Data Visualization: The Current Situation of African Internet

Zhaoqing Wang
School of Information & Communication Engineering, Beijing Information Science & Technology University, Beijing 100101, China
mj741561@163.com

Abstract. The current Internet penetration in African is significantly below the world average level, while African has the second largest population in the world. This article aims at analysing the current situation of the African Internet to see the insights behind the massive related datasets, containing the design and implementation processes of a resulted platform. The platform consists of three web pages for showing the results, databases, and webpage controller implemented by Java. After collecting available datasets online, we now need a visualized tool to analyse the reasons behind those raw data, to explore the attributes and environment of the complex network, hence we derived a web-page platform for the task, where researchers can query detailed information according to certain requirements. The first page of our platform shows the structure of the African network topology, the second web page allows users to interact with the data, generating corresponding graphic results, and the third web page handles queries, request with certain manner defined by users.

1. Introduction

African is the second most-populous continent in the world, with 1.2 billion people in land, yet comparing to the world, the Internet and Telecommunication Industries are limited by lower penetration rate, significantly below the average level as the above figure shows.[1] In recent years, researchers have been collecting large-scale datasets to investigate why this is the case and trying to understand the nature and characteristics of infrastructure in the nation.

Data Visualization, viewed by many disciplines as a modern tool of data analysis, involving the creation and study of the visual representation of data with attributes or variables for the units of information, can be used to get the insights of the current situation of African Internet.
A primary goal of data visualization is to communicate information clearly and efficiently via statistical graphics, plots and information graphics. Numerical data may be encoded using dots, lines, or bars, to visually communicate a quantitative message. Effective visualization helps users analyse and reason about data and evidence. It makes complex data more accessible, understandable and usable. Users may have particular analytical tasks, such as making comparisons or understanding causality, and the design principle of the graphic (i.e., showing comparisons or showing causality) follows the task. There are some methods such as: tables are generally used where users will look up a specific measurement, while charts of various types are used to show patterns or relationships in the data for one or more variables.

2. Literature Review

Data Visualization has long been considered as strong tool to structure plentiful network data, to visualize the network as collections of nodes and edges, there are some related previous work done in recent development.

2.1. Technologies in Webpage Development

Hypertext Markup Language (HTML) is the standard markup language for creating web pages and web applications. With Cascading Style Sheets (CSS) and JavaScript, it forms a triad of cornerstone technologies for the World Wide Web. In this project, we use it as the representation layer of the results.

AJAX, short for “Asynchronous JavaScript and XML” is a technology used to create fast and dynamic web pages. It can communicate with servers and maintain data without refreshing the entire page, based on standardized and widely supported technology, it does not need to download browser plug-ins or small programs, but needs the client to allow JavaScript to execute on the browser.

2.2. Fundamental Frameworks

Framework is a reusable design of the whole or part of the system, which represents a set of methods of interaction between an abstract component and a component instance.

The Spring Framework is an application framework and inversion of control container for the Java platform. The framework's core features can be used by any Java application, but there are extensions for building web applications on top of the Java EE (Enterprise Edition) platform. By using Spring MVC, objects created by the container are also called managed objects or beans. Besides, object can be obtained by means of either dependency lookup or dependency injection.

Hibernate is an open source object relational mapping framework, which carries out a very lightweight object encapsulation to JDBC (Java Database Connectivity). It establishes a mapping relationship between POJO and database tables. It is a fully automated ORM (Object Relational Mapping) framework. Hibernate can automatically generate SQL statements and automatically execute them, so that Java programmers can use object programming thinking to manipulate databases as they want. [2]
3. Design and Implementation

3.1. Database Design
Because this project is a visualisation of the data in the website, it is very important to define a visualization tool before the website development, according to the comparison of the strengths and weaknesses of the three generic tools (Echarts, Highcharts, and D3.js), Highcharts is selected to develop the data visualization of the project.

Based on the requirement of web development, by using MySQL as data development platform. For draw Entity-Relationship (E-R) diagrams, two entities are designed: connectivity_traces and hop. The “connectivity traces” entity has attribute that source country code name, destination code name and the count of route hops. The “hop” entity’s attribute not only include the origin and target code name but also has IP destination, probe date, target date and the RTT (Round-Trip Time) value.

3.2. Web Page Design
The functional frame introduces the module of this web site, and the main function that is implemented in the web site. In the first webpage, put the Africa map and import geographic data then implement the topology structure. At the same time, the map should allow downloaded by user. The second page will develop to an interactive page that allow user explore with data. Considering the users need to generate a line chart between different country. And it allows users download the diagram generated as well. Last a query page will be developed for users’ enquiry, and back return a table with subordinate page. It will help users for analysing data.

3.3. Topology Map Implementation
First of all, generating Africa map on the page is necessary. According to the highcharts’ document, highmaps.js is need to load for forming a map and loading the data.js for importing geographic data. And then obtain a map object by object oriented of characteristic of javascript “var mapdata = Highcharts.maps['custom/africa'];". Next initialize the map through $('#container').highcharts('Map', {title,series,datalebls,color}) to set the detail of the map and implement it by <div id="container" ></div> in the label <body> of the HTML.

1. Method setNode():
Once the map initializes, it will load the map data including the position of each country in the Africa. In this method, according to location data of each country, declare a corresponding variable “img”, it is an image and set its CSS properties to match the location of the country on the web page. Finally append the this “img” to the body of HTML.

function setNode(){

Fig.2 Hibernate model, retrieved from https://howtodoinjava.com/hibernate-tutorials/
2. Method drawLine():
An array is created to store the coordinates of every node which is marked in the map. Before that this project prepares two lines vector graph by Photoshop. One is from upper left to lower right (like this: "\"\"), the other one is from upper right to lower left (like this: "\"\") Then determine the position relationship of two points, according to the length and width of the distance and direction, set the properties of the CSS of the line graph. Then append it to the body of HTML as well. Thus, the topology structure is formed as Figure below.

![Fig.3 Topology Map](image)

3.4. Data Display Page Implementation
The first page displays the performance data stored in the database through appropriate charts. This page is developed around RTT value, because RTT(Round Trip Time) is an important index of health of network. The first choice is the bubble chart, which produces bubbles on the Africa map according to the document in highcharts.com.
4. Results and Discussion

The finished work has been shown in Section 3, where the charts are available. From the Topology Map, we can discuss pros and cons about the platform.

- **Strengths:** Optional box and refresh button effectively display the data
- **Weaknesses:** The shortcoming in this webpage is that the routing structure cannot be displayed as hop-by-hop, which weakens the expression of the routing structure. Because there are hops in the original data that are not clearly recorded, a complete routing link cannot be formed. Which can be further improved.

The Bubble Chart also indicates the advantage and disadvantage about the platform:

- **Strengths:** Bubble charts and line charts are generated smoothly and data is displayed clearly
- **Weaknesses:** Bubble chart lacks interactivity. It can be added function when the user clicks on the bubble, the detailed network performance of the country is displayed.

With the foundation of the visualization web site, we can do further data analysis. The platform divides Africa into five parts according to the east, south, west, north and central. Kenya, South Africa, Garner, Egypt and Central African were selected as research subjects to conduct data analysis according to the functions of three web pages.

We now have a general understanding of the topological structure of these five countries. Figure 6 Topology Structure based on Five Countries shows that all these countries have good network topology and intuitively present their mutual connection.
Then when we redirect to the second page. In the generated Bubble chart we can clearly find that the largest bubble on the map is belongs to the Central African. It’s average RTT value is reach 625ms, which means the network performance in this region is very poor, and it greatly reduce the degree of user experience. This is also understandable because the physical environment of the African network is surrounded by 12 long submarine cables. And the in-land optical fibre is still fragmented.

Next, we can explore the RTT values of five countries and record them in the generated results, we can see that South Africa has the best network performance, followed by Egypt. The result is in line with the current network construction environment in these two countries.

5. Conclusion
This project implements a web site contain three pages. The main function of these pages is that it allows users interact and explore with two kind of network data, which is Topology data and performance data. Users can intuitively understand and research the Internet data in Africa on the basis of visual data charts.

The results indicated this web site has mass of data to be queried and the response is faster, the process of generating the variable graph is very smooth. However, the deficiency is the display of routing topology is not exhaustive in the first page. The routing communication between two countries should present the nodes it contains, connecting the nodes they implement a complete topology route.

This will become the primary task of this web site optimization in the further work. The other part for further improvement is to optimize the result of the third page return and add other types of charts to enrich the degree of data visualization.

References
[1] Africa Internet Users, 2018 Population and Facebook Statistics. www.internetworldstats.com/stats1.htm.
[2] Himansu, et al. “Hibernate Tutorial.” HowToDoInJava, 11 July 2018, howtodoinjava.com/hibernate-tutorials/.
[3] S. Song.(2015, May). African Undersea Cables. Retrieved May, 2015, Form Web site: https://manypossibilities.net/african-undersea-cables/
[4] R. Fanou, G. Tyson, P. Francois, and A. Sathiaseelan. (2015, May). Technical Report: African Content Measurement Web site: https://fourier.networks.imdea.org/external/techrep_cdma/index/
[5] Oracle. January (1, 1999). "1.2 Design Goals of the Java™ Programming Language" Retrieved January 14 2013, From web site: http://www.oracle.com/technetwork/java/intro-141325.html