MAPPING THE AFRICAN INTERNET TOPOLOGY

Maganyase Elias Sihlangu

Shlmag002@myuct.ac.za

ABSTRACT

This literature review focuses on the mapping and visualization of internet infrastructure and interconnections, with a specific focus on the African region. The review highlights different approaches to mapping internet infrastructure, such as utilizing BGP datasets and traceroute measurements to infer logical paths between networks and countries and mapping the physical and logical topology onto a map. Additionally, the review explores the diversity of physical and logical routes between networks and evaluates the utility of visualization platforms for this purpose. The literature review draws from research papers that cover topics such as ISP interconnectivity, BGP-multipath routing, topology visualization tools for national research and education networks, and cross-AS (X-AS) internet topology mapping. The insights and methods presented in these papers could inform the data analysis, visualization, and evaluation aspects of similar projects. Overall, this literature review provides a valuable resource for researchers interested in mapping and visualizing internet infrastructure and interconnections, particularly in the African region.

I. INTRODUCTION:

As the internet continues to evolve and expand, understanding the underlying infrastructure that supports it has become increasingly crucial. Mapping the physical and logical topology of the internet infrastructure can provide valuable insights into the network's connectivity, performance, and potential bottlenecks [1]. Over the years, researchers have utilized various techniques such as BGP datasets, traceroute measurements, and network simulation to map the internet infrastructure. As a result, several studies have been published on mapping the internet infrastructure in different regions worldwide.

One of the most critical aspects of mapping the internet infrastructure is the identification and visualization of the relationships between internet service providers (ISPs) and ASes across different countries and regions [2]. Understanding the relationships between ISPs is essential for identifying points of interconnectivity, potential bottlenecks, and potential areas for improvement in network connectivity [1] [2].

Despite the various approaches to mapping the internet infrastructure, there is still a need for further research, particularly in Africa. Mapping the internet topology in Africa presents unique challenges due to the continent's diverse geography, limited infrastructure, and the prevalence of mobile networks [3]. Therefore, there is a need for a comprehensive overview of mapping the internet topology in Africa that focuses on techniques used to infer and visualize relationships between ISPs across the continent [3] [4].

The primary objective of this research project is to infer and visualize relationships between Internet Service Providers (ISPs) across Africa. By analyzing Border Gateway Protocol (BGP) datasets and Traceroute measurements, we aim to map the logical paths between networks and countries. Additionally, we will map the physical and logical topology of the African internet infrastructure, including cables and Internet Exchange Points (IXPs), and visualize them on a map. This research aims to evaluate the diversity of physical and logical routes between networks and assess the utility of such a visualization platform.

Mapping and visualising the internet topology is becoming increasingly important due to the rise of global connectivity and the increasing dependence on the internet for commerce, and communication, among other things. Accurate mapping and visualization of the internet infrastructure can help policymakers, researchers, and network operators understand network interconnectivity, identify potential points of failure or congestion, and optimize network routing. This is particularly important for Africa, where internet connectivity and access are still developing, and where understanding the internet's physical and logical topology is crucial for improving access and connectivity.

This literature review aims to provide a comprehensive overview of research on mapping the internet infrastructure. The review will explore the diverse approaches used to map the physical and logical topology of the internet infrastructure and evaluate the utility of such visualization platforms, including tools used to map and visualize the infrastructure.

The subfields of internet topology analysis, network route discovery, IP alias resolution, IP geolocation, traceroute, BGP, and visualization techniques are interconnected, with each contributing to a better understanding of the internet topology. Network route discovery and topology analysis provides a foundation for understanding the physical and logical connections between networks, while geolocation and traceroute techniques help to identify the geographic location of network devices.

II. NETWORK ROUTE DISCOVERY

Discovering and mapping network routes is an important aspect of understanding the logical topology of networks. This involves identifying the paths taken by data packets as they traverse from one network to another, and ultimately reach their destination. The process of route discovery involves several steps, which include [2]-[4]:

1. Traceroute: Traceroute is a commonly used tool for discovering network routes. It works by sending packets with increasing Time To Live (TTL) values towards the destination, and monitoring the response times and source IP

addresses of the intermediate routers. By analyzing the responses, traceroute can infer the path taken by the packets through the network.

2. BGP: The Border Gateway Protocol (BGP) is used by ISPs to exchange routing information and determine the best path for traffic to take through the internet. By analyzing BGP routing tables, it is possible to map out the logical paths between networks and identify the relationships between them.
3. Network Topology Discovery: Network topology discovery involves actively probing the network to discover the devices and links that make up the network. This can be done using tools such as SNMP and LLDP, which can extract information from network devices such as switches and routers.

Challenges in discovering and mapping routes include the complexity of modern networks, which often involve multiple paths and redundant links, as well as the presence of network address translation (NAT) devices that can obscure the true source and destination of traffic. Additionally, some ISPs may deliberately obfuscate their routing information for security or business reasons [2]-[4].

Previous literature has explored various techniques for route discovery and mapping, including the use of traceroute, BGP, and network topology discovery. For example, a study by Motamedi et al. [2] used traceroute and BGP data to infer the interconnections between networks in the internet, while another study by Nur and Tozal [5] proposed a method for mapping the topology of the internet using cross-AS (autonomous system) measurements.

However, there are limitations to the accuracy and completeness of these approaches, particularly when it comes to mapping the physical topology of the network. For example, traceroute can only provide information about the logical paths taken by packets, and may not reflect the physical paths taken by traffic due to network optimization techniques such as traffic engineering. Additionally, BGP data may not capture all routes taken by traffic, particularly in cases where traffic is redirected due to network congestion or routing policy changes [2]-[5].

III. IP ALIAS RESOLUTION TECHNIQUES

The process of alias resolution plays a crucial role in creating Internet topology maps using traceroute. This task involves identifying the IP addresses that belong to a specific router. Accurately identifying these addresses is vital for creating an accurate representation of the network topology. If there are any inaccuracies in the alias resolution, the topology map will be incomplete or incorrect, which can result in misleading insights and analysis. Therefore, it is essential to ensure the accuracy of alias resolution to create reliable and useful Internet topology maps [6][7].

Keys et al. [6] proposes a scalable and efficient algorithm for resolving IP aliases. The approach attempts to determine whether connectivity patterns between IP addresses can reveal whether they belong to the same interface. MIDAR is used to analyse network-level features and infer aliasing relationships. MIDAR obtains IP ID time series data from various locations and applies the Monotonic Bounds Test (MBT) to identify IP addresses that are likely aliases of the same router. The MBT utilizes monotonicity rather than proximity, resulting in a high PPV and a low false positive rate. MIDAR employs multiple probing methods to increase target responsiveness and expand the coverage of IP ID-based techniques. It also uses a sliding-window scheduling algorithm and multiple vantage points for probing scalability [6]. To evaluate the performance of their approach, the Keys et al. [6] conducted experiments on a large-scale dataset of addresses. They compared MIDAR to previous methods and showed that their approach achieves high accuracy in identifying aliases. The results indicate that MIDAR is capable of identifying aliases with high precision and recall [6].

In contrast Gunes and Sarac [7] proposes a method for identifying IP aliases based on analysing the hop-by-hop delays of traceroute measurements. The paper's methodology consists of two main steps: (1) identifying potential alias sets and (2) resolving the aliases. In the first step, Gunes and Sarac et al. [7] use a clustering algorithm to group together IP addresses that have similar hop-by-hop delay patterns. The resulting clusters represent potential alias sets, which may contain multiple IP addresses that correspond to the same physical interface. In the second step, the authors use additional measurements to verify the presence of aliases within each set. Specifically, they send multiple traceroute packets to each IP address within a set and analyse the resulting delay patterns. If two or more IP addresses within a set exhibit similar delay patterns, they are considered to be aliases of each other.

Gunes and Sarac [7] evaluate their method using a dataset of over 15,000 traceroute measurements from 138 Autonomous Systems (ASes). They compare their results to those obtained by a previous method that used reverse DNS lookups to identify aliases Gunes and Sarac [7] report that their method outperforms the previous method in terms of accuracy, precision, and recall. They also demonstrate the effectiveness of their method in building Internet maps, showing that it can significantly improve the accuracy of inferred AS-level topologies.

IV. IP Geolocation Techniques

The resulting data from traceroute probes include IP address, geographically locating these IP addresses is an important task in network topology mapping. The representation of the network devices associated with the resulting IP addresses on a geographic map form part of topology visualisation. In this section, various techniques used to geographically locate IP addresses are explored.

One popular technique for geolocation is based on the use of publicly available geolocation databases such as MaxMind and IP2Location. These databases contain information about the location of IP addresses based on various criteria such as IP range, country, region, and city. The accuracy of these databases varies depending on the quality and frequency of the data updates. Studies have shown that the accuracy of these databases can range from a few kilometres to a few hundred meters depending on the region and the level of granularity required [4][8][9].

Another technique for geolocation is based on the use of network measurement tools such as traceroute and ping. These tools send packets to the target IP address and measure the time it takes for the packets to reach their destination. Based on the time taken and the routing information obtained, the geographic location of the target IP address can be estimated. However, this technique can be affected by several factors such as network congestion, routing changes, and the presence of firewalls and other security measures [10][11].

Machine learning algorithms have also been used for IP geolocation. These algorithms use various features such as latency, routing information, and network topology to build models that can accurately predict the geographic location of IP addresses. Studies have shown that machine learning algorithms can achieve high accuracy in geolocation, with errors of less than 10 kilometres in some cases [11]. The following machine learning techniques to geolocate IP addresses have been studies in previous literature:

1. The Dan et. [11] propose a new method for IP geolocation that leverages the reverse DNS mapping of IP addresses to identify their corresponding geographical location. Reverse DNS mapping works by mapping an IP address to a domain name, which can then be resolved to obtain information about the domain's location. Domain names often contain information about their geographical location, such as the country code top-level domain (ccTLD) or the organization's name or location in the domain name. The IP addresses and domain names obtained from the DNS look-up is then used to develop a machine learning model that predicts the geographic location of IP addresses based on their associated domain names [11]-[13].
2. Another involves using a dataset of traceroute measurements from a large number of Internet hosts. The traceroute measurements are used to build a network topology map of the Internet, which is then used to train a machine learning model. The model is trained to predict the physical location of routers based on their IP addresses and network topology information [14].

There are also hybrid approaches that combine multiple techniques to improve the accuracy of IP geolocation. For example, some studies have proposed using a combination of geolocation databases and network measurement tools to improve accuracy Gupta et al [3].

V. IP to AS Mapping

IP to AS mapping is necessary in internet topology visualization as it allows the identification of the Autonomous System (AS) that controls a particular IP address. ASes are critical components of the internet infrastructure that provide connectivity and routing services to their networks and customers. Understanding the relationships between ASes and their interconnections is crucial for visualizing the internet topology accurately. By mapping IP addresses to their corresponding ASes, it is possible to determine the network pathways between ASes and the routes that data takes when traversing the internet. Most of the methodologies explored in this literature review leverage IP to AS mapping to achieve topology visualisation [2][4].

In the study by Dodge and Kitchin [15], the authors examine various approaches for mapping internet infrastructure, including IP to AIP to AS mapping is challenging due to the dynamic nature of the internet and the potential for multiple ASes to be associated with a single IP address. The authors explore several methods for IP to AS mapping, including WHOIS and BGP data, and conclude that a combination of techniques is necessary for accurate mapping.

Li et al. [16] proposed the use of BGP multipath routing to map IP addresses to AS numbers. They argued that the traditional BGP routing approach leads to suboptimal routing and proposed the use of multipath routing to increase network performance and improve IP to AS mapping accuracy. Claffy et al. [1] discussed the evolution of internet mapping from art to science, highlighting the need for more accurate and efficient mapping techniques.

Claffy et al. [1] also discuss the challenges of mapping the internet, including IP to AS mapping. They describe a framework for internet mapping that involves multiple steps, including data collection, cleaning, and analysis. The authors highlight the importance of collaboration between researchers and network operators in order to improve the accuracy and completeness of internet maps.

VI. The use of traceroutes and BGP data in network topology visualisation.

Traceroute datasets provide information on the path that packets take through the network from a source to a destination. It works by sending a series of packets with increasing time-to-live (TTL) values, which causes the packets to be dropped at each successive hop in the path. The traceroute tool then sends additional packets with the same TTL value, allowing it to build a map of the path that the packets take through the network. This information can be used to map the topology of the network by identifying the ASes and links between them. The accuracy of the mapping depends on the number and location of traceroute vantage points used to collect data, as well as the processing techniques used to analyse the data [7].

One approach to traceroute mapping is to use machine learning algorithms to classify the links based on their characteristics, such as latency, jitter, and loss rate. This approach was used in the study by Yang et al. [17] to develop a topology visualization tool for National Research and Education Networks in Africa. The tool was designed to help network administrators visualize the topology of their networks and identify potential performance issues.

BGP datasets, on the other hand, provide information about the routes that traffic takes between different networks on the internet. The protocol BGP is used by routers to exchange information about available routes and to determine the best path for traffic to take between networks. provide information on the routing policies and interconnections between ASes in the Internet. BGP datasets can be used to infer logical paths between networks and countries, and to identify potential bottlenecks or points of congestion [2][4][15].

Researchers typically rely on public BGP routing tables, which are collected and maintained by regional Internet registries and academic research networks. The data is then cleaned and processed to obtain information about the AS-level connectivity between networks, and the resulting graph is used to infer the logical topology of the Internet. BGP data is collected from routers and exchanged between ISPs to determine the most efficient path for traffic between ASes [2][4][15].

However, BGP data is not always complete or accurate, as some ASes may not publicly advertise their routes or may use private peering agreements that are not visible to external vantage points. Despite these limitations, BGP data has been widely used in network topology mapping studies due to its availability and the valuable insights it can provide [2][4][15].

Once the traceroute and BGP data has been processed, the resulting graphs can be visualized using a variety of tools. The graphs can be mapped onto geographical locations to create a visualization that shows the logical and physical topology of the network. This can be done using tools such as Gephi or Cytoscape, which provide network visualization capabilities. The resulting visualizations can help researchers gain insights into the topology of the network, identify key network infrastructure, and assess the robustness of the network to failures or attacks [4][17][18].

VII. Design considerations in developing topology visualisations tools.

Network topology visualisation is an essential tool for understanding and managing complex computer networks. The design considerations for creating effective network topology visualisations can be broadly categorised into three areas: data representation, layout algorithms, and user interaction.

Data representation is a fundamental aspect of network topology visualization as it determines the ease of understanding of the underlying data. In visualizing network topology, various types of data can be displayed, including node and link properties such as location, type, and bandwidth. It is crucial to consider the data properties to display and how to represent them visually to create an informative and understandable network topology visualization. For instance, color-coding can be used to represent different types of nodes or links, such as green for routers and red for servers, while link thickness can indicate link bandwidth or traffic intensity. Additionally, node size can be used to represent the importance of a node or to indicate the degree of connectivity, where larger nodes represent higher connectivity or significance. The choice of data properties and visualization techniques must be made with the target audience in mind to ensure the visualization's effectiveness in communicating the intended information [17][19].

According to Chen et al. [20], effective data representation is crucial in designing network topology visualization tools. The authors suggest using a combination of techniques, such as colour, shape, and size, to represent different types of data accurately. Additionally, they recommend using a clustering approach to group nodes based on their functional roles to reduce clutter and improve the visualization's readability. Similarly, Nesbitt and Friedrich [21] propose using the Gestalt principles of visual perception, such as similarity, proximity, and closure, to enhance the visualization's readability and minimize the cognitive load of the viewer. In summary, the design considerations for network topology visualization are multi-faceted and must be tailored to the specific needs of the target audience to ensure that the visualization effectively communicates the intended information.

When it comes to creating effective network topology visualizations, the choice of layout algorithm plays a critical role. Layout algorithms are responsible for determining the position of nodes and links in the visualization, which can greatly impact the clarity and usefulness of the resulting image.

The choice of layout algorithm is also critical in creating effective network topology visualisations. Layout algorithms determine the position of nodes and links in the visualisation. There are several different types of layout algorithms available, each with its own strengths and weaknesses. One popular type is the force-directed layout, which simulates physical forces such as repulsion and attraction between nodes to create a visually balanced and aesthetically pleasing network map. This type of layout is often used for visualizing large and complex networks [20][21].

Tree-based layouts are another option, which organize nodes in a hierarchical structure similar to a family tree. This type of layout is often used for visualizing organizational charts or other hierarchical data.

Hierarchical layouts are similar to tree-based layouts, but they allow for more complex branching and can accommodate multiple levels of hierarchy. This type of layout is often used for visualizing complex relationships between entities.

Ultimately, the choice of layout algorithm will depend on the specific requirements of the network topology being visualized. It's important to consider factors such as the size and complexity of the network, the desired level of detail, and the intended audience when selecting a layout algorithm [20][21]. With careful consideration and the right choice of algorithm, a network topology visualization can be a powerful tool for understanding and communicating complex relationships and systems.

User interaction is another important consideration in network topology visualisation design. Users should be able to interact with the visualisation to explore and understand the network topology better. This can include features such as zooming, panning, filtering, and searching [20][21]. It is essential to design user interaction features that are intuitive and easy to use to ensure users can navigate and understand the network topology effectively.

Other design considerations for network topology visualisation include scalability, performance, and aesthetics. Visualisations should be scalable to accommodate large and complex networks. Performance is also crucial, and visualisations should be designed to ensure fast and responsive rendering. Aesthetics are also important, and visualisations should be designed to be visually appealing and easy to read.

VIII. PINNING INFERRED INTERCONNECTIONS AND IP-GEOLOCATIONS ON A GEOPHYSICAL MAP.

After inferring interconnections and locations of network devices, there are various methods to pin them on a geographic map. One common method is to use Geographic Information Systems (GIS) software to plot the inferred interconnections and IP-geolocations on a map. This involves converting the inferred data into a format that can be read by GIS software and then overlaying the data onto a geographic map [22][23].

Geographic Information Systems (GIS) software is a powerful tool for visualizing and analyzing spatial data. It is widely used in various fields, including geography, urban planning, and environmental science. To use GIS software to pin the inferred interconnections and IP-geolocations on a geographic map, the inferred data needs to be converted into a format that can be read by the GIS software.

The conversion process typically involves transforming the inferred data into a common format, such as a shapefile or a KML file[22][23]. This process may involve using specialized tools and software to ensure that the data is properly formatted and that the spatial information is correctly georeferenced.

Once the inferred data has been converted into a GIS-compatible format, it can be imported into the GIS software and overlayed onto a geographic map. This process allows for the inferred interconnections and IP-geolocations to be visualized in a spatial context, providing valuable insights into the physical and logical topology of the network [22][23].

GIS software also provides a range of tools for analysing and manipulating spatial data, allowing for further exploration and interpretation of the inferred interconnections and IP-geolocations. For example, network analysts can use GIS software to calculate distances between network devices or to identify areas of the network that may be particularly vulnerable to disruption [22][23].

Another approach is to use web-based mapping tools such as Google Maps [17][24]. Web-based mapping tools provide an alternative approach for visualizing inferred interconnections and IP-geolocations. These tools can be accessed through a web browser and offer a variety of features for creating and customizing maps. Users can create custom maps and add markers or other graphical elements to represent the inferred interconnections and IP-geolocations.

One advantage of using web-based mapping tools is that they are generally more accessible and user-friendly compared to GIS software, which can be more complex and require specialized skills to use. Additionally, web-based tools often have built-in functionality for sharing and collaborating on maps, which can be useful for collaborative research projects.

However, web-based mapping tools may have limitations in terms of the amount and complexity of data that can be displayed on a single map. In addition, the accuracy and reliability of IP-geolocation data may vary depending on the source of the data and the specific method used for inference. Therefore, it is important to carefully evaluate and verify the accuracy of inferred data before visualizing it on a map [17][24].

Custom mapping tools have been developed by various research studies to visualize network interconnections and IP-geolocations. One such example is the topology visualization tool for National Research and Education Networks (NRENs) in Africa, developed by Yang et al. [17][18]. The tool overlays network topology data onto a map of Africa, providing a comprehensive visualization of the interconnections between various networks in the region. This tool can aid in assessing the diversity of physical and logical routes between networks, as well as evaluate the utility of such a visualization platform.

IX. Discussion

The existing literature employed different approaches and methodologies in visualising the African internet topology. For instance, Yang et al. [17] propose a topology visualization tool for National Research and Education Networks (NRENs) in Africa. They note that NRENs are critical for supporting research and education in Africa, and mapping their infrastructure can help identify areas for improvement. The tool developed by the authors allows for the visualization of NREN topology, including the location of nodes and links. A user cantered approach has been adopted in the development of this tool. Multiple rounds of user testing involving computer science honours students as participants has been carried out [17][18].

The ability of a participant to obtain specific data when performing a task is the defining factor in successful task completion. Accuracy in answering questions and visual queries was used to determine success, with the identification of the country with the most fiber cables being the most accurately answered visual query. Tasks 1 and 2 had a 100% successful completion rate, while Tasks 3, 4, 5, and 6 had lower completion rates due to the complexity of the subtasks. The overall success rate for all questions and tasks across all participants was 79.55%, which is above the average task-completion rate of 78% [17][18].

The average System Usability Score (SUS) was 67.8, which is similar to the average SUS score reported in other studies. However, there were some limitations to the visualization, including a mismatch between the model and reality due to constraints of the Maxmind GeoLite Cities database [16][18].

In contrast Gilmore et al. [4] proposes a holistic approach to map the African internet. They proposed a pair of techniques for charting the African Internet, along with three approaches for illustrating these maps. By utilizing these strategies in tandem, a comprehensive view of the African Internet was created. Gilmore et al. [4] utilized the Walrus graphing package to display the AS structure of the African Internet and employed two- and three-dimensional visualizations to represent the geographical locations of its global and Africa components [4][25]. The router level maps reveal the layout of the physical connections, while the AS level map provides insights into the relationships between them. Note that this approach relies heavily on BGP datasets as a data source [4]

The technique was effectively executed. Nonetheless, there is a downside to this approach as a lot of routers have disabled their source routing capabilities to avoid misuse. however, a few routers still have source routing activated and locating them would enhance the precision of the mapping approach. To further enhance the accuracy of the geographical mapping, it is recommended to acquire more precise commercial geolocation databases [4][20].

Overall, these studies have shown that mapping the African internet topology requires a combination of techniques and approaches, including the use of trace routes, BGP datasets, and visualisation tools. These studies also highlight the importance of user-cantered design and testing for effective visualisation tools.

Although there is a wealth of literature done on internet topology visualisation, there is minimal research done in the African context. Hence only the literature that strongly pertain to the topic of internet topology visualisations in the African context are discussed in this section.

X. Conclusions

Based on the existing literature discussed in this literature review, it can be concluded that mapping the African Internet topology is a challenging task due to the lack of comprehensive and accurate data on the network infrastructure in Africa. However, advances in technology and the availability of new data sources have made it possible to develop tools and approaches to visualize and analyse the topology of the African Internet.

The paper by Gilmore et al. [4] presents an approach to mapping the African Internet topology using traceroute data collected from various points in the network. The authors found that the African Internet was characterized by a high degree of fragmentation and asymmetry, with many paths passing through international links. The paper highlights the need for more extensive data collection efforts to better understand the topology of the African Internet and to support the development of more efficient and robust network infrastructure.

In contrast, the paper by Yang et al. [17] presents a topology visualization tool called NetVizura that was specifically designed to visualize the topology of National Research and Education Networks (NRENs) in Africa. The authors applied the tool to several NRENs and identified areas for improvement, including areas of low connectivity and single points of failure. The paper demonstrates the potential of visualization tools for identifying areas for improvement in NRENs and suggests that such tools could be used to inform network planning and decision-making.

In summary, the existing literature highlight the importance of mapping the African Internet topology to support the development of more efficient and robust network infrastructure. While challenges remain, advances in technology and new data sources are making it increasingly possible to visualize and analyse the African Internet topology, providing valuable insights for network administrators and policymakers. Although the existing literature contributes to the field of research, several challenges and gaps remain in the literature. One challenge is the lack of standardized methodologies for mapping internet infrastructure and topology. This can make it difficult to compare results across studies. Another challenge is the rapidly changing nature of the internet and the need to keep up with changes in infrastructure and topology.

One gap in the literature is the lack of studies on internet infrastructure and topology in developing regions, such as Africa, where internet access and infrastructure can be limited. This makes the proposed project particularly valuable in filling this gap.

The literature also presents opportunities for further research. For instance, there is a need to explore the impact of different routing protocols on internet topology, as highlighted in Li et al.'s [16] paper. Moreover, with the increasing importance of the internet and the growing number of devices connected to it, there is a need for more research into visualizing internet infrastructure and topology at a global scale.

Overall, the proposed project to infer and visualize relationships between ISPs in Africa has the potential to contribute to this field of study by providing insights into the diversity of physical and logical routes between networks in Africa as well as evaluating the utility of a visualization platform for mapping internet infrastructure and topology. Furthermore, the project fills a knowledge gap in the literature concerning internet infrastructure and topology in developing countries.

References

1. Kimberly Claffy, Young Hyun, Ken Keys, Marina Fomenkov, and Dmitri Krioukov. 2009. Internet Mapping: From Art to Science. *2009 Cybersecurity Applications & Technology Conference for Homeland Security*. https://doi.org/10.1109/catch.2009.38
2. Reza Motamedi, Bahador Yeganeh, Balakrishnan Chandrasekaran, Reza Rejaie, Bruce M. Maggs, and Walter Willinger. 2019. On Mapping the Interconnections in Today’s Internet. *IEEE/ACM Transactions on Networking* 27, 5: 2056–2070. https://doi.org/10.1109/tnet.2019.2940369
3. Arpit Gupta, Matt Calder, Nick Feamster, Marshini Chetty, Enrico Calandro, and Ethan Katz-Bassett. 2014. Peering at the Internet’s Frontier: A First Look at ISP Interconnectivity in Africa. *Passive and Active Measurement*: 204–213. https://doi.org/10.1007/978-3-319-04918-2_20.
4. Gilmore, J., Huysamen, N., & Krzesinski, A. 2007. Mapping the african internet. In *Proceedings Southern African Telecommunication Networks and Applications Conference (SATNAC), Mauritius*.
5. Abdullah Yasin Nur and Mehmet Engin Tozal. 2018. Cross-AS (X-AS) Internet topology mapping. *Computer Networks* 132: 53–67. https://doi.org/10.1016/j.comnet.2018.01.011
6. Keys, K., Young Hyun, Luckie, M., and Claffy, K. 2013. Internet-Scale IPv4 Alias Resolution With MIDAR. *IEEE/ACM Transactions on Networking 21*, 2, 383–399.
7. Gunes, M.H. and Sarac, K. 2009. Resolving IP Aliases in Building Traceroute-Based Internet Maps. *IEEE/ACM Transactions on Networking 17*, 6, 1738–1751.
8. Poese, I., Uhlig, S., Kaafar, M.A., Donnet, B., and Gueye, B. 2011. IP geolocation databases. *ACM SIGCOMM Computer Communication Review 41*, 2, 53–56.
9. Hyunsu Mun and Youngseok Lee. 2017. Building IP geolocation database from online used market articles. *2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)*. https://doi.org/10.1109/apnoms.2017.8094175
10. Brian Eriksson, Paul Barford, Joel Sommers, and Robert Nowak. 2010. A Learning-Based Approach for IP Geolocation. *Passive and Active Measurement*: 171–180. https://doi.org/10.1007/978-3-642-12334-4_18
11. Ovidiu Dan, Vaibhav Parikh, and Brian D. Davison. 2021. IP Geolocation through Reverse DNS. *ACM Transactions on Internet Technology* 22, 1: 1–29. https://doi.org/10.1145/3457611
12. 1.Manaf Gharaibeh, Anant Shah, Bradley Huffaker, Han Zhang, Roya Ensafi, and Christos Papadopoulos. 2017. A look at router geolocation in public and commercial databases. *Proceedings of the 2017 Internet Measurement Conference*. https://doi.org/10.1145/3131365.3131380
13. Alexander Gamero-Garrido, Elizabeth Belding, and David Choffnes. 2022. Using reverse IP geolocation to identify institutional networks. *Proceedings of the 22nd ACM Internet Measurement Conference*. https://doi.org/10.1145/3517745.3563021
14. Armand Prieditis and Gang Chen. 2013. Mapping the Internet: Geolocating Routers by Using Machine Learning. *2013 Fourth International Conference on Computing for Geospatial Research and Application*. https://doi.org/10.1109/comgeo.2013.17
15. Dodge M, Kitchin R. Examining different approaches to mapping Internet infrastructure, 2001. Retrieved March 10, 2023, from UCL working paper series paper 39: https://discovery.ucl.ac.uk/id/eprint/174014/1/paper39.pdf
16. Jie Li, Vasileios Giotsas, Yangyang Wang, and Shi Zhou. 2022. BGP-Multipath Routing in the Internet. *IEEE Transactions on Network and Service Management* 19, 3: 2812–2826. https://doi.org/10.1109/tnsm.2022.3177471
17. Chantal Yang, Hussein Suleman, and Josiah Chavula. 2016. A topology visualisation tool for National Research and Education Networks in Africa. *2016 IST-Africa Week Conference*. https://doi.org/10.1109/istafrica.2016.7530605
18. Chantal Yang. 2015. Visualising the Network Topology of National Research and Education Networks in Africa.
19. David Howard and Alan M. MacEachren. 1996. Interface Design for Geographic Visualization: Tools for Representing Reliability. *Cartography and Geographic Information Systems* 23, 2: 59–77. https://doi.org/10.1559/152304096782562109
20. Wei Chen, Fangzhou Guo, and Fei-Yue Wang. 2015. A Survey of Traffic Data Visualization. *IEEE Transactions on Intelligent*

Transportation Systems 16, 6: 2970–2984. https://doi.org/10.1109/tits.2015.2436897
21. K.V. Nesbitt and C. Friedrich. 2002. Applying Gestalt principles to animated visualizations of network data. *Proceedings Sixth International Conference on Information Visualisation*. https://doi.org/10.1109/iv.2002.1028859
22. Ramakrishnan Durairajan, Subhadip Ghosh, Xin Tang, Paul Barford, and Brian Eriksson. 2013. Internet atlas. *Proceedings of the 5th ACM workshop on HotPlanet*. https://doi.org/10.1145/2491159.2491170
23. Arnaud de la Losa and Bernard Cervelle. 1999. 3D Topological modeling and visualisation for 3D GIS. *Computers & Graphics* 23, 4: 469–478. https://doi.org/10.1016/s0097-8493(99)00066-7
24. 1.Rick Hofstede and Tiago Fioreze. 2009. SURFmap: A network monitoring tool based on the Google Maps API. *2009 IFIP/IEEE International Symposium on Integrated Network Management*. https://doi.org/10.1109/inm.2009.5188876
25. 2023. MaxMind Server IP Addresses. *MaxMind Server IP Addresses | MaxMind Developer Portal*. Retrieved from https://dev.maxmind.com/geoip/geolite2-free-geolocation-data?lang=en
26. B. Huffaker, D. Plummer, D. Moore, and K. Claffy. 2002. Topology discovery by active probing. *Proceedings 2002 Symposium on Applications and the Internet (SAINT) Workshops*. https://doi.org/10.1109/saintw.2002.994558
27. R. Govindan and H. Tangmunarunkit. 2000. Heuristics for Internet map discovery. *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*. https://doi.org/10.1109/infcom.2000.832534