconnections at an IXP are instrumental in delivering a superior and more reliable online experience to end-users, enhancing their satisfaction and trust in the services provided.

Impact on Business Operations and Service Providers

Beyond improving end-user experience, network interconnections at an IXP also have far-reaching implications for business operations and service providers. ISPs and content providers can leverage these interconnections to optimize their network efficiency and reduce operational costs. By exchanging traffic directly at an IXP, service providers can circumvent the need to route traffic through costly long-haul links, which can translate into substantial savings in terms of bandwidth costs. Additionally, the enhanced connectivity options provided by an IXP facilitate the rapid deployment of new services and applications, fostering innovation and competitiveness in the market. For businesses, these interconnections enable them to reach a broader audience more effectively and provide seamless access to their products and services. In sum, network interconnections at an IXP are a cornerstone of efficient and cost-effective business operations, offering competitive advantages and opportunities for growth in an increasingly connected digital landscape.

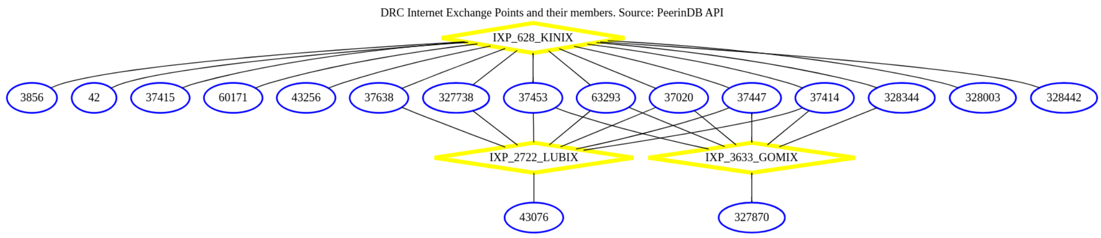

Figure 5.4: DRC IXPs' AS members.

Internet Exchange Points in the DRC and their AS members

Given the value IXPs bring to a country's Internet ecosystem in terms of resilience, our subsequent efforts resulted in associating the ASNs that leverage such Internet infrastructure. As shown in Figure 5.4 below, the nodes marked in a yellow diamond ring are the 3 existing IXPs in the country, whereas the nodes in blue oval rings are the DRC ASNs having an interconnection with at least one of the three IXPs. Clearly, the distribution of the number of AS members per IXP is uneven, whereby KINIX in the capital city appears to have 15 members, whereas LUBIX and GOMIX serve 8 and 7 members each respectively.

Another observation is that, without considering the networks that might be having private peering relationships with each other, only 38% (17 out of 44) of DRC ASNs (as of May 2021) leverage the IXP facilities available to them. However, we see the dominating eyeball networks: Orange (AS37447), Vodacom (37453), and Celtel-DRC (AS37020) are all connected to the three IXPs, which is a good sign as far as end-users quality of experience concerns are concerned.

5.1.3 BGP-level data

Each AS uses BGP to announce which IP addresses they are responsible for and which other ASes they connect to. BGP routers take all this information from ASes around the world and put it into databases called routing tables to determine the fastest paths from AS to AS. When packets arrive, BGP routers refer to their routing tables to determine which AS the packet should go to next. Figure 5.5 shows the peers or neighbors of each AS in the DRC as seen in the

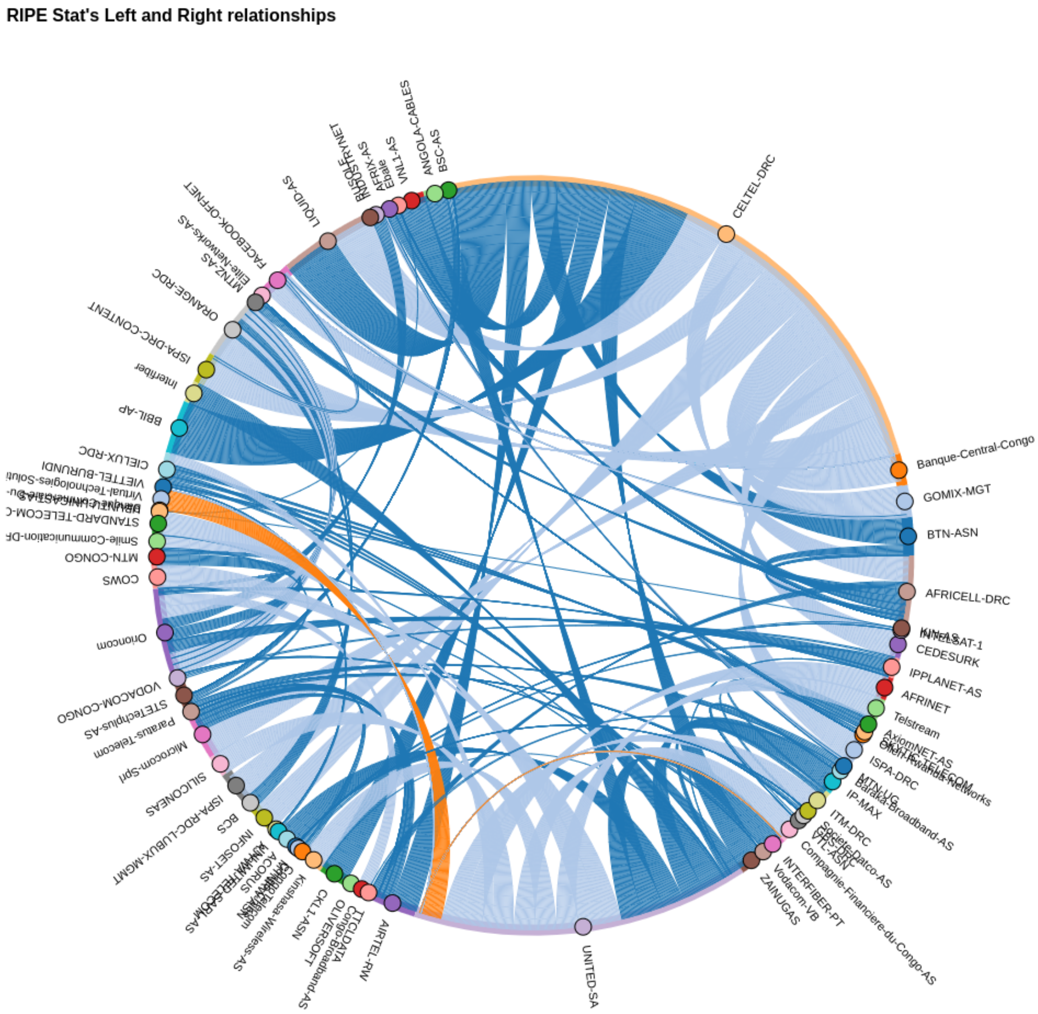

Figure 5.5: DRC ASes' Left and Right Neighbours.

Routing Information System. The terms "right neighbor" and "left neighbor" used refer to the ASes that are directly connected to a particular AS.

The left neighbors are represented by the Sky blue color, whereas the rights are in Blue. The number of neighbors tells us how many AS numbers were on the "left side" of the queried AS number in the BGP AS-path (closer to the origin of the announced prefixes), and how many were on the "right side". If an ASN has only left ones, this means it is an originating-only ASN. In the case of DRC ASNs, we find that Celtel-DRC and UNITED-SA appear to be on the left and right of the majority of ASNs. This is evident because they both have the most number of customer networks (7 and 10 respectively), as we saw previously in

Figure 5.2. Hence the announcement of the prefixes of those networks given their position of being their providers.

5.1.4 AS Hegemony

The Internet consists of over 70,000 networks. Each of these connects to the rest of the Internet via one or more connections to other networks. An article by RIPE Labs [85] argues that a very rough way of categorizing these networks would be: networks with users ("eyeball networks"); networks with servers ("content networks"); and networks that provide connectivity between networks ("transit networks"). Using the same analogy, in Figure 5.6 we visualize the interconnectivity of DRC's eyeball networks specifically and how they connect to other networks both inside and outside of the country. Figure 5.6 shows the interconnection between networks in DRC (with red nodes) and other networks either inside or outside DRC (with the latter represented as green nodes). Tier1 networks are added to the graph as yellow nodes. The size of a node is determined by its 'importance' for Internet routing for end-users - where importance is here measured in terms of the extent to which end users in DRC depend on this node for reaching the rest of the Internet. i.e. in terms of between-ness centrality [86] as estimated by AS Hegemony [79]. It is important to note that the main focus of this visualization is to show the interconnections that exist between the networks shown. As for traffic volume flowing over these interconnections, the only indication of this in the visualization is the size of the nodes, as nodes with a higher estimated betweenness centrality will likely have a higher traffic volume. We don't have data about the capacity of these links or the business (i.e. transit/peering) relationships between networks.

We can observe that the Internet comes into DRC through AS37677 SCPT (Société Commerciale des Postes et Telecommunications), which receives the connection from the ACE (Africa Coast to Europe) undersea cable and interconnects with Angola Cables, Interfiber, and other foreign ASes. This same AS then pro-

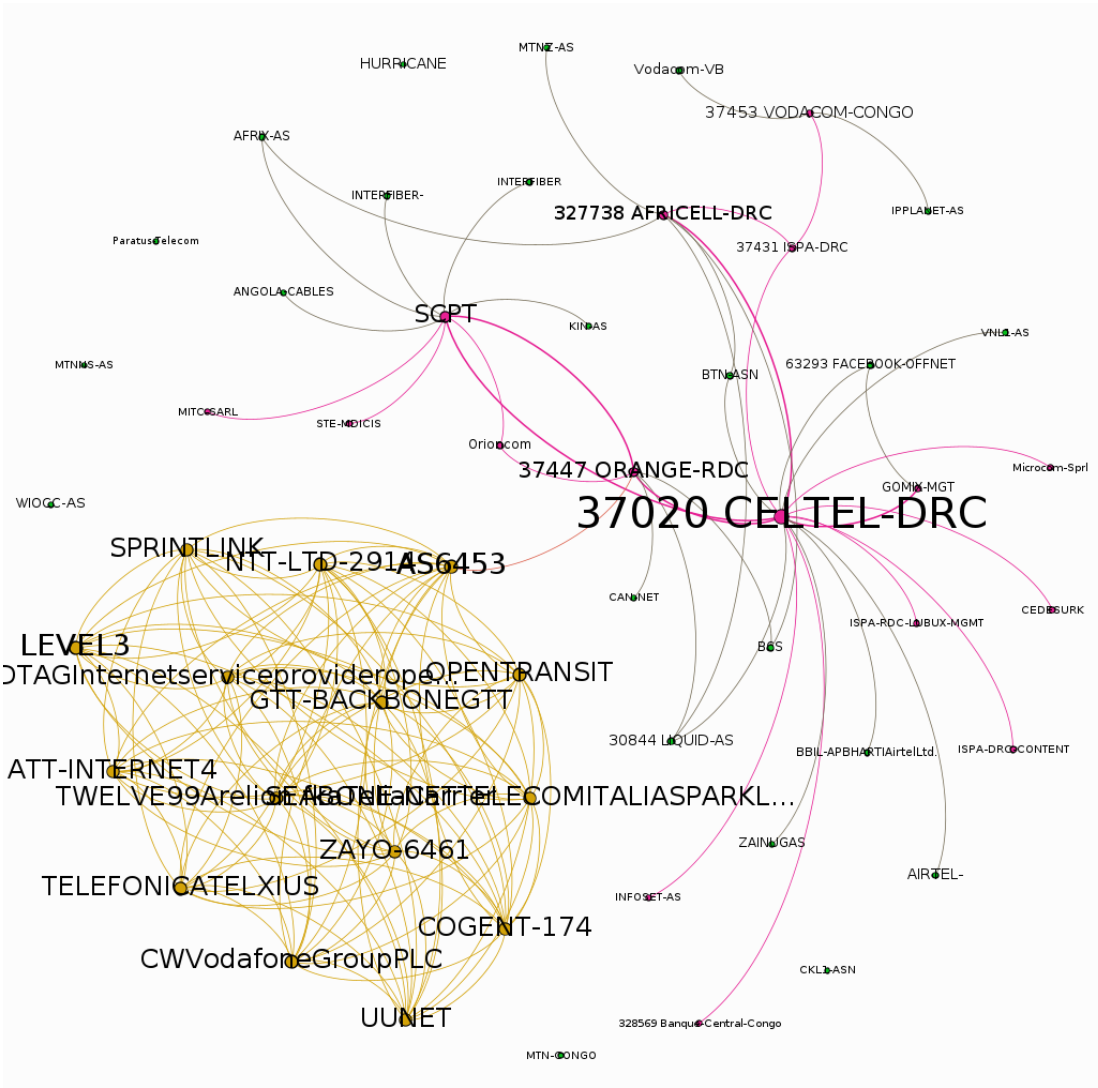

Figure 5.6: DRC Eyeball networks interconnections.

vides connectivity to bigger Eyeball Networks such as Orange-DRC(AS37447) and CELTEL-DRC(AS37020). It is noticeable that AS37020 has the biggest node size (the largest local AS hegemony score), hence a sign of high dependence of the majority of Congolese end-users on its services. We see the same behavior for the connections to foreign ASes: CELTEL-DRC(37020) has more edges to the green nodes.

Due to the mechanism put in place by ISPA-DRC (the association of Congolese ISPs), we can see eyeball networks such as Africell-DRC, Vodacom-CONGO, and Celtel-DRC sharing a common interconnection node (AS-37431 ISPA DRC). When we look at the IXP and members structure in 5.4 and identify the ASN to which the Kinshasa Internet Exchange Point belongs, we indeed find that it is AS-37431 ISPA DRC. This clearly demonstrates that these network business operators do agree on the importance of exchanging traffic between the "downstream" users of each network, instead of accessing them through foreign paths.

While Tier1 networks are a fundamental part of how the Internet is interconnected, we can see that presently, their usage is declining due to the flattening of the Internet [87], which refers to the observation that the Internet has become less hierarchical and more interconnected over time. In the case of the DRC, only Orange-DRC (AS37447) is connected to one Tier1 network, namely AS6453 (Tata Communications). Clearly, this is because there are more and more interconnections bypassing them, despite them being known as the core of the Internet ecosystem.

5.1.5 Conclusions

In conclusion, we found that the DRC has a limited number of ASNs with a high local AS hegemony score. This indicates that the country relies on a small number of external networks for connectivity. This limited number of providers may make the DRC vulnerable to outages or other disruptions in its Internet infrastructure.

The AS market structure in the DRC is also heavily concentrated in a few key players. UNITED-SA, Celtel-DRC, and the government's commercial company of post and telecommunications all have a significant number of customer networks, while the remaining ASNs have only one or zero customers. This concentration of ASNs in a few key players could potentially lead to issues of monopolization and lack of competition in the market.

Overall, the analysis of the AS market structure in the DRC shows that the distribution of ASNs is uneven and the market is heavily concentrated in a few key players. This limited connectivity to external networks and concentration in a few players could potentially impact the stability and competitiveness of the DRC's Internet infrastructure. It would be important for policymakers to consider these factors in planning for the development and regulation of the country's Internet infrastructure.

Chapter 6

MOBILE BROADBAND INTERNET PERFORMANCE FROM ACTIVE MEASUREMENTS

This chapter presents the results of the active measurement activities conducted with research participants in the Democratic Republic of Congo (DRC) to assess the quality of service (QoS) of mobile broadband networks and users' quality of experience (QoE) on the web within the country, as well as the status of traffic destination between local networks. The active measurements were conducted using a combination of client-based tools, such as Ookla Speedtest and MLAB's NDT to gather data on various performance metrics including download and upload throughput, and latency. In addition, through custom python and shell scripts executed on users' computers, Web Quality of Experience (QoE) metrics such as DNST, TTFB, TCPT, and Total Time were measured to give insights into the user experience when accessing popular websites and services.

6.1 QoS

6.1.1 Dataset

Our sample consists of measurement instances conducted by 335 unique participants across the five cities in DRC. As illustrated in Figure 6.1, We chose three focus cities, one on the West side of the country (Kinshasa), another on the

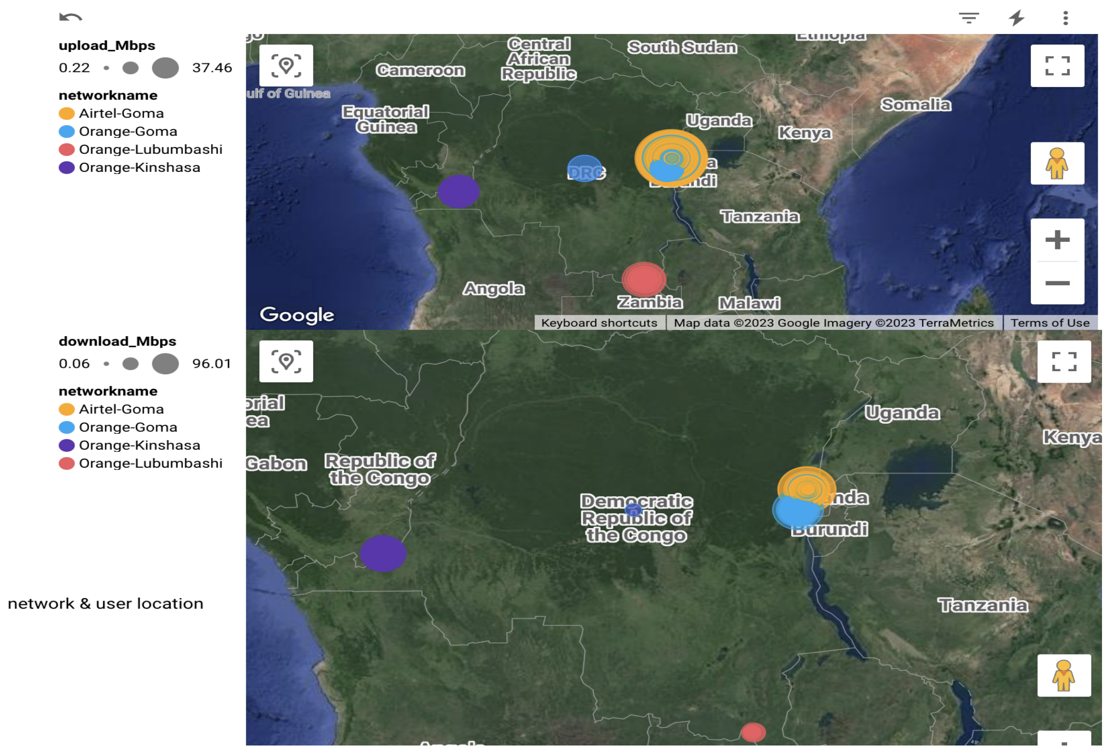

Figure 6.1: Congolese cities used as vantage points for Network QoS measurements

East side (Goma), and the other on the South-East Side (Lubumbashi). This comparison is designed to capture the effects of DRC's late deployment of Internet Exchange Points and its biggest cities in terms of economic development and population density. Kinshasa the capital city happens to have the first ever installed IXP in the country (KINIX), followed by LUBIX in Lubumbashi and lately Gomix in Goma. The sociodemographics of our focus cities reflect those of the West and East Sides of the country, including a disparity in household broadband connectivity, as discussed in we'll see in subsections 6.1.2 and 6.1.3.

b) Upstream throughput Irrespective of the mobile broadband network participants are subscribed to, there are 205 tests conducted in Goma, 81 in Kinshasa, and 10 in Lubumbashi. The remaining 39 tests are distributed to other cities like Bukavu, Kisangani, Beni, and Bunia respectively. In our analysis, we group all tests into the "network-region" category. This strategy was used to capture the variation of Internet performance across the city, as a point of comparison with our focus networks per region. Our analytic approach thus conducts singlewise comparisons between three cities: Kinshasa, Lubumbashi, and Goma with respect to the network a test was conducted from.

Due to the fact that our measurements are continuous and ongoing over three months, the data itself has some discontinuities. This is the result of many factors, such as when participants temporarily stop running throughput tests on their phones or use a hot spot to can conduct the tests, and so forth. To minimize the effects that these kinds of discontinuities could have on our analysis (e.g., averages over two distinct throughput tiers, or over periods where data was missing or zero), we selected a two-month period that had minimal discontinuities from which we performed our analysis. After exploring the distribution of active participants across all three months of measurement, we chose the time period with the most continuously active participants, which were the months of August and September 2022. For this reason, the period of analysis in this paper reflects the time period of August 1 to September 28, 2022.

In this section, we present the results of our active measurements and analysis. We focus first on comparing both downstream and upstream throughput across three IXP-present cities (Section 5.1) before turning to latency (Section 5.2). The presented results are based on our own active measurement dataset and MLAB's historical Internet performance data.

6.1.2 Active Ookla Speedtest

a) Metric: minimum, lower, mean, upper, and maximum quartiles. When comparing throughput across networks and regions in this study, we faced two challenges: (1) Due to the monopoly of a single network in one of the cities selected for the measurements, the collected data faces a risk of being unrepresentative as far as the comparison to other broadband countries in the same city is concerned and (2) throughput can vary over time, for a variety of reasons.

To account for these uncertainties and variabilities, we rely on the distribution of throughput and skewness by displaying the throughput percentiles and averages.

This technique gives a five-number notation summary of throughput representing the p25, p50, p75, and p100 where p(i) is the ith percentile measurement from that network within a specified city. Intuitively, these percentiles capture how a throughput measurement is distributed over the course of longitudinal measurements, hence making the identification of mean values, the dispersion of the data set, and signs of skewness quick.

In the context of this subsection of the dissertation, the chosen method of data presentation is the boxplot. A boxplot, also known as a box-and-whisker plot, is a graphical tool used to summarize and visually represent the distribution of a dataset. It is especially useful for providing a concise overview of key statistical characteristics, aiding in the understanding of the data's central tendency and variability.

The fundamental components of a boxplot include the box, whiskers, median line, and outliers. The box itself spans the interquartile range (IQR), stretching from the lower quartile (Q1) to the upper quartile (Q3). This portion of the plot encapsulates the middle 50% of the data, showcasing the spread of this central portion of the distribution. Extending from the box are the whiskers. The lower whisker commences at a defined range below Q1, often set at 1.5 times the IQR. The upper whisker extends to a specified range above Q3. These whiskers signify the overall range of the dataset, excluding any outliers that may lie beyond these limits.

Contained within the box, a vertical line represents the median (Q2), which serves as a measure of central tendency. It indicates the data point that falls exactly in the middle when the dataset is arranged in ascending or descending order.

Outliers, if present in the data, are data points that fall well outside the whiskers. These data points deviate significantly from the main distribution and are often depicted as distinct symbols, such as the green circles mentioned in the current study. Outliers may warrant further examination due to their potential to influence the interpretation of the dataset.

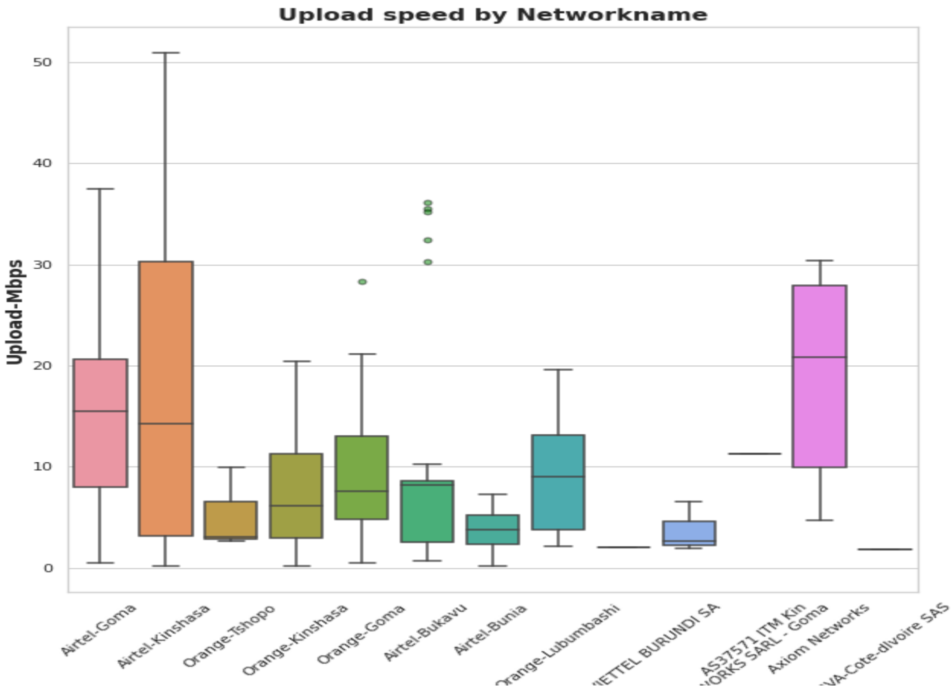

Figure 6.2: Congolese networks' Upstream throughput in Megabits per Second using Ookla Speedtest

In essence, the use of a boxplot in this study is a strategic choice for conveying a comprehensive summary of the dataset's key statistical characteristics. It allows for a quick assessment of data variability, central tendency, and the identification of any potential outliers, all of which are crucial aspects of data analysis and interpretation within the context of this thesis.

b) Upstream throughput Figure 6.2 shows that the distribution of the upload throughput is positively skewed in that the median upload throughput at the majority of the network tends to be closer to the lower end of the box.

At a network level, we see that users are susceptible to experiencing upload throughput close to 0Mbps. This holds true for Airtel and Orange irrespective of the city in which a user is located. At the 50th percentile, Airtel shows the highest median upstream throughput (15Mbps), close to double the median upload throughput subscribers on Orange experience (9Mbps). At a city level, participants in Goma (East) tend to experience better upload throughput than those in the capital city Kinshasa (West) irrespective of the network they are subscribed to. We argue this by looking at the pair (Airtel-Goma, Orange-Goma) versus (Airtel-Kinshasa, Orange-Kinshasa). Notice that we purposely omit Lubumbashi because it does give the perspective from one single network (Orange-Lubumbashi), which nevertheless outperforms the previous two cities with a median upload throughput close to 9Mbps.

At the 75th percentile, we do not only notice changes at both the network and city levels but also much higher disparities between them: 75% of Airtel subscribers in Kinshasa (West) can reach an upload throughput less than 30Mbps in contrast from 20Mbps for Goma (East) subscribers to the same network. Orange on the other hand hardly reaches 15Mbps at the 75th percentile in both Goma, Lubumbashi, and Kinshasa cities. This represents close to half less upload throughput compared to Airtel. As far as maximum upload throughput is concerned, Airtel subscribers in Kinshasa can experience almost double and half more upload throughput compared to those from Orange. However, we notice some outliers for Orange subscribers in Bukavu (East) that show as much maximum upload throughput as Airtel subscribers from Goma City.

b)Downstream throughput Figure 6.3 shows the distribution of download throughput per network and city. While the throughput is skewed similarly to upload throughput in figure 6.2 above, it is shown that users from all the networks represented can also experience minimum download throughput close to 0Mbps. At the 50th percentile, Orange shows the highest median downstream throughput (22Mbps), slightly better than Airtel which shows 16Mbps approximately. The lowest median download throughput of 2Mbps is recorded at Airtel in Bunia (North-East). From a city perspective, once again participants from Goma (East) tend to experience slightly better performance than those in Kinshasa (West) but with more than double the download throughput of participants from Lubumbashi (South East). At the 75th percentile, Orange performs better with download throughput close to 40Mbps in both Goma and Kinshasa, whereas Airtel hardly

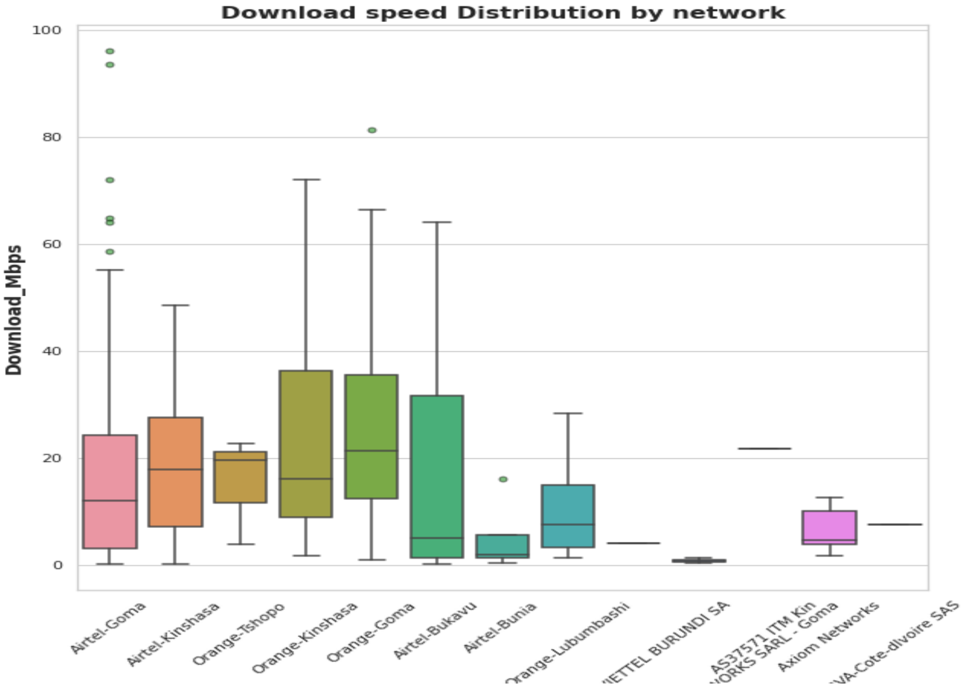

Figure 6.3: Congolese networks' Downstream throughput in Megabits per Second using Ookla Speedtest

reaches 25Mbps in the aforementioned cities. A similar trend is shown when the maximum download throughput is considered where Orange in Kinshasa and Goma outperforms Airtel with download throughput exceeding 70Mbps versus 55Mbps. However, outliers are mostly identified at Airtel Goma with a 10Mbps difference from the outlier for Orange, in the same city, Goma.

Latency We present latency results in terms of an absolute increase between the minimum, mean, and maximum latency values for each Ookla server location. As opposed to throughput, we present results in terms of an increase, since larger values of latency (as opposed to smaller values of throughput) represent worse performance. We do not, however, normalize the reduction, since latency values are well-known and do not necessarily correlate with throughput tier— giving us the ability to both measure a baseline and construct groups that are large enough to compare without having to first, normalize. As with throughput, each result shows the min (lowest whisper), median (line in the box), and max (highest whis-

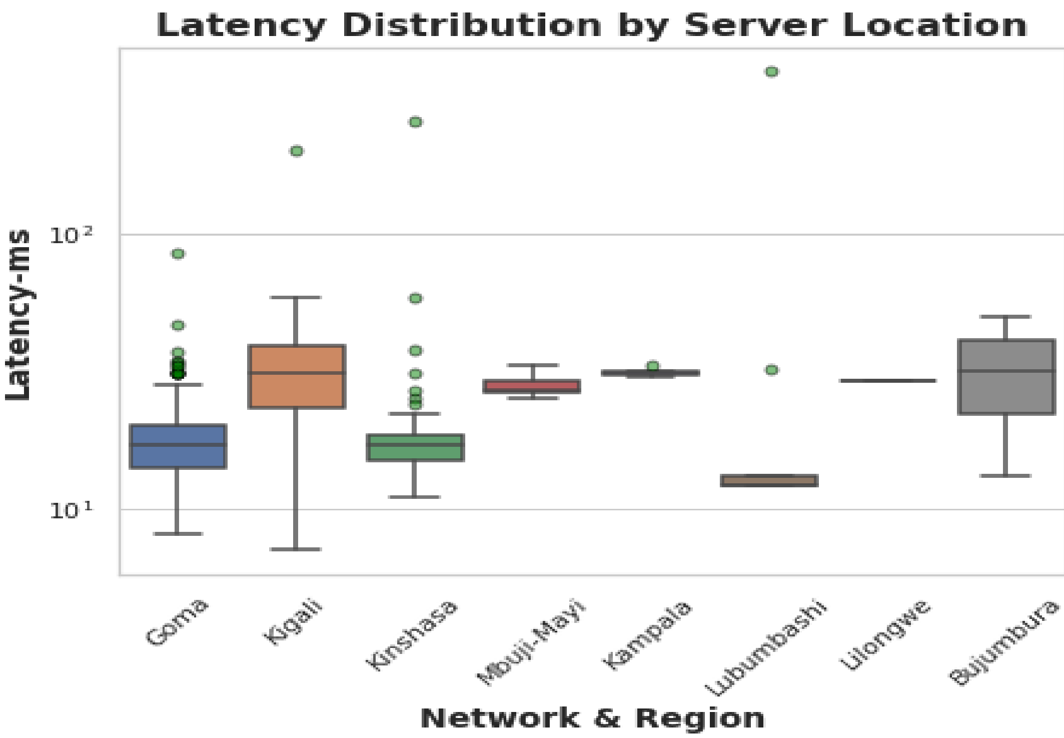

Figure 6.4: Users' experienced Latency in milliseconds from Congolese Networks to Ookla servers

per) latency values per server location within their respective percentile range.

Figure 6.4 shows a similar trend as throughput: Goma city is shown with the lowest latency value of 7 milliseconds as far as local server selection is concerned. At the 50th percentile, both Goma and Kinshasa show the same mean latency of 20 milliseconds. Foreign Ookla servers on the other hand perform a bit worse with a mean latency of 50 milliseconds at the 50th percentile for tests going to Kigali, Kampala, and Bujumbura, while the maximum latency gets close to 80 milliseconds. Outliers are mostly recorded at local Ookla servers, the highest ranging from 90 ms in Goma to 150 ms in Kinshasa and 190 ms for Lubumbashi.

6.1.3 MLAB NDT speedtest

The results of the active measurements presented in the section above take into account two-thirds of the most influential cellular networks in DRC. The remaining third accounts for Vodacom had firewalls installed that blocked all Ookla throughput test measurements. To cater to this, we rely on historical data from

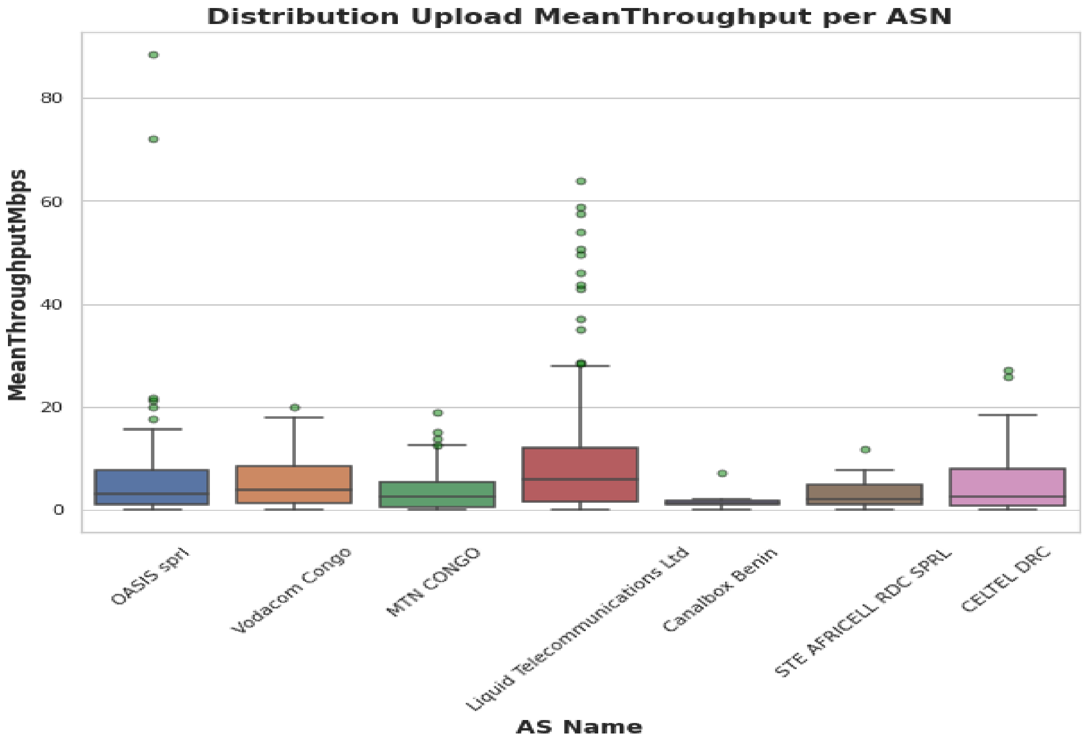

Figure 6.5: MLab crowd-sourced data: Mean Upload throughput in Mbps by DRC Autonomous System

MLAB datasets to present the performance of this network, while making comparisons to those covered in Section 6.1.2. In contrast to Ookla measurements, in this section, we represent the individual networks by their respective Autonomous System numbers.

Upload Mean throughput per ASN

6.5 shows the distribution of upload throughput per ASN. It is shown that clients at all the represented networks can reach a minimum mean upload throughput close to 0 Mbps. Vodacom (AS37453) shows a mean upload throughput of 3Mbps, similar to Orange (AS37447) at the 50th percentile whereas Airtel (AS37020) shows slightly worse performance (2Mbps). The network that outperforms the rest is an ISP, Liquid Telecom with an upload throughput close to 6Mbps for fifty percent of the tests conducted, nearly double of previously mentioned cellular networks. At the 75th percentile, Airtel, Orange and Vodacom users seem to

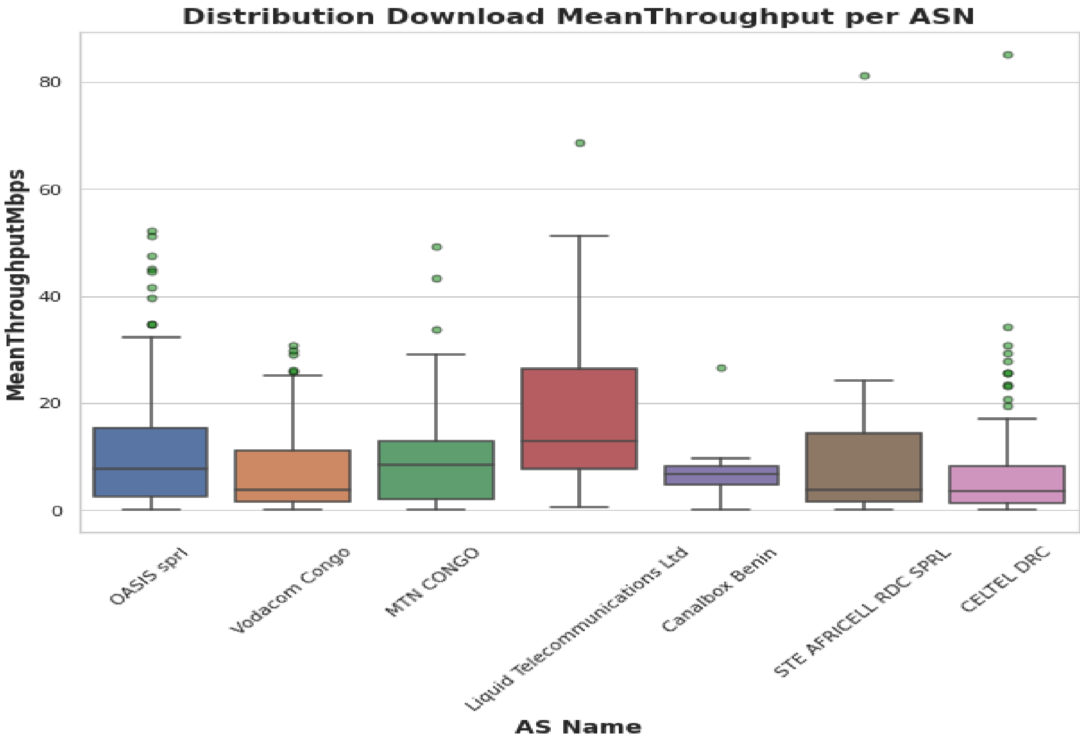

Figure 6.6: MLab crow-sourced data : Download throughput by Congolese network

experience an equal performance of 8 Mbps, still below that of Liquid Telecom clients who get 12 Mbps. In terms of the maximum upload throughput for each network represented in the queried dataset, Airtel shows a higher throughput (18Mbps) than Vodacom (17Mbps) and Orange (15Mbps). Liquid Telecom on the other hand shows uploads throughput close to 30Mbps, not surprising because it is a fiber optic-based network provider.

Download Mean throughput per ASN

Similar to the upload throughput, download throughput test throughput can get closer to zero Mbps. At the 25th percentile, the performance is lower than 3 Mbps for all cellular networks represented. The differences are noticeable at the 50th percentile where Airtel (AS37020) users' experience (4Mbps) is the same as those of Vodacom (AS3753). Orange (AS37477) on the other hand doubles them with a mean download throughput of 8Mbps.

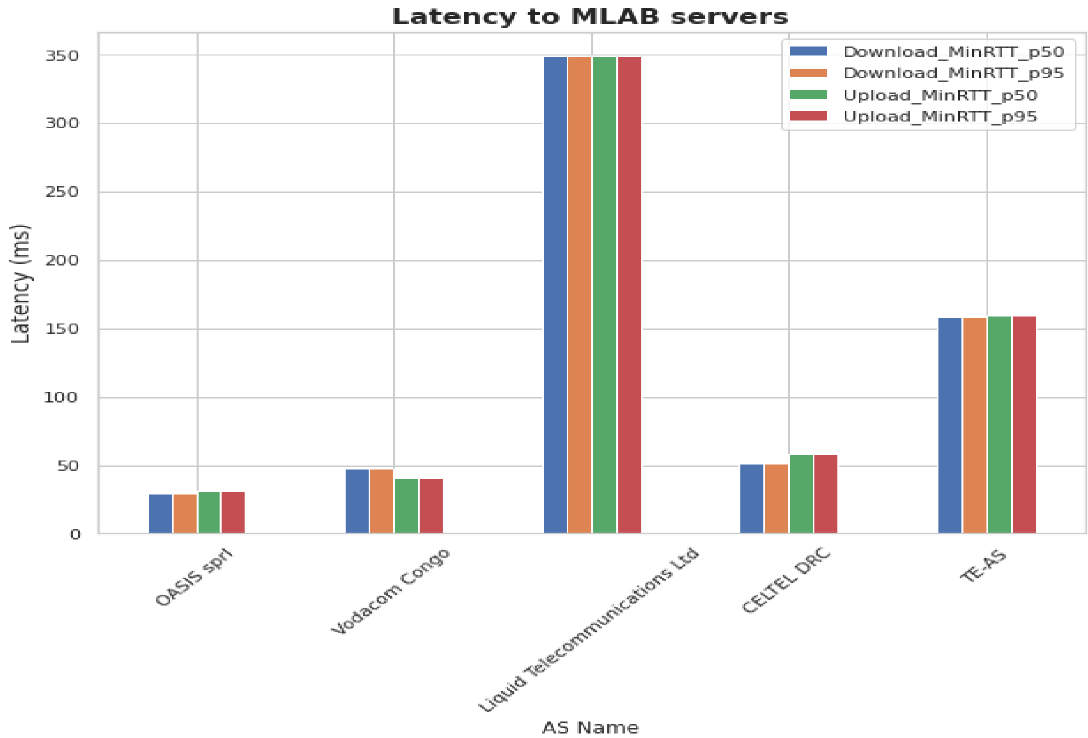

Figure 6.7: MLab crow-sourced data: Latency from Congolese Networks to MLab Servers in foreign networks and countries

Latency

Metric: Absolute increase. We present latency results in terms of an absolute increase between the 50th and 95th percentile latency values for each network. As opposed to throughput, we present results in terms of an increase, since larger values of latency (as opposed to smaller values of throughput) represent worse performance. We do not, however, normalize the reduction, since latency values are well-known and do not necessarily correlate with throughput tier giving us the ability to both measure a baseline and construct network tests that are large enough to compare without having to first normalize. In contrast with throughput, each result that shows the 50th and 95th percentiles per network represents a percentile of throughput tests conducted within that network, so one could read the 50th percentile of absolute increase for a network as the median increase that a client would see for that network.

In contrast to Ookla, MLAB servers are located outside the client's network and its test selects the nearest server to measure the performance. For the case of DRC, most tests were going to servers in Kenya and Nigeria. Figure 6.7 shows the latency experienced by clients in the specified Congolese networks towards the nearest servers between 1st January and 1st October 2022 at the 50th and 95th percentile for both upload and download directions of the measurements. It is also shown in the same Figure 6.7 that at the 50th percentile of download tests, clients from the three cellular networks: AOSIS Sprl (new ORANGE DRC), Vodacom, and CELTEL-DRC (Airtel) experienced a round-trip time of less than 50 ms, with ORANGE outperforming the rest with a latency value of 30 ms. We notice a similar performance at the 95th percentile. Upload tests, on the other hand, 50 percent of the tests from the three cellular networks show that Celtel-DRC (AS37020) experienced the highest latency value of 70ms followed by Vodacom (40ms) and lastly Orange with the least (32ms). Ideally, the network with the smallest latency value reflects better performance as far as the connection to MLAB servers is concerned. At the 95th percentile of tests in the upstream direction, similar to download tests, there is no difference noticed in latency values from the 50th percentile.

6.1.4 Bridging the Gap: Aligning Active Network Measurements with User Perceptions for Improved Internet Performance

The active measurements conducted within the surveyed networks provide valuable insights that can be linked to users' perceptions of internet performance covered in Chapter 4. In the context of upstream throughput, Figure 6.2 in Section 6.1.2 illustrates that the upload throughput distribution is positively skewed, indicating that a substantial portion of users tends to experience lower upload speeds. This aligns with the observations of internet users who, regardless of the network or city, are susceptible to experiencing upload throughput close to

0Mbps. However, the data also reveals variations in performance among networks and cities. Airtel consistently outperforms Orange in terms of median upstream throughput, with Airtel users experiencing close to double the upload speeds compared to Orange users. This positive correlation between active measurements and user perceptions emphasizes the importance of upload throughput as a key factor influencing user satisfaction.

Furthermore, at the 75th percentile, significant disparities emerge not only between networks but also across cities, indicating that users' internet experiences vary widely. For instance, Airtel subscribers in Kinshasa can reach upload speeds significantly lower than those in Goma, further emphasizing the influence of geographical location on internet performance. These findings corroborate users' perceptions, as 75% of Airtel subscribers in Kinshasa can experience upload throughput less than 30Mbps, significantly lower than their counterparts in Goma. Orange, on the other hand, consistently falls behind Airtel in terms of upload throughput, reinforcing users' dissatisfaction with Orange's performance.

In the context of downstream throughput, as illustrated in Figure 6.3, similar trends are observed. Users across all networks can experience minimum download throughput close to 0Mbps, aligning with the notion that low download speeds contribute to user dissatisfaction. Orange consistently outperforms Airtel in terms of median downstream throughput, which directly correlates with users' perceptions of Orange providing slightly better internet performance. However, the presence of outliers, such as those identified in Airtel Goma, indicates that real-world performance can sometimes deviate from the median measurements. In conclusion, the active measurements align with users' perceived internet performance, emphasizing the significance of both upload and download throughput in shaping user satisfaction and highlighting the need for network providers to address disparities to improve overall user experience.

6.2 Mobile vs Fixed Internet performance comparison between Central African Countries

In the pursuit of understanding the dynamics of internet performance in the DRC, our research delves into the second part of the third research question, which aims to compare the internet performance of the Democratic Republic of Congo (DRC) with other Central African nations. This comparison provides valuable insights into the state of digital infrastructure and connectivity within the region. To address this question, we turn our attention to Ookla Speedtest's Internet Performance Global Index [2], a comprehensive dataset widely regarded for its extensive coverage and methodology [88].

6.2.1 Dataset

Ookla Speedtest's Internet Performance Global Index [2] collects and analyzes data from millions of user-initiated throughput tests conducted worldwide in mobile and fixed broadband networks. The global index uncovers not only how countries perform but also how their urban centers fare in terms of internet throughput. Therefore, the power of this dataset to encompass diverse geographical regions with a large sample size, makes it a reliable resource for assessing internet performance on a global and urban scale.

6.2.2 Mobile broadband Internet Download throughput

Our analysis begins with a focus on mobile internet download throughput in select Central African countries, including the DRC. Figure 6.8 illustrates the median mobile internet download throughput for each country, measured in megabits per second (Mbps). This data sheds light on the effectiveness of mobile networks in delivering data to users, a crucial aspect of digital infrastructure. Rwanda and Uganda emerge as regional leaders with median download throughput of approx-

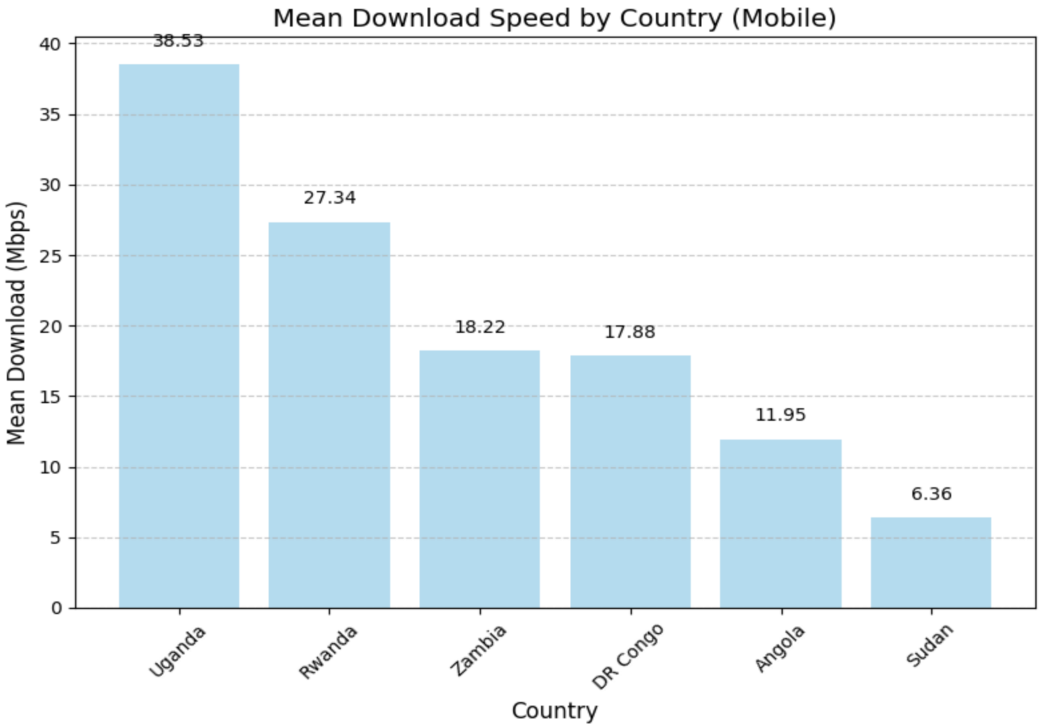

Figure 6.8: Internet Download throughput comparison by Central African Country: Mobile broadband

imately 27.34 Mbps and 38.53 Mbps, respectively. These nations exhibit robust mobile network infrastructures capable of delivering high-throughput internet access to their populations. Meanwhile, the DRC, represented as "DR Congo" in the graph, demonstrates a median mobile download throughput of 17.88 Mbps, positioning it within the mid-range among the featured countries. Sudan and Angola exhibit relatively lower median download throughputs of 6.36 Mbps and 11.95 Mbps, respectively, indicating potential challenges in their mobile network infrastructure. These findings provide insights into how the DRC compares to its African counterparts in terms of mobile internet performance.

6.2.3 Fixed Broadband Internet Download throughput

Shifting our focus to fixed broadband internet download throughput, we continue to explore internet performance within the same select Central African countries. Fixed broadband networks typically offer higher throughput and stability, making them essential for various applications, including online education and

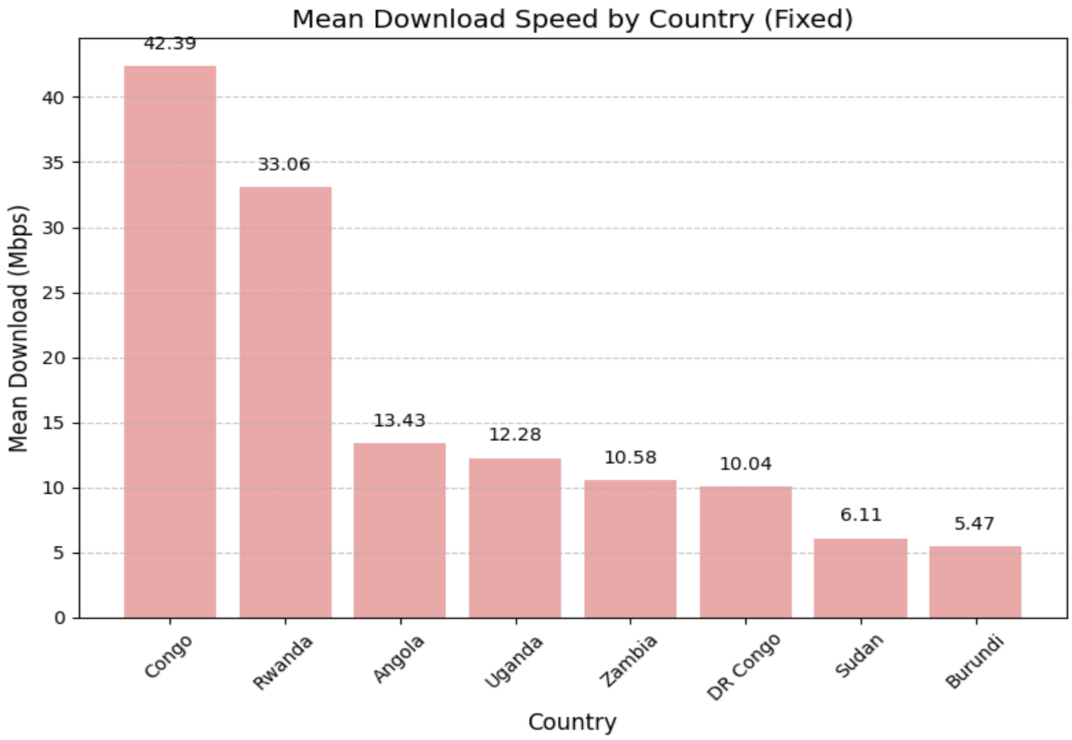

Figure 6.9: Internet Download Throughput Comparison by Central African Country: Fixed broadband

telecommuting. As seen in Figure 6.9, Congo (Brazzaville) leads the group with a remarkable median download throughput of 42.39 Mbps for fixed broadband, indicating a robust fixed broadband infrastructure. Rwanda follows closely behind with a median download throughput of 33.06 Mbps, showcasing the country's commitment to improving its digital infrastructure. In contrast, the DRC presented as "DR Congo" in the graph, records a median fixed broadband download throughput of 10.04 Mbps, positioning it among the lower end of the featured countries. Burundi reports the lowest median download throughput for fixed broadband at 5.47 Mbps, suggesting potential challenges in its fixed broadband network development.

6.2.4 Conclusion and Recommendations

Our examination of internet performance throughput in the selected Central African countries, including the Democratic Republic of Congo (DRC), has shed light on both achievements and areas for improvement within the realm of digital infrastructure. As we conclude our analysis, we offer recommendations tailored to each country, acknowledging that enhancing internet performance is a multi-faceted endeavor, influenced by various factors.

Democratic Republic of Congo (DRC)

The DRC exhibits commendable progress in mobile internet download throughput, surpassing some of its Central African counterparts. However, to further bolster its digital infrastructure, we recommend:

Firstly, Investment in Network Infrastructure: Expanding and upgrading mobile network infrastructure to ensure consistent high-throughput connectivity across regions is crucial. Collaboration between the government and telecom providers can facilitate these improvements. Secondly, Promotion of Digital Inclusion: Initiatives aimed at increasing digital literacy and accessibility to mobile internet services should be prioritized, particularly in rural areas. Thirdly, Regular Performance Monitoring: Continual monitoring of internet performance through tools like Ookla Speedtest can provide valuable insights for further improvements.

Rwanda

Rwanda stands out as a leader in both mobile and fixed broadband internet throughput among the featured countries. To maintain and advance its position, we recommend:

1. Continued Investment: Sustaining investment in network infrastructure, including 5G rollout and fiber optic expansion, will further enhance internet throughput and coverage.
2. Digital Skills Development: Fostering digital literacy and technology skills among the population ensures that high-throughput internet is effectively leveraged for economic growth and social development.
3. Collaboration with Stakeholders: Collaborative efforts between the government, private sector, and international organizations can accelerate digital trans-

formation and innovation.

Sudan and Angola

Sudan and Angola face challenges in mobile internet performance, indicating the need for strategic improvements:

1. Infrastructure Enhancement: Investing in network infrastructure upgrades and expanding coverage in underserved areas can boost mobile internet throughput.

2. Regulatory Support: Regulatory reforms that promote healthy competition among telecom providers can lead to better services and affordability.

Other Countries

Each country has its unique context, and our recommendations are tailored to the specific insights drawn from the data. Future developments should consider the following overarching strategies:

1. Data-Driven Decision-Making: Regularly assess and analyze internet performance data to identify areas for improvement.

2. Cross-Sector Collaboration: Collaborate with relevant stakeholders, including government, private sector, and civil society, to collectively address challenges and opportunities in digital infrastructure.

3. Digital Inclusion: Prioritize initiatives that promote digital literacy and ensure that high-throughput internet access benefits all segments of society.

4. Investment in Innovation: Explore innovative technologies and solutions to advance internet connectivity, such as fiber optics, 5G, and satellite internet.

In conclusion, our analysis underscores the significance of high-throughput internet connectivity in driving economic growth and social development. By implementing these recommendations and maintaining a commitment to improving internet performance, these countries can pave the way for a digitally inclusive and prosperous future.

Future work for this section of this research may involve an in-depth examination of the challenges and opportunities in executing these recommendations, along with a focus on the impact of improved internet performance on various sectors and communities.

6.3 Web QoE

6.3.1 Dataset

In addition to the quality of service of the selected Congolese networks in the previous section, we also measured the experience of the users on the web. Web traffic via protocols such as HTTP2/ HTPP3 undoubtedly serves as the number one means of accessing the Internet through websites, hence the importance of measuring how fast do users in the DRC can access the Web.

Using the Semrush [80] dataset, we obtained a list of the most accessed websites in DRC as of 15th October 2022. We considered only websites whose traffic is generated by both desktop and mobile user devices. Using that, we measured users' Web QoE metrics according to LibCurl's approach [82]. These Web QoE metrics include DNST, FTTB, TCPT, and Total Time.

After applying the filter to the raw data, our final list was comprised of 32 unique websites covering five industry categories which include news, social media, entertainment, betting, and pornography sites. Additionally, the considered websites are not limited to global top-level domains (.com, .net, .org) but also, those that have DRC's ".CD" TLD are included. Therefore, we rely on a website's industry category and its coverage (global or local as far as a country' TLD is concerned) to present the results of our Web QoE measurements. The measurement activity was distributed through a GitHub repository [89] where users (preferably technical) cloned and executed a python script during a period of eight days, three times per day. A CSV file named by date and time was output at each execution of the script, then we asked each participant to send a final Zip file that contained all the resultant files of the measurements for further analysis. In the following

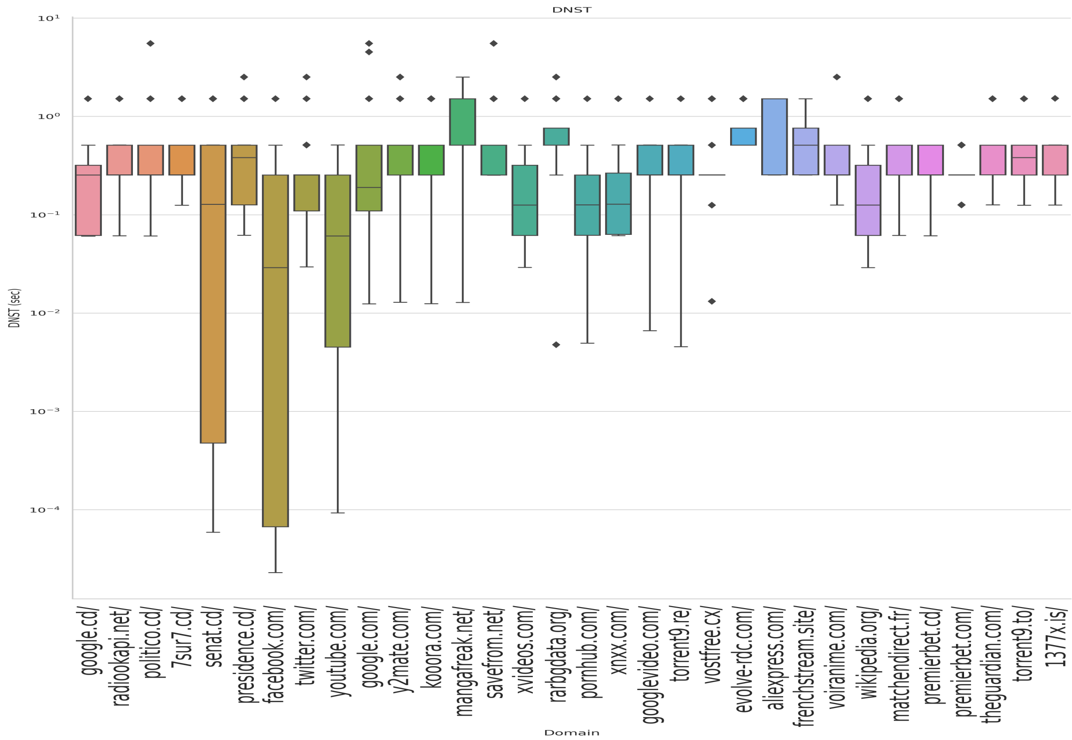

Figure 6.10: DNS Resolution Time: The total time in seconds from the start until the name resolution of a website is completed

subsection, we present the results of the measured DRC Internet users' Web QoE.

6.3.2 Results

DNS Resolution Time (DNST)

figure 6.10 displays the distribution of DNS resolution times (DNST) on a log scale, broken down by different top-level domains (TLDs). The plot shows that the majority of TLDs have median DNST values that are below or equal to 0.5 seconds. However, there are two exceptions to this trend: the ".net" and ".site" TLDs. These TLDs appear to have DNST values that are significantly higher (above 1 second). Additionally, we notice that the ".org" TLD has DNST values above 0.5 seconds but less than 1 second.

These results suggest that DNS resolution times for the ".net" and ".site" TLDs are significantly slower than other TLDs. This could be due to a variety of factors, such as poor server performance, high traffic, or issues with the DNS

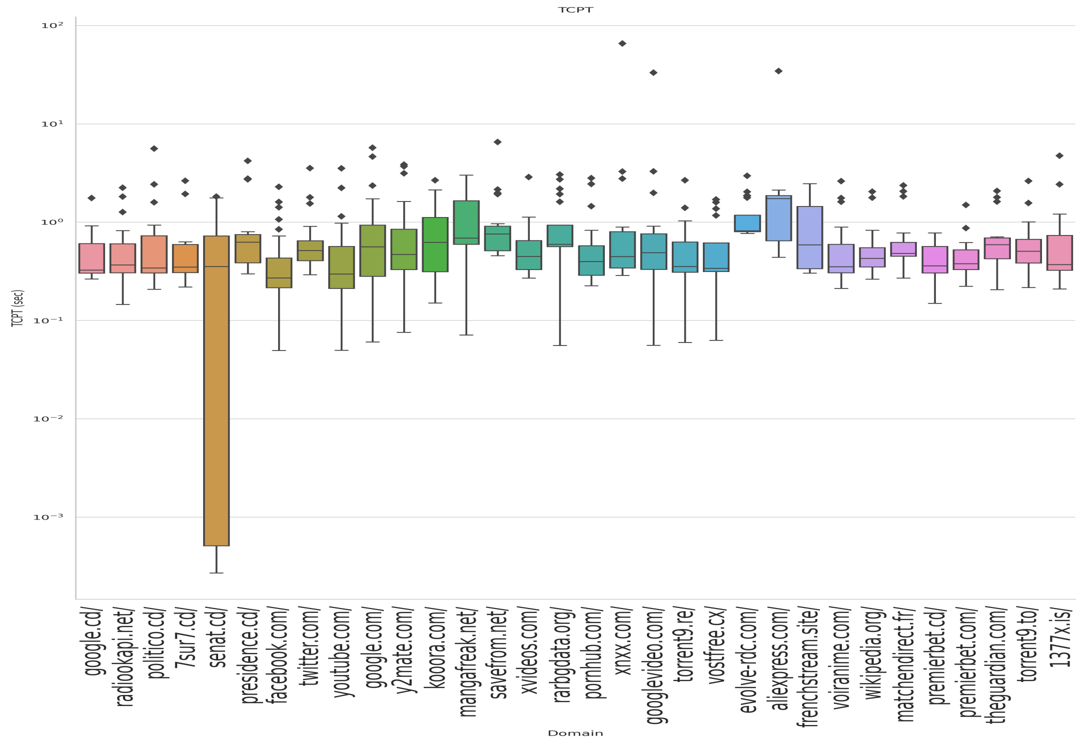

Figure 6.11: TCP Time by Domain: The total time in seconds from the start until the connection to the remote host (or proxy) is completed

infrastructure. The slower resolution times for the ".org" TLD, while not as severe as the ".net" and ".site" TLDs, still indicate a potential issue that should be investigated.

Overall, these results highlight the importance of monitoring DNS resolution times for different TLDs. By identifying and addressing issues with slow resolution times, organizations can ensure that their websites and online services are accessible to users in a timely manner.

TCP Time

Figure 6.11 presents the distribution of Transmission Control Protocol (TCPT) times on a log scale, grouped by different top-level domains (TLDs). The y-axis values at the first quartile range between 0.2 and 0.5 seconds for the majority of the domains represented in the plot, with the exception of the domain senat.cd, hosted in Switzerland, which shows significantly better performance compared to all the other domains represented (polico.cd, presidence.cd and google.cd), with a

TCPT time below 0.003 seconds. At the second quartile, all ".CD" domains have a similar median TCPT of approximately 0.5 seconds. This suggests that while some websites on the ".CD" TLD have better performance than others, overall, the majority of websites on this TLD have similar TCPT times despite being hosted in foreign locations such as Worcaster UK for presidence.cd, politico.cd in the USA. We, therefore, notice a bad practice happening in the management of the .CD TLD server for the following reasons: First, hosting the TLD server outside of a country can create security risks. TLD servers are crucial for the functioning of the internet and contain sensitive information about a country's domain name system (DNS). Having this server hosted in a foreign country makes it vulnerable to cyber-attacks and data breaches, which could have serious consequences for the security and stability of the internet in the country.

Second, hosting the .CD TLD servers outside of a country can reduce the reliability and throughput of internet services for users within the country. The server's physical location will affect the latency and response time for DNS queries, leading to slower and less reliable internet connections. This can have negative impacts on businesses and individuals who rely on the internet for communication, commerce, and other critical activities.

Finally, the .CD TLD servers located in foreign countries may create political and legal issues. The hosting country may have laws and regulations that are different from those in the country whose TLD server is being hosted, which could lead to conflicts over issues such as privacy, censorship, and intellectual property rights.

For these reasons, we recommend Congolese policymakers host the .CD TLD servers within DRC borders, where they can be better protected and managed, and where they are subject to the laws and regulations of the country.

Figure 6.11 also shows that entertainment, search engine, and social media websites such as youtube.com, google.com, and twitter.com have a similar minimum TCPT of 0.08 seconds as pornographic sites such as pornhub.com and xvideos.com which are hosted in New York, US, and Amsterdam, Netherlands respectively. However, in the second and third quartiles, the TCPT for these websites is variable, mostly below 1 second. This suggests that while some users may experience a quick loading time for these websites, others may experience slower loading times.

Another set of websites represented in the far right section of the subplot is betting, streaming and torrent downloading sites (premierbet.cd hosted in Toronto-Canada, torrent9.to in Neward-USA, and matchendirect.fr in Roubaix-France), which show better and faster TCPT than google.com, twitter.com, and youtube.com at the first quartile. However, in their performance at the second and third quartiles, users appear to experience a TCPT time as long as that of porn websites mentioned above. This suggests that while these websites may have a fast loading time initially, users may experience slower loading times as they navigate deeper into the website.

Overall, these results suggest that some websites on the ".CD" TLD have better performance than others and that the majority of websites on this TLD have similar TCPT times. Additionally, the analysis of other TLDs reveals that while some websites may have a fast loading time initially, users may experience slower loading times as they navigate deeper into the website. Hence, monitoring TCPT times for different TLDs and addressing issues with slow loading times ensures that users have a positive experience when accessing online services.

Time To First Byte (TTFB)

Figure 6.12 presents the distribution of Time To First Byte (TTFB) values on a log scale, grouped by different top-level domains (TLDs) for users in the Democratic Republic of Congo (DRC). Similar to the TCPT results presented in the previous subplot, the domain with the lowest TTFB for users in DRC is seen at senat.cd with a value lower than 0.003 seconds, whereas the maximum time does not exceed 5 seconds for all domains and TLDs combined, except for outliers. This domain is an exception compared to its sisters with the ".CD" TLD, given

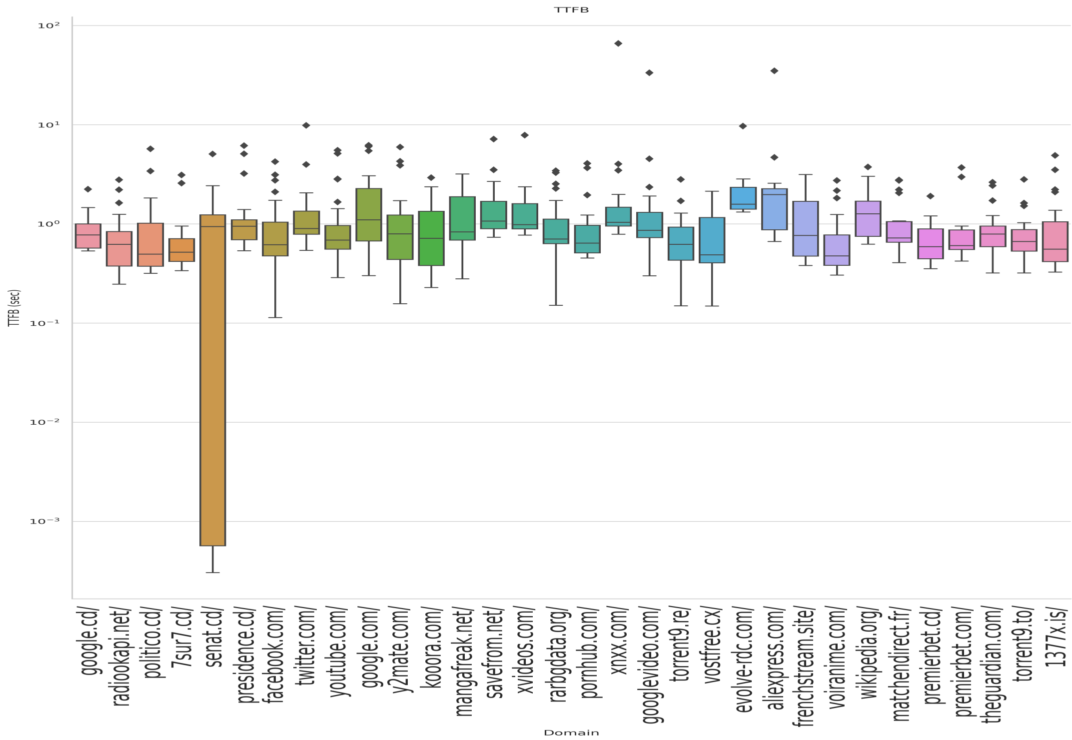

Figure 6.12: Time To First Byte Time by Domain: The time in seconds, it took from the start until the web request was just about to begin

that their best performance as far as TTFB is concerned ranges between 0.5 and 1 second for users in DRC.

However, it is observed that the median TTFB value varies per domain for users in DRC, regardless of the TLD a website is registered to. While the median TTFB values range between 0.4 and 4 seconds for users in DRC, it is noticed that two websites share the lowest median value, namely ”vostfree.cx” and ”voiranime.com” which are both movie-streaming websites for users in DRC. On the other hand, the highest median TTFB is seen at the ”frenchstream.site” website for users in DRC, suggesting the difference in performance users in DRC can experience when consuming streaming content.

Furthermore, the variance in median TTFB performance is also observed for popular websites such as youtube.com (0.8 sec), facebook.com (0.7 sec), twitter.com (1 sec), and google.com (2 sec) for users in DRC. This suggests that even popular websites with high traffic and resources can have varying TTFB times for users in DRC, and it is important for these websites to monitor and address any issues

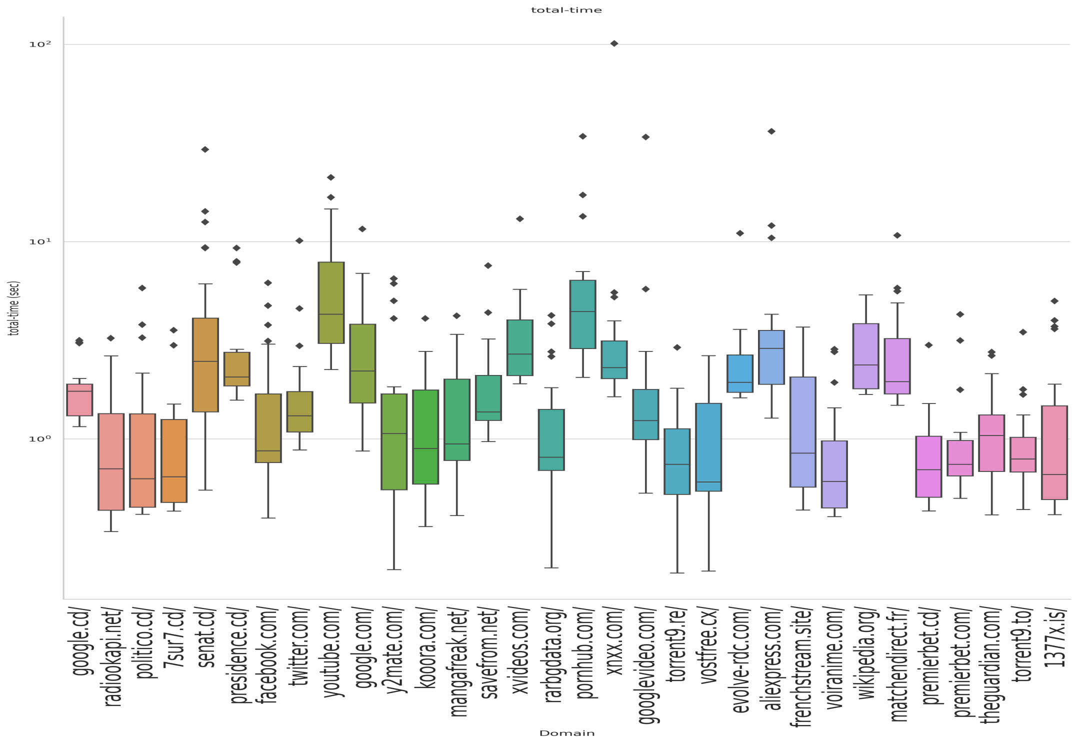

Figure 6.13: TCP Total Time by Domain: Total time in seconds for previous transfers, including name resolving, and TCP connections

that may arise for users in DRC. In conclusion, the results presented show that while the domain with the lowest TTFB is seen at senat.cd with a value lower than 0.003 seconds for users in DRC, the median TTFB value varies per domain for users in DRC, regardless of the TLD a website is registered to. This variance is observed for popular websites such as youtube.com, facebook.com, twitter.com, and google.com for users in DRC, suggesting that even popular websites with high traffic and resources can have varying TTFB times for users in DRC. These results highlight the importance of monitoring TTFB times for different TLDs and addressing issues with slow loading times to ensure that users in DRC have a positive experience when accessing online services.

Total Time

Figure 6.13 presents the distribution of total time values in seconds per domain category for users in the Democratic Republic of Congo (DRC). The boxplot representation of the data shows the median, quartiles, and outliers of the data.

The subplot shows that the maximum total time experienced by users in DRC is seen at the "youtube.com" website with a value close to 12 seconds, whereas the minimum total time of approximately 0.2 seconds, hence the fastest experience is seen when users access three sites: namely y2mate.com, rarbdata.org, torrent9.re and vostfreee.cx.

At the 50th percentile (median), the majority of websites with the .CD Congolese TLD do take less total time than large and popular sites such as youtube.com, google.com, and facebook.com, with a 0.2 to 4 seconds range difference. This shows the advantage of leveraging local Internet infrastructure for content that is meant to be local for a better Quality of Experience (QoE) for users in DRC.

It is worth noting that the median total time for websites meant to provide relatively large content like videos for download and streaming varies greatly. For instance, we notice the highest total times reaching 8 and 9 seconds at both youtube.com and pornhub.com in the third quartile. This scenario is different for websites whose content appears to be mainly text-based, such as news and betting sites (see the far left and right of the subplot) whose maximum total time does not exceed 4 seconds and whose median total time is below 0.5 seconds for 98 percent of them.

Overall, Figure 6.13 presents a detailed analysis of the distribution of total time values for different domains for users in the Democratic Republic of Congo. The results show that while large and popular sites such as youtube.com, google.com, and facebook.com have longer total time values, the majority of websites with the .CD Congolese TLD have shorter total time values. This highlights the advantage of leveraging local Internet infrastructure for content that is meant to be local for a better Quality of Experience (QoE) for users in DRC. Additionally, the subplot also shows that the median total time for websites meant to provide relatively large content like videos for download and streaming varies greatly, with text-based websites having a much shorter total time compared to video-based websites. These results highlight the importance of monitoring total time values for different domains and addressing issues with slow loading times to ensure that users in DRC have a positive experience when accessing online services.

6.3.3 Conclusion

In conclusion, this analysis provides an in-depth understanding of the key metrics related to web performance for users in the Democratic Republic of Congo. DNS Resolution Time (DNST) is an important metric as it measures the time taken for a domain name to be resolved to an IP address. Monitoring DNST can help identify and address issues related to domain name resolution, which can affect the overall web experience for users.

Transmission Control Protocol Time (TCPT) is another important metric as it measures the time taken for data to be transmitted between the client and server. Monitoring TCPT can help identify and address issues related to data transmission, which can affect the overall web experience for users.

Time To First Byte (TTFB) is an important metric as it measures the time taken for the first byte of a response to be received by the client. Monitoring TTFB can help identify and address issues related to server response time, which can affect the overall web experience for users.

Total Time is an important metric as it measures the time taken for a page to fully load. Monitoring total time can help identify and address issues related to page load time, which can affect the overall web experience for users.

It is important to monitor these metrics regularly and take appropriate actions to address any issues that may arise. It is also recommended to leverage local internet infrastructure, as it can provide a better Quality of Experience (QoE) for users. Additionally, it is important to understand the specific needs of different types of websites, such as the difference in performance for text-based and video-based websites. By monitoring and addressing issues related to web performance, organizations can ensure that users have a positive experience when accessing online services.

6.4 Traffic destination and routes

Having measured the quality of service of Congolese networks and the quality of experience of the users on the web, we sought to understand the traffic paths these networks traverse for packets targeted at a local destination IP. The aim is to identify the presence of circuitous routes between Congolese networks and open doors or a correlation between their performance as presented in sections 6.1 and 6.3 and the length of the path taken by packets that are presumably expected to be local. We executed the traceroute tool with 22 different destination IP addresses taken from the top 10 Congolese prefixes ranked by their market share. A dataset containing each network with its traceroute output was constructed. Next, we parse each network's output and retrieve the hops through which the packets were traversing before they reached the destination IP. During the data cleaning process, we ignored the hops that have disabled ICMP packets for we cannot identify them nor their location or their ASN. Essentially, these are hops whose traceroute output contains a star (*) symbol. Next, we retrieved geo-localization and AS name info of each traversed hop to help us conclude whether there are any circuitous routes being taken by each network or not.

In our analysis, we first consider path length between eyeball networks. These include Celtel-DRC(AS37020), ORANGE (AS37447), and Vodacom (37453) for they represent the majority of the customer population distributed across all the regions of the country. Furthermore, in our discussion, we include the path to other DRC ASes that appear to have a customer base in a specific region or province. These include Africell DRC SPRL (AS327738), MTN, and TIGO.

6.4.1 Celtel DRC: AS37020

Celtel DRC was presented by our hegemony graph in Figure 5.6 as the most influential network in DRC given its number of customer ASes and subscribers on the last mile. In terms of routing of traffic towards local ASes, we found the following.

Celtel to Vodacom

We found that packets going to a Vodacom IP in DRC take a path to four unique foreign ASes located in six different countries before they reach their destination. The path involves AS37020 (Celtel-DRC) to AS37124 (Tigo Rwanda) to AS3636 (Airtel Kenya) to AS9498 (BHARTI Airtel Ltd) in India to AS37662 (WIOCC-AS) in Nigeria to again AS37662 (WIOCC-AS) in Kenya, to AS36994 (Vodacom-VB) in South Africa to AS37453 (Vodacom-CONGO).

In the traceroute output, we notice a median latency value of 1057 milliseconds between the hop at WIOCC-AS in Kenya and Vodacom in South Africa. While this is very high latency, we have to point out the presence of two "Request timeout" responses between the two ASes. This potentially reflects a block of ICMP packets at the two devices along the path. The last AS path segment from Vodacom-VB in South Africa and Vodacom-Congo (the destination) suffers a median latency of 343 ms and a maximum of 505 ms. That said, when we look at the interconnection state of the ASes as illustrated in Figure 5.4, we notice that both CeltelDRC and Vodacom DRC are connected at the same IXPs, but clearly, it does not necessarily mean that are peering with each other. This leaves room for more investigation on why this is the case between two eyeball networks that could potentially have mutual business and customer satisfaction benefits if they leveraged fully the opportunities offered at the IXP.

Celtel to Africell

Africell DRC SPRL (AS327738) is another cellular network with a considerable percentage of the customer population in DRC (5%) despite being limited to one city (Kinshasa) in terms of geographical coverage in contrast to the others. Traceroute data to an AS327738 IP show a path involving four unique foreign ASes along the path located in five different countries. The path is seen as follows: AS37020 to AS37124 (Tigo Rwanda) to AS3706 (Liquid Telecommunications Rwanda) to AS30844 (Liquid Telecommunications Ltd) with consecutive hops in Kenya, United Kingdom, Zimbabwe, and Zambia respectively, to AS36962 (MTN Zambia) to AS28698 (UUNET Africa Zambia) to AS327738 (Africell-DRC) the destination. Along the path, we see the packets reaching Liquid Telecom's hope in Europe (UK) before it's handed to a Zimbabwean IP from the same AS. Surprisingly enough, while the median latency recorded between these two hops is over a hundred (150 ms), it is still not bad enough for such a distant route. Liquid Telecom is a fiber optic company, and hence an illustration of the difference in performance modern fiber connections can offer in contrast to the older data transmission medium. Before reaching the destination, the packets pass through Zambia suffering a median latency of 153 milliseconds, which in contrast to the Vodacom destination of packets from Celtel explained above, is better performance despite packets traveling to a different continent. Once again, in terms of peering, the source and destination ASes discussed are common members of two IXPs in DRC, namely KINIX and LUBIX (Kinshasa and Lubumbashi Internet Exchange Points) as shown in Figure 5.4.

6.4.2 ORANGE AS37447

6.4.3 Orange to Celtel

Orange represents the second biggest network in DRC in terms of the customer population. By tracing its packets to a Celtel IP (41.243.13.222) as the destination, we found that the traffic takes the AS Path 30844 (Liquid Telecommunications Operations Limited) to 3491 (PCCW Global Inc) to 37662 (West Indian Ocean Cable Company) to 37020 (Celtel DRC). While the AS count between Orange and Celtel-DRC appears to be short in number (three), we notice that the packets get out of the country through Liquid with AS Number 30844 in DRC (Kinshasa) and then traverse six more hops from the same AS but located in two different countries with an inter-hop latency not exceeding 50 milliseconds: first in South Africa (Johannesburg), then in Mauritius (Quatre Bornes) and back to South Africa again (Johannesburg). Next, Liquid Telecom in South Africa passes the packets to AS3491 (PCCW Global Inc) which passes its neighbor AS37662-West Indian Ocean Cable Company (still in South Africa) until they finally reach to Celtel-DRC in Kinshasa with a mean latency of 233 milliseconds.

Orange to Vodacom

Similar to destination Ips from Celtel DRC, Orange's traffic towards Vodacom-congo leaves the country via Liquid Telecom(AS30844) which again forwards the packets its hops in South Africa and Mauritius and back to South Africa again, in Cape Town city. Next, the packets are seen to take a route to Europe (London, England) through a sibling AS in Cape Town (AS30445) with a latency of 209 ms, which in turn passes the packets to AS3356 (Level 3 Parent, LLC) until they travel back to Vodacom-Congo in Kinshasa, DRC facing a median latency of 235 milliseconds.

Orange to Africell

With an estimated 746,212 subscribers, we considered Traces of packets from Orange to one of Africell's two routers' IP addresses in DRC. Results show a similar path to that of Celtel-DRC and Vodacom destinations (through 1 local and 6 foreign IPs from AS30844) in the beginning, passing to AS3491 (PCCW Global Inc) also in South Africa before being passed to AS33763 (Internet Technologies Namibia) and AS37273 (Bandwidth and Cloud Services Group Ltd) in Kenya which in the end delivers it to Africell in DRC by paying a median latency of 226 milliseconds.

6.4.4 Conclusion

In conclusion, the analysis of the traffic destination and routes of networks in the Democratic Republic of Congo (DRC) showed that packets traveling to Vodacom and Africell take different paths involving multiple ASes located in various countries. The traceroute data revealed a high median latency between the hop at WIOCC-AS in Kenya and Vodacom in South Africa, with the presence of "Request timeout out" responses indicating a potential block of ICMP packets. The source and destination ASes are common members of two IXPs in DRC, KINIX, and LUBIX, but the lack of peering between them remains a topic for further investigation. In the case of Orange, the packets to Celtel and Vodacom take different paths, with a low inter-hop latency not exceeding 50 milliseconds. The findings suggest the importance of a well-optimized network structure for ensuring efficient and fast packet delivery and highlight the opportunities for improvement and growth in the DRC's telecommunication sector.

6.5 Chapter Conclusions

In conclusion, in this Chapter, we presented results from the active measurement activities conducted in this research which provided valuable insights into the Quality of Service (QoS) and Quality of Experience (QoE) of end-users in the Democratic Republic of Congo (DRC). The results of the measurement of QoS showed that the download and upload throughput and latency of networks in the DRC, as measured by the Ookla throughput test tool and MLAB crowdsourced data, indicate room for improvement.

The results of our Web Quality of Experience (QoE) measurements indicate that websites use the DRC's.CD top-level domain (TLD) tend to outperform larger global websites in various key metrics, including Domain Name System Time (DNST), Time to First Byte (TTFB), Transmission Control Protocol Time (TCPT), and Total Time. However, this performance disparity is closely tied to the nature of the content hosted on each website and the complexity involved in maintaining them, which is often influenced by their popularity.

Furthermore, our analysis of network traffic patterns in the Democratic Republic of Congo (DRC) reveals an interesting phenomenon. Despite the presence of three

Internet Exchange Points (IXPs) within the country, network traffic frequently takes routes through remote international destinations before ultimately arriving at its intended destination within the DRC. This routing behavior persists even when it appears that networks in the DRC could establish direct peering connections through these IXPs.

In light of these findings, we propose a set of recommendations to enhance the Quality of Service (QoS) for networks in the DRC. It is crucial to focus on improving both download and upload throughput, as well as reducing latency. Additionally, we suggest that the servers hosting the DRC's .CD TLD websites should be located within the country's borders. This move would facilitate better security, management, and compliance with local laws and regulations. Furthermore, it could encourage website owners in the DRC to prioritize optimization efforts, particularly for the efficient delivery of high-bandwidth content such as video streaming.

Lastly, we emphasize the need to optimize network traffic routing within the DRC. This optimization can be achieved by leveraging the available local peering infrastructure and implementing best practices within the peering ecosystem. Such measures will ultimately result in an improved quality of experience for network users and customers in the DRC.

Chapter 7

DISCUSSIONS

The discussions chapter presents an in-depth analysis of the key findings from the various research activities conducted for this study on Internet transparency in developing African regions, with a focus on the Democratic Republic of Congo (DRC). This chapter draws together the results from the secondary dataset analysis, online survey, and active measurement activities to provide a comprehensive understanding of the current state of Internet connectivity and performance in the DRC. The chapter will begin by summarizing the main findings from each of the research activities and highlighting the key trends and patterns identified. It will then proceed to discuss the implications of these findings for the broader context of Internet transparency and connectivity in developing African regions. Additionally, it will explore the potential barriers to improving Internet performance in the DRC and identify areas where further research is needed. The chapter will conclude by providing recommendations for policymakers, regulators, and other stakeholders on how to address the challenges identified and improve Internet connectivity and transparency in the DRC and other developing African regions.

7.1 Users perceived Vs Experienced performance

The results presented in the previous chapter 6 tell us that there exists a correlation between a user's perceived Internet performance and their experienced Internet throughput. In other words, if a user perceives their Internet as performing well, they are likely experiencing a fast Internet throughput. Similarly, if a user perceives their Internet as performing poorly, they are likely experiencing a slower Internet throughput. TThis applies to the 89 Congolese Internet users in our survey. Among them, 45% expressed satisfaction or dissatisfaction with their network provider's Internet performance. Specifically, for 85% of the networks evaluated using the Mlab NDT test at the 75th percentile, the download throughput did not exceed 18 Mbps. A similar trend was observed for upload throughput, where dissatisfaction correlated with 85% of the mobile broadband networks failing to reach a 10 Mbps throughput at the 75th percentile. The remaining 15% of the networks represented cover for a fiber optic network provider that shows double the performance of the measured mobile broadband networks for download and upload throughput at all percentile ranges.

When taking Ookla throughput test data into account, the measurements show better performance for the same participants who perceive their Internet performance as unsatisfying. While the networks represented in these measurements are fewer in numbers compared to the MLAB crowdsourced and historical data, we noticed a significant increase in download and upload throughput for all the networks represented. This leads us to another important note that perceived Internet performance is not always directly related to experienced Internet throughput. There are other factors that can affect a user's perception of Internet performance, such as the quality of the website they are visiting, the performance of their own device, the quality of their Internet service provider (ISP), the technology used to connect to the Internet, and more importantly the distance between the user and the source of the Internet connection. The correlation made when considering Congolese users' measurements towards MLAB servers, which are located outside the ISP's network, in foreign countries like Kenya and Nigeria presents a different story compared to the users' measurements towards Ookla servers which most often reside in the client's ISP network. This is because the path between the client's device to MLAB servers will always cross through multiple networks maintained by different operators. Therefore, depending on what aspect of the performance is being measured and the goal of the measurements, the question of choosing the right tool holds relevance. If the goal is to measure whether advertised throughput is equal to the measured throughput or to know whether connection measurements meet a specific definition of broadband, then Ookla's results are what needs to be considered. In contrast, if the goal is to check how is the connection to outside networks (because the Internet is a network of networks), then MLAB results should be the most prevalent.

Additionally, a user's perceived Internet performance can be influenced by their expectations and past experiences. For example, if a user is used to a very fast Internet connection, they may perceive a slightly slower connection as poor performance, even if it is still considered fast by most standards.

ISPs typically advertise a certain level of Internet throughput, such as "high-throughput" or "gigabit," which indicates the maximum throughput that the user can expect to experience under ideal conditions. However, the actual Internet throughput that the user experiences may be lower than the advertised throughput due to a variety of factors, such as network congestion or the user's distance from the ISP's network equipment.

The technology used to connect to the Internet can also affect the user's experienced Internet throughput. For example, a wired connection, such as Ethernet, typically offers faster and more stable Internet throughput than a wireless connection, such as Wi-Fi. Additionally, newer technologies, such as fiber optic connections, can provide even faster Internet throughput than traditional cable or DSL connections.

Finally, the user's distance from the source of the Internet connection can affect their experienced Internet throughput, specifically for mobile broadband networks. In general, the further the user is from the source of the wireless connection, the slower their Internet throughput will be. This is because the signal has to travel a longer distance, which can result in signal loss and other forms of degradation.

7.2 Equity in Internet access

The internet is a crucial aspect of modern life, providing access to information, communication, and entertainment. However, the quality of internet service can vary greatly depending on where end-users are located with respect to where the connection is originating from. As part of this research, we have explored the differences in internet throughput between cities in the DRC, and how these differences affect users' perceptions of internet performance.

To begin with, in Section 7.1 we presented how internet throughput can vary greatly depending on a variety of factors. As a result, it is not surprising to see differences in internet throughput between cities in the same country. The results of our survey and active measurements in sections 4 and 6 have highlighted these differences and showed that internet throughput in some cities was significantly faster than in others. For example, users in Kinshasa the capital city reported a median download throughput of 15 Mbps, while users in Goma reported an average download throughput of only 10 Mbps. These differences in internet throughput can have a significant impact on users' experiences with the internet. For example, users in cities with slower internet throughput may experience longer download times for files, lower-quality video streaming, and difficulty accessing certain websites or applications. In addition, users in cities with slower internet throughput may be more likely to experience frustration and dissatisfaction with their internet service.

Another interesting finding from our measurement activity was that subscribers from the same network can experience disparities in internet throughput depending on the city they are located in. For example, users who were subscribed to the same ISP in Kinshasa and Goma reported significantly different internet throughput. This suggests that factors other than the user's choice of ISP may be contributing to differences in internet throughput between cities.

One potential explanation for this phenomenon is that internet infrastructure and connectivity can vary from city to city. For example, some cities may have more advanced or extensive networks of fiber optic cables, which can support faster internet throughput. In contrast, other cities may have older or less developed infrastructure, resulting in slower internet throughput.

Another possible explanation is that the amount of congestion on the network can vary between cities. For example, a city with a high population density and a large number of internet users may experience more congestion on the network, resulting in slower throughput. In contrast, a city with a lower population density and fewer internet users may have a less congested network, resulting in faster throughput.

Latency

In addition to variations in internet throughput across different cities, our analysis reveals a noteworthy alignment between latency patterns and observed throughput. Latency, defined as the time it takes for data to travel from the user's device to the server and back, plays a pivotal role in internet performance, directly influencing website and application responsiveness.

Our results underscore that users in networks with higher internet throughput consistently experience lower latency when connecting to servers within local networks. This phenomenon is attributed to the reduced data travel distance, resulting in swifter response times. Conversely, users in networks within cities characterized by slower internet throughput reported heightened latency when accessing local network servers, indicating that slower internet speeds could act as a limiting factor in overall internet performance.

In contrast, when accessing servers located in foreign networks, the disparity in latency among Congolese networks was less pronounced. The extensive distance data must traverse in such scenarios diminishes the influence of local internet throughput. Nevertheless, users in networks with superior internet speeds tended to report marginally lower latency even for connections to foreign servers.

In summary, our findings underscore the substantial impact of internet throughput disparities between networks on internet performance, especially concerning latency, when connecting to local network servers with shorter data travel distances. Further research is warranted to delve deeper into the factors contributing to these variations and to explore potential solutions aimed at enhancing internet performance and elevating the overall user experience.

Web performance

a)Impact of Hosting Country's Top-Level Domain Servers Abroad

The geographical location of a country's top-level domain (TLD) server can significantly influence the experience of users within that country. In our investigation focused on the most frequently accessed websites in the Democratic Republic of Congo (DRC), we discovered that all .cd (DRC's TLD) websites are hosted and maintained outside the country. To extend this inquiry, we also assessed the status of other global websites on the list.

One notable consequence of hosting TLD servers abroad is the effect of distance on server performance. The farther the server is from the user, the longer it takes to resolve DNS queries, resulting in delayed website access. This leads to slower loading times, ultimately diminishing the user experience.

Additionally, the performance of a foreign-hosted TLD server can be compromised by network congestion or other issues within the foreign network. High traffic levels or performance challenges in the foreign network can impact the throughput and reliability of DNS queries, causing further delays in website loading times and an overall suboptimal user experience.

Furthermore, hosting and managing a TLD server in a foreign country can raise concerns about security and privacy for users. This includes apprehensions regarding the safety of personal data and the confidentiality of DNS queries when handled by a foreign network. Such concerns can erode trust in the TLD server and the websites utilizing it, potentially affecting user adoption and engagement.

In summation, the decision to host a country's TLD server abroad carries various implications for internet performance and security within that country. It is imperative for nations to carefully assess these impacts and implement measures to optimize user experiences for their citizens.

b) Variations in User Experience Based on Website Hosting Location

Our research underscores significant variations in user experiences when accessing websites hosted locally versus those hosted globally. These disparities depend on multiple factors, including the nature of the website and the user's geographical location.

For instance, users accessing websites in the adult content category generally encounter faster loading times and higher-quality video streaming when connecting to locally hosted websites. The reduced data travel distance results in lower latency and improved performance. Conversely, users accessing globally hosted adult content websites often experience slower loading times and lower-quality video streaming.

A similar pattern is observed when accessing news websites. Users typically enjoy faster loading times and more responsive navigation when visiting locally hosted news websites due to reduced data travel distance. Conversely, globally hosted news websites often lead to slower loading times and less responsive navigation.

In the realms of social media, betting, and entertainment websites, variations in user experience also prevail. Users accessing locally hosted websites in these categories tend to experience quicker loading times and more responsive interactions. Conversely, globally hosted websites in these categories often result in slower loading times and less responsive interactions.

In conclusion, our research underscores substantial disparities in user experiences contingent on whether websites are hosted locally or globally. These variations are influenced by the website type and the user's location. Consequently, it is vital for websites to meticulously consider these distinctions and implement optimization strategies to deliver the best possible user experiences.

7.3 Peering and performance

We explored the state of network interconnections in DRC, we measured the delays between them and the destination of network traffic that is meant to be local between Congolese ASNs. While there is a presence of three Internet Exchange Points in the country, we noticed that the majority of traffic still travels to foreign networks, sometimes outside the continent before they reach the targeted destination hope that we are located in DRC. We discuss below the potential reasons for such behaviors. One potential reason why packets may travel through circuitous routes in foreign countries before reaching their destination IP, even when the most influential network service providers in a country are connected to the same internet exchange points, seems to be the reality of peering between networks.

Peering is the practice of connecting two or more networks together in order to exchange traffic and provide internet services to users. Peering is an important aspect of the internet, as it allows ASNs to share resources and improve their overall performance. However, peering agreements between networks can be complex and can vary depending on a variety of factors.

One of the key factors that can influence peering agreements is the relative size and influence of the networks involved. Networks that are larger and more influential may be able to negotiate more favorable peering agreements, including lower transit fees and better routing options. In contrast, smaller or less influential ASNs may have to accept less favorable peering agreements, including higher transit fees and less direct routing options.

This dynamic can lead to situations where packets may travel through circuitous routes in foreign countries before reaching their destination IP. For example, if a smaller network is sending traffic to a larger network, the traffic may have to pass through one or more intermediate networks in order to reach its destination. This can result in longer routing paths and higher latency, which can negatively impact the overall performance of the internet.

To address this issue, network operators want to consider reviewing and negotiating their peering agreements in order to improve their routing options and reduce transit fees. This can help to improve the overall performance of the internet and enhance users' experiences with the internet. Additionally, service providers want to consider joining internet exchange points, which can provide a central hub for interconnecting networks and facilitating peering agreements. This can help to improve the efficiency of the internet and reduce the need for circuitous routing paths.

There can be several reasons why networks present at the same internet exchange points in the same country may forward packets destined for local peers to foreign networks before they reach their destination. One possible reason is that the networks involved may not have direct peering agreements with each other.

Another possible reason is that network providers may have peering agreements, but the agreements may not be optimized for local traffic. For example, they may have negotiated lower transit fees and better routing options for international traffic, but not for local traffic. In this case, the service providers may choose to route local traffic through foreign networks in order to take advantage of the more favorable peering agreements.

A third possible reason is that network service providers may be experiencing congestion or other performance issues on their local networks. In this case, they may choose to route traffic through foreign networks in order to avoid congestion and improve performance. This can help to improve the overall performance of the internet, but it can also result in longer routing paths and higher latency for local traffic.

To address these issues, service providers may want to consider reviewing and negotiating their peering agreements in order to optimize them for local traffic. This can help to improve the performance of the internet and reduce the need for circuitous routing paths. Additionally, network operators may want to monitor the performance of their local networks and take steps to address any congestion or other issues that may be impacting the performance of the internet.

Chapter 8

CONCLUSIONS

In conclusion, this research aimed to understand the state of Internet performance in developing African regions, specifically in the Democratic Republic of Congo (DRC). Through the analysis of secondary datasets, an online survey, and Internet measurement on personal devices, the research found that there is a lack of transparency in the DRC's Internet structure, with certain networks exerting a significant level of influence on users' Internet connectivity. The historical QoS of networks in the country was found to be subpar, with low upload, and download throughput compared to other Central African Countries.

Participants in the survey reported a low level of satisfaction with the services offered by their ISPs. This indicates that there is a need for improvements in the quality of service offered by ISPs in the DRC. Moreover, the analysis of the primary data collected from the Internet measurement performed on personal devices, showed that the QoS and Web QoE in the DRC are far from being optimal, which is in line with the results obtained from the survey.

It is clear that the DRC faces significant challenges when it comes to Internet infrastructure and connectivity. This is a major issue, as access to the Internet is becoming increasingly important for economic and social development. The Internet is an essential tool for education, communication, and commerce, as well as for accessing information and services. Without reliable and accessible Internet, individuals and businesses in the DRC are at a disadvantage.

In light of these findings, it is recommended that the DRC government and private sector invest in the development of Internet infrastructure and work to increase transparency in the country's Internet structure. This could include measures such as increasing competition among ISPs, providing support for the development of new technologies and networks, and implementing regulations to ensure that ISPs are held accountable for the quality of service they provide. Additionally, more research is needed to understand the specific challenges faced by the DRC in terms of Internet infrastructure and connectivity, in order to develop targeted solutions.

Overall, this research contributes to a better understanding of the state of Internet performance in developing African regions, with more emphasis on the Democratic Republic of Congo. The results of this research can be used by policymakers, researchers, and the private sector to inform decisions related to Internet infrastructure and connectivity in the DRC, and as a basis for future research. With the increasing importance of the Internet in today's world, it is crucial that efforts are made to improve Internet access and quality in developing regions such as the DRC.

8.1 Key findings

In our second question, we attempted to address the unawareness issue by involving users into Internet measurement activities by providing users with the information they need to accurately assess the performance of their Internet service and potentially make informed decisions about which ISP to choose in the future. Additionally, the results of our Internet measurements can provide ISPs with valuable data that can help them improve their performance and compete more effectively in the market. We found that subscribers from the same network are experiencing disparities in Internet speeds depending on the city they are located in. Further research is needed to better understand the factors that contribute to these differences and to explore potential solutions that could help improve Internet speeds and enhance users' experiences with the Internet.

We have shown that there are significant variations in Internet speeds between cities in the same DRC and that these differences affect users' perceptions of

Internet performance. The causes of these differences, and to explore potential solutions that could help improve Internet speeds and enhance users' experiences with the Internet.

How do users perceive Internet performance in DR Congo? The research found that users in the DRC have a low level of satisfaction with the Internet performance in the country. The participants in the online survey reported that Internet connectivity is often slow, and unreliable and that the services offered by their ISPs were not meeting their needs. This indicates that there is a need for improvements in the quality of service offered by ISPs in the DRC.

To what extent are Autonomous systems in the DRC interconnected and how does this affect network performance? The research found that the Autonomous systems in the DRC are not well interconnected, which leads to poor network performance. The analysis of secondary data revealed that certain networks have a significant level of influence on users' Internet connectivity. This lack of transparency in the country's Internet structure is a major obstacle to improving network performance.

What quality of service and quality of experience do broadband networks in the DRC provide and how does it compare to other Central African countries? The primary data collected from the Internet measurement experience performed on personal devices showed that the QoS in the DRC is far from being optimal given the low position the country takes in terms of overall Internet download throughout as seen in Section 6.2. While latency data was not collected in other countries we compared DRC to, its impact is indeed relative to the end-users proximity to the server against which the latency measurement is executed. As far as the quality of experience on the Web is concerned, it has been highlighted that access to locally hosted websites is quicker compared to internationally hosted websites.

In addition, the QoS of broadband networks in the DRC is lower compared to other Central African countries, as presented in Section 6.2 based on metrics such as Upload and download throughput. It also revealed that there is a lack of transparency in the DRC's Internet structure, with certain networks exerting a significant level of influence on users' Internet connectivity.

The above findings indicate that there is a need for increased transparency and improved Internet infrastructure in the DRC, in order to better serve the needs of its citizens and support economic and social development. The lack of Internet transparency and quality of service is a major issue that hinders the development of the country, and it is crucial to address it, in order to improve Internet access and quality in the DRC.

8.2 Future work

In future research, there remains ample room for an in-depth exploration of various facets, including the refinement of our findings in Chapter 4 to encompass a granular breakdown based on subscription type, specifically post-paid versus pre-paid, and differentiation between capped and uncapped subscribers. This endeavor would provide a more nuanced understanding of how diverse user profiles interact with and experience Internet connectivity in the DRC.

Additionally, the augmentation of the Internet measurement aspect of our study presents an exciting prospect. This augmentation can be realized through the incorporation of advanced metrics and the expansion of vantage points. For instance, the study could delve deeper into the quality of service and quality of experience measurements, utilizing innovative metrics such as packet loss, jitter, and bufferbloat. This refined approach would yield a more comprehensive grasp of Internet connectivity performance in the DRC and in the Central African region, affording more precise insights.

An innovative avenue for future research lies in the application of machine learning techniques to scrutinize the extensive Internet measurement dataset. This data-driven approach offers the potential to unearth intricate patterns and trends that might elude human observation, thereby enriching our comprehension of the state of Internet connectivity in Central Africa.

Furthermore, a captivating line of inquiry involves the exploration of the societal impact of Internet access and quality in the DRC. This multifaceted exploration could encompass in-depth case studies of individuals and communities profoundly affected by suboptimal Internet access and quality. Moreover, it should entail a rigorous analysis of critical social indicators, such as education, healthcare, and poverty, to gauge the profound influence of Internet connectivity on social development within the nation.

Lastly, future research endeavors could be dedicated to the development of a real-time monitoring system tailored to monitor Internet connectivity in the DRC. This proactive system could harness Internet measurement data and other relevant information sources to provide timely updates and insights, thereby offering valuable support to stakeholders and policymakers.

Embracing these avenues of investigation will not only contribute to a deeper understanding of Internet connectivity in the DRC but also empower us to address the unique challenges and opportunities inherent in this dynamic landscape.

Bibliography

[1] World Stats. *Internet usage statistics.* (accessed: 01.08.2023). URL: https://www.internetworldstats.com/stats.htm.

[2] Ookla SpeedTest. *Speedtest Global Index.* (accessed: 01.08.2023). URL: https://www.speedtest.net/global-index.

[3] OONI. *Internet Censorship by OONI.* (accessed: 01.08.2023). URL: https://ooni.org/post/.

[4] Lord Kelvin. *Electrical Units of Measurement.* Popular Lectures and Addresses vol. 1. Popular Lectures and Addresses, 1883. ISBN: 9780198520115.

[5] D. MUNORIYARWA MOYO. "'Data must fall': mobile data pricing, regulatory paralysis and citizen action in South Africa, Information, Communication Society". In: 2021, pp. 365–380.

[6] ITU. *Global Connectivity Report.* (accessed: 01.09.2023). URL: https://www.itu.int/hub/publication/d-ind-global-01-2022/.

[7] Internet Society Michael Kende. "Moving Toward an Interconnected Africa:The 80/20 Initiative". In: *Moving Toward an Interconnected Africa:The 80/20 Initiative.* 2021.

[8] D. Thakur T.Woodhouse. "Improving Mobile Broadband Quality of Service in Low and Middle-Income countries". In: *Improving Mobile Broadband Quality of Service in Low and Middle-Income countries.* A4AI WWW Foundation. 2017, pp. 1–23.

[9] A4AI WW Foundation. *Advancing Meaningful Connectivity, Towards Active and Participatory Digital Societies.* (accessed: 03.04.2021). URL: https://a4ai.org/wp-content/uploads/2022/02/ExecsummaryENGLISH.pdf.

[10] Amreesh Phokeer, Melissa Densmore, David Johnson, and Nick Feamster. "A first look at mobile internet use in township communities in south africa". In: *Proceedings of the 7th Annual Symposium on Computing for Development*. 2016, pp. 1–10.

[11] National Research Council et al. *The Internet Under Crisis Conditions: Learning from September 11*. National Academies Press, 2003.

[12] National Research Council. *The Internet Under Crisis Conditions: Learning from September 11*. Washington, DC: The National Academies Press, 2003. ISBN: 978-0-309-08702-5. DOI: 10.17226/10569. URL: https://www.nap.edu/catalog/10569/the-internet-under-crisis-conditions-learning-from-september-11.

[13] Musab Isah, Amreesh Phokeer, Josiah Chavula, Ahmed Elmokashfi, and Alemnew Sheferaw Asrese. "State of internet measurement in Africa-A survey". In: *International Conference on e-Infrastructure and e-Services for Developing Countries*. Springer. 2019, pp. 121–139.

[14] Josiah Chavula, Amreesh Phokeer, Agustin Formoso, and Nick Feamster. "Insight into Africa's country-level latencies". In: *2017 IEEE AFRICON*. IEEE. 2017, pp. 938–944.

[15] J. Livingood N. Feamster. *Measuring Internet Speed: Current Challenges and Future Recommendations*. (accessed: 03.04.2021). URL: https://cacm.acm.org/magazines/2020/12/248801-measuring-internet-speed/fulltext.

[16] MyBroadband. *My Broadband Tool*. URL: http://mybroadband.co.za. (accessed: 25.06.2020).

[17] Jeroen Van Der Ham. "Ethics and Internet measurements". In: *2017 IEEE Security and Privacy Workshops (SPW)*. IEEE. 2017, pp. 247–251.

[18] OONI. *Why measure internet censorship?* URL: https://ooni.org/support/faq/#why-measure-internet-censorship. (accessed: 21.04.2021).

[19] LLC Ookla. "Speedtest global index". In: *Online. Accessed April* (2020).

[20] RIPE NCC. *RIPE Atlas: A global internet measurement network Internet Protocol Journal*. (accessed: 01.05.2020). URL: `https://atlas.ripe.net/`.

[21] Marshini Chetty, Srikanth Sundaresan, Sachit Muckaden, Feamster Nick, and Enrico Calandro. "Measuring broadband performance in South Africa". In: *Proceedings of the 4th Annual Symposium on Computing for Development*. 2013, pp. 1–10.

[22] SpeedChecker Ltd. *SpeedChecker - Helping make the Internet faster for everyone*. (accessed: 22.05.2020). URL: `https://www.speedchecker.com/`.

[23] MLAB. *Measrement LAB*. (accessed: 01.05.2021). URL: `https://www.measurementlab.net/`.

[24] Tidbits. *TUse Apple's network quality Tool to Test Internet Responsiveness*. (accessed: 11.06.2022). URL: `https://tidbits.com/2022/04/22/use-apples-networkquality-tool-to-test-internet-responsiveness`.

[25] Srikanth Sundaresan, Sam Burnett, Nick Feamster, and Walter De Donato. "{BISmark}: A Testbed for Deploying Measurements and Applications in Broadband Access Networks". In: *2014 USENIX Annual Technical Conference (USENIX ATC 14)*. 2014, pp. 383–394.

[26] *Netrics*. (accessed: 13.08.2022). URL: `https://github.com/chicago-cdac/nm-exp-active-netrics`.

[27] A.S. Asrese. *WebLAR: A Web Performance Measurment Tool*. (accessed: 17.05.2021). URL: `https://github.com/alemnew/weblar`.

[28] Alemnew Sheferaw Asrese, Ermias Andargie Walelgne, Vaibhav Bajpai, Andra Lutu, Özgü Alay, and Jörg Ott. "Measuring web quality of experience in cellular networks". In: *International Conference on Passive and Active Network Measurement*. Springer. 2019, pp. 18–33.

[29] A.S. Asrese. *MONROE : Measuring Mobile Broadband Networks in Europe*. (accessed: 17.06.2021). URL: `https://www.monroe-project.eu/`.

[30] Samknows. *Samknows Platform.* (accessed: 19.05.2021). URL: https://samknows.com.

[31] Yasir Zaki and Jay Chen. *Measuring Web Latency and Rendering Performance: Method, Tools Longitudinal Dataset.* (accessed: 09.07.2021). URL: https://zenodo.org/record/2547512#.YOgTuJozZqM.

[32] Alemnew Sheferaw Asrese, Ermias Andargie Walelgne, Andrua Lutu, Özgu Alay, and Jörg Ott. "Measuring web latency in cellular networks". In: *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS).* IEEE. 2018, pp. 1–2.

[33] International Computer Science Institute. *Netalyzr: From Diagnostic to Research Tool.* URL: https://www.icsi.berkeley.edu/icsi/gazette/2011/09/netalyzr-diagnostic-research-tool. (accessed:03.07.2020).

[34] Iljitsch Van Beijnum. *BGP: Building reliable networks with the Border Gateway Protocol.* " O'Reilly Media, Inc.", 2002.

[35] Cloudfare. *What is BGP? — BGP routing explained.* (accessed: 11.08.2021). URL: https://www.cloudflare.com/learning/security/glossary/what-is-bgp/.

[36] Internet Society. *Internet Society: IXPs make the Internet faster and more affordable.* URL: https://www.internetsociety.org/issues/ixps/. (accessed: 09.02.2022).

[37] Nikolaos Chatzis, Georgios Smaragdakis, Anja Feldmann, and Walter Willinger. "There is more to IXPs than meets the eye". In: *ACM SIGCOMM Computer Communication Review* 43.5 (2013), pp. 19–28.

[38] Bernhard Ager, Nikolaos Chatzis, Anja Feldmann, Nadi Sarrar, Steve Uhlig, and Walter Willinger. "Anatomy of a large European IXP". In: *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication.* 2012, pp. 163–174.

[39] Brice Augustin, Balachander Krishnamurthy, and Walter Willinger. "IXPs: mapped?" In: *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement*. 2009, pp. 336–349.

[40] Gillwald Patel P. Chirwa. "The State of Internet Connectivity in Africa". In: *Research ICT Africa*. 2019. URL: https://www.researchictafrica.net/publications/the-state-of-internet-connectivity-in-africa/.

[41] Internet Society. *Measuring the State of the Internet in Africa*. (accessed: 11.03.2021). URL: https://www.internetsociety.org/resources/doc/2016/measuring-the-state-of-the-internet-in-africa/.

[42] Stiemerling Woundy Huitema C. Römer K. "Internet Peering in Africa". In: *Internet Society*. 2014. URL: https://www.internetsociety.org/resources/doc/2014/internet-peering-in-africa/.

[43] PeeringDB. *PeeringDB API*. (accessed: 19.05.2021). URL: https://www.peeringdb.com/apidocs/#operation/list%20fac.

[44] CAIDA. *CAISA AS Relationship dataset*. URL: https://publicdata.caida.org/datasets/as-relationships/serial-1/. (accessed:03.02.2021).

[45] Ookla. *Speedtest by Ookla Global Fixed and Mobile Network Performance Maps*. (accessed: 21.04.2021). URL: https://registry.opendata.aws/speedtest-global-performance/.

[46] A4AI Web Foundation. *2020 Affordability Report*. URL: https://a4ai.org/affordability-report/report/2020/. (accessed:03.06.2021).

[47] Alexa. *Top Sites in Democratic Republic of Congo*. URL: https://www.alexa.com/topsites/countries/CD. (accessed: 25.06.2020).

[48] Alexa. *What is Alexa Rank?* URL: https://blog.alexa.com/marketing-research/alexa-rank/. (accessed: 25.06.2021).

[49] AFRINIC. *African Network Information Centre*. (accessed: 03.07.2020). URL: https://afrinic.net/about.

[50] Agustin Formoso, Josiah Chavula, Amreesh Phokeer, Arjuna Sathiaseelan, and Gareth Tyson. "Deep diving into africa's inter-country latencies". In: *IEEE INFOCOM 2018-IEEE Conference on Computer Communications.* IEEE. 2018, pp. 2231–2239.

[51] J Gilmore, N Huysamen, and A Krzesinski. "Mapping the African internet". In: *Proceedings Southern African Telecommunication Networks and Applications Conference (SATNAC), Mauritius.* Citeseer. 2007.

[52] Arpit Gupta, Matt Calder, Nick Feamster, Marshini Chetty, Enrico Calandro, and Ethan Katz-Bassett. "Peering at the internet's frontier: A first look at isp interconnectivity in africa". In: *International Conference on Passive and Active Network Measurement.* Springer. 2014, pp. 204–213.

[53] Josiah Chavula, Nick Feamster, Antoine Bagula, and Hussein Suleman. "Quantifying the effects of circuitous routes on the latency of intra-Africa internet traffic: a study of research and education networks". In: *International Conference on e-Infrastructure and e-Services for Developing Countries.* Springer. 2014, pp. 64–73.

[54] Bradley Huffaker and Evi Nemeth. "Tools to visualize the internet multicast backbone". In: *International Networking Conference (INET).* 1999.

[55] Chantal Yang, S. Hussein, and Chavula J. "A topology visualisation tool for National Research and Education Networks in Africa". In: *2016 IST-Africa Week Conference.* IEEE. 2016, pp. 1–11.

[56] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira. "Avoiding traceroute anomalies with Paris traceroute". In: *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement.* 2006, pp. 153–158.

[57] M Azizur Rahman, Algirdas Pakstas, and Frank Zhigang Wang. "Network topology generation and discovery tools". In: *Proc. of the 7th EPSRC An-*

nual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting (EPSRC PGNet 2006). 2006, pp. 26–27.

[58] Matthew Luckie, Young Hyun, and Bradley Huffaker. "Traceroute probe method and forward IP path inference". In: *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*. 2008, pp. 311–324.

[59] Yasir Zaki, Jay Chen, Thomas Pötsch, Talal Ahmad, and Lakshminarayanan Subramanian. "Dissecting web latency in ghana". In: *Proceedings of the 2014 Conference on Internet Measurement Conference*. 2014, pp. 241–248.

[60] Présidence de la République Démocratique du Congo. "Plan National du Numérique Horizon 2025 Pour une RDC connectée et performante". In: *Plan National du Numérique Horizon 2025*. Présidence de la République Démocratique du Congo. 2019, pp. 1–130.

[61] World Bank. *République démocratique du Congo - Vue d'ensemble*. (accessed: 08.14.2020). URL: https://www.banquemondiale.org/fr/country/drc/overview.

[62] World Food Programme. *Rapport Annuel Observatoire du Marché de la Téléphonie Mobile*. (accessed: 13.06.2021). URL: http://arptc.gouv.cd/wp-content/uploads/2020/06/Rapport-observatoire-de-marche%5C%CC%5C%81-T4-2019.pdf.

[63] ISPA. *Internet Service Provider Association of DRC*. (accessed: 13.05.2021). URL: http://www.ispa-drc.cd/index.php?option=com_content&view=article&id=24&Itemid=26.

[64] LiquidFb. *Liquid Technologies Facebook Partner to Lay 2,000 Km of Fibre in DRC*. (accessed: 07.07.2021). URL: https://kenyanwallstreet.com/liquid-technologies-to-lay-fibre-in-drc/.

[65] techradar. *What is FTTx, FTTP, and FTTN?* (accessed: 07.07.2022). URL: https://www.techradar.com/news/what-is-fttx-fttp-and-fttn.

[66] Airtel. (accessed: 07.09.2021). URL: https://www.airtel.cd/.

[67] Vodacom. (accessed: 07.09.2021). URL: https://www.vodacom.cd/.

[68] Orange. (accessed: 07.09.2021). URL: https://www.orange.cd/.

[69] Africell. (accessed: 07.09.2021). URL: https://www.africell.cd/.

[70] APNIC. *AS Customer Population*. URL: https://stats.labs.apnic.net/cgi-bin/aspop?c=CD. (accessed: 12.06.2021).

[71] ARPTC. *Observatoire dU Marche de la Telephonie Mobile :Rapport du 2ème Trimestre 2021*. URL: https://arptc.gouv.cd/app/uploads/2022/02/ARPTC_Rapport-du-2eme-trimestre-2021.pdf. (accessed: 12.06.2021).

[72] Routeviews. *Routeviews project*. URL: https://routeviews.org/. (accessed: 21.04.2022).

[73] RIPE NCC. *RIS*. (accessed: 01.05.2021). URL: https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris.

[74] Archipelago. *Archipelago (Ark) Measurement Infrastructure*. URL: https://www.caida.org/projects/ark/. (accessed:03.02.2021).

[75] Stacey Omeleze and Hein Venter. "Architectural requirements specifications for designing digital forensic applications". In: *European Conference on Cyber Warfare and Security*. Academic Conferences International Limited. 2016, p. 405.

[76] Aric Hagberg and Drew Conway. "NetworkX: Network Analysis with Python". In: *URL: https://networkx. github. io* (2020).

[77] Vasileios Giotsas and Shi Zhou. "Improving the discovery of IXP peering links through passive BGP measurements". In: *2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE. 2013, pp. 121–126.

[78] IIJ Research Lab. *Internet Health Report,* (accessed: 11.04.2022). URL: https://ihr.iijlab.net/ihr/en-us/api.

[79] Romain Fontugne, Anant Shah, and Emile Aben. "The thin bridges of as connectivity: Measuring dependency using as hegemony". In: *International Conference on Passive and Active Network Measurement*. Springer. 2018, pp. 216–227.

[80] semrush. *Trending websites by country*. (accessed: 11.11.2022). URL: `https://www.semrush.com/trending-websites/cd/all`.

[81] Curl. *Libcurl API's* $curl_easy_getinfo-Extractinformationfromacurlhandle$. (accessed: 11.09.2022). URL: `https://curl.se/libcurl/c/curl_easy_getinfo.html`.

[82] Curl. *Libcurl getinfo options*. (accessed: 11.09.2022). URL: `https://curl.se/libcurl/c/easy_getinfo_options.html`.

[83] Dieudonne Munganga. *Survey on DRC Internet users' perceptions on Internet performance*. (accessed: 01.10.2022). URL: `https://survey.cs.uct.ac.za/limesurvey/index.php/634178?lang=fr`.

[84] CAIDA. *AS Rank*. (accessed: 01.05.2021). URL: `https://asrank.caida.org/`.

[85] RIPE Labs. *How Is Russia Connected To The Wider Internet?* `https://labs.ripe.net/author/emileaben/how-is-russia-connected-to-the-wider-internet/`. (accessed: 15.04.2022).

[86] *Betweenness Centrality*. URL: `https://en.wikipedia.org/wiki/Betweenness_centrality`. (accessed: 28.03.2022).

[87] Timm Böttger, Gianni Antichi, Eder L Fernandes, Roberto di Lallo, Marc Bruyere, Steve Uhlig, and Ignacio Castro. "The elusive internet flattening: 10 years of IXP growth". In: *CoRR* (2018).

[88] Ookla SpeedTest. *Ookla Ranks Internet Performance in the World's Largest Cities, Plus Read 34 New Ookla Market Reports*. (accessed: 01.08.2023). URL: `https://www.ookla.com/articles/global-index-cities-announcement`.

[89] Dieudonne M. *Measurement script for TCPtimings*. (accessed: 11.06.2022). URL: `https://github.com/monsieurpapa/TCPtimings`.

8.3 Appendix

LimeSurvey

Charger un questionnaire non terminé   Finir plus tard

Ce questionnaire n'est actuellement pas activé. Vous ne pourrez pas sauver vos réponses. ×

Accès à Internet :La qualité d'experience des utilisateurs de la RDC

Bienvenue et merci de faire partie de cette recherche.

L'objectif de cette enquête est de déterminer la perception de la qualité d'expérience des internautes congolais lors de l'accès à Internet à travers les réseaux mobiles aux quels ils sont connectés. Cette enquête vise également à identifier le niveau auquel les utilisateurs d'Internet en RDC sont conscients de l'importance du mesurage d'Internet , et leur offre des connaissances sur les pratiques de mesurage d'Internet, et comment ils peuvent contribuer à la comprehension de l'Etat d'Internet dans le pays grâce aux données collectées.

Welcome and thank you for beng part of ts research. The aim of this survey is to determine the perception of the quality of experience of Congolese Internet users when accessing the Internet through the mobile networks to which they are connected. This survey also aims to identify the level at which internet users in the DRC are aware of internet metering and provide them with knowledge about internet metering practices, and how they can contribute to understanding the state of internet metering. Internet in the country thanks to the data collected.

Cette recherche a été approuvée par le comité d'éthique de la faculté des sciences de l'Université du Cap en Collaboration avec un chercheur de l'Université de Kinshasa, avec le code d'approbation : FSREC 002-2020.

This research was approved by the University of Cape Town Faculty of Science Ethics Committee in collaboraton with a reseacher from the University of Kinshasa, with approval code: FSREC 002-2020.

Il y a 27 questions dans ce questionnaire.

DÉMOGRAPHIE

Demographie du participant

Nom & Post-nom /Firstname and Surname

☐

*Genre /Gender

Cochez la ou les réponses

☐ Masculin

☐ Féminin

☐ Autres

À quelle tranche d'âge appartenez-vous ?/ Which age group do you belong to?

Cochez la ou les réponses

☐

☐ 18-25

☐ 26-35

☐ 36-45

☐ 46-55

☐ 56+

✱Dans quelle province residez-vous ? / In which provide do you reside?

Cochez la ou les réponses

- [ ] Kongo Central
- [ ] Kwango
- [ ] Kwilu
- [ ] Maï-Ndombe
- [ ] Kasaï
- [ ] Kasaï-Central
- [ ] Kasaï-Oriental
- [ ] Lomami
- [ ] Sankuru
- [ ] Maniéma
- [ ] Sud-Kivu
- [ ] Nord-Kivu
- [ ] Ituri
- [ ] Haut-Uélé
- [ ] Tshopo
- [ ] Bas-Uélé
- [ ] Nord-Ubangi
- [ ] Mongala
- [ ] Sud-Ubangi
- [ ] Équateur
- [ ] Tshuapa
- [ ] Tanganyika
- [ ] Haut Lomami
- [ ] Lualaba
- [ ] Haut Katanga
- [ ]

✱De quelle ville de la DRC participez-vous ? / From which town in DRC are you participating?

✱Quelle est votre occupation PRINCIPALE ? / What is your PRIMARY occupation?

Cochez la ou les réponses

- [ ] Etudiant/Student
- [ ] Travailleur Independant/Self-employed
- [ ] Employé/Employed
- [ ] Chomeur/Unemployed
- [ ] Autre :

SPÉCIFICATIONS DE L'APPAREIL / DEVICE SPECIFICATIONS

Cette partie couvre les details sur comment vous accedez à Internet / This part covers your connectivity means to the Internet

***Quel(s) appareil(s) utilisez-vous le plus pour accéder à Internet ?/ Which device do you mostly use to access the Internet?**

Cochez la ou les réponses

- ☐ À la fois smartphone et ordinateur portable / PC and Smartphone
- ☐ Bureau / Desktop
- ☐ Ordinateur portable /Laptop
- ☐ Téléphone intelligent /Smart phone
- ☐ Tablette /Tablet

***Quelles applications utilisez-vous le plus pour acceder à Internet ? / What applications do you use the most to access the Internet?**

Cochez la ou les réponses

- ☐ Navigateur Chrome / Chrome Web Browser
- ☐ Navigateur Mozilla / Mozilla Web Browser
- ☐ Navigateur Safari / Safari Web Browser
- ☐ Youtube
- ☐ Facebook
- ☐ Whatsapp
- ☐ Twitter
- ☐ Instagram
- ☐ Snapchat
- ☐ Tiktok
- ☐ Zoom
- ☐ Autre : ____

*À quelle fréquence utilisez-vous Internet pour les activités suivantes: / How often do you use the internet for the following activities:

	Souvent /Often	Quelquefois /Sometimes	Jamais /Never
	○	○	○
Emails	○	○	○
Information	○	○	○
Achats en Ligne - Online Shopping / Auctions	○	○	○
Travail /Business/ Etudes - Work / Business / Studies	○	○	○
Jeux - Games	○	○	○
Téléchargement des chansons, films - Downloading Music, Films, etc	○	○	○
Regarder des vidéos diffusées en direct - Watch lives-creamed videos	○	○	○

*Les quels de ces aspects influencent le plus votre qualite d'experience sur Internet

*Les quels de ces aspects influencent le plus votre qualité d'experience sur Internet

	Extremement important /Extremely important	Important	Pas important / Not important
Le temps de chargement d'un site web / The loading time of a website	○	○	○
Le nombre de gels pendant la vidéo en direct / The number of freezes during live video	○	○	○
La fréquence des décrochages lors d'un appel vocal / The frequency of dropouts during a voice call	○	○	○
La fréquence des décrochages lors d'un appel vidéo /The frequency of dropouts during a video call	○	○	○
Le Temps d'envoie/reception des messages / Sending/Receiving time of messages	○	○	○

Comment evaluez-vous votre experience avec les applications ci-dessous / How would you rate your experience with the apps below

	Excellent /Excellent	Tres bon /Very Good	Bon / Good	Mauvais / Bad	Tres Mau-vais / Very Bad	je ne sais pas/ I don't know	Sans réponse
Navigateur Chrome / Chrome Web browser	○	○	○	○	○	○	●
Navigateur Mozilla / Mozilla Web browser	○	○	○	○	○	○	●
Navigateur Safari / Safari Web browser	○	○	○	○	○	○	●
Youtube	○	○	○	○	○	○	●
Zoom	○	○	○	○	○	○	●
Facebook	○	○	○	○	○	○	●
Whatsapp	○	○	○	○	○	○	●
Twitter	○	○	○	○	○	○	●
Instagram	○	○	○	○	○	○	●
Tiktok	○	○	○	○	○	○	●
Snapchat	○	○	○	○	○	○	●

DETAILS DE LA CONNECTION / CONNECTION DETAILS

Dans cette partie nous cherchons à savoir par quel moyen vous vous connectez à Internet

***Par quel moyen de connectivité accédez-vous principalement à Internet ?/ By what connectivity means do you mostly access the Internet?**

Cochez la ou les réponses

- [ ] Wi-Fi
- [ ] Réseau cellulaire(forfait de données de téléphonie mobile) /Cellular Network (mobile phone data plan)
- [ ] Fibre Optique /FiberCable
- [ ] ADSL/DSL : Asymmetric digital subscriber line
- [ ] Satellite
- [ ] Fixe sans fil/Fixed wireless
- [ ] Je ne sais pas / I don't know

*A quel fournisseur de réseau êtes-vous abonné ?/ To which network service provider are you subscribed to?

Cochez la ou les réponses

- [ ] Airtel
- [ ] Vodacom
- [ ] Orange
- [ ] Supercell
- [ ] Autres/Other

*Pouvez-vous nous dire la nuture de votre connectivité à Internet ? Avez-vous un accès Internet plafonné (une quantité définie de données mobiles est disponible) ou non plafonné (un nombre illimité de données est disponible) ? / Please tell us the nature of your connectivity to the Internet. Do you have capped (a set quantity of data is available) or uncapped (unlimited data is available) internet access?

Cochez la ou les réponses

- [ ] plafonné / capped
- [ ] non plafonné / uncapped
- [ ] Pas certain / Not sure

*A quel niveau êtes-vous concerné de savoir si votre fournisseur d'accès Internet fournit la vitesse Internet qu'il promet ? / How concerned about knowing if your Internet provider provides the advertised Internet speed?

Cochez la ou les réponses

- [ ] Très concerné/Very concerned
- [ ] Concerné / Concerned
- [ ] Pas concerné / Not concerned
- [ ] Je m'en fiche / I don't care

*Veuillez choisir une option: Compte tenu de la performance, du coût et de tout autre facteur liés à la connectivité, êtes-vous satisfait de votre fournisseur de services Internet ?Quel facteur vous affecte le plus? / Please choose one option: Considering performance, cost, and any other factors – how satisfied are you with this Internet service provider? Which factor affects you the most?

Ajoutez un commentaire seulement si vous sélectionnez la réponse.

- [ ] Très satisfait Very satisfied ____________
- [ ] Satisfait / Satisfied ____________
- [ ] Neutre / Neutral ____________
- [ ] Pas satisfait / Dissatisfied ____________
- [ ] Très insatisfait / Very dissatisfied ____________

*Avez-vous été confronté à des interruptions d'Internet aujourd'hui ? Si oui, à quelle heure de la journée ? / Have you faced internet disruptions today? If yes, what time of the day?

Cochez la ou les réponses

- [ ] Non / No
- [ ] Matin / Morning
- [ ] Après-midi / Afternoon
- [ ] Soir / Evening
- [ ] Tout le temps / All the time

***À quelle fréquence êtes-vous confronté à des interruptions d'Internet par semaine (/7) ? / How often do you face internet disruption per week (/7)?**

Cochez la ou les réponses

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7

***En pensant aux applications que vous utilisez principalement pour accéder à Internet, quels sont les problèmes techniques les plus courants auxquels vous êtes confronté ? / Thinking of the Apps you mostly use to access the Internet, what are the most common technical issues you face?**

Cochez la ou les réponses

- ☐ Retards d'accès aux sites Web / Delays in accessing websites
- ☐ Impossible de regarder des vidéos en direct/ Unable to watch live videos
- ☐ Impossible de télécharger/d'envoyer une image ou une vidéo / Unable to download and Send an image or video
- ☐ Impossible de faire un appel vidéo / Unable to have a video call
- ☐ Retard dans l'envoi des messages / Delay in sending text messages
- ☐ Autre

***À votre connaissance, quelle vitesse d'Internet votre fournisseur d'accès Internet vous promet-il environ ? / To the best of your knowledge, approximately what speed does your Internet service provider promise you?**

NOTE : Mbps : Méga octets par Seconde / Megabits per second

Cochez la ou les réponses

- ☐ 1 à 3 Mbps / 1 to 3 Mbps
- ☐ 5 à 10 Mbps / 5 to 10 Mbps
- ☐ Moins de 25 Mbps / Less than 25 Mbps
- ☐ Plus de 1Gbps / More than 1Gbps
- ☐ je ne sais pas /I don't know

***Au meilleur de votre connaissance, environ, À quelle vitesse Internet pensez-vous que votre fournisseur vous sert réellement ? / To the best of your know-ledge, approximately, What Internet speed do you think your provider actually serves you?**

Cochez la ou les réponses

- ☐ 1 à 3 Mbit/s / 1 to 3 Mbps
- ☐ 5 à 10 Mbit/s / 5 to 10 Mbps
- ☐ Moins de 25 Mbps / Less than 25 Mbps
- ☐ Plus de 1Gbps / More than 1Gbps
- ☐ je ne sais pas / I don't know

***Combien de fois ou combien de temps utilisez vous Internet par jour / How often or how long do you use the internet per day?**

*Combien de fois ou combien de temps utilisez-vous Internet par jour / How often or how long do you use the internet per day?

Cochez la ou les réponses

☐

MESURAGE D'INTERNET

Cette partie cherche à savoir votre connaissance sur le mesurage d'Internet et son utilité / This part of the survey seeks to know your level of awareness of Internet measurements and its importance

***Avez-vous déjà entendu parlé du mesurage de l'Internet / Have you ever heard of Internet Measurements before?**

Cochez la ou les réponses

☐ Oui / Yes

☐ Non / No

☐ Peut-etre / Maybe

***Avez-vous déjà mesuré la vitesse de votre Internet auparavant ? Si oui,Quel outil aviez-vous utilisé ? Have you ever measured the speed of your Internet before? What tool did you use?**

Ajoutez un commentaire seulement si vous sélectionnez la réponse.

☐ Oui / Yes

☐ Non / No

☐ Je ne me souviens / I don't remember

***Est-ce important pour vous de connaître la vitesse d Internet que vous recevez de votre fournisseur ? Si oui, pourquoi ?veuillez ajouter un commentaire supplementaire si possible. / Is it useful for you to know the speed of your Internet? If yes, why? please select add nay additional comments**

Ajoutez un commentaire seulement si vous sélectionnez la réponse.

☐ Pour comparer la vitesse de mon fournisseur à celle de ses concurents / To compare my provider's speed to other providers'

☐ Pour comprendre pourquoi mon Internet est lent / To understand why my Internet is slow

☐ Suggérer à un nouvel arrivant à quel fournisseur de réseau il doit s'abonner / To suggest a newcomer which network provider he/she should subscribe to

☐ Autre / Other

***Considérons que vous êtes intéressé à savoir la vitesse de votre Internet, quelles actions souhaiteriez-vous entreprendre ? Let's consider that you are interested in knowing the speed of your Internet, What actions would you wish to take?**

Cochez la ou les réponses

☐ Installer une application sur mon appareil et lancer moi-même les mesures / Install an App on my device and launch measurements myself

☐ Installer une application sur mon appareil qui mesure la vitesse en arrière-plan / Install an App on my device that measures the speed in the background

☐ Installer un programme sur mon ordinateur portable qui mesure automatiquement la vitesse d'Internet / Install a program that measures the Internet speed automatically on my laptop

- [ ] Brancher une sonde matérielle sur mon routeur domestique/bureau / Plug a hardware probe on my home/office router
- [ ] Accéder à un site Web qui exécute les mesures une fois ouvert / Access a website that executes the measurements once opened

Seriez-vous intéressé à participer à une interview pour une autre expérience sur votre qualite d'experience sur Internet ? Si oui, veuillez laisser vos coordonnées ci-dessous / Would you be interested in participating in an interview for another experiment on your quality of experience on the Internet? If yes, please leave your contact details below

Envoyer