



UNIVERSITY OF CAPE TOWN



DEPARTMENT OF COMPUTER SCIENCE

COMPUTER SCIENCE HONOURS
FINAL PAPER
2015

Title: AfriNREN: An Interactive Visualization Dashboard for African NRENs

Author: Robert Passmore

Project Abbreviation: AfriNREN

Supervisor: Josiah Chavula and Hussein Suleman

Category	Min	Max	Chosen
Requirement Analysis and Design	0	20	5
Theoretical Analysis	0	25	0
Experiment Design and Execution	0	20	20
System Development and Implementation	0	15	0
Results, Findings and Conclusion	10	20	20
Aim Formulation and Background Work	10	15	15
Quality of Paper Writing and Presentation	10		10
Adherence to Project Proposal and Quality of Deliverables	10		10
<i>Overall General Project Evaluation (this section allowed only with motivation letter from supervisor)</i>	0	10	0
Total marks	80		80

AfriNREN: An Interactive Visualization Dashboard for African NRENs

Robert Passmore
Department of Computer Science,
University of Cape Town
Rondebosch 7701, Cape Town
pssrob001@myuct.ac.za

ABSTRACT

African national research and education networks (NRENs), exchange network data through a variety of collaborative network applications from video conferencing to large research data transfers. Researchers have shown that up to 75% of traffic originated in and destined for these African institutions is routed circuitously through other continents, increasing latencies of network traffic [6] and therefore affecting the collaborative efforts of African NRENs. This circuitous routing can be avoided through peering of NRENs at Internet exchange points (IXPs) in Africa [8] but the managers of the NRENs first need to make an argument for peering to acquire funding and support. Information visualization can be helpful in augmenting cognitive capabilities of professionals [9] these network managers in preparing an argument for peering at African IXPs. Currently there is a lack of research into the use of visualisation tools to assist the network managers and policy makers of African NRENs in this task.

In this paper, we describe the design, implementation and evaluation of the AfriNREN Web application – an interactive visualization dashboard intended to help African NRENs identify the biggest communication partners that transfer traffic through their network and the characteristics of that traffic to support their aims of improving network performance for collaborative research and education applications.

The experimental results show that a basic information visualization dashboard successfully assists users in identifying the Autonomous Systems (ASes) that communicate the most through a network; determining the amount of traffic transferred between communicating ASes; identifying which continents these ASes are located on; and determining the type of traffic being sent from large sources of traffic.

CCS Concepts

- **Human-centred computing** → **visualisation**

Keywords

Visualization; Dashboard; NREN; Network Traffic Visualization; National Research and Education Networks; Information Visualization

This report was completed as part of a Bachelor of Sciences (Honours) Computer Science degree at the University of Cape Town in 2015. For more information about this project, please visit <http://people.cs.uct.ac.za/~snbros001/AfriNRENWeb2/>

1. INTRODUCTION

Up to 75% of network traffic originating from and destined to African research and education institutions follows intercontinental circuitous routes [6]. These longer routes double the latencies of traffic exchanged within Africa [6], effecting the network performance experienced by users at research and education networks. The network managers and engineers of these African institutions seek to improve the performance of information exchange between each institution to facilitate network-intensive collaborative work done between them [7; 16]. One way to improve latencies is to peer African networks and Internet Service Providers at Internet Exchange Points (IXPs) [8] in Africa so that the traffic can travel more directly from source to destination rather than circuitously to other continents. Gupta et al. [8] discusses how African ISPs don't do this and instead peer at larger European IXPs. Network managers could improve the performance for intra-African traffic by peering their networks with other African networks with which they exchange a significant amount of information. To do so, network managers require information about the volumes and characteristics of traffic being exchanged between African networks to argue for peering between the networks at IXPs.

Information visualizations – interactive visual interfaces designed to facilitate data communication and exploration – are powerful tools to augment the cognitive capabilities of users [9] such as network managers and can help the network managers in exploring the traffic information and building an argument for network peering at African IXPs. Currently, there is a lack of research into effective network data visualization tools in the context of the UbuntuNet and African National Research and Education Networks (NRENs). This research is needed to help develop tools to facilitate network design, planning and operation and, in the distinct context of NRENs, aid policy makers in the determination of whether peering between two autonomous systems (ASes) is viable.

Optimizing network performance between African research and education institutions is important because high latencies retard the ability of these institutions to collaborate using the Internet [6]. Furthermore, data creation and transfer between institutions will continue to increase in the future with the Square Kilometer Array project and other data intensive research projects. This will result in high data transmission costs if these projects – and African NRENs in general – continue to transmit data circuitously [6]. African network managers can lower these costs and high latencies in internetwork data transfer by peering ASes. In order to make a case for peering between ASes, a network manager would need to know how significant the link between the ASes is (i.e. how much traffic is exchanged between the pair relative to

other pairs) and the characteristics of the traffic such as the type of traffic being transferred.

The aim of this work is to design and evaluate a dashboard of data visualizations to assist network managers in identifying the amount of traffic exchanged between different networks in the past as well as the type of traffic being exchanged.

This paper follows a user centered design approach with multiple iterations of an analysis, design and evaluate cycle. The dashboard will be presented as a Web application utilizing the D3.js JavaScript framework [4] to implement the visualizations. The visualizations will be designed to display data retrieved from NetFlow logs of traffic captured at African NRENs. The system is evaluated using focus groups for early feedback and user evaluations for further feedback to determine the success of the dashboard in achieving the aims. Success is determined by testing users' ability to complete specific tasks and based on the feedback from representative users through the System-usability Scale and open-ended feedback questions.

1.1 Research Question

The research question here is "can a dashboard of interactive non-geospatial visualizations of NetFlow data effectively communicate network traffic information about a network to the network managers?"

The tasks we test to determine the dashboard's effectiveness are: identifying the Autonomous Systems (ASes) that communicate the most; determining the amount of traffic transferred between communicating ASes; identifying which continents these ASes are located on; and determining the type of traffic being sent from large sources of traffic.

2. BACKGROUND

2.1 Network data visualizations

Visualizations of network topologies and traffic have been explored before for purposes including the monitoring of large networks for network security events. A few examples are worth mentioning. Boitmanis et al. [3] utilizes a network graph (node and link) to communicate the structure of the Internet at the Autonomous Systems (AS) level. In general, Node and link graphs are applicable to the context of network visualisation [2] since they show the connections and relationships between entities. A force-layout approach – stress majorization – is used for the graph layout. The network graph is a 'hairball' where the number of links between nodes creates a dense nest of indistinguishable connections. Network diagrams commonly suffer from this [2]. Lad et al. [10] also uses a network graph to visualize the change in routing between Autonomous Systems (ASes).

FloVis [14] attempts to solve the occlusion and the 'hair balling' problem of network diagrams by 'bundling the connections and aggregating the nodes' of the diagram by aggregating nodes with the same network prefix [14]. This is effective in displaying the number of links created between nodes in a network but would not be as effective at showing the relationship between AS nodes given AS nodes cannot be aggregated in such a fashion.

[NetFlow data visualization based on graphs] is a project with similar goals to bridge the gap between visualizations that aggregate details of a network's NetFlow data and those that provide simple tables of records of NetFlow data.

Commercial applications such as SolarWinds' NetFlow Traffic Analyzer have also presented options for network managers. The summary view of SolarWinds' demo uses pie charts and tables along with a comprehensive filtering feature to give the user breakdowns of how much traffic different applications, users, and endpoints are responsible for as well as the country and AS involved in the flow. The visualization of the information is basic and only provides a high level aggregation of the NetFlow data.

2.2 Evaluating Information Visualizations

Information visualisations are not easily evaluated. Visualisations are tools used in the creative activity of discovery and to observe users using the visualisation in environments that are foreign to their normal work environments would not produce realistic observations [13]. Users regularly collaborate with other users, look at the same data on many different occasions and formulate and answer questions they didn't originally have. Furthermore, engagement between the observer/developer and the user are key to improving the tool and in helping the domain-expert become an expert user of the tool. Shneiderman and Plaisant advise that evaluations of information visualisations take the form of hands-on and intensely interactive engagements and observations of the users in their normal work environment. Their recommended approach is the 'Multi-dimensional In-depth Long-term Case Study' (MILCs) approach to evaluating information visualisations. In their paper [13], they also recommend following a guideline for ethnographic observations to minimize the likelihood of the observations going wrong and causing the experiment to fail. The guideline involves detailed preparation work, intensive observation through collection and documentation, accurate analysis of the data and appropriate reporting. This approach would not work for this project given the limited timeframe and limited access to the domain experts.

Alternatively, information visualisation shares similarities with empirical human computer interaction (HCI) research and general evaluation techniques can be used in an evaluation of an information visualisation system [6].

The System Usability Scale (SUS) [5] is a popular and free to use evaluation questionnaire used to assess a system's usability. It can be used on any interface or technology given the general wording [11]. It asks a user to respond to 10 statements that alternate in being positive or negative. The responses scale from 1 ('I strongly disagree') to 5 ('I strongly agree'). The statements relate to their experience with the system and the scale is scored from 0 to 100 and can be compared to a global collection of SUS scores to get a comparative idea of the system's usability [11]. The SUS can be used to measure perceived usability on a small sample (8 – 12 participants) [15].

3. DESIGN

This project focused on the design and development of a Web application, a dashboard of data visualizations, the preparation of data for input to the visualizations and an experiment to evaluate the dashboard.

3.1 System Architecture

The AfriNREN dashboard is a Web application, served to a Web browser from a laptop acting as a Web server. Django – the Python based Web application development framework – was used to develop the system along with a PostgreSQL database backend. The data – once prepared with pre-processing shell scripts and shell programs – is loaded into the database to be later

fetched, processed and stored as JSON files on the server. The JSON data is later delivered to the client upon request. The data is accessed through API calls either from the user (when they visit an API formatted URL) or from API calls made from HTML and JavaScript files.

The visualizations – implemented using HTML, CSS and the D3.js JavaScript framework – would call the API and use the data to generate the visualizations for rendering at the client side. Interactions with the visualizations only manipulate the data on the client’s browser; there are no subsequent calls to the API once the page is first completely loaded.

3.2 Data Preparation

The data processing behind AfriNREN’s dashboard followed the flowchart displayed in Figure 1. The data preparation process was not automated through development of a data processing pipeline but instead preparation was done ad hoc to the dataset available.

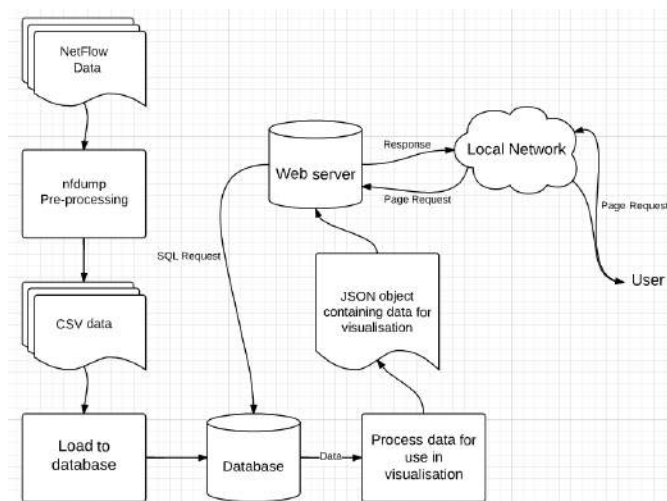


Figure 1: Data processing flow chart

3.2.1 The Data

The data used in this project was historic network traffic data captured and stored as NetFlow data. The data was given to us by an African NREN and had been anonymised while maintaining subnet information.

3.2.2 Data Pre-processing and Storage

The raw NetFlow data was first processed using nfdump – a NetFlow display and analysis command line tool – to extract the desired data into a comma-separated format. The NetFlow data is not stored in human readable format and nfdump is required to extract the desired fields from the data. Each entry in a NetFlow file is a flow – a communication session between a source and destination IP address and port number. Nfdump can be used to do basic aggregation by flow which sums all the fields for each unique flow (unique source and destination address pair). The data pre-processing step used the flow aggregation option and wrote the following fields to a csv file: date first seen, flow duration, source IP address, source port, destination IP address, destination port, protocol, number of flows, and total bytes transferred.

The csv files were loaded into a database for use by the Web application through Django’s ORM and database API which was used to perform further aggregation operations on the data.

3.2.3 JSON processing

In Django, a model-view-controller (MVC) pattern is followed and so when a page is requested, the request is handed to a controller which fetches the required data from either JSON objects stored in files on the server or from the database. If fetched from the database, the controller first performs the required aggregation operations on the data to prepare it for use in a visualization. The data is then manipulated into a JSON object and passed that to the “view” – the HTML and JavaScript that makes up the webpage – which uses the data to generate the visualizations. The JSON object includes all the data needed by the dashboard’s visualizations.

3.3 Visualization Design

The design of the visualizations and dashboard followed an iterative user-centered design approach. The approach included two iterations each with three activities: analysis, design and evaluation [17].

The first iteration began with a task and context analysis by surveying a collection of NREN CTOs and network managers within Africa. The tasks identified were used to choose appropriate visualization types as identified in the literature review and to design a low-fidelity paper prototype dashboard. This prototype was evaluated by a local focus group of representative users with experience in networking theory and practice. A focus group is useful as an early-stage feedback evaluation to make sure the design is not fundamentally flawed [17].

The second iterations each involved analysis of the previous iteration’s evaluation feedback and a re-design of the visualizations. The new design was implemented as a high-fidelity interactive visualization and evaluated using the system-usability scale as a usability test. This usability test was given to a small group (20 people) of users trained in networking theory for the final evaluation.

This iterative design process helped discover design problems but did not ensure a successful visualization design [17]. The final evaluation was used to determine the success of the project’s design.

4. IMPLEMENTATION

The development of the dashboard followed a user-centred design approach with two iterations of the [design, implement, evaluate] cycle.

4.1 Iteration 1

The early envisioning phase included task and context analysis using a survey of NREN CTOs and network managers within Africa. The response rate to the survey was poor with only two full responses. To gather requirements further interviews were conducted with a network manager and technician from the University of Cape Town’s ICTS. The interviewee provided insight into the tools he used to monitor the University’s network and suggested features that might be of use to an NREN manager. The interviewee ranked the most important tasks for a NREN manager starting with “classifying the traffic by protocols”, followed by “identifying time-series information of network utilization” and “clustering IP addresses by the AS to which they belong”. These tasks were used to create the initial requirements for the project but were subsequently changed to remove the time-series task as this was seen to be more a priority for a manager

involved with operational management rather than a manager investigating long term plans to peer with other NRENs at IXPs.

The requirements stated that the dashboard should allow a domain expert to identify the Autonomous Systems (ASes) that communicate the most through a network being analysed; determining the amount of traffic transferred between the communicating ASes; identifying which continents these ASes were located on; and determining the type of traffic being sent from large sources of traffic.

A design was developed using examples from the literature survey and implemented as a low-fidelity paper prototype as seen in Figure 2.

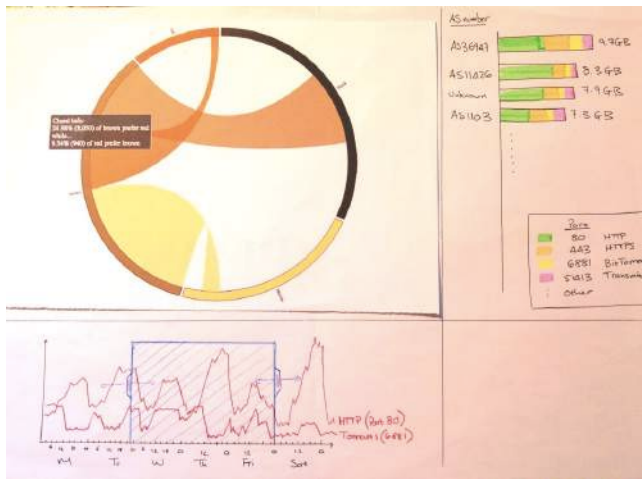


Figure 2: Paper prototype of the AfriREN Dashboard

The paper prototype was evaluated through a focus group of five: two networking professionals and three postgraduate students with experience in networking theory and practice. The prototype in Figure 2 featured a chord diagram, a row chart as well as a time-series graph – all of which were populated with example or random data. The chord diagram was designed to display the relationship between ASes and how much traffic each AS sent to each other AS and how much traffic they received from each other AS. The chord diagram’s data was taken from a dating website where each user indicates their own hair colour and the hair colour that they prefer on a partner. As displayed in Figure 2, each arc segment represents a hair colour and is designated the colour that matches the hair colour in the data (blonde, brown, red, black). Each chord’s width is different at the start of the chord (where the chord is the same colour as the attached arc segment) and the end of the chord. As an example, refer to the red chord that connects the red arc segment to the brown arc segment in Figure 2. The width at the start of the chord is encoded with the quantity of users with red hair that prefer partners with brown hair. The width at the end of the chord (where the red chord meets the brown arc-segment) encodes the quantity of brown haired users that prefer partners with red hair.

When used in the context of the project, each arc segment would represent an AS and would be designated a unique colour. Each chord in the diagram would be given a colour of one of the ASes which the chord connects. The width of the chord’s different ends would encode the number of bytes of information sent from one AS to the other.

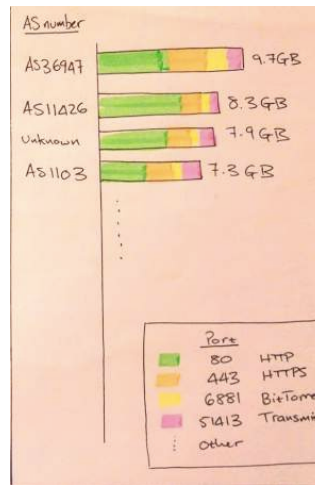


Figure 3: Paper prototype rowchart

The rowchart and time-series graph were populated with random data. The rowchart Figure 3 displays a bar for each AS in the top N sources of data. Each bar is broken down by the ports that sent the most data from within the AS. Only the top few ports (by bytes sent) were shown (HTTP, HTTPS, etc.). The time-series graph displayed the quantity of data sent from a selected AS over time. Each line in the time-series graph represented the quantity sent from a particular port in the selected AS.

The feedback revealed that the chord diagram was difficult to understand unless a detailed description and exercise were given before hand. The focus group recommended that an alternative visualisation be used that is easier to understand. The feedback on the rowchart and the time-series graph was neutral although the focus group did agree that the rowchart would be useful to an NREN manager but did not show as much interest in the time-series graph.

4.2 Iteration 2

The design of the dashboard used feedback from the focus group of the previous phase and was implemented as a hi-fidelity Web application following the design outlined in the design section. The duration of the implementation stage was longer than expected due to the time required to filter and format the NetFlow data for use with the D3.js visualization framework. The implemented dashboard included a subset of the desired features.

The Web application was used in the experiment to evaluate the usability of the dashboard and to retrieve feedback from the experiment participants. The final dashboard design is presented in section 5 and the experimental design is presented and discussed in section 6.

4.3 Challenges & Difficulties

Despite having motivating research for this project, we were not able to involve real users of the system in the interview and evaluation process. As Carpendale mentions, a difficulty in evaluating an information visualization is accessing domain experts and getting them to volunteer their time [6]. There was therefore concern over whether the proposed features of the system were realistic and whether the feedback received from our substitute users would be representative of feedback from real users.

The plan of this project also neglected the importance of data analysis and preparation. Designing a dashboard to satisfy

requirements and assist in task completion needs to involve a proper investigation of the data, the format it is received in and the format that is required by the visualization system. We did not scope the project correctly and most of the implementation time was spent on data preparation.

The system architecture was designed with the idea of fetching data from the database everytime the dashboard was requested from the client; however, the time taken to process the data requests was longer than expected and would have negatively impacted user experience. To avoid this a change was made to the data pre-processing. The data was saved to JSON formatted files on the server and then forwarded to the client on request, successfully avoiding the data processing bottleneck that would otherwise affect the user experience.

The D3 visualisation JavaScript framework also presented challenges in the development process. The framework has a steep learning curve especially when developing a custom visualisation. However, D3 was still the optimal choice as it was the only visualisation framework that allowed a low enough level of control to customise the whole visualisation. Other visualisation frameworks such as Sigma.js and DC.js were explored but these higher-level frameworks only allowed a limited set of customisations such as changing colours of a pre-defined chart type.

In this project, the feature scope of the dashboard was reduced to handle the unanticipated time required by the data preparation. This was unavoidable as we could not simply generate random data to use for visualization because a correct visualization design required designing around realistic data in order to create a usable tool for professionals.

5. FINAL DASHBOARD DESIGN

The final design of the AfriNREN dashboard featured two visualizations as seen in Figure 4. The figure is based on NetFlow data collected at an African NREN over a weeklong period between the 12th and 18th March 2015. The Figure 5 shows the left panel of the dashboard – titled ‘Busiest AS Communication Partners’ – and features a network graph (node and link diagram) showing the Autonomous Systems (ASes) that have sent and received the most traffic through the NREN as nodes and the links represent the amount of data transferred between each connected node. The graph follows a force-layout layout algorithm and each node can be repositioned by the user and the graph will recalculate the position of each other node in respect to the changed node. Furthermore, the user can access details of each AS by hovering their mouse over a node in the graph. A tooltip appears over the node with information of the AS’s AS number, the organization’s name, the AS’s country and continent of location, and the number of bytes sent and received as shown in Figure 5. The dataset used to test and develop the dashboard features anonymised data which may have altered the AS mappings. This would affect how accurately the visualization represented the realistic scenario.

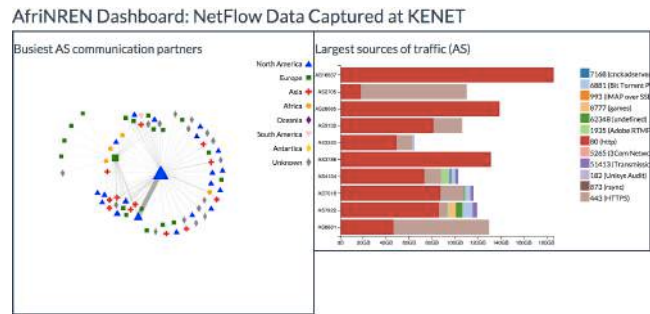


Figure 4: AfriNREN Web Application Dashboard

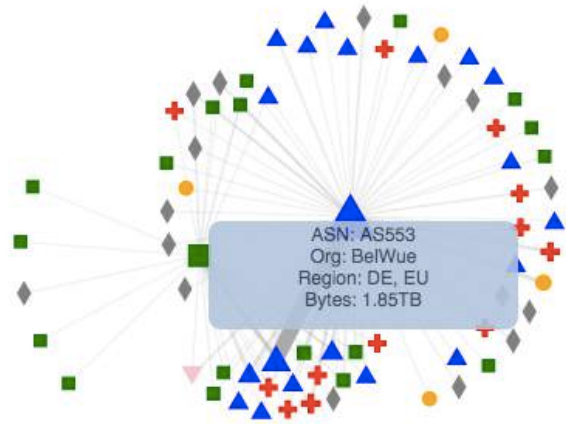


Figure 5: Node Tooltip in ‘Busiest AS Communication Partners’ Network Graph

The right panel of the dashboard – titled ‘Largest Sources of Traffic’ – gives a row chart displaying the ASes that were the largest sources of traffic sent through the NREN. Each row in the row chart is a separate AS and the AS number can be read on the y-axis by the base of the bar. Each row shows the total traffic in bytes sent from the related AS to all destinations and the total is broken down by the port from which the traffic originated. The port number is linked to the ‘type of traffic’ in the legend on the right panel of the dashboard as seen in Figure 4. Although port numbers do not directly correlate to the traffic type, it is a useful indicator for network managers when analyzing network traffic. Users can access finer-grained detail on the traffic breakdown by hovering their mouse over the sections of a bar in the row chart to show a tooltip containing information of the AS number of the AS from where the data was sent, the port number of the source of the data of the currently selected section, and the total number of bytes of data sent from the selected port of the selected AS.

6. EVALUATION

The evaluation combined a task completion test as well as a usability questionnaire – the System-Usability Scale. It was conducted with 20 student users.

6.1 Experimental Methodology

6.1.1 System Usability Scale (SUS)

The SUS questionnaire was used to evaluate the usability of the dashboard.

This scale cannot diagnose a system’s problems so it must be used in conjunction with other tests to get feedback on where to

improve a system. The SUS was given to users along with three open-ended questions for more detailed feedback. These questions were “What features did you find useful?”, “What features do you feel are missing?”, and “What general feedback could you give?”. The responses to these questions could help direct subsequent design phases.

6.1.2 Task Completion Test

We included a task completion test in our usability test to receive feedback on whether or not users were able to complete the tasks that a network manager of an African NREN would complete when exploring a dataset. Failure to complete the tasks would indicate the system’s lack of a feature, while incorrectly completed tasks would indicate a poorly designed feature. This feedback could be used to direct improvement on the features of the system.

Thirteen questions were created to test whether a user using the dashboard could correctly complete the tasks of a network manager as outlined in the research question. These tasks are: identifying the Autonomous Systems (ASes) that communicate the most; determining the amount of traffic transferred between communicating ASes; identifying which continents these ASes are located on; and determining the type of traffic being sent from large sources of traffic. Each question tests the effectiveness of the dashboard in completing of one or more of these tasks.

The task completion test included the following 13 questions:

- 1) Which two Autonomous Systems (ASes) exchange the most traffic? Please write down their ASNs.
- 2) Which AS is the busiest (sent or received the most traffic) in Europe? Please write down the ASN.
- 3) How much traffic was exchanged to and from the busiest AS in Europe?
- 4) Which AS is the busiest?
- 5) Which African AS transferred the second most traffic with AS174?
- 6) Are there more ASes represented from Africa or Asia?
- 7) Which AS in Africa is the busiest?
- 8) How much port 80 traffic was sent from the above AS?
- 9) Which AS is the second largest source of traffic?
- 10) Which AS is the largest source of HTTPs traffic?
- 11) How much traffic was sent from port 6881 from AS7922?
- 12) List the ASes which sent more HTTPS traffic than HTTP traffic.
- 13) How much port 80 traffic (in gigabytes) was sent from the source of the most port 80 traffic?

The mapping of which tasks were tested in which question is shown in Table 1.

Table 1: Task to Question Mapping

Task description	Questions that test task
Identify nth ‘busiest’ AS	2, 3, 4, 7
Identify the continent of an AS	3, 5, 6, 7
Identify nth ‘busiest’ communication partners by traffic exchanged	1, 4, 5

Identify nth largest source of traffic	9, 10, 13
Identify the type and quantity of traffic sent from an AS	3, 8, 10, 11, 12, 13

6.1.3 Users and Experiment Process

The experiment was a user evaluation held in 30-minute long timeslots over 2 consecutive days. 20 users participated in the experiment who ranged from 2nd year Computer Science students with familiarity of computer networking to Honours level students with formal computer networking education and practical experience. The user evaluation included a 13 question task completion test and the SUS questionnaire appended with 3 open-ended feedback questions. The users were compensated at the University’s recommended hourly rate for student experiment participants. The appropriate ethical clearance was received from the Science Research Ethics Committee and Department of Student Affairs for using students in our evaluations.

After an introduction from an experiment facilitator and an assurance of anonymity of the feedback they might provide, the users were asked to sign a consent form detailing the anonymity of the test as well as the compensation for their participation. Next they were given a printed document containing a walkthrough guide which explained the features of the dashboard and how to use it. The document gave the user instruction to visit a URL address in the Web browser, to perform tasks using the dashboard and write down answers to 13 different questions. The users were informed that they were allowed to ask the facilitator for assistance if they ran into a problem or wanted clarification on a task but should not share answers with one another. Once the users completed the tasks, they were asked to complete the SUS survey online and provide written feedback in response to three open-ended questions about the features of the dashboard.

The users conducted the tasks using the AfriNREN Web application and the Chrome Web browser. The dashboard was set to use the same dataset for all participants and did not change during the experiment.

The evaluation took place in a computer laboratory that was co-occupied with students not participating in the evaluation. Noise level and potential distraction were not controlled; however, this setting may closer resemble the working environment of the intended users of the system where distractions from co-workers are likely. Shneiderman makes an argument for evaluations performed in settings that are not noise and distraction controlled because they provide more realistic feedback to the researcher as this setting more closely resembles a work environment in which the visualization would be used [13].

6.1.4 Processing of results

The answers to the task completion questions were checked and each answer was marked either correct or incorrect. First a task completion ratio was calculated on each user’s answers (ratio of the number of questions correctly answered to the number of questions asked). The task completion ratio should be as close to 100% as possible however Sauro found the average to be 78% [11].

We calculated the average (ratio of correctly answered questions to total answers given) for each question for use in determining whether the task tested by the question was achievable using the dashboard or not. Where questions had an average below 78%, the incorrect answers were inspected to find common incorrect

answers for use in determining whether the task question was ambiguous or whether the dashboard provided ambiguous answers. The questions that tested the same tasks were averaged to determine the dashboard’s effectiveness in task completion.

The SUS responses were scored accordingly [5] and the results are discussed in section 7.

The feedback responses to the questions appended to the SUS survey were manually encoded and analysed for recurring themes to determine the most commonly requested improvements and the most commonly offered criticisms.

7. RESULTS & DISCUSSION

The results from the task completion test, the System-Usability Scale survey as well as the open-ended feedback questions are presented and discussed below.

7.1 Task Completion Test

The questions from the task completion test were marked and the percentage of correct answers for each question are presented in Figure 6.

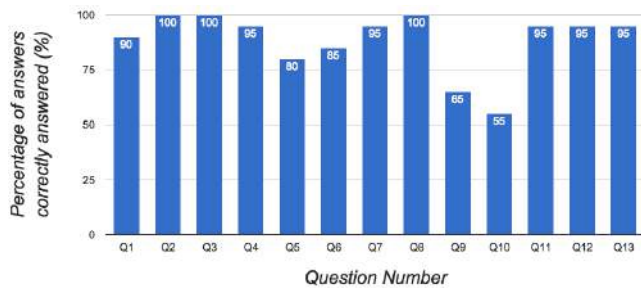


Figure 6: Task Completion Test Results

The results show 11 out of the 13 questions (84.6%) in the task completion test had been answered correctly by users with high success rates (> 80%). These results reflect positively on the usefulness of the dashboard in supporting the tasks given to the users.

The results of the task completion test highlight problems with questions 9 and 10 of the task. Figure 6 shows a 65% and 55% successful completion rate of each question respectively. On inspection of the answers given to question 9, we discovered that 4 of the 7 incorrect answers given (57.1%) were the same. The question was “Which AS is the second largest source of traffic” and although worded correctly, there was no emphasis placed on “source”. If not read correctly, this could lead to the use of the wrong visualization in the dashboard to answer the question. Judging from the results, this was the largest cause of error. The network graph (which shows combined sent and received totals) was incorrectly used to answer the question where the row chart (which shows just the source total) would have given the correct answer.

Inspection of the answers given to question 10 revealed that 7 of the 9 incorrect answers (77.8%) were the same. Question 10 was “Which AS is the largest source of HTTPs traffic?” and the most common incorrect answer given was the AS that was the largest source of HTTP traffic. It appears that the largest cause of error for question 10 was a misreading of the question.

These questions are mapped to the “Identify the nth largest source of traffic” and “Identify the type and quantity of traffic sent from an AS” tasks. Analysis of the feedback given in the open-ended feedback questions highlights the theme of confusion over the

wording of the questions of the task test, further supporting our theories about the low success rate for questions 9 and 10.

The results from the task completion test were grouped by the question to task mapping and the new results displayed in Figure 7.

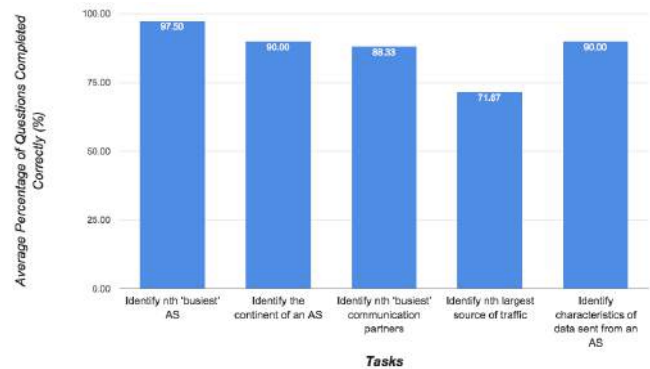


Figure 7: Task Completion Test Results Grouped by Task

Of the five tasks, four tasks were completed correctly with high rates (> 80%).

The low success rate of task 4 (“Identify nth largest source of traffic”) can be traced back to the problems with question 9 and 10 discussed earlier. Questions 9 and 10 made up 2 of the 3 questions that mapped to task 4. The remaining question mapped to task 4 was question 13 which had a 95% success rate.

Figure 7 identifies task 3 (“Identify nth ‘busiest’ communication partners”) as the task with the second lowest rate of successful completion. The questions that mapped to this task (1, 4 and 5) required the user to identify the thickest “links” between the nodes of the network graph. As highlighted in the theme analysis of the feedback given, the users found this task difficult to perform. The results shown in Figure 7 support the feedback and future design iterations of the dashboard should focus on supporting this task.

The high successful completion rate of the tasks (ranging from 71.67% to 97.50% with an average of 87.5%) was at first glance supporting of the dashboard’s effectiveness. However, on reading the feedback given from the users, and identifying common recurring themes within the feedback, it was clear that the design of the dashboard was not satisfactory. Given that we designed both the dashboard and the questions of the task completion test, we suspect that the questions were not a fair test of the dashboard’s effectiveness but instead only tested the features that were strong and avoided features that were weak.

7.2 System-usability Scale

The results from the scored SUS answers are presented in Figure 8.

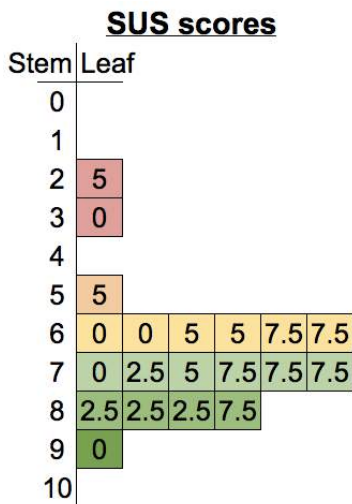


Figure 8: Stem and Leaf Plot of SUS Scores

The average SUS score for our usability tests was 68.5. This is close to the average SUS score of 68 found by Sauro in their study of 500 SUS scores [11]. However, as Figure 8 highlights, our results include two concerning scores (25 and 30) which show that our system was rated very poorly by some users.

Bangor et al. [1] performed a study that found a strong correlation between SUS scores given by users and the adjective those users used to describe the overall usability of the system [1]. They then created a mapping of SUS scores to a scale of adjectives in order to bring more descriptive meaning to the SUS score given to a system. SUS scores that fall within a range can be described using the adjective related to that range. The average SUS score of our evaluation (68.5) fits the range described by “Good” as seen in Figure 9. However, individual scores in our result set fit the “Poor” description with one of them sitting on the border between “Poor” and “Worst Imaginable”. It is clear that the usability of the dashboard requires further design and implementation iterations based on the current feedback.

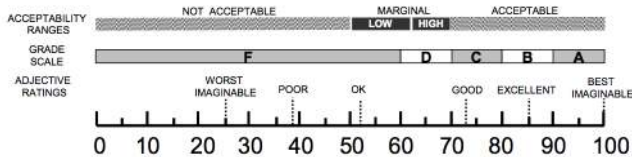


Figure 9: SUS Score Adjective and Grade Scale

7.3 Feedback Theme Analysis

Appended to the SUS questionnaire were three open-ended feedback questions that were optional to answer. Findings from the manual theme analysis of the responses to the open-ended feedback questions are presented in Figure 10.

Three questions were asked: “What features were useful?”, “What features are missing?” and “Please give any general comments”. From the 20 users tested, 16 responses were recorded for question 1 and 3, and 15 responses were recorded for question 2. The response rates for the feedback questions were 0.8, 0.75 and 0.8 respectively.

The responses had a median word count of 11.5, 15 and 16 respectively. Although this doesn’t provide much insight itself, a

notable outlier of a response was a 297 word response to question 2 from a user who gave the dashboard a SUS score of 25.

The feedback was mined for recurring themes by manually colouring terms and concepts within the feedback where similar terms and concepts were coloured with the same colour. It must be noted that “similarity” here is determined subjectively and so ambiguity or misinterpretation could lead to some error of encoding.

Once the terms were coloured, the recurring colours were counted and the rate of their appearance was calculated per question.

This basic thematic analysis allows us to identify common feedback suggestions and criticisms to guide future design iterations.

The results highlighted the most common feedback themes were: the need for integration between the visualizations on the dashboard; the ability to sort the row chart by different measures; and a filtering feature.

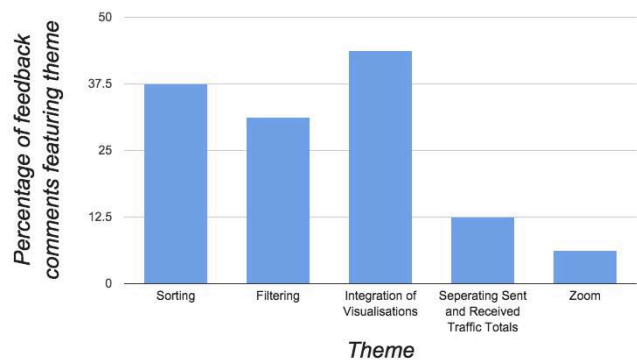


Figure 10: Theme analysis results from user feedback

These features weren’t ignored in the design thinking but the priority of producing a testable dashboard meant that only a subset of the desired features were implemented.

7.4 Discussion

The users involved in the experiment were students and not domain experts. Carpendale highlighted this concern along with the use of small datasets and simple tasks in evaluations of information visualizations as common reasons for why the results for these evaluations are not convincing [6]. The experiment performed for this project has similar characteristics to Carpendale’s scenario in that the experiment featured simple tasks using one dataset where in reality many diverse datasets would be expected. The results from the experiments can therefore not be rationally generalized to indicate the success of the project’s aims and the dashboard’s usefulness to African NREN network managers as domain experts. However, the results do have use for others seeking to further explore the use of interactive information visualisations in assisting the development of African NRENs.

8. CONCLUSIONS

This work set out to answer the research question “can a dashboard of interactive non-geospatial visualizations of NetFlow data effectively communicate network traffic information about a network to the network managers?”. Through a usability evaluation and task completion test, we found that the final design of the dashboard was able to effectively communicate a network’s traffic information to the network managers. The experiment and user feedback showed that features implemented in the AfriNREN dashboard were usable and assisted in the completion of tasks

specific to the NetFlow data represented – namely: identifying the Autonomous Systems (ASes) that communicate the most; determining the amount of traffic transferred between communicating ASes; identifying which continents these ASes are located on; and determining the type of traffic being sent from large sources of traffic.

However, the dashboard does not support all the necessary features and tasks as highlighted by the feedback received from users in the user evaluation performed. The dashboard requires additional design efforts to correct the current problems with usability. Furthermore, additional features would have to be designed, implemented and evaluated before the dashboard could be used to determine latency information and network structure between Autonomous Systems traversed by network traffic in Africa.

8.1 Future Work

This project has shown that there is a need for a larger feature set on the dashboard. Future work would include involvement of real users to gather requirements, feedback on design changes and evaluation to determine the effectiveness of the dashboard. With the involvement of domain experts, the results from an experiment would convincingly measure the success of the system.

The features should be expanded to include filtering, ordering and linking between the visualisations of the dashboard. Furthermore, the dashboard should be evaluated using multiple sources of NetFlow data to evaluate the versatility and generalizability of the dashboard.

9. REFERENCES

- [1] Bangor, A., Kortum, P., and Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4.3 (2009), 114-123.
- [2] Becker, R., Eick, S., and Wilks, A. Visualizing network data. *Visualization and Computer Graphics, IEEE Transactions on*, 1 (1995), 16-28.
- [3] Boitmanis, K., Brandes, U., and Pich, C. Visualizing internet evolution on the autonomous systems level. In *Graph Drawing*, Springer, Berlin Heidelberg, 2008, 365-376.
- [4] Bostock, M., Vadim O., and Jeffrey H. D³ data-driven documents. *Visualization and Computer Graphics, IEEE Transactions*, 17.12 (2011), 2301-2309.
- [5] Brooke, J. SUS: a quick and dirty usability scale. In Jordan, P.W., Thomas, B., Weerdmeester, B.A. and McClelland, I.L. ed. *Usability Evaluation in Industry*, Taylor and Francis, London, 1996, 189-194.
- [6] Carpendale, S. 2008. Evaluating information visualizations. *Information Visualization*. Springer Berlin Heidelberg, 19–45. http://doi.org/10.1007/978-3-540-70956-5_2[9]. Carr, D. A. 1999.
- [7] Chavula, J., Feamster, N., Bagula, A., Suleman, H. Quantifying the Effects of Circuitous Routes on the Latency of Intra-Africa Internet Traffic: A Study of Research and Education Networks. in *Proceedings of the AFRICOMM conference*, (Uganda, 2014), AFRICOMM, 1-10.
- [8] Fryer, T. Research and Education Networks around the World and their Use. in *Segunda Conferencia de Directores de Tecnología de Información*, (Lima, 2012), Red Clara, 1-13.
- [9] Gupta, A., Calder, M., Feamster, N., Chetty, M., Calandro, E., and Katz-Bassett, E. Peering at the Internet's frontier: A first look at ISP interconnectivity in Africa. in *Passive and Active Measurement Conference*, (Los Angeles, 2014), Springer International Publishing, 204-213.
- [10] Knight, C. Visualisation effectiveness. In *International Conference on Imaging Science, Systems, and Technology*, (Las Vegas, 2001), CISST, 249–254.
- [11] Lad, M., Massey, D., and Zhang, L. Visualizing internet routing changes. *Visualization and Computer Graphics, IEEE Transactions on* 12, no. 6 (2006): 1450-1460.
- [12] Sauro, J. What is a good task-completion rate? MeasuringU, March 21, 2011. Retrieved October 15, 2015, from MeasuringU: <http://www.measuringu.com/blog/task-completion.php>
- [13] Sauro, J., Lewis, J. When designing usability questionnaires, does it hurt to be positive? *Proceedings of the Conference in Human Factors in Computing Systems CHI*, (Vancouver, 2011), ACM, 2215-2224.
- [14] Shneiderman, B., Plaisant, C. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. in *AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, (Venice, 2006), ACM, 1-7.
- [15] Teryl, T., Paterson, D., Glanfield, J., Gates, C., Brooks, S., and McHugh, J. Flovis: Flow visualization system. In *Conference For Homeland Security, CATCH'09. Cybersecurity Applications & Technology*, (Washington, DC, 2009) IEEE, 186-198.
- [16] Tullis, T. S., Stetson, J. N. A comparison of questionnaires for assessing website usability. in *Usability Professional Association Conference*. (Minneapolis, 2004), Usability Professionals Association. 1-12.
- [17] UbuntuNet Alliance. What is UbuntuNet Alliance? UbuntuNet. Retrieved April 25, 2015 from UbuntuNet: http://www.ubuntunet.net/sites/default/files/What_Is_UbuntuNet_v1_20141030.pdf
- [18] Wassink, I., Kulyk, O., and Van Dijk, B. Applying a user-centered approach to interactive visualization design. *Trends Interact. Vis.* (2009), 175–199. DOI:<http://dx.doi.org/10.1007/978-1-84800-269-2>