UNIVERSITY OF
LINCOLN

Visualising Network Traffic Data From Air Traffic Control Radio Systems

A thesis submitted in fulfilment of the requirements for the degree of

MSc Computer Science by Research

in the

School of Computer Science

Author:
Adam Walker

Supervisors:
Mrs. Yvonne James
Dr. Chris Headleand

November 19, 2019

Declaration of Authorship

I, Adam WALKER, declare that this thesis titled, "Visualising Network Traffic Data From Air Traffic Control Radio Systems" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at the University of Lincoln.
- This work and associated artefact have not been submitted for any other degree at the University of Lincoln or any other institution.
- Where I have consulted the published work of others, this is always clearly attributed.
- I have acknowledged all main sources of help.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

Signed:

Date:

Abstract

Visualising Network Traffic Data From Air Traffic Control Radio Systems

by Adam WALKER

In recent years the aviation industry has begun to embrace digital technology for Air Traffic Control (ATC) radio systems. This change has created challenges not only for the industry but also for personnel. However, this implementation offers many improvements over older systems; more precise control, straightforward integration with other ATC systems and a more efficient way to provide software updates. The challenge for personnel is to develop a new skillset enabling a learning transition from analogue to digital systems, with a specific emphasis on computer networking skills.

This project was undertaken in collaboration between the University of Lincoln (UoL) and Park Air Systems (PAS), an industry-leading provider of Air-Space communication solutions. A system has been developed to find a mechanism to monitor and visualise network traffic. The use of graphs provides a direct interface for the end-users, enabling a mechanism for identifying performance issues to meet the transitional challenges from analogue to digital. An easy to use interface has been designed, which will enable non-technical users to interact effectively with the system.

Considerable testing was undertaken to investigate the system usability concerning the practical application for users with limited networking experience. A survey provided a range of quantitative and qualitative data which was further analysed to scrutinize user perspectives on system usability. This involved engineers from PAS and postgraduate students from UoL to compare results between direct industry personnel and unaffiliated participants.

Acknowledgements

I would first like to thank my supervisor Mrs Yvonne James for all of her support and guidance throughout this project, it would not have been possible without it.

I would also like to thank Dr Chris Headleand for his additional help and support.

To those at Park Air Systems, I would like to thank you for your help and collaboration during this project. In particular Darren Wallace, Neil Upton and Danny Milligan.

To my friends Neville, Charlie, Tom and Joe, thank you for all your help and inspiration, and also for listening to my complaints for the last year and not getting too annoyed.

Lastly, To Meike, thank you for all of your encouragement and support to finish this project. It has helped more than you know.

List of Abbreviations

ARP	Address Resolution Protocol
ATC	Air Traffic Control
ATM	Air Traffic Management
CCNA	Cisco Certified Networking Associate
CMV	Coordinated and Multiple Views
CIDR	Classless Inter-Domain Routing
DCHP	Dynamic Host Configuration Protocol
HCI	Human Computer Interaction
ICMP	Internet Control Message Protocol
IDS	Intrusion Detection System
IP	Internet Protocol
ISO	International Organisation for Standardisation
GDPR	General Data Protection Regulations
MAC	Media Access Control
MCP	Multiple Comparisons Problem
MIB	Management Information Base
NIC	Network Interface Controller
NMS	Network Management System
OSI	Operating Systems Interconnection
RFC	Request For Comments
RTP	Real-Time Transport Protocol
SAM	Student Activity Monitor
SES	Single European Skies
SIP	Session Initiation Protocol
SIEM	Security Event and Information Management
SNMP	Simple Network Management Protocol
STEM	Science, Technology, Engineering and Mathematics
SUMI	Software Usability Management Inventory
SUS	System Usability Scale
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UMUX	Usability Metric for User Experience
VLAN	Virtual Local Area Network
VoIP	Voice over Internet Protocol

List of Figures

List of Tables

Table of Contents

Chapter 1

Introduction

1.1 Background

In airports across the world, Air Traffic Controllers make use of radio systems for communication between the ground and air traffic. These systems are some of the most vital assets at the airport, and they must be correctly installed, configured and maintained. The radios create a computer network of connected devices, each operating on different frequencies or serving as a backup for another device.

In recent years the aviation industry has been in the middle of a transition towards digital systems and away from the analogue systems which have traditionally been used. This transition mirrors that of many other industries which have also sought digital integration between their hardware, software and IT infrastructure, and will provide benefits over previous systems. One such benefit, for the aviation industry, will be the ability to provide improved frequency management of radio communications, allowing for increased communications capacity an increasingly crowded airspace [1]. This particular example is pertinent due to the trend in air traffic growth, which is expected to put a strain on and therefore increase demand for, aviation infrastructure. The Single European Sky (SES) initiative to unify the airspace over the European Union has a performance target for Air Traffic Management (ATM) systems to handle a three-fold increase in capacity post-2020 [2].

However, this transition is not without its problems. The engineers who have previously worked with analogue systems will require retraining to develop a new area of expertise. Where engineers could previously make use of electronic equipment to measure signal changes in an analogue system to determine how it is functioning, this is no longer possible with digital 'black box' systems. As a result, engineers who may have years of experience may now be potentially at a significant disadvantage if they do not have a strong understanding of computer networking.

This research is in collaboration with Park Air Systems (PAS) and the University of Lincoln. Park Air Systems is an industry leader in the development and manufacture of Air-Space communication solutions for Air Traffic Control systems, with both civilian and military clients. The engineers who are employed at PAS are undergoing retraining to understand digital radio networks. The problem affecting Park Air Systems is that to verify a radio network is functioning correctly, they need to conduct tests and monitor the system for indicators of poor performance. Furthermore, each system is bespoke designed according to the specific requirements from the client, meaning that each radio network has a unique set of requirements. This problem is too significant for any single solution; however, two key challenges have been identified.

Firstly, it is difficult to present a clear and concise overview of the network and its performance for the engineers working on the network. Secondly, to provide a system capable of meeting this challenge while also remaining small, easy to use for non-experts, and that is not restricted by commercial licensing. Given this challenge, Park Air Systems have presented a set of requirements for the research and development of a piece of software which can assist their engineers.

1.2 Aims and Objectives

This research project aims to develop a system that can meet the challenges mentioned above. The requirements are that a visual representation of the network is provided at a level that can be used to aid engineers in completing their confidence checks.

The research also investigates the use of network monitoring systems, integrating principal features into a system which makes use of visualisation techniques to help present the information in a more transparent way than is typical of expert-level systems.

Research Question

Can a high-level, bespoke visualisation system of ATC radio networks be usable for engineers with limited expertise in computer networking?

Research Objectives

These research objectives have been set as key milestones to provide a way to measure the state of progress throughout the research.

1. Develop a Network Monitoring System for Park Air Systems

The primary objective of this research project is to incorporate different elements of network capture and monitoring systems, and then to visualise the information in a novel way to provide a potential solution to the problem faced by Park Air Systems.

2. Investigate the usability of different presentation mechanisms

The usefulness of the Network Visualisation System is predicated upon its' ease of use and relative simplicity compared to alternatives; therefore, a study to investigate how different techniques affect the usefulness is to be conducted.

1.3 Main Contributions

The main contributions of this research have been the ways that this project has looked at to present network information more clearly and efficiently to users of a visualisation and analytics system. It is especially important to do this, given that the systems already available are often designed by and for professional network engineers, and as such are not very friendly for non-expert users. Therefore by focusing on ease of use for a system typically reserved for expert users, the barrier to entry can be lowered and allow users to become competent at a system much sooner than they might otherwise be.

The original intention of the project was to collect the network data live and display it in real-time, however, due to concerns about providing access to live data or environments from Park Air Systems, the focus of the project shifted to the analysis and presentation of data that had already been collected. It is this focus on using analysis and visualisation techniques to maximise the ease of understanding for limited experience users in the field that this research makes contributions.

Data analysis plays a significant role in condensing and presenting information in a way that is useful and actionable. There is often more data than can be understood in any meaningful way on its own, however by making use of data analytics techniques and algorithms; valuable information can be learned without a need to see the raw data. This is why it was essential to ensure that the system developed would isolate the required data and provide that as necessary, and remove any extraneous or corrupted data wherever possible to ensure that only the most relevant data was analysed and presented. Furthermore, the use of data analytics allows for more efficient interactions with the data, as there is no need to manually investigate pieces of information when they can be calculated ahead of time and presented as analysed results. This process of automating analysis is beneficial as it both saves time for the users as well as lowers the standard of training required to operate a system and understand the data. This work allowed for data visualisation to be much more straightforward than it may have otherwise been and was beneficial with ensuring the system was as simple to use as possible.

The contributions of this research are in the areas of data visualisation and human computer interaction. With regards to data visualisation, the focus of the project shifted towards this area after discovering the problem of users who lack experience, while simultaneously learning to use expert level software. It is also essential to understand how different ways of measuring this user interaction can affect the results of such studies and their interpretations. There is an International Organisation for Standardisation (ISO) standard definition of usability which is explored further in chapter 2.5. This definition is the starting point for many surveys designed to measure perceived usability. Many different metrics can be measured in user research, from objective metrics such as task time and error rates to the previously mentioned usability surveys, which can be more subjective [3].

1.4 Outline of Thesis

The structure of this thesis mirrors the process undertaken to complete the research project, beginning with a thorough investigation of the background material and the current state of the field for both academic and commercial work. Details regarding the content in each chapter and their structure are provided below.

- **Chapter 2 - Related Work** contains background research and information about previous work done in related fields which may be relevant to this project. The primary topics covered are related to networked systems and how they are monitored, data analytics, visualisation systems and also the usability of the software.
- **Chapter 3 - Methodology** documents the methods used throughout the research project. The details of a user study undertaken to investigate system usability are also described in this chapter.
- **Chapter 4 - Development** details the processes and practices followed to create the visualisation system as initially described by Park Air Systems and further used in the usability study as detailed in chapter 3 methodology. The chapter is divided into the design of the system, including initial project requirements, challenges and changes to solve them, and the implementation of said designs. The practical challenges, technical details and process of development is detailed in this section.
- **Chapter 5 - Results** contains the data collected from the usability study and is presented alongside a discussion and interpretation of the results and their context. Additional findings from throughout the project are also included and discussed here.
- **Chapter 6 - Conclusion** presents the overall conclusion to the research project and an evaluation of the project results concerning the initial research aims and objectives. A summary of the work done, any challenges, issues and changes are discussed and how they impacted the project, as well as avenues for potential future work, are also included.
- **References & Appendices** are provided in the sections following chapter 6. A list of references to the work of others throughout this thesis is presented in full. The appendices contain copies of study materials, sets of results too long to include in chapter 5 or any other relevant materials that could not be included in the main body of content.

Chapter 2

Related Work

The development of a system for monitoring the performance and configuration of a network through data analytics and visualisation of connected radio systems is a novel technical challenge. Individually, each of these features has been created and used in existing systems in the past; however a bespoke system capable of this functionality in a focused and targeted manner, aimed specifically for use by Park Air Systems, does not currently appear to be available in the commercial or open-source communities. Furthermore, by combining these features from larger and more complex systems into something smaller with a focus towards non-expert users, the resulting system may prove to be more useful than the sum of its parts. This chapter will review related work done in the fields of *Network Systems*, *Data Analytics*, *Data Visualisation* and *System Usability*.

2.1 Background of Air Traffic Control Radio Systems

This section provides information relating to the general context and background details which are relevant to the research project as a whole. It contains information directly from Park Air Systems as a first-hand source of information relating to the aviation industry at large, and more directly the technical and business details of their company as they relate to this research. Accordingly, this information from PAS may not be accompanied by citations to academic literature as much of the information is regarding standard practice and general contextual information about the field and not any specific research that can be cited.

2.1.1 Radio Networks

Radio systems in the Air Traffic Control industry are moving from 4-wire E&M and E1 based telecommunication systems to Session Initiation Protocol (SIP) [4] and Real-time Transport Protocol (RTP) [5] over Internet Protocol (IP) [6]. Traditionally systems would be tested with electronic test equipment. The radio systems designed and manufactured by Park Air Systems, make use of a proprietary data transmission protocol which they refer to as MARC, this protocol handles much of the radio specific functions and operates using the Transmission Control Protocol (TCP) [7] protocol on port 5001.

Although the radio systems communicate wirelessly with other radios, the local radio networks themselves are all using wired connections, which is the exclusive focus of this project.

2.1.2 Park Air Systems

As part of their process, before shipping a radio system to a client for installation, the system is constructed in the Park Air Systems facility to undergo testing and validation. In addition, the client is invited to come to view their product at this stage to further verify its completeness. The use of a system to provide a clear overview of the network could be beneficial to engineers in need of a simpler tool to see the network at a high level to determine where they need to focus their effort more efficiently at a low level. Moreover, a tool designed explicitly for ease of use could potentially be used to present information to clients, who may not have a great deal of experience with network analysis.

2.2 Network Systems

The primary focus of this project is to develop a piece of bespoke software for visualising information about the radio networks created by Park Air Systems for their clients. Therefore, this section presents work related to the field of networking, as a thorough understanding of the work done regarding network systems, tools and principles are essential to this project.

2.2.1 Network Monitoring

Network monitoring is the process of monitoring a computer network for abnormal behaviours, such as unresponsive devices or slow data transfer rates. Network monitoring is done through the use of a network monitoring system. These systems are a standard tool for providing the network administrators with a view of what is happening on the network by monitoring the various devices and components and the data being sent across it. The range of functionality that a Network Management System(NMS) can perform varies from system to system, but common functions include; monitoring bandwidth usage across the network, identifying faulty and incorrectly configured devices or detecting network topologies. The information collected by these systems is then used to inform and alert the user to any issues in the network.

It is crucial that as systems designed to aid users with analysing and understanding (potentially large and complex) networks, they need to be able to present the information clearly to their users so that the information can be actionable and thereby useful. Typically, this need for clarity is not a significant issue, as the users of these systems are usually network engineers or administrators and are therefore qualified and experienced in networking. In addition, the network(s) that they monitor can often be the same one(s) for an extended period, such as their corporate workplace network.

One issue with these systems is that they are very focused on expert-level users, who have the level of experience and training to understand and make use of the system and especially its advanced features. This means, however, that any non-expert-level users are potentially left at a disadvantage, as they will likely need more time to learn and understand the system before they can use it as effectively.

This issue is one of the primary motivations behind the proposal from Park Air Systems to develop a clearer and more straightforward system that is capable of displaying network information without the expert level complexity of a more extensive system. In having a system capable of supporting not only the technical engineers but potentially any other staff or clients, in understanding how the network is operating at a high level, the tests and demonstrations that they undergo could be made a little easier to conduct.

2.2.2 Key Components of A Network

To develop a bespoke system for analysing and then visualising captured network traffic, as well as other information about these radio networks, it is necessary first to clarify which elements are crucial to a network and network monitoring systems, and how they relate to this project.

The field of networking is enormous, with many areas relevant to developing a system like the one proposed for this project. This section will discuss a few key concepts of computer networking that are essential to an understanding, and monitoring, a network.

2.2.3 Network Topologies

A network topology can be defined as "the arrangement or relationship of the network devices and the interconnections between them" [8]. This topological information can be used to examine the design of the network, which in turn can reflect its purpose [9]. Network administrators can also use this information to analyse the structure of the network to locate faults, such as bottlenecks, and make performance predictions [10], [11]. Also, this information can allow the administrator to get a much clearer understanding of the network status, thereby allowing them to manage the network more effectively [12]. Finally, the purpose of a network topology is to provide a representation of the network, and this can be done as either a physical or a logical topology.

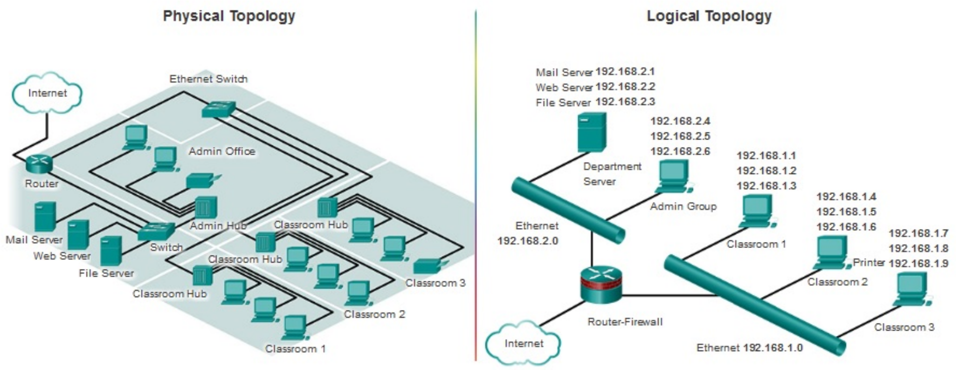

FIGURE 2.1: Physical and Logical Topologies [8]

Physical Topology

The physical topology refers to the physical details of the network, such as the placement and interconnection of various network components. This can be expressed in terms of the Operating Systems Interconnection (OSI) model of networking as a focus on devices from the data link layer (L2) such as switches, rather than network layer (L3) devices such as routers [10]. As such, with a physical topology, we are more concerned with the physical implementation, structure and location of the network than with its logical flow. Physical topologies are usually represented as one of several standard layout types: Mesh, Star, Bus, Ring or Tree.

Each of these reflects a different layout of interconnections between the networked devices, and each comes with benefits and drawbacks to consider, typically cost and performance. The benefits of physical topologies are for those who need to understand the capabilities and design are of the network, the specification of components, the cost of cabling and physical layout of the network are all useful pieces of information that a physical topology can contain. Although some of the physical details from this type of topology may be useful for this project, they are not likely to be essential to the operation of the system. A physical topology, therefore, is not as relevant to this project as a logical one.

Logical topology

In contrast to the physical topology, a logical topology is concerned not with the physical implementation of the network, but rather how signals act on the network [8]. This topology illustrates how data moves within the network, showing details such as the arrangement of connections between devices on the network, how they transfer data from node to node, the location details of each node and representations of OS services [8].

Since the flow of data through the network is what makes up the logical topology, this means that it is possible to have two networks which are physically distinct, having different connection interfaces and being different distances apart for example, but still have the same logical topology. Table 2.1 details several components of a network and whether they would be grouped as physical or logical components.

Network Component	Topology
Traffic flow	Logical
Routing domains	Logical
IP addressing schemes	Logical
Subnets	Logical
Virtual Local Area Network (VLAN)	Logical
Device location	Physical
Cable Installation	Physical

TABLE 2.1: Network Components grouped by topology

Topology Discovery

The act of determining the topology of a network is known as topology discovery, and there are many tools available which can achieve this. These tools can make use of several different techniques to determine the topology of a network and often rely upon standard network protocols designed for such tasks, with two of the most commonly used is Simple Network Management Protocol (SNMP) and Internet Control Messaging Protocol (ICMP) [9], [11].

The use of SNMP for topology discovery works by making requests to all devices which support the protocol, and through the data stored in the Management Information Base (MIB) it can build up a knowledge of devices which have connected on the network. This process relies upon the devices on the network supporting SNMP, which, although common, is not guaranteed for every kind of device. An alternative technique is to use ICMP to discover network devices; this is done by sending ICMP echo queries (pings) and subsequently monitoring for replies [11]. These requests are often made as *ping sweeps* to check for hosts on the network at regular intervals.

These techniques are, however, not without their problems. As previously mentioned, the use of SNMP can only work for devices on the network which support the protocol. Furthermore, according to [12], many studies conducted into topology discovery using SNMP or ICMP have failed to address several key issues:

1. *Discovering Device Type* - With the growth in the number of connected devices and types of devices on a modern network, the ability to detect each device and its type is difficult, given that devices operate at layers two and three of the OSI seven-layer model. Switches are associated as layer two data-link layer devices, and routers as layer three network layer devices. The radio systems have a limited number of devices in the configuration which is typically

sourced from the same manufacturers. This means device discovery can be tailored according to the radio system physical configuration

2. *Topology Visualisation* - This issue relates to a lack of tools which support topology visualisation well, as this project is to develop a system for visualising network information, it is safe to say this issue is being addressed - see section 2.4 below.

3. *Complete Topology Discovery* - There has been less work conducted into topology discovery for interconnecting L2 and L3 topologies compared to discovering each layer separately.

[13] present a discovery algorithm based on SNMP, designed to combine the topologies of the *data link* (L2) and *network* (L3) layers, and improve discovery of the complete network.

2.2.4 Relevant Protocols

Several key protocols are particularly relevant to this project. These protocols are detailed in this section to give some context to how they function, what purpose they serve, and why it is important to this project.

TCP

Transmission Control Protocol [7] is a session layer (L4) communication protocol which handles how packets are transferred from one device to another. It is also one of the primary protocols that make up the TCP/IP protocol suite, which dictates how internet traffic is packaged, routed, transmitted and received.

A large portion of network traffic is transmitted using this protocol, including, SNMP data, and so it is crucial to understand how and why it is used as part of a network. TCP provides a way for transmitting packets of data in an ordered structure when all of the data must be transmitted and received correctly, such as for transferring a file. The protocol initiates a 3-way handshake (shown in figure 2.2) with the target device to create a connection, before transmitting data as a stream of ordered packets. The use of ordered packets allows the receiver to know if it is missing anything and request a re-transmission of that packet.

UDP

User Datagram Protocol (UDP) [14] is another session layer (L4) protocol. It is essential for uses such as streaming audio and video data, which do not need to verify every packet and instead rely on making up for any loss with a constant stream of new data. This is directly relevant to the project, as this is often done with protocols such as Real-time Transport Protocol (RTP), which is used by the radio systems for Voice over Internet Protocol (VoIP) functionality.

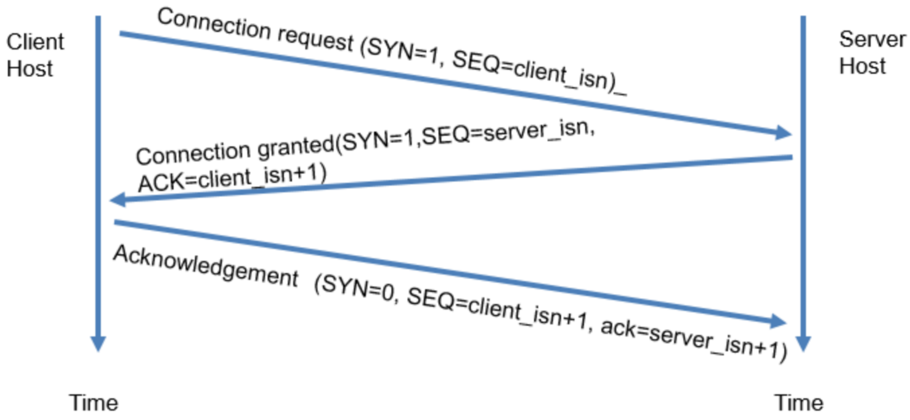

FIGURE 2.2: TCP Three Way Handshake

SNMP

Simple Network Management Protocol (SNMP) [15] is an application layer (L7) protocol in the OSI model, which was developed in order to provide a mechanism for managing device nodes such as servers, routers, switches or workstations on an IP network [16]. The use of the protocol, therefore, allows administrators to monitor performance and solve problems across the network. [17] call it "the *de facto* standard for TCP/IP networks management".

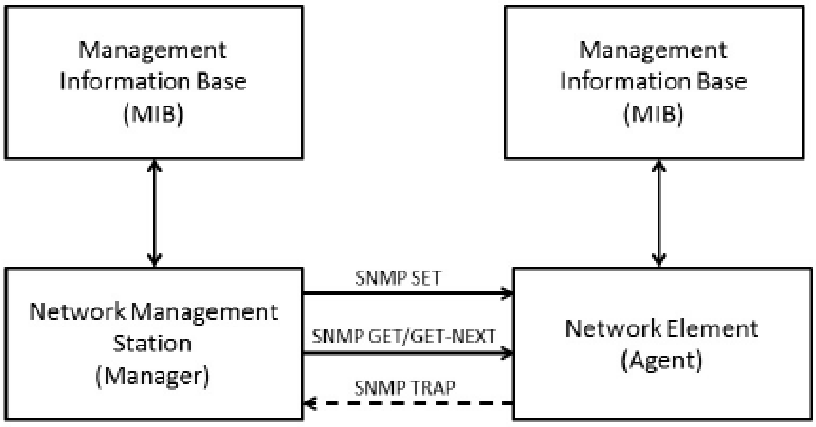

FIGURE 2.3: SNMP Architecture [18]

SNMP operates using a model of distributed agents running on each device and a central manager to control and coordinate them (shown in figure 2.3. An agent is any managed node on a network which is equipped with an SNMP agent software module. The manager is an entity which is responsible for communicating with network devices which have implemented SNMP agents; this is normally a machine running a NMS. This manager can use *get* and *set* actions to receive information or make changes to an agent.

ARP

Address Resolution Protocol (ARP) [19] is a data-link layer (L2) protocol which is responsible for resolving IP addresses into MAC addresses. If there is traffic on a network that is addressed to a device that does not have a corresponding entry on the ARP table, a request will be sent out to resolve this new device [16].

SIP

Session Initiation Protocol (SIP) [4] is one of the most common protocols for enabling multimedia communication sessions, used as part of video, voice, text or other communication applications and services. The protocol is used to initiate, manage and terminate the real-time sessions involving multiple endpoints on IP networks, and is often used to initiate communications for RTP. This means that any real-time communication, such as streaming audio over a network is likely done over a connection that is initiated and managed via SIP.

RTP

Real-time Transport Protocol (RTP) [5] is used for real-time communications over IP, sending data over the network via UDP. The protocol is used in conjunction with SIP, which initiates and manages a connection between two devices, enabling RTP to stream audio and/or video data between the connected devices.

2.2.5 Physical and Logical Addressing

For devices to be able to communicate with one another on a network, they must have some address on the network such that other devices can find them and know where to send data packets. These addresses that identify a device on a network come in two different forms: physical and logical.

Physical Addresses

A physical address for a device is a unique identifier tied to the physical hardware of a specific device. This address, commonly known as a Media Access Control (MAC) address, is assigned to each Network Interface Controller (NIC) by the manufacturer and does not change. MAC addresses are globally unique; however, it would be impractical to try and search the entire internet for a target address, so they are only used internally within their network for communication. A logical address is also required to communicate externally from a network (via a router).

Logical Addresses

Logical addresses, also known as Internet Protocol (IP) addresses, are assigned to a device when it is connected to a network, this is typically done automatically by the

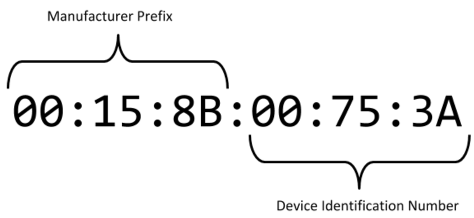

FIGURE 2.4: MAC Address Structure

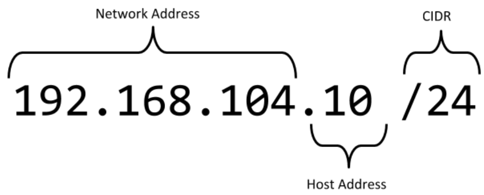

FIGURE 2.5: IP Address Structure

Dynamic Host Control Protocol (DHCP). These IP addresses contain a network and a host portion which enables devices to communicate outside of their network. IP addresses use an identifier known as Classless Inter-Domain Routing (CIDR) to identify the subnet mask, which indicates the portion the address assigned to the network of the host. Logical addresses are (usually) not permanent and can be changed or reassigned to other devices when no longer required by the current device.

2.3 Data Analytics

Data analytics is the process of analysing data to extract, interpret and draw conclusions from the information it contains [20]. Organisations continue to collect and analyse increasingly large and complex sets of data to drive their decision making and improve their internal operations [21].

Concerning this project, the use of analytics is to enable more information to be extracted from the data being presented to the user. The potential benefits are numerous, for example, automating everyday tasks, comparing different datasets more effectively or providing more detail in an overview style display.

Data analysis is commonly divided into distinct categories, each with their use cases and challenges [22]:

1. *Prescriptive* - This type of analytics shows what actions are recommended to be taken. This is one of the most common kinds of analysis, as it can provide much value to an organisation.

2. *Predictive* - This analysis shows what outcomes are likely to happen in the future based upon past actions. This tends to be done through machine learning models, which are not relevant to this project.

3. *Diagnostic* - An analysis of past data to determine what happened and why. As the name suggests, this is primarily for diagnostic use, and therefore of lesser relevance to this project.

4. *Descriptive* - Also known as *data mining*, this is an analysis of what is currently happening based upon incoming data. Given the nature of this project, this seems to be the most suitable type of analytics.

2.3.1 Techniques for Analytics

Techniques for data analysis can vary depending upon factors such as the type and amount of data available, the expected outcome and the analysis task. While the specific techniques may vary, the overall workflow does not change significantly. for getting from raw data to results that can be presented to a user, as it relates directly to this project, can be summarised from the *data science process* shown in figure 2.6. Raw data is collected, (pre)processed and the clean data can then be further processed by algorithms, before visualisation for the user.

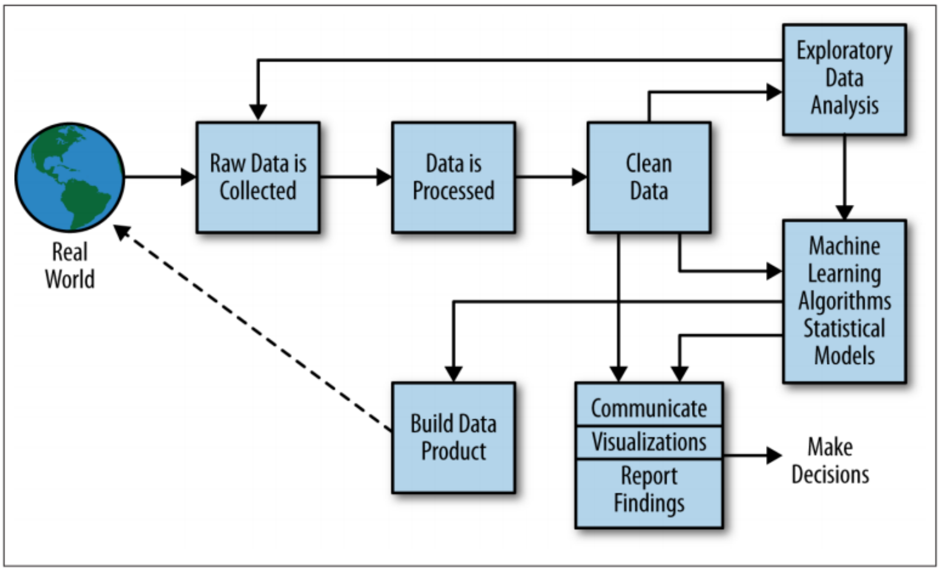

FIGURE 2.6: The Data Science Process [23]

Data Preprocessing

Also known as *cleaning* the data, preprocessing is an essential stage of the process. The purpose of data preprocessing is to remove unwanted or invalid data from the dataset, such as noise or outliers. This stage of processing the data may also involve aggregating or clustering the data together from different sources [24].

[25] describe preprocessing as filtering, transforming and statistically summarising data. Some examples in the literature of data preprocessing include; cleaning and aggregating relevant attributes from the data based upon specific project requirements in order to present data in time intervals to the user [26]. Filtering data, as described by [27], is a technique to decrease the amount of data required for later visualisation. In order to visualise changes in logging infrastructure at Twitter, [28]

use regular expressions and a heuristic algorithm to aggregate specified occurrences in log files before storing in a database for later visualisation.

Finally, according to [24], data often has an element of uncertainty that must be accounted for in some way. This is due to data not always being perfectly precise, or some invalid data was removed, leaving unknown gaps in the dataset. According to [29], some techniques to present uncertain data include; showing data ranges rather than exact values, error bars, or using colours to represent degrees of certainty.

Data Processing

The process of data analytics (also known as *Data Mining*) is intended to 'help support the user in their exploration process' [24]. Techniques for this are best described according to the analysis task or technical challenge they attempt to complete. For example, when performing analysis on real-time data, there are additional technical challenges related to performance, as all calculations must be done in real, or near-real, time to keep up with the data stream. This means that it is crucial to have effective compression and feature extraction techniques to pull the critical data from the stream [30]. Other performance considerations such as latency reduction have been discussed by [26], [31] who suggest that addition distributed or parallelised processing capabilities could improve performance and thereby minimise latency.

Techniques for detecting or predicting unusual usage patterns are common and tend to make use of sophisticated statistical or machine learning models [32]–[34]. This use of predictive analytics is found in many research papers, [35] surveyed 285 papers relating to 'information visualisation for network and service management', and found many whose methods contained some machine learning or genetic algorithms.

Statistical methods used by [26] include time series analysis through the use of histograms and graphs. These methods operate upon clustered or otherwise aggregated data, relying on techniques for structuring the data such as dimensionality reduction [20].

2.3.2 Visual Analytics

Visual analytics can be described as *"the science of analytical reasoning facilitated by interactive visual interfaces"* [36]. The need for visual analytics arose from our capability to gather more data than we can analyse. The use of visual analytics there can allow for the interpretation of more massive sets of data that could otherwise be understood in a raw form. A more specific goal for visual analytics is provided by Keim *et al.* [20], who describe it as combining "automated analysis techniques with interactive visualisations for an effective understanding, reasoning and decision making based on vast and complex datasets". Again, this description is exploring the idea that large and complex datasets are difficult to understand without some form of abstraction into a more understandable form.

They further define the goal of visual analytics as to the creation of tools and techniques which enable the following [20]:

1. Synthesise information and derive insight from massive, dynamic, ambiguous, and often conflicting data.
2. Detect the expected and discover the unexpected.
3. Provide timely, defensible and understandable assessments.
4. Communicate assessment effectively for action.

The use of visual analytics or some form of interactive visualisations could be incorporated into the system for this research project, as it will have to display large sets of captured network data in some way that is clear and intuitive for users to understand.

2.3.3 Challenges with Data Analytics

Following a series of interviews with data analysts in 25 organisations, Kandel *et al.* [21] state that their respondents can typically be grouped into one of three categories: *Hackers*, *Scripters* and *Application Users*. These three groups describe the level of capability and flexibility with the roles of each analyst. The hacker is the most flexible user, who is capable of working with any dataset, can write scripts and perform different statistical techniques. The Scripter has less flexibility but typically a deeper domain knowledge than the hacker, and so can perform deeper analysis on larger datasets. Finally, the Application User has a minimal ability to write scripts or work with complex datasets, they typically work with simpler programs such as Excel rather than using more dedicated analysis tools, and as a result, have the least flexibility of the three groups.

This group model of analyst types applies well to this research project, as the information provided by Park Air Systems suggests that their workforce is divided mostly into the latter two groups of people with very specific domain knowledge (Scripters) or people without the flexibility to operate their dedicated tools (Application Users). This project, therefore, presents an opportunity to address some of these issues as they apply to the analysis and visualisation of network information. Similarly, the paper itself concludes that the use of tools which improve the quality and speed of data analysis can empower and enable these employees to work and collaborate more effectively [21].

Another challenge with data analytics, and in particular, when working with network data, is the quality of the input data [37]. With network systems, many potential faults can lead to less than ideal data quality, issues such as; sudden traffic spikes, outliers or noise can lead to problems when attempting to process the data.

2.4 Data Visualisation

Data visualisation is the field of study focusing on the visual representation of complex data. Human beings can perceive patterns and relationships in data more effectively when represented visually, [27], [38] and thus the use of data visualisations enables for a clearer perception of a large set of data by rendering it in a more straightforward form.

The purpose of data visualisation is as a tool to allow the user to gain a greater understanding of the data they are working with by representing it in a way that takes advantage of our innate ability to perceive patterns rather than through data in its raw form.

The field of data visualisation has existed in one form or another for a long time, though it is the advances in modern digital technology which have necessitated the most growth and change in the field. As datasets continue to grow ever more substantial, the need for visualisation techniques capable of making sense of such large volumes of data has also grown. This has led to new ways of both visualising and interacting with data.

It may initially appear that data visualisation and visual analytics are the same things, however, that is not the case. Visual analytics is more than just visualisation, and can be seen as *an integral approach combining visualisation, human factors and data analysis* [30]. The scope of visual analytics extends across multiple different disciplines, such as statistical analysis, cognitive science, data visualisation and interaction [30], [36].

2.4.1 Static vs Interactive Visualisation

Static visualisations are any visualisation that does not make use of dynamic or interactive elements and therefore remain static, an example of such a visualisation could be any standard visual element such as a bar chart. The use of interactive mechanisms with visualisations is common for many use cases, particularly with large volumes of data. However, by pre-rendering the content and not changing it, an alternative use for static visualisations can be to represent a story through the data. Though it is also possible to do this with interactivity as well, it can be easier to do so without, additionally, one common way of doing this is by using an infographic, a form of static visualisation often dedicated to telling a story through the data [38].

The use of data visualisation is an essential part of the process for understanding and analysing data, however, it is often no longer sufficient to merely visualise a set of data. Another vital part of the process of visualising data is to provide an element of interactivity with the visualisation. This use of interaction matches the description of visualisation as given by Salvador and Granville [17], who describe the technique as being composed of visual representation, and also commonly including an associated interaction mechanism. They go on to define the visual representation as "a form of mapping attributes of an abstract data structure onto visual attributes, usually at a higher level of abstraction than the raw data". The point of this interaction is to allow users to manipulate the visual representation to explore the dataset more efficiently. The typical way this is done is by selecting a new subset of data to display, based upon information learned from the previous subset [27].

Given the nature of the data collected by network monitoring systems, data transmitted between different devices, using different protocols and techniques at different times it makes a single way of visualising the data difficult, as much of the value of the data can be found by cross-referencing it with other data or analysing the relationships between different data points. This type of analysis and cross-comparing of data lends itself to interactive visualisation very well, as it tends to require manipulation of the data to be able to make these comparisons.

2.4.2 Data Visualisation Challenges

Scalability

The problem of scalability crops up often with regards to data analysis and visualisation [30], [37], also referred to as the *information overload problem*, it is simply a result of our capability to collect and store data being greater than our ability to analyse it. As datasets continue to grow larger, they also increase the need for a good visual representation to be able to make sense of them. Additionally, the amount of data that must be processed and rendered can become overwhelmingly large, which can lead to longer response times for user queries, making the very techniques designed to improve the way we visualise and interact with data, more difficult to create and use [27].

There has been extensive effort to solve this problem in recent years, and as such, there are several innovative solutions to the problem. Some have suggested that increasingly high-resolution displays will be able to display larger datasets [39], however, these suggestions are flawed for several reasons. First, and the most obvious is that this 'solution' is not scalable with the growth of datasets, and it still relies on the ability to load the entire dataset at once, which is not practical given the amount of data that is common for *big data* applications. Finally, this can lead to the issue of visual complexity, which can arise from visualisation systems that present too much information or functionality to the user and become overwhelming to them [40].

Alternative attempts to solve this problem have focused less on displaying the entirety of the data, and rather on abstracting the data into a format that is easier to understand, for example through the use of colours to more clearly differentiate between elements of data [41], [42]. The apparent solution to this problem, and what has become one of the most commonly used techniques for dealing with large sets of data is use interactivity to allow the user to filter, compare or otherwise modify the way that the data is presented to them [40], [43], [44].

Multiple Comparisons Problem

There is a problem found when analysing data that can lead to patterns or relationships being found in the data that in reality, do not exist. This Multiple Comparisons Problem (MCP), as it is known in statistics, is caused by the assumptions about the data we may already have or the results we are interested in seeing, begin to be compared with the actual data we have to analyse. As more data is examined and more comparisons made, the likelihood of spotting something 'interesting looking' will increase, thus leading to incorrect insights being drawn from the data [45].

User Agency & Acceptance

The user needs to be able to interact with the data being presented to them to be able to manipulate or filter it to suit their needs. To support user agency, visualisations should be interactive and responsive [44]. In their paper [46] demonstrated that using elements such as colour, grouping or motion can help with user perception and thereby improve their performance. [47] investigated the effects of latency on user interaction and their perception of a system; their results suggest that any latency above 500ms response time has a significant impact on user performance and efforts should be taken to minimise it.

User acceptance is a challenge to the design of visualisation systems, as it is possible to develop a new system that has excellent functionality and meets all of the requirements, however, if the design is too different from the norm, it can be hampered by its users refusing to adapt their working routines to it [30], [48].

2.4.3 Visualisation Techniques

The process of data visualisation can, at a high level of abstraction, be described simply as *bringing data into memory and then applying a visualisation algorithm to it* [27]. Much like techniques for analysing data, the nature of those visualisation algorithms can vary greatly depending upon the type of data or visualisation. The type of data can influence the visualisation as there are well-documented examples of visualisations best suited to different kinds of data [49]. This section will discuss some common areas of visualisation, and some techniques that are used in those areas that may be of relevance to this project.

Data Overview

Providing an overview of the dataset is one of the first and most fundamental techniques for visualising data, Shneiderman's visual information seeking mantra [49] begins with an overview: *"Overview first, zoom and filter, then details on demand"*. As such, the overview allows for an essential starting point for more complex processes of visual exploration [39]. The overview can be used to present a large amount of data at once by utilising high levels of abstraction to present fewer individual elements, thereby displaying more data overall while simultaneously removing visual clutter [50].

Some examples of a data overview in visualisation software include the dashboard interface, such as the Student Activity Monitor (SAM) by Duval [51], the network visualisation tool Gephi [52], NetMiner [53] and dashboard visualisation interfaces by [54]. Also, a survey of papers by [35] cites 154 examples of research papers exploring the use of overview for information visualisation.

The use of an overview to provide a clear starting point and removal of visual clutter are both directly useful to the purpose of this research - providing a clear and usable visualisation system.

Data Comparison

Data comparison is an essential task in interactive visualisation and makes up a fundamental part of exploratory analysis. Many papers have explored different designs for providing comparative visualisations, such as comparing data between groups in scatterplots [55] or designing visual elements to improve user performance with designs that the human brain can more easily understand [46]. This latter research is interesting as it tests different visual elements, such as colour and motion, and how they can assist the user to process and recall the data being displayed more effectively, thereby enabling more effective comparative exploration. Although visualisation can help with comparison, it cannot directly address it, for example by relying upon the user's memory, the best way to support data comparison is through tools specifically designed to do so [56].

In a survey of over 100 different tools for comparative data visualisation, Gleicher *et al.* [57] grouped visual designs for tools designed explicitly for comparison into three different categories:

1. *Juxtaposition*, wherein the objects to be compared are placed separately in space or time, relying upon the viewer's memory to help compare the objects. A traditional example of this type would be two graphs placed side by side to one another.

2. *Superposition*, which differs form juxtaposition by overlays the data from multiple objects in the same visualisation, for example, a line graph with multiple lines.

3. *Explicit Encoding* is different from the other two categories, as it calculates the relationship between the objects and visualises that rather than the objects themselves. An example of this is the UNIX diff tool, which shows the differences between two files but not the contents of the files.

As an example of *juxtaposed* comparative visualisations, [24] provides a summary of research into Coordinated and Multiple Views (CMV) for data visualisation and comparison. This paper raises some concerns around the lack of tools to more easily develop CMV systems and the relatively few systems that engage in meaningful usability testing of their systems.

Temporal data is characterised as one of seven basic data types in Shneiderman's Task by Data Type Taxonomy [49], though this was initially presented in 1996 it is still applicable today, as representing data with respect to time is a prevalent task in visualisation [58]. Further considerations and techniques for visualising data involving time include; displaying points in time versus time intervals, or whether the time is linear (it has a start and stop point) or cyclic (such as seasons) [58].

2.4.4 Visualising Network Data

There are numerous systems available for visualising different aspects of network data. A simple online search will yield dozens of commercial solutions for monitoring home or business networks. These systems, however, tend to be very generic systems designed for home or small business use, which is not relevant to this situation. At the other end of the spectrum are large enterprise-grade systems which often incorporate many advanced features such as integration with a Security Incident and Event Management (SIEM) or Intrusion Detection System (IDS) to provide security and compliance benefits to an organisation, and many of these enterprise systems are moving into the cloud [59] to provide greater coverage for the organisation's assets. These are all very suitable to their environment, however for the use case for monitoring and visualising these radio systems; there are not many available systems that are sufficiently capable and easy to use that will meet the needs of Park Air Systems.

Due to the different properties and dynamic nature of network data, it can be challenging to visualise the information effectively, as different visualisation elements may be necessary for each type of data [60]. Displaying the bandwidth utilisation patterns requires different techniques than displaying packet data, for example.

In a paper reviewing network visualisation techniques, Witall *et al.* [61] suggest that current network performance visualisation techniques can be grouped into three classes:

1. *Geographic Visualisations*, which present the data concerning the physical location of the nodes.

2. *Abstract Topological Visualisations*, which present the relationship between data as the point of focus.

3. *Plot-Based Visualisations*, which usually focus on a single point in the network often represented with respect to time.

Using the requirements list as provided by Park Air Systems, the types of network information that may need to be visualised as part of this project are *Bandwidth Utilisation*, *Protocol Usage Network Topology*. Using these elements as a guide to the types of information that should be visualised, this section will briefly discuss techniques on how they might be visualised and any similar examples of such work in the past.

Network bandwidth information is about showing how much data is transmitted across the network in any given interval of time. According to Witall *et al.*, a visualisation of this would be considered a *Plot-Based Visualisation*. [62] presented this information in multiple ways; the first was as part of a network graph showing connections between host devices; this graph used a colour scale to represent the amount of bandwidth being used on the connection. This is a visually clear representation of resource usage between specific devices; however, to view the information concerning time, it is displayed as a more traditional histogram displaying packets sent over time. These types of visualisation are common for displaying resource usage, and as such, can also equally apply to protocol usage as well.

Representing protocol usage on the network, regarding how often each protocol is used, and between which devices, provides the user with an operational network overview. This information can be presented in a variety of ways; for example, by comparing the usage of different protocols by device, by the amount of data transmitted or usage over time.

Network Topologies, as a representation of the network itself concerning relationships between devices and data flow, can be best shown as network graphs which are capable of displaying the interconnections between each of the nodes.

Additionally, this is particularly well suited to interactive visualisation as, with all but the simplest of networks, it can be challenging to understand the data without any manipulation or control over how it is presented. The use of an interactive network graph can allow the user to change the point of focus, move between nodes to gain an appreciation for the traversal routes between nodes and many other interactions depending upon the level of engagement that the element offers.

By providing information about the devices present on the network, the user can verify the physical configuration. Maintaining a list of what devices are on the network and their average performance levels can help with identifying performance issues and the inclusion of unexpected devices.

2.5 System Usability

Usability is defined in the Oxford English dictionary as "*The degree to which something is able or fit to be used*" [63]. This is a good overall description of the concept; however it does not provide any mechanism through which we can measure or quantify the usability of any given system. Conveniently, the concept of usability has an international standard definition, part 11 of ISO 9241 (1998) [64] defined usability as [1]:

> *The extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*

This definition is much more applicable than the dictionary definition, as it provides specific goals that can be more effectively quantified, or otherwise measured. Interestingly, these goals are remarkably similar to another oft-cited description of usability by Jakob Nielsen. In his book *Usability Engineering* [66], good usability is described by providing a list of essential 'dimensions' (attributes) which contribute towards the usability of a product or service. These dimensions are: *Learnability, Efficiency, Memorability, Error tolerance and prevention* and *satisfaction*.

Furthermore, the ISO defines usability as the *"extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"* [65]. These two separate definitions of usability both share similar attributes, both expressly stating that a system should be efficient and satisfying for its user. The third attribute in the ISO definition: effectiveness, is also not too different from the Nielsen attributes of learnability and error tolerance, as it would be difficult to use a system effectively if it was challenging to learn and did not handle mistakes well.

Therefore, given the similarity of both definitions, as well as the formalised nature of the ISO standard, the definition of usability for discussion in this thesis is the one presented in ISO 9241-11:2018 ([65]).

2.5.1 Measuring Usability

As usability is inherently subjective, it can be challenging to measure objectively. To solve this problem, there are many different surveys available for different use cases to help researchers measure how a user perceives the usability of a system.

System Usability Scale

One potential survey for measuring system usability is the System Usability Scale(SUS). It was initially presented in 1996 as a "quick and dirty usability scale" [67], the SUS was intended to provide an easy and low-cost solution for measuring perceived usability. However, in the years since its initial publishing, it has become possibly the most widely used survey for measuring perceived usability in a system [68], [69]. The survey has been cited in over 1200 publications, according to a reflection written by the original author [70]. The reflection also states that the survey has begun to

[1] This standard has since been updated in 2018 [65] to expand upon some earlier concepts and further relate to a more modern understanding of usability.

be referred to as an 'industry standard', a description which is backed up by a study of unpublished industrial usability studies, which found the SUS accounted for 43% of post-study questionnaire usage [3].

The structure of the survey is comprised of ten statements that the participant must grade on a five-point Likert scale based upon their level of agreement with the statement from their experience using the system. The statements alternate between positive and negative tones, to counteract a loss of attention when completing the survey and force each user to read the statements carefully [67].

The SUS is popular due to its ease of application and grading, as well as the reliability of its results, which also be easily interpreted and compared [69], [71], [72]. The survey calculates the final score by manipulating the results of each question to produce a value from 0-4, then multiplying the sum of these values by 2.5 to produce a single value between 0 and 100 that represents the usability of the system. One of the benefits of the SUS is that this single result is straightforward to understand for a wide range of people [71]. However, it is also important to remember that the final score produced by the SUS does not itself have any meaning and that the value of the result is in comparison to results from other systems.[68].

One criticism of the scoring method for the SUS is that it can be difficult to objectively tell how good any particular system is from its score. The values are easy to compare to one another, but without knowing what a good score is, it can be difficult to interpret the results. To address this issue, there have been attempts to determine what any given SUS score means. These attempts have produced different rating scales which are more familiar measures of performance to describe the survey scores. Bangor *et al.* [73] propose an adjective rating scale which encompasses the range of SUS scores and allows for a clearer understanding of the result. The paper also suggests a second and third rating scale using letter grades to denote performance, and ranges of acceptable/not acceptable values (see figure 2.7).

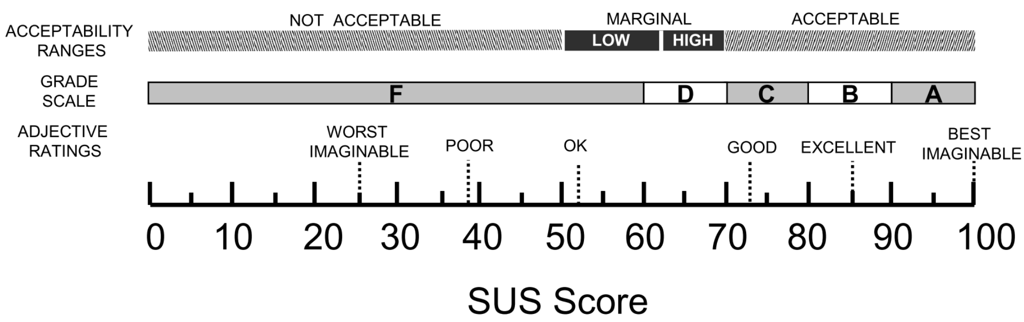

FIGURE 2.7: SUS Score Rating Scales [73]

Usability Metric for User Experience

A newer alternative measurement survey for usability is the Usability Metric for User Experience (UMUX) [74], which is designed to produce similar results as the SUS and is based on the ISO 9241-11 [64] definition of usability. The survey makes use of fewer questions than the SUS, only four instead of ten, each of which is in "closer conformity with the ISO 9241 (1998) definition of usability" [75] and graded on a seven-point Likert scale. The purpose of the UMUX is "to provide an alternate metric for perceived usability for situations in which it was critical to reduce the number of items while still getting a reliable and valid measurement of perceived usability" [72]. The final score is calculated using a similar technique as the to produce a score that falls between 0 and 100. The scores returned by the survey have been shown to have a strong correlation with scores from the SUS [74], [76].

A simpler alternative to the UMUX survey is the UMUX-LITE, an amended version of the original survey to use only two questions to be as efficient as possible [76]. Furthermore, in their paper Lewis *et al*, state the purpose of the survey is to provide a "promising alternative to the SUS when it is not desirable to use a 10-item instrument". This purpose is similar to that of the UMUX itself; however, at only two questions, it does provide a strong case for favouring its use when a very brief survey is essential to the study.

Software Usability Measurement Inventory (SUMI)

The Software Usability Measurement Inventory (SUMI) [77] is a much longer usability survey than the those which have already been discussed. The survey consists of 50 questions, designed to assess the concept of usability as defined by ISO 9241. This aspect of the survey is very beneficial as it applies to the same definition of usability that has been adopted for use in this thesis. However, unlike the other measurement surveys discussed so far, SUMI is not freely available for use and instead requires purchasing in one of two different versions: basic education or full professional.

Table 2.2 provides an overview of several usability surveys.

Survey	No. of Questions	Availability
SUS	10	Non-Proprietary
UMUX	4	Non-Proprietary
UMUX-LITE	2	Non-Proprietary
SUMI	50	Proprietary

TABLE 2.2: Usability Study Comparison Table

2.5.2 Effects of Experience on Perceived Usability

A vital element of the research in this project is to develop a system not only capable of providing a clear representation of the radio network but also to do so in a way that is accessible to users who are not necessarily networking experts. To this end, it is crucial to understand how users perceive systems depending upon their relative expertise or experience.

In a study of 262 participants using two different systems, McLellan *et al.* found that users with a higher degree of experience using the product tended to rate the usability of the system higher, by as much as 15% with the SUS [78]. Furthermore, in 2013, a study by Kortum and Johnson [79] suggested a similar correlation between the level of experience and perceived usability. Bjurling *et al.* [80] wished to investigate which of three groups of employees from a company were best suited to conduct software usability testing, with the groups differing in their level of domain knowledge. The study concludes that participants with the lowest levels of domain knowledge were the ones who found the most severe issues with the system, indicating that domain experts may be overlooking issues which most cause novice users to struggle.

The importance of experience influencing how a person perceives the usability of a system is important, as these studies suggest that greater experience leads to a greater sense of usability, however, there is also evidence that this greater sense of usability may be predicated on their experience allowing them to overlook some potential issues with the system. To address this, [80] suggest that usability testing of a system be done with both novice and expert users throughout the process to have a higher chance of finding issues, and gaining a more extensive range of feedback.

2.5.3 Usability of Visualisation Systems

As has briefly been mentioned in section 2.4, the usability of visualisation systems is not always taken very seriously. [24], [25] both conducted reviews of visualisation systems and both conclude that very few of the systems in question performed meaningful usability studies on their systems.

The need for usability in systems is apparent; the very act of using a product is dependent upon how usable the system is to a given user. This project focuses on the need to develop a novel implementation of a network monitoring and visualisation tool designed for usability by non-experts.

[57] describes visual complexity in information visualisation as coming from too many objects, objects with too many subparts or heavily abstracted data. These decrease the ability of the user to understand the visualisations, and therefore affect the usability of the system. [50] makes a similar point about visual clutter diminishing the potential usefulness of the visualisation, placing a great deal of importance on clutter reduction techniques such as sampling, filtering and clustering data.

2.6 Summary of Literature

The research conducted in this literature review has covered critical areas of relevance, and each of those areas has provided some important information that applies to this project.

Park Air Systems have a desire for a user-friendly system capable of providing monitoring and visualisation functionality for radio network data. The system will be used to assist company engineers in their responsibilities to test their radio systems as part of the installation and configuration stages, to ensure that they are functioning correctly, and potentially to demonstrate to their clients that they are operating as they should.

The radio networks operate using a custom operating system built from the Linux kernel, ensuring that the network operations will be the same as any other equivalent network. Networking itself is a vast field that could not possibly be summarised here; however, some critical information that is important when attempting to understand how to monitor and visualise a network has been provided. There are many systems which are capable of providing network monitoring and visualisation, from both commercial and open-source providers. Existing systems have been surveyed and summarised by [25], [37], [59]. In addition, there have been many papers which discuss techniques for monitoring or visualising network data [26], [35], [62], [81].

Data analysis is a field that focuses on taking raw data and returning knowledge. The data is cleaned, transformed and structured in such a way that statistical techniques, machine learning models or other algorithms [20] can be applied to the data and produce an output that enables greater understanding from the user. Other essential techniques, such as performance improvements to prevent latency [26] can also improve the users' experience and ability to understand the data.

Data visualisation as a technique has existed for a long time; however, it has become vital as our modern world continues to generate more and more data. To handle the increasingly vast amounts of data generated, techniques such as interactivity [43], [44], [82], allows users to explore and manipulate the dataset far more effectively than they would otherwise be capable of doing. Despite the increase in interactive visualisations designed to help users become more efficient analysts, there are concerns that not enough effort is being placed on ensuring that the systems are usable [24], [83].

Usability is an area of Human Computer Interaction (HCI) that relates to how well a system or product enables its use. The current standard definition for usability comes from the ISO 9241-11:2018 [65]. There are many surveys available to measure how users of a system perceive its usability [67], [74]. There has been a great deal of research done using these surveys that indicates they have a high degree of reliability. Furthermore, comparisons between the results of these surveys generally share a similar result for testing the same systems, indicating that they are similarly effective. The SUS is considered the *de facto* industry standard for measuring software usability, due to its system agnosticism, easy application and free availability [79].

Chapter 3

Methodology

3.1 Research Methods

This chapter details the research methods and techniques employed throughout this project, focusing on how the system was influenced by information gathered during the background investigation, and how the aims and research question were tackled. The specific details of the technical implementation and development of the system used for the study can be found in chapter 4.

The primary study in this research project was designed to investigate how users of this particular piece of software interact with and control the mechanisms for presenting data from the radio network systems. Given that the system has been created with the expectation of use by technically minded engineers, the study uses only participants with a technical background, and when possible, direct staff from Park Air Systems.

The research methods employed throughout the project have gathered both quantitative and qualitative data, making use of a mixed methods approach for collection. Although the primary study has been the predominant focus of data collection for this project, there have been smaller meetings with Park Air Systems including some unstructured interviews in order to gather more information or inform design choices during the design and development phases. Furthermore, during the implementation stage, performance metrics and testing were conducted to validate the system and determine if the approach being taken was appropriate from a technical standpoint.

3.1.1 Unstructured Interviews

Throughout the project, several meetings with staff at Park Air Systems took place to have short unstructured interviews in order to investigate the problem more thoroughly and to inform the continuing work at that stage. These interviews served primarily as fact-finding investigations and consisted of pre-written questions relating to critical topics and making notes of the answers. These notes have been summarised below.

1. *Engineer Training & Experience* - The level varies, but they are predominantly software developers or electrical engineers. There are not many specifically trained network engineers, though many staff are currently undergoing Cisco Certified Networking Associate (CCNA) networking training

2. *Standard Network Size / Configurations* - Each network is completely custom designed for the client and their specific needs, and as such, there is not any standard network configuration

3. *Current Practice for Confidence Checks* - Radio networks are built to client specification and tested at full scale against requirements, using a variety of tools.

3.2 Usability Study

The purpose of this primary study is to determine if users prefer one interaction scheme more than another for this use case.

In order to evaluate the usability and design implementation of the visualisation software prototype developed during the project, a study to measure the usability of the software from the perspectives of both non-expert users and also the direct end users of such a system.

The study was designed to compare two different mechanisms for interacting with and controlling the visualisation software. The two versions both presented the same dataset to the user through the same visualisations. The variable being tested here is the mechanism for how the visualisations appear to the user. One version of the system utilises a structured dashboard type interface with elements grouped by category, the second version is initially empty and presents the user with a control bar which can be used to add and remove visualisation elements from the screen.

3.2.1 Participants

Sample Size

The number of participants for this study was twenty-four, with twelve (half) of the participants being direct technical staff from Park Air Systems, and the other twelve being comprised of current postgraduate students of Science, Technology, Engineering and Mathematics (STEM) subjects at The University of Lincoln.

The small sample size was predominantly due to the desire to have a group of participants who were as directly related to the project as possible, thus ensuring they are the most useful subjects for the study. It is for this reason that the twelve participants were chosen from the Park Air Systems technical staff, comprising of network engineers, system developers and test engineers, all of whom are aware of the radio systems and networks as they related to this project, and therefore can understand the purpose of the system in that context. Furthermore, they are the most relevant test subjects as they are the exact members of staff who will be developing, testing and using any version of the system in its final form after handing over to Park Air Systems, and as a result, their feedback and test results are the most useful.

The reason for only twelve participants is a pragmatic one, in that it was not possible to get access to more engineers to complete the study. Since they each have their work responsibilities to complete, it would not be reasonable to ask a busy company to temporarily release a large portion of their technical staff from their jobs for a length of time.

The second set of twelve participants was chosen to have a comparison set of people not from Park Air Systems, such that any results might be compared to determine how the contextual understanding of the user may affect their perception of the system. Given the sample size of twelve engineers from Park Air, the second set of non-company test subjects also contained twelve people to have an evenly sized split and allow the results to be as comparable as possible.

Recruitment

The participants from Park Air Systems were recruited by corresponding with the member of staff at the company who served as the point of contact throughout the project. The number of participants was discussed after establishing the amount of staff available who could reasonably take part in the study and also were able to take time off from their work to do so. These members of staff were contacted by the company contact and provided a short explanation of the project and asked if they would be willing to participate. This process was done approximately two weeks before the study would be taking place, giving them time to ask any initial questions or find replacements if needed, though this was not necessary.

The study participants in the second group, the students and recent graduates from the University of Lincoln were recruited in person approximately one week in advance so time could be arranged for the experiment to be carried out. The participants were given a short explanation of what they would be asked to do as part of the study, however as with the participants from the company this explanation did not contain specific details of the study to avoid giving the participant any impression of the study before they took part.

The second group of participants was made of students and recent graduates from the University for practical purposes. The study was conducted in person using a laptop computer preconfigured with the correct software, and as such, it would not be possible to perform the study remotely. Also, it took time to prepare and organise the participants from the first group as a trip to their workplace was required, and to do that for individual participants would be highly infeasible. Finally, given the requirements for taking part in the study (see *Restrictions* section below), the most readily available pool of potential recruits was from the student body.

Consent

At the start of the study each participant was provided with a written information sheet (see appendix A); detailing the purpose and methodology of the study, what is expected of them, what data they will be providing and how it will be securely stored. Each person was given a chance to ask any questions they may have at this stage, as well as leave before the study begins if they no longer wanted to take part in the study.

Once each participant had read the information sheet and had a chance ask any questions, they were asked to complete and sign a form (see appendix A) indicating they were providing their informed consent to participate in the study.

Restrictions

The limiting conditions to be a participant in this study were that each person has some amount of experience in a technical field, as the even though the system is designed for non-expert network users it is still for technical and engineering staff. The participants from Park Air Systems were recruited from the technical staff, which in this case is defined as working either directly as a member of the software or hardware engineering teams.

As the second set of participants was recruited from students, the main limiting factor was their field of study; each student must be from a technical field (i.e. a STEM subject). Requiring that each participant have a technical background is for the same reason that only technical staff from the company were chosen; this piece of software was developed with the expectation that such a system is predominantly for use by users with a technical background.Finally, every participant was required to be over the age of eighteen. This was for two primary reasons; the first was that the system would not likely be used by anyone under the age of eighteen and so they are not beneficial to the research. Secondly, it is essential that they are over eighteen to be able to provide proper informed consent as required to participate in the study

3.2.2 Method

The requirements for the study were to investigate the difference in usability between two different versions of the visualisation software and how the users of the system engage with the control mechanisms. The equipment used for the study was a Dell XPS laptop running Linux with the required dependencies installed to run the code, for a full description of the development environment and software used see chapter 4.

Section 2.5 describes several different surveys that have been developed to measure the usability of a system. One survey considered for this project was the SUMI, a well-recognised survey form for measuring system usability; however, it would be impractical for this study as the length of the survey is far too long and would require each participant take far longer than the time available. The survey requires the purchase of a commercial license for use, which makes it unsuitable for this project.

Finally, the SUS was chosen because it is freely available and provides a 'de facto' industry standard in measuring system usability, and in particular for comparative analyses of multiple systems. The survey results are calculated into a final score which can be used as a value to compare against other systems, which is precisely what is required from this study. The alternative UMUX was rejected as it is primarily a shorter alternative for the SUS, and lacks the significant body of prior use in research to establish its credibility.

The study was conducted in one of two different locations depending on the participant. The first location was at the Park Air Systems facility in the village of Market Deeping; this location was used for the tests involving Park Air Systems employees. The second location was the University of Lincoln, in an isolated room away from interruptions and distractions which might disrupt the experiment.

Each participant was given the same introduction and explanation of the software through the use of printed informational sheets, to ensure consistency throughout the experiment. Moreover, once the study began, the only communication between the researcher and the participant was during the change between versions of the software and completing the related survey form.

The method for conducting the study was to provide each test participant with both versions of the program and give them a series of tasks to complete with each version and record their performance in completing these tasks as well as asking them to complete the SUS after each version.

At the end of the test, each participant was asked to participate in a short semi-structured recorded interview where they were asked to discuss their feelings regarding the system overall and how they felt the different versions compared against each other.

The tasks provided for the participants varied slightly from the two systems to ensure they could not merely remember the previous answers. The tasks were designed to test the participants to use each of the different visualisation elements of the system to ensure they get to experience the whole system.

The dataset used as part of the experiment was a simulated network traffic capture from the MARC radio network simulator system used at Park Air Systems for simulating network builds. The same packet capture file from the radio simulators was used for both versions of the system used in the test to ensure the same data was presented to every participant. This dataset was used for the experiment as it is the closest possible approximation to the real data which would be used for the system which can be captured without having access to a full sized radio network. This was changed from the initially intended live capture functionality due to internal security concerns from Park Air Systems who determined they would be unable to allow live data capture from their equipment. As a result, the network capture parts of the system were removed from the prototype, and any further work was stopped.

Each participant was introduced to the controls and functionality of the system before beginning the test, to ensure they had an understanding of how to operate it. The introduction followed the structure of an explanation sheet which was provided and also available throughout the test to refer back to if necessary (see appendix whatever). The introduction began before the test for each version of the system and took a few minutes to explain the system and how each of the interactive elements could be controlled and manipulated.

3.2.3 Ethical Concerns

As with any research project, care must be taken to minimise potential ethical concerns and mitigate those that cannot be avoided. Given the nature of this research project and the structure of the study, there are not expected to be as many ethical concerns; however, this does not mean that there are not any at all, and steps must be taken to minimise them if at all possible. Table 3.1 details key concerns and steps taken to mitigate them for this study.

This research was undertaken to investigate how users interact with, and what their opinions are of, a piece of software created to visualise network traffic data from radio systems. Though there are not many concerns, there were still things to consider and effort was taken to minimise these potential issues.

Concern	Mitigation
Potentially identifiable data is collected from network traffic capture	Data will be stored anonymously, and no attempts to identify personal details will be made
How to use and store data in a secure and anonymised way	All data stored or used for this project will be done so in accordance with the General Data Protection Regulations (GDPR).
Loss of study data	All data will be anonymised and stored on an encrypted, password protected machine. Any paperwork will be stored in a locked cabinet.

TABLE 3.1: Ethical Concerns

3.2.4 Data Collection

The range of data collected in the study is of both a qualitative and quantitative nature. The qualitative data was taken as observations during the experiment and a short semi-structured interview at the end of the study to get the participant's opinions on the test and any additional feedback they may have. Additionally, the bottom of the SUS form offered space for the participant to optionally add any unstructured notes or additional details as further information.

An important factor with the use of these tasks is that each question has an objectively correct answer, rather than a subjective one. This distinction allows for the task sets to be marked regarding how many of the tasks were completed correctly. In addition to recording whether the task was completed correctly, each participant had their performance times measured per task to analyse how much time they took to complete each task and the complete set. These measurements provide quantitative data which can be analysed after the study is completed to compare the performance of the participants across the different systems.

During the experiments, as well as taking quantitative measurements of performance, observations were noted regarding the individual performance of the participant. These observations focused more on the state of the participant throughout the experiment and any other specific details relating to their performance which were noteworthy at the time. They were taken to provide notes which could be used to provide direction and discussion material for the semi-structured interview following the conclusion of the tests. Moreover, they served to provide contextual information to help with the analysis of the other results.

In addition to the performance measurements taken for each task as part of the experiment, each participant was asked to complete a SUS questionnaire regarding their experience after completing each set of tasks.

The System Usability Scale questionnaire consists of ten agree/disagree statements measured on a five-point Likert scale. The survey focuses on the experience of each participant after having used a system, and the result of the survey is a final score measured on a scale from 0 - 100.

The results of this survey also allow for a comparative analysis of the systems.

Following the SUS and as a final roundup to the experiment, each participant was engaged in a short semi-structured interview with the results recorded and transcribed. Each discussion centred around the general experience of the participant, their thoughts regarding the system and any other comments they may have to accompany the direct observations made during the experiment. This semi-structured interview also discussed the particular version of the system concerning the current workflow being used by the Park Air engineers to perform their confidence checks on the radio systems.

3.2.5 Data Analysis

Quantitative Data

The timing data for how long it took each participant to complete each task was collected to provide additional information alongside the primary data from the SUS study and unstructured interviews. The data were grouped and statistically analysed for features to help describe and contextualise the rest of the data. Information such as the average time to complete each task helps to provide further insight into the participant responses. Moreover, as quantitative data, it is possible to more easily draw conclusions from objective facts rather than the more subjective interpretation of qualitative data.

To make use of the results from the answers to each task the participants were given, they need to be marked for accuracy. This accuracy data was used to provide further insight into the other results from the study, much like the timing data. Knowing whether a particular task or element from the system is easier or harder than expected may be part of the qualitative participant responses, however, if the correctness of their answers to a particular task is different, then it is further evidence, especially when compared to the time taken to complete those tasks.

The SUS is scored on a five-point Likert scale, for every odd-numbered question, one is subtracted from the value on the scale, and for the even numbered questions, the scale value is itself subtracted from five. The total sum of these adjusted values is then multiplied by 2.5 to produce a value from 0 to 100. Though this value is not of any use individually, the purpose for utilising the SUS was for its benefit in comparing systems, which is made easy with this one value.

Qualitative Data

The primary source of qualitative data for this study was through semi-structured interviews with study participants were conducted at the end of each experiment. These interviews, which ranged from approximately 3 - 6 minutes in length, were each transcribed to be further analysed later.

The use of semi-structured interviews was chosen to give a minimum level of structure to the interview while allowing it to remain open to a more informal conversation with the participant rather than a formal question and answer interview with no ability to go into further detail without making one interview different from the others. The purpose of the discussions was to gain a better understanding of the participant's experience of each system they used during the study and how they felt about them, in their own words.

To achieve this, each participant was asked three questions designed to be open enough to allow the participant to share their experience and encourage further back and forth discussion before moving onto the next question.

- **How do you think that went overall?** - This question was to gain an understanding of how the participant felt about the whole experience, such as if there were any particular difficulties, unusual occurrences or other relevant pieces of information regarding either system or the study in general.

- **Do you think you would you find it helpful?** - With the Park Air Systems participants, this question allowed them to discuss their current workflow and how such a system may be useful as part of it. The non-company participants discussed more hypothetical use cases for the system, as they had no direct need in their work for it.

- **How did you find the first version compared to the second one?** - A prompt to get the participants to compare the two versions and see what additional thoughts they have about either one.

The method chosen in this study to analyse the qualitative interview data was thematic analysis. This technique is one of the most widely used approaches to analysing qualitative data and is used for *"identifying, analysing and reporting patterns (themes) within data"* [84].

A thematic analysis was conducted on the transcribed interview data to gain a better understanding of the important themes and ideas as they were provided by each participant. Thematic analysis was chosen because it is a well documented and useful technique for understanding and contextualising qualitative data such as the interview responses from this study. The results of such an analysis also offer the kind of insights into the data that was desired in this case.

3.2.6 Pilot Study

To validate the study methodology, as well as serve as a 'trial' run to ensure that the timings are appropriate and no large design flaws exist, an initial beta test version of the study was performed. This study involved conducting the experiment as it was originally designed, with a small number of participants, and then evaluating the results.

Method

The pilot study was conducted with four participants, all students from the Computer Science department at the University of Lincoln, and followed the same structure as the main study; a set of tasks to complete, followed by a usability survey, for each version of the system. This initial study contained only five tasks to complete, and made use of the SUS survey.

Observations

- During the introduction phase of the experiment, it became apparent that the explanations of the software as given verbally were not consistently timed or descriptive, and tended to vary on whether the participant appeared to be following easily or not.

- Recording the time taken for participants to complete each task raised an issue. Each participant would take different times to read the question and write their answers on the answer sheet.

- The number of tasks being completed by each participant (five) was not as much data as initially hoped, in addition these tasks were being completed faster than expected meaning the study was running shorter than intended.
- Participants often had their own comments or feedback about the experiment or system after completion, and without taking notes on these it would potentially be missed data.

Changes made

After the conclusion of the test study and evaluating the observations, the following changes were made to the study design;

The number of tasks to complete by the participants would be extended to ten, this would increase the amount of performance data collected, take more time for the study to complete thereby giving the participants more experience with the system which would help inform their perception of its usability.

To record the time of each participant, the recording was started when they visibly stopped reading the task sheet, and stopped when they began to write an answer. This is not a perfect solution, however it is consistent for each participant and without a complex system to monitor their focus and behaviour it is difficult to be perfectly precise in this measurement.

To minimise the variance in introduction to the system, a script was written that had an explanation of each different visualisation type on the system. This script was used to provide a clear and standard explanation of each system to ensure that every participant had a consistent baseline of understanding.

Finally, it was noted that some participants had comments on the system after completing the experiment, and so to give each participant a chance to discuss their feelings and experience, a short semi-structured interview would be recorded with each person at the end of the study.

Chapter 4

Development

4.1 Requirements

The requirements for this project, as provided by Park Air Systems, have been broken down into two groups. The first set contains the essential requirements which the project must achieve in order to be considered complete. The second set are additional desired requirements which are offered as potential further extensions for the project, after having completed the essential requirements. The essential requirements which upon completion will yield a functional system for monitoring the radio network are as follows:

Essential

The completed system must be able to be attached to a network, via a passive tap or using a SPAN/Mirror port on a managed switch, in order to monitor and analyse network traffic. The device should also be completely passive and not send any traffic onto the network.

The device should allow for basic network misconfiguration such as conflicting IP addresses, incorrectly configured subnets or broadcast storms to be detected and highlighted to the user in a clear way.

Desirable

Further potential requirements for the program beyond the essential functionality are as follows:

- Create a resolved list of all equipment vendors on the network (via MAC Address)
- Monitor protocol activity per device
- Monitor bandwidth utilisation (globally, per tap, per device)
- Detect unusual activity (Protocols, bandwidth, activity changes)
- Live monitoring of voice streams
- Display traffic within different VLAN

4.2 Design

The primary design requirement of the system is to be clear and without excessive complexity, otherwise the system is not an improvement on existing options. The intended use case for the system is not to replace more complex and mature tools used for monitoring and visualisation of network data, such as *Wireshark* [85], but instead to provide an additional tool for engineers when they need a more accessible way to view the network as a whole rather than specific detail with more complex tools.

The design process involved paper prototyping and sketches of interface designs for the system. These initial designs were influenced (both positively and negatively) by other monitoring and visualisation systems such as *Snort* [86], *Gephi* [52] or *NetMiner* [53]. After experimenting with several different ideas for the structure of the system, it was decided to use a simple dashboard style interface that would be able to provide a clear overview of the network to the user. This interface would have different visualisation elements that present a live view of the data collected by the system, data such as bandwidth and protocol usage. The initial topology discovery would be conducted at the initialisation of the system, to collect the required data regarding devices on the network and then present the resultant topology diagram on the interface.

4.3 Implementation

4.3.1 Network Monitoring Prototype

This section will discuss the design and implementation of the initial prototype monitoring system that was developed and used to experiment with designs, before changes were made that affected the specification of the system and necessitated a rework into the visualisation system that was used for the study.

Traffic Capture

Network traffic capturing is composed of two primary components; packing capturing (also known as packet sniffing) and parsing, to extract the contents and make use of the data. When packets are captured they are then added to a database in the the system to be retrieved later for visualisation.

Initially, the program made use of the Python *socket* library to enable the program to sniff for packets on the given network interface. This was done using RAW_SOCKETS and was the first prototype implementation of this functionality, and it required more coding than using a library would. This manual development was done because it was initially thought that greater control would allow for easier integration with other features of the system, however it was soon realised this was not the case and changed in favour of using an open source library instead.

The second implementation made use of the *pcapy* library which allows for collecting an assigned number of packets on any given interface. The library cut down the amount of code written manually while simultaneously increasing the reliability of the code given its open source community development.

Parsing

The original implementation of packet capturing also included manual parsing of packet data, extracting information from the frame header such as source and destination addresses as well as the protocol number. This was also changed to use the third party library *impacket*, for the same reasons given above.

Topology Discovery

As discussed in section 2.2.3, there are many topology discovery algorithms that can be used on networks. For the purposes of this project however, a simple implementation is sufficient because the purpose is to present the data to the user, and not develop a novel discovery algorithm.

Using the *pysnmp* library, an SNMP walk gathers information about the devices on the network, the information on the SNMP agents is then collected and used to help build a database of devices on the network, which can then be visualised using a network graph library such as *networkx*.

Data Visualisation

Visualising the data collected from the network was initially done using the *matplotlib* library, to take in the captured network data and generate plots to display it to the user. Although using it to generate initial graphs did provide useful feedback on how the information could best be displayed, the library is typically used for visualisations of scientific data and following those initial tests it was found to be a unsuitable choice for this use case.

Plotly was chosen to provide the new functionality as it is a popular library for generating interactive graphs and visualisations for various types of data. In addition, *Dash* is a framework provided by *Plotly* to quickly and easily develop an analytics dashboard for interactive visualisations. The use of *Plotly* and *Dash* allowed for easier implementation of interactive visualisations, which are important for this project as they allow the user to explore the dataset more easily. Figure 4.1 shows an example of a graph made using *Plotly*, with a interaction taskbar in the top right corner.

The process of visualising live data is handled by *Dash* and *Plotly*, the system is configured to automatically update and query the database at a predefined interval, typically every ten seconds. This allows the new data to be added to the database as it is captured, to be rendered when the system updates with new data.

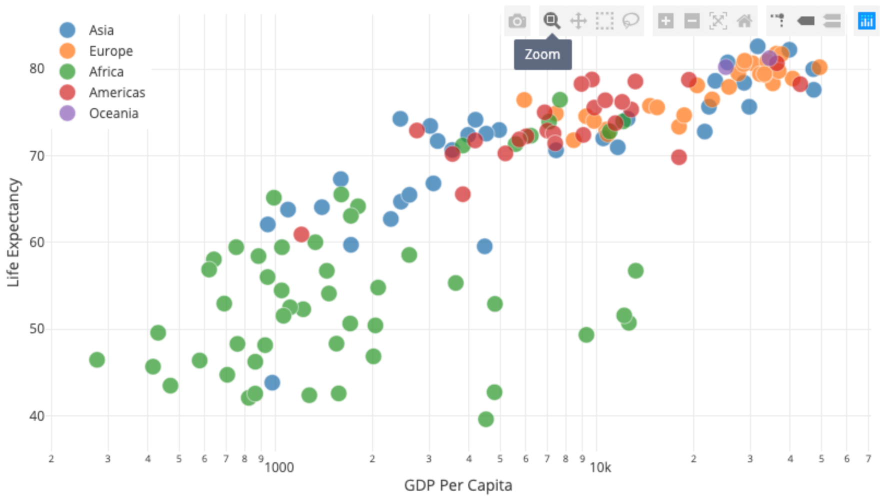

FIGURE 4.1: Plotly visualisation example

An additional benefit of using *Plotly* is that it has a large amount of native interactivity which can be configured for the visualisations. This was something that was very important to the design of the system, as it provides the end user with more control over the data, enabling a more usable system.

Evaluation

Following the first demonstration of the system at a meeting with Park Air Systems, using an early prototype, it was determined that going forwards the program would no longer be required to capture any data from the radio networks itself. This was because there are viable alternative systems for capturing data already, the company would prefer that the system only operate using data which had already been captured and would simply provide the analysis and visualisation of that data. Additionally, it would not be possible to provide direct access to the radio systems and so a dataset of simulated network traffic would be provided to use for any testing or academic studies involving the system.

In response to these new changes, the live capture, topology and network monitoring functionality would no longer be developed as they would not have any impact out the outcome of this research and the primary focus of the system pivoted to the visualisation of the data itself.

4.3.2 Network Visualisation System

This section contains the final design and implementation details of the network visualisation system that was created from the prototype monitoring system. This is the system that was used in the user study as described in chapter 3, the results of which can be found in chapter 5.

The program was developed using Python 3.6 [87] due to its easy of development and wide range of code libraries that can easily integrate functionality without writing everything from scratch. Table 4.1 shows the list of libraries used in this project.

Library	Purpose
pcapy	Packet Capture and Parsing
pandas	Data analysis
dash	Building Visualisation Dashboards
dash-core-components	Interface components for dash
dash-html-renderer	Use HTML to create page structure
dash-table-experiments	Create interactive data tables
plotly	Create Data Visualisations

TABLE 4.1: Python libraries used.

Design Changes

Following evaluation of the prototype monitoring system, some changes to the design of the system were made.

The biggest change was in the purpose of the system, originally intended for displaying network data as it was captured to provide a clear visual representation of the network. The new requirements for visualising captured network data introduced new design challenges; the amount of data to visualise in unknown and the amount of visual and other display elements is also unknown. As the system is intended for use by Park Air Systems, either directly or as a prototype to aid with their own development, the way in which the user finds and interacts with the data will be important.

Because of this, it was decided to create two versions of the system with the same visualisations but utilising different control mechanisms to find and interact with the data, and conduct a study to investigate if the different mechanisms have an effect on the usability of the system. The first version utilised the same dashboard overview as the prototype, with visualisations grouped into related sections, thereby serving as a central hub for information that the user may want to see. The second version was intended to contrast with the first, by initially displaying no data, instead providing a searchable menu bar that the user can use to present any of the visualisations, in any combination or order they want.

Network Data

With the packet capture being handled by Wireshark, each packet is parsed at the point of capture and stored as a .pcap file, which is converted into a csv file to improve read times when starting the program down to a few seconds from several minutes. This csv file also contains fewer data attributes than the original file, these additional data points (such as alternative time formats) are removed to improve performance and because they are not necessary for the visualisation system.

Analysis

After the data is loaded into the program it is stored in a dataframe before processing and analysis can begin. The data is first cleaned by checking for corrupted or malformed packets, which are then removed from the dataset. To prepare for visualisation, the data is grouped together into new dataframes by attributes that are necessary for displaying, such as protocols, byte size or timestamp.

To display the manufacturer information about each physical device on the network, the system populates a dataframe with each unique MAC address. These addresses are sent to an online API provided by *macvendors.co* [88], which will return the manufacturing company and address to be added to the dataframe.

Visualisation

The techniques for visualisation focus heavily on providing interactivity with each element on screen, this is due to the research suggesting that interactive visualisation provides the user with greater control and usability, both of which are highly desirable goals for this system. Moreover, as is frequently cited in research papers, Shneiderman's visualisation mantra of *overview first, zoom and filter, then details on demand* [49] can be handled by each interactive visualisation, through the use of *Plotly*, which enables zooming, filtering, and displaying specific data points when selected.

Each visualisation in the system contains its own dataframe, taken from the full dataset and grouped with only the relevant data to enable performance and limit the complexity of any given graph. The visualisations used to present different data types include: histograms, line graphs, time series, pie charts and tables. The selection was chosen to best represent the types of visualisation that would most likely be used on a full version of the system.

Figure 4.2 shows the full length of the interface for version 1 of the system. This combined set of screenshots shows what the version looks like to the user as they scroll through the page. The visualisations shown are grouped together by relevant subject, such as bandwidth utilisation or protocol breakdowns. This was to help provide a clear way for users to understand what information that would find in a given section and allow for easier navigation.

Figures 4.3 and 4.4 show two different configurations of the interface from the second version of the system. The menu to choose each visualisation is clearly visible in both images, and demonstrates the way in which a user would select what to display.

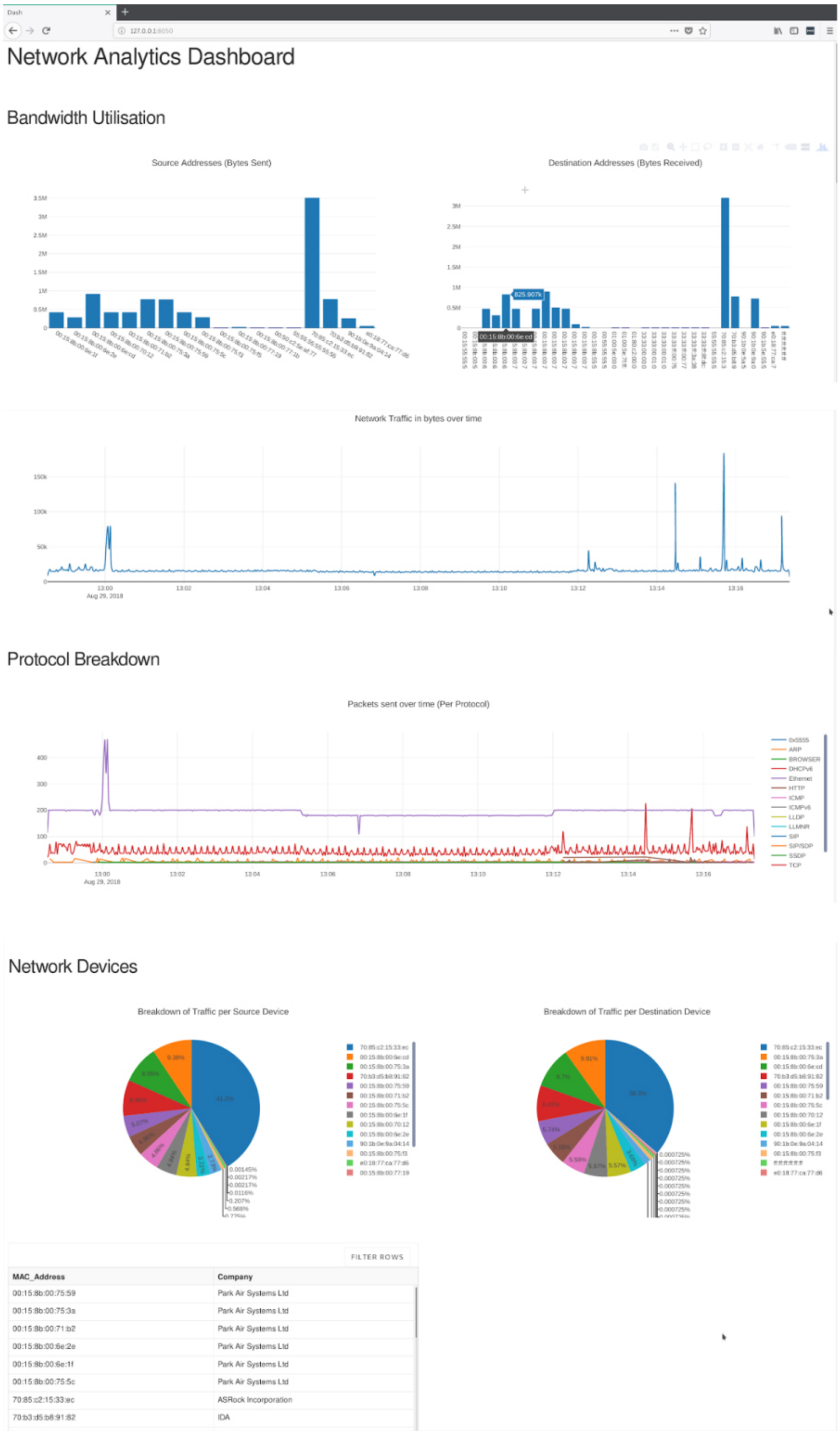

FIGURE 4.2: Dashboard Interface (V1)

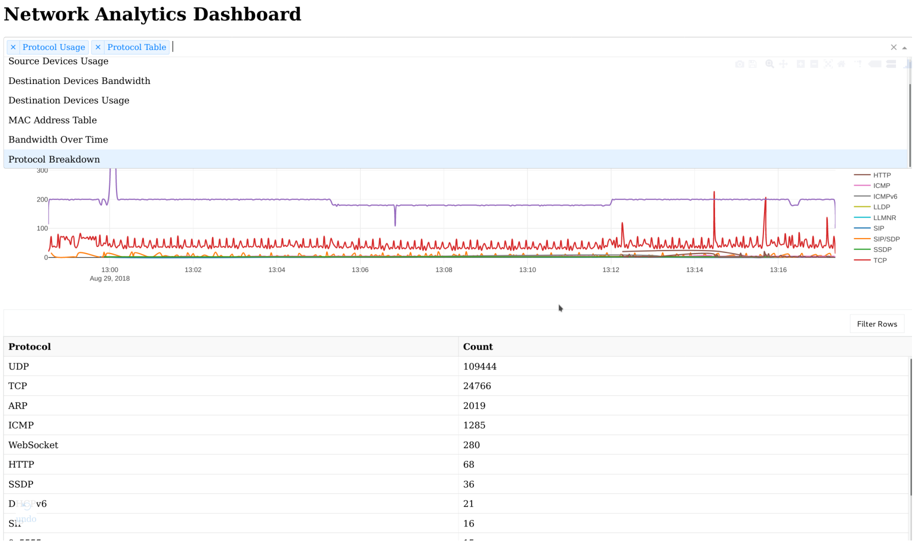

FIGURE 4.3: Dropdown Menu (V2)

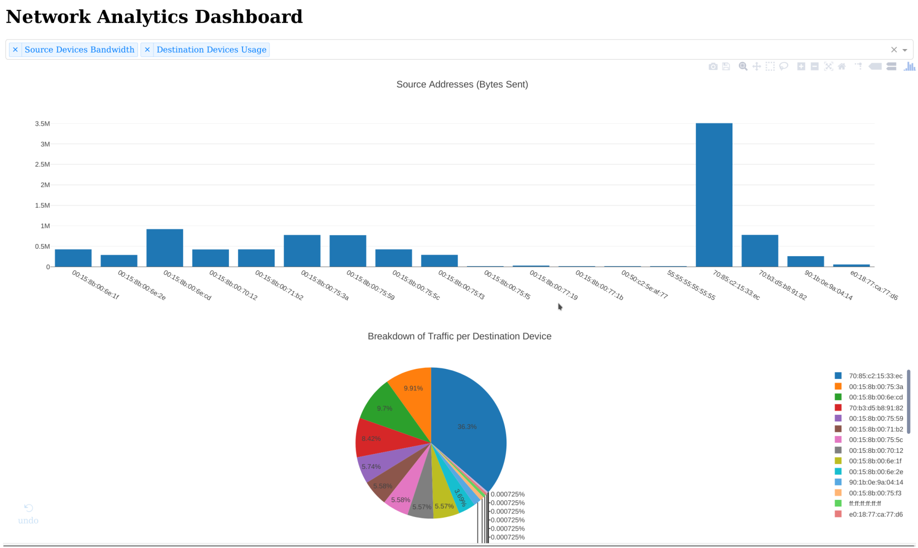

FIGURE 4.4: Alternative Visualisations (V2)

Chapter 5

Results

5.1 Quantitative Results

5.1.1 Task Times

Typically, when presenting task times as results from a usability study, the arithmetic mean can be an unreliable indicator of the centre of the distribution, as times can be positively skewed due to longer times that cannot be matched by equivalently short ones [89]. In an effort to determine the best alternative, [89] compare multiple different statistical techniques for determining an average from time data with differing sample sizes. Their conclusion was for a sample size of under 25 that the geometric mean was the best estimation and had a lower error rate than the sample median, therefore for this study with *N=24* the reported average times are using the geometric mean.

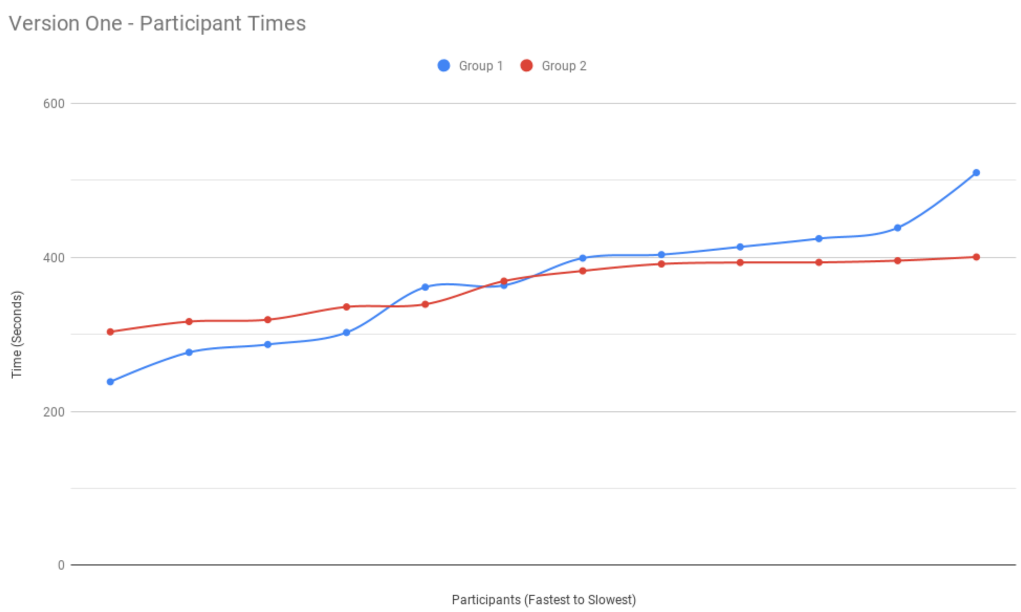

FIGURE 5.1: Total Time - All Tasks (V1)

Figure 5.1 shows the times taken by each participant to complete all of the tasks for the first version of the software during the experiment. The figure clearly shows a much broader range of times from the participants in the first group in comparison to the second group whose times are more consistently similar. Furthermore the figure shows that seven of the twelve pairs of times were from the first group of participants, slightly above 50%.

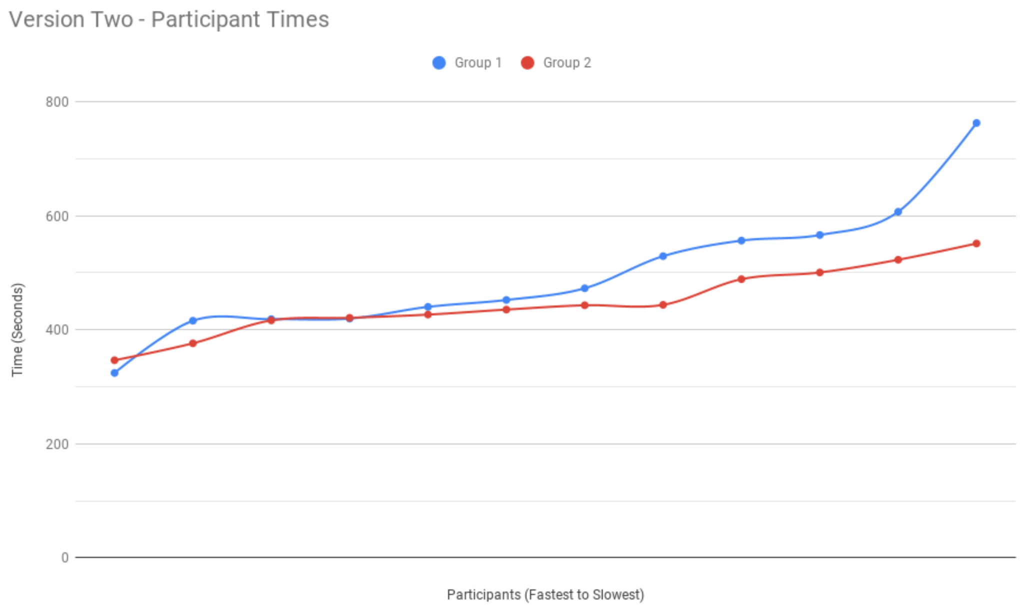

FIGURE 5.2: Total Time - All Tasks (V2)

In comparison, the times taken for the second version of the software, for each participant to complete all tasks are shown in figure 5.2. The graph shows that once again, the participants from group one had a greater range of times than group two, who appear to have again been more consistent in their times. The figure also shows that eight of the twelve participants from group one took longer to complete the tasks than their group two counterparts, compared with seven for the first version.

When we compare these graphs to figure 5.3, we can see more clearly this range of values between groups one and two, and for both versions of the system. The exact performance figures can be found in table 5.1.

Version	Group	Min	Max	Average	SD
1	1	238.41	509.97	360.03	79.01
1	2	316.56	400.41	359.87	36.29
2	1	324.38	762.86	485.73	115.37
2	2	346.64	551.43	444.20	59.17

TABLE 5.1: Participant Times - Key Figures

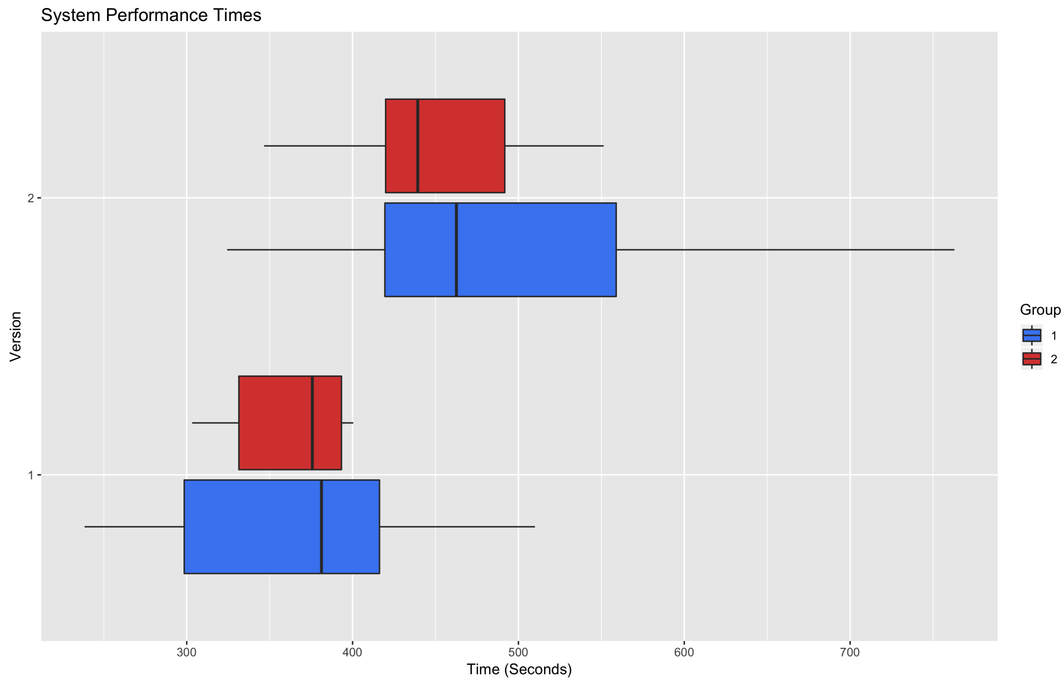

FIGURE 5.3: Participant Times Per Group and Version

These results show that although the fastest overall time came from the first group, so too did the slowest, leading to a much higher range of times than in the second group, and this is true for both versions of the system. The average (geometric mean) of both groups is very similar for the first version of the software, which is interesting as it indicates that the performance times for group two were very consistent to maintain such a similar average, whilst having a smaller range of times. This is confirmed with the standard deviation of the first group being approximately double that of the second.

Table 5.1 shows that again the first group had the fastest and slowest times to complete the tasks, with a range of times that is much greater greater than the second group. In this version of the software however, the average times to complete the set of tasks were much further apart, with the group two participants performing approximately 40 seconds faster. The standard deviations show that there was a much higher spread amongst the results of the first group, which is consistent with the other reported metrics in the table. Furthermore, the standard deviation for the times of group one participants is almost double that of the group two participants.

A Linear Mixed Model test was performed on the performance data to investigate whether there is a relationship between the performance times and the version of software used or group the participant belonged to. This test was chosen over using a t-test as it also looks for an interaction effect between the two factors being tested, which would yield additional insight that may not be present from other tests.

Analysis of the histograms (see appendix B) show that the data is approximately normally distributed, and was determined to be sufficient for this test, given its lack of sensitivity to non-parametric data. Additionally, the time data was log transformed before performing the test, as mentioned in section 5.1, to ensure the average value is calculated using the geometric mean to counter any positive skew of the data. This also ensures that all reported averages for timing data are done so consistently. One final note is that given the participant demographics (see table 5.6), the age of the participants was included as a variable in the calculation to ensure that this difference is accounted for.

Table 5.2 contains the results of this test, showing that there is a statistically significant difference (p = 0.00381) between the performance times on each *Version* of the system. The difference between the times of the two participant groups however is not significant (p = 0.70828), and neither is the difference between the ages of the participants which was another (albeit difficult to control for) difference between the groups. Finally, the interaction effect between the two variables has a p value of 0.25443 which is also well above the threshold for statistical significance, meaning we cannot infer any interaction between the two variables.

	Num DF	Den DF	F Value	p
Age	1	21.000	1.0779	0.31097
Version	1	22.000	10.4615	0.00381
Group	1	36.950	0.1422	0.70828
Version:Group	1	22.000	1.3695	0.25443

TABLE 5.2: Linear Mixed Model Results for Participant Times

5.1.2 Task Accuracy

To complete each task, the participant was required to write the answer on their task sheet, which was then used to mark the results of the participants. This provided additional information which can be compared against the time it took to complete the tasks. Table 5.3 shows the mean accuracy per group on each version, and the standard deviation across the participants within each group.

Version	Group	Mean	SD
1	1	85.83%	13.64%
1	2	86.67%	12.55%
2	1	86.67%	10.54%
2	2	87.50%	11.28%

TABLE 5.3: Accuracy for All Tasks

5.1.3 Usability Survey Scores

SUS scores are graded on a scale from 0 - 100 to present a single value that represents how usable the system is. The results of these surveys, presented in figure 5.4 shows that the study participants consistently preferred the first version of the system to the second.

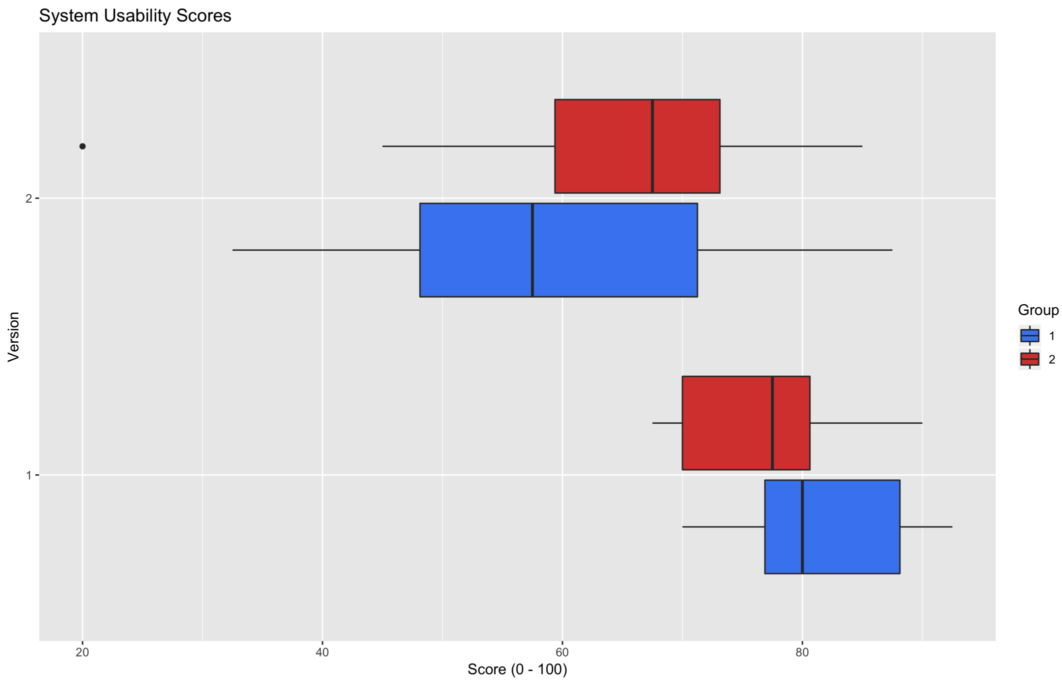

FIGURE 5.4: Usability Scores Per Group and Version

Figure 5.4 also shows that there was a high degree of consistency to results from both groups regarding the first version of the software. This is further confirmed by the values in table 5.4 which show similar standard deviations of 8.01 and 6.85 respectively. That is is sharp contrast to the usability scores shown, which indicate that there was a much broader range of results as well as much higher standard deviations, indicating less of the data is spread around the mean.

Finally, figure 5.4 shows the distribution of results in the dataset, which again clearly shows that the first version of the system was reported to be more usable by participants from both groups. The average scores from both groups for the first version are higher than either score for the second version by at least ten points, and both plots for version one are much more compact and have shorter whiskers than their version two counterparts. For the full details, including standard deviation, see table 5.4.

SUS Mean Score Results

The results of using another Linear Mixed Model to test for statistical significance in the SUS results, as well as an interaction effect between the two factors are shown

Version	Group	Min	Max	Average	SD
1	1	70	92.5	81.25	8.01
1	2	67.5	90	76.67	6.85
2	1	32.5	87.5	58.96	16.6
2	2	20	85	63.13	17.52

TABLE 5.4: Average SUS Scores Per Group and Version

in table 5.5. The significance score for the system version was p = 0.01458, which is highly significant. However, a score of p = 0.27078 for group is not at all significant, and neither is the interaction between the software version and the participant group (p = 0.24992). Like before, the age of the participants also did not have any effect on the results.

	Num DF	Den DF	F Value	p
Age	1	21.000	0.0049	0.94492
Version	1	22.000	7.0307	0.01458
Group	1	29.292	1.2600	0.27078
Version:Group	1	22.000	1.3966	0.24992

TABLE 5.5: Linear Mixed Model Results for Participant SUS Scores

5.2 Qualitative Results

5.2.1 Thematic Analysis

A thematic analysis was conducted on the interview transcripts to evaluate the responses and attempt to extract essential information and ideas from the data. This analysis was exploratory in nature as the intent was to review the data for themes, rather than searching specifically for predetermined themes in the data.

The following themes have been taken from the interview transcripts and are helpful for understanding the context and reasoning behind the perception of usability of each system by the participants. The survey data is enriched by the participants' explanations of their experiences, and is discussed below with excerpts [1] for reference.

Control

The actions a participant can take with the system and the degree of agency that they feel they have in using the software are important factors in the perception of

[1]Some quotations may include additional information for context within []

the system, and have a profound impact on whether the user felt able to complete their task. The level of control that users felt they had, or would like to have, was mentioned by most of the participants, primarily as a reason why they may prefer one system over another. This theme also includes any other feedback relating to the way the system is controlled for each different version.

> *So as a new user, the first one is definitely the easier one to use, however I suspect that as I became more used to the systems, then I'd probably appreciate the second system more because I'd have more of a degree of control.*
>
> Group One Participant

This statement, by a participant from the Park Air Systems group, shows an appreciation for what they describe as the greater control afforded by the second version of the system. The quote was in response to the hypothetical situation of having a much larger set of data to work with and it shows that the participant may find the second system preferable given the right circumstances, despite the quantitative results demonstrating a overall preference for the first version amongst both groups. Moreover, this suggests that the amount of control afforded by the system, or at least the perception of control, is a significant factor in determining how the user views its utility.

> *Overall I think the option to choose certain specific things rather than everything in one page was a benefit.*
>
> Group One Participant

The ability to make a conscious choice regarding what information to display when using the system was mentioned several times by participants. It is clear that the freedom to control the data themselves, is an important concern for some participants. This was further discussed during the interview by asking if believe that they would feel the same way using the system with a larger dataset and greater scale.

> **Researcher** - *How do you think that if the system were scaled up [into a full version] you'd feel comparing the two then?*
>
> **Participant** - *During the testing I found the feature of the second program very useful, that I had the option to just select the things I wanted to see and compare to find the thing I was looking, for rather than navigating through the whole set to try and find what I wanted instead.*
>
> Group Two Participant

This participant mentioned that they found the ability to select what they were looking for to be more useful than having the information already presented to them. They also went further and added that this benefit would only increase if you had to work with a larger set of data.

> *So with the second one I thought it was better to just show what I wanted, and for a larger system with more data I think that would be even more useful.*
>
> Group One Participant

> *I would say I prefer the second version somewhat to the first because I think it's more convenient to navigate to things directly, rather than having a page full of extra stuff.*
>
> Group Two Participant

These responses are interesting because they present a view that is contrary the what the quantitative results seem to suggest, which is that the first version of the system is more efficient and more preferable for users in both participant groups. Furthermore, other participants felt that the more efficient way of controlling the system was to have direct control of what elements were on the screen.

These answers given during the post-study interviews however, appear to suggest that the results are not as clear as this. One potential explanation is that the results are true of the system at this scale, however they may not necessarily remain that way at a much larger scale.

> *I think the one, the second one, where you select what you need - under the assumption that you know exactly what it is that you're looking for is a lot quicker to just select what you want rather than scrolling through everything.*
>
> Group One Participant

This quote introduces a new idea, the assumption that the participant knows what they are looking for, as a key element to being able to use the system effectively. The theme of control is often presented as a positive thing by participants who mention it, whereas any negative aspects that were brought up about control tended to be seen as lesser issues that are worth putting up with in return for greater freedom of control.

Learning to use the System

The theme of learning how the system works is prevalent throughout the dataset. Many participants discuss their initial thoughts of how they should interact with the system and how they thought it worked. The impressions of participants were, broadly speaking, positive towards understanding the first system. They stated how their impressions of the system were immediately positive and clear, regarding the way to interact with the software.

> *I found the first one very easy to use, everything was there to look at straight away, and the second one I had to click around a lot and kept forgetting which table was which.*

Group Two Participant

This quote presents the users' immediate impression of the two systems and how they felt that they compared to one another. The initial impression of the system is important to consider, as it is intended to be easy to use for people who do not necessarily have the level of experience or training with other similar pieces of software. As a result, the ability for a user to quickly understand what they're doing is very important.

> **Researcher** - *Do you think that [your preference] would change over time after you had a chance to use it more and get over that learning curve?*
>
> **Participant** - *Oh yes, I think over time you definitely would yes, but I still prefer that the first version is a lot more visual by showing everything on the page, and you can just scroll through it and find what you're looking for a lot quicker.*
>
> Group One Participant

This user suggests that the first version is the better one because it is simpler to understand, and that even with additional time to learn the system it is still a clearer and more visual interface. A similar sentiment is expressed by another participant who found the layout of the first system to be easier than the second version.

> *Overall I think that went well, it was quite easy to use. Though I definitely preferred the first system as all the data was there and I didn't need to understand what each [graph] title meant. I could just look and find what I needed.*
>
> Group One Participant

Both of these quotes express a preference for the first version because they feel it requires less effort to understand and use when compared with the second version. Though not all participants are in agreement on this, as one user felt that the fact it does display everything is not necessarily a benefit.

> *In hindsight I think it's difficult to say which one's better, the first one does have everything there but it can get complex at times, whereas this one [the second version] could probably be tailored for a specific use case*
>
> Group Two Participant

In this excerpt, the user has brought up the possibility that the benefit of the second version is in less general use cases, and rather the fact it can be customised would allow it to serve users better who may have very specific requirements.

So whilst the initial impressions for the first version of the system may have been favourable, the second version of the system had more complex responses.

Although interestingly, upon asking participants to reflect upon their opinions and how a change in circumstances may affect them, several stated how they would likely prefer the second version of the system if they were working with larger datasets or had more time to learn and work with it, thereby overcoming their initial difficulties. These responses are similar to those previously discussed within the theme of *control*, and provide further evidence to suggest that the preference for the first version may not be so certain.

Challenges

One of the most common themes discussed by participants was the challenges and what they struggled with, both with the systems they were using as part of the experiment, but also with the work that these systems are designed to help with.

> *I'd open one [of the graphs] up and find out I'd got the wrong one and have to close it again. Sometimes I'd accidentally click on something in the table which opened a sub-table and I'd forget where I'd got to and have to close it all down.*
>
> Group One Participant

This shows how the participant struggled to use a particular feature of the system, in this case referring to a feature on the second version of the software. This difficulty in finding the correct table that they were looking for was caused by uncertainty towards which was the correct table, as well as becoming disoriented when making a mistake.

The first issue, searching for the correct table, is noted repeatedly in the observations of participant activity taking during each experiment (see section 5.2.2). This is useful information with regards to changes in the system as part of an iterative development process.

The second issue, becoming disoriented after opening the incorrect tables and having to close them all down to start again from a position of understanding, demonstrates that although some participants have stated they would find the extra control more useful, their actions suggest they may need more time to acclimatise themselves with that extra functionality. This is a sentiment expressed by at least one other participant when discussing this issue.

> *I think to be able to use the second one fully you need to know all the sections and table names already because otherwise you could get a bit caught up trying to understand where things are*
>
> Group Two Participant

This idea, that in order to make better use of the second version it would require more understanding of the system, is potentially at odds with the desire to use a simple system that can be understood and used immediately, which is something the participants themselves state that they like about the system.

> *I thought the second was a bit more clunky, having to search and click through each of the options to find what you want.*
>
> Group Two Participant

Several participants suggested that they found the second version to be more complicated to use initially when compared to the first version, although many also suggested that this additional complexity provided greater control which was desirable.

It is important to note though, that whilst the responses from the *Understanding the System* section may give the impression that the first version was universally preferred by the participants, there were still some who found it to be challenging, and brought up some difficulties they found with it.

> *Although I found the first one easy, easier, to use. Initially when I first opened it I found it quite daunting, because there was so much there, you just kept scrolling down and down past loads and loads of graphs*
>
> Group One Participant

This participant found the display showing all of the data immediately to be a little overwhelming at first, though it was still easier than the second version.

Suggestions for Improvements

Making suggestions for how the system could be improved or what changes it would benefit from was an expected point of feedback going into the interviews, and it provides a good way to have users self-report on areas of weakness in the system and how they feel it is lacking. The responses that were relevant to this theme often overlapped with comments from the *struggles* theme, as participants often began by discussing an issue, and followed with how they felt it could be improved upon.

> *There's a couple of times it would be nice, you know, asking questions such as like "what is the highest amount of traffic on this specific interface" it would be nice if that was put in a legend somewhere on here*
>
> Group One Participant

This suggestion appears to be a request for more information to be included on top of what is already available, by using simple requests about the data that could produce pre-calculated results. This is a good example of functionality that could help less experienced users to take advantage of the system for their work, without needing a lot of training.

> *Being able to make the graphs display whatever I specifically needed rather than just these premade ones*
>
> Group Two Participant

As previously discussed in the *control* theme, the degree of agency that a user feels they posses is a critical part of how they perceive the usability of a system, and in suggesting that the system would benefit from allowing users to manually create graphs, there appears to be an implicit request that the user is granted greater freedom to select and control the data as they wish.

Group One Participant

5.2.2 Participant Observations

Throughout the duration of the study, each participant was observed for noteworthy behaviour which may be used to provide contextual information or additional detail later during analysis of the study results. The following section contains a series of key observations provided as short bullet point style notes, taken verbatim from the written notes made during the study. Additionally, many observations were repeated across multiple participants, such as particular difficulty completing a certain task, as a result these selected observations have been chosen to best represent the most common or noteworthy observed details, the rest of the observations and comments are available in appendix B.

- Struggling to isolate individual traces[2]
- Difficulty resetting the graphs to their original layout
- Finding the correct graph from the search bar is difficult

The observations frequently observed participants from both groups struggling to find understand the table names from the search bar in the second version. This is supported by the participants who raised the issue in their interviews, as previously discussed in the thematic analysis above. The repeat observations as well as explicit mentions by participants suggest that it was one of their primary issues with operating the second version of the system.

5.3 Discussion

5.3.1 Participant Demographics

The participant demographics are presented below to provide contextual detail to the study. All questionnaires provided to participants in the study, including the demographic sheets, are available in appendix A

Table 5.6 shows the age ranges of the participants from each group for the study. The first group refers to the participants from the technical staff at Park Air Systems

[2]Traces refers to a single set of data in a graph showing multiple data points (an individual network protocol for example) in the graphs

and group two the student participants who were unrelated to the company. The age ranges vary between the two groups, with the Park Air staff containing several people over the age of thirty (30), which seems appropriate given the participants are drawn from their workforce. Whereas group two is entirely comprised of people aged between 18 and 30, which is also to be expected as the participants were recruited from current and recently graduated students at the University of Lincoln.

Group	18 - 30	31 - 40	41 - 50	51 - 60	61+
1	7	1	4	0	0
2	12	0	0	0	0

TABLE 5.6: Participant Age Ranges

The gender distribution amongst the participants of each group is identical with only a single female participant out of twelve people. Although this ratio is highly skewed towards males, the state of engineering and technology fields means this is not a particularly abnormal proportion. Moreover, since it has been already shown that a person's gender has no significant relationship to their perception of a system's usability [71], [90], there is no need to make comparison between the results of each gender for this study and therefore it is not an issue.

Group	Male	Female	Other	Prefer Not To Say
1	11	1	0	0
2	11	1	0	0

TABLE 5.7: Participant Genders

The final piece of demographic information collected from each participant was a job title. This information was to provide some extra information about the background of the participants, particularly relating to seniority of their position (and thereby experience) and also their direct line of work. The job titles are presented in table 5.8 below.

Group	Engineer	Developer	Manager	Student
1	9	2	1	0
2	1	2	0	9

TABLE 5.8: Job Titles of test Participants

The job titles have been grouped together by title similarity for ease of presentation, of the engineers and developer from group one, four had some degree of seniority in their titles, *Principle Engineer or Senior Software Engineer* for example. This level of seniority also corresponded with ages over 30 which also indicates a greater level of experience than younger, more junior participants.

The participants from group two were all university students, or recent graduates, in STEM subjects. Of the participants who listed their job title as *student*, all except one were postgraduate level.

5.3.2 Task Performance

Findings from the study show that the dashboard based system was more popular with the participants from both groups, the average time to complete all of the tasks was also lower for the first version, which appears to fit with the usability scores. The results of the Linear Mixed Model also show that the version of the system made a difference to the performance and perceived usability of the system by the participants. This indicates that a reason the participants preferred the first version was that they were able to use it more efficiently to find what they were looking for, doing so quicker than with the second version.

The results also show that the group that a participant belonged to, and thereby the differences between those two groups had no effect on their performance. The primary difference between the two groups is of course their backgrounds, group one consists of technical staff from the company Park Air Systems, whereas the second group comprises of students and recent graduates from University, all in STEM subjects. This difference in the backgrounds of the participants is interesting given the nature of the results, as a prior assumption is that the participants with relevant experience and knowledge of the field would perform better in the tests than those who did not possess such knowledge. The evidence however appears to point towards the opposite of this, suggesting that the lack of relevant experience in the field either improved the participants scores, or did not sufficiently hinder participants who are perhaps predisposed to the kinds of tasks used in this study.

An alternative explanation is that the demographic differences between the groups can explain how the second group performed equal to, or better than, the first group on average. As has already been noted, there is no difference in the gender distribution of the two groups, therefore it is not a variable that could affect the outcome of the test. However, the age distribution between the two groups does vary, which could indicate that the age of the participants had an impact on the results more than the variation in their backgrounds. Though the majority of the participants in the first group are young (aged 18 - 30), over 40% of the group reported their ages as over 30 (see table 5.6). Comparatively, the second group are more homogeneous with 100% of participants falling into the 18 - 30 age range.

The difference between the ages in the participant groups may be the reason for the results being as they are. With an average age of 18 - 30, each person from the group has grown up with technology as a part of their life in a way that members of an older generation have not, and it is possible that even with knowledge and experience in the field, the older members of the Park Air Systems participants are unable to match the same level of passive comprehension that accompanies familiarity with technology from childhood.

When initially conducting the study, it was expected that the accuracy of the participants when answering the tasks would produce an interesting outcome, however the data shows that both groups achieved almost identical average accuracy rates for both systems. Furthermore, these accuracy rates were very high (85+%) which suggests that only a small fraction of mistakes were made, and since the values are so similar it does not appear to provide an interesting comparison to other data such as performance times or usability scores. It is because of this that no further analysis of the accuracy results has been performed, as this would distract from the primary results presented in this thesis, and do so without any apparent conclusions to draw or worthwhile information learned.

5.3.3 Usability of the Systems

The results of the usability survey for the two systems demonstrate a clear preference for the first version of the system, by a margin of at least 10 points on the SUS survey for both groups. Looking at the results more directly shows that the results of the SUS for group one show that the difference between the average usability scores was 22 points higher for the first system in comparison to the second, a difference of almost 32%. These results are supported by the task times showing that the fastest average times were for the first version rather than the second. This is relevant as it has already been demonstrated by [91] that there is a positive correlation between performing usability tasks more efficiently, and a higher likelihood of rating the usability of the system positively.

If we compare the results from the SUS survey, against the objective rating scale proposed by [92] we can see that version one of the system is rated as *good* by both groups, whereas the second system is only considered *OK*. The grading scale provides more specificity than the objective scale, and the complete set of ratings can be seen below in table 5.9.

Version	Group	Average Score	Adjective	Grade
1	1	81.25	Good	B
1	2	76.67	Good	C
2	1	58.96	Ok	F
2	2	63.13	Ok	D

TABLE 5.9: SUS Scores on the Adjective and Grade Rating Scales

This table shows that participants from group one, had the strongest opinions regarding the usability of the systems. They found the first version to be the most usable, and they rated it higher overall than the second group. They also rated the second version lower than the other group, suggesting they have a stronger overall preference between the two.

5.3.4 Participant Interviews

One takeaway from the participant interviews is that there is a fine balance in each of the two versions of the system that will determine whether a participant is more likely to prefer a given version. The first version must find a balance between providing enough information to be relevant and easy to understand, but not so much information that it is overwhelming and becomes impractical to use. Conversely, the second version of the software must find a compromise between providing the user with more control over what exact information they wish to see, and also maintaining a degree of simplicity such that the system remains easy to understand for most users.

Chapter 6

Conclusion

6.1 Findings

The findings from the study show that both groups of participants found the first version of the system to be more usable in their testing. This was indicated by the faster times on average to complete the set of tasks by both groups for this version of the system. Further analysis of this data showed a statistically significant difference in the participant times between the two versions, indicating that the first version is easier to use than the second because the difference in the design enables easier use, thereby resulting in better performance.

Additionally, the results SUS survey were aligned with these finds, as they showed a distinct preference for the first version of the system from participants in both groups. Statistical tests showed that there is a significant difference between the usability scores for the different versions of the system. Again showing that the difference between the two systems had an impact on the way the users perceived them.

The final usability scores for the dashboard version of the system were rated as *good* and between a *B* or *C* grade for usability, which suggests that the system has been successful in its original objective to produce a visualisation system that is more usable than other options. Finally, the performance results by the group two participants indicate that the system is just as capable of being used by those who have no experience with, or knowledge of, radio networks or the system itself.

6.2 Summary of Thesis

6.2.1 Work done

The work conducted as part of this thesis was to design and create a piece of software which could be used to visualise network data captured from radio networks manufactured by Park Air Systems. A shift in the aviation industry has meant a transition to a digital infrastructure that has created problems for the company and their engineers, who require a system capable of providing insights into the status of their networks, without the extensive functionality and complexity of larger network tools or systems. The desired solution to the problem was to determine how to create a system capable of presenting the network information without requiring extensive training or experience with the system or the data.

This problem led to the background research into networking systems and tools, data analytics and visualisation techniques in order to understand the previous work completed in those areas and how it may pertain to this project. The related work presented in chapter 2 included prior academic research as well as commercial and industrial systems that currently exist to solve similar problems.

The design of the system was influenced by various tools, especially by systems that provide a dashboard style overview layout. The requirements as provided by Park Air Systems, detailing required and desirable functionality were used as a guide to determine what functionality was to be developed in order to create a system capable of testing and use in an academic study within the timeframe of the project.

The system was initially developed with the aim of producing a system that would both capture and visualise the data from a live radio network. Unfortunately due to challenges with providing access to the radio networks for testing and use as part of the study, the live capture functionality of the system was stopped and the focus shifted towards analysis and visualisation of a pre-captured set of data. After the focus of the system moved away from real-time data, and onto post-hoc analysis and visualisation, the system was developed with a more limited scale in order to enable a functional prototype to be created sooner than it would have otherwise been. The system was created using Python and made extensive use of standard libraries to implement functionality; this was done for ease of development and future maintenance. The language was also chosen because the nature of the program meant it would be more useful to create a functional demonstration of the system for conducting research and testing, rather than a much slower process to develop a more optimised implementation with a lower level language.

Two versions of the system were created, each using a different control scheme to present the data to the user. The different systems represented two different approaches to a visualisation system; one which provides an overview and allows the user to explore as they wish, the other requires the user to know what they want but offers a potentially more efficient worlflow.

The study conducted in the project was intended to investigate the usability and effectiveness of the system in general and the two control mechanisms in particular. With the focus on how the users interacted with the system and what effect the different control mechanisms had on how they perceived the usability of the system as well as how their performance changed between each version. The results of the study indicated that the dashboard overview system was universally preferred for usability as well as performance times.

6.2.2 Objectives Completed

Initially the objectives of this project were to develop a system capable of capturing and analysing network data from radio systems for Park Air Systems, and to conduct a study into how the users of such a system interact with and perceive it.

Though there have been some changes to the specifics of the system and its exact functionality, these objectives have not significantly changed, and have in fact been met.

Objective One

Objective one was to create a system which can be used for capturing and visualising network information. This objective was subject to some changes throughout the course of the project as the initial requirement for the system was to capture live data from the network and then visualise it for the user. This changed during the project into a system that would visualise data that had already been captured, and as a result any further work into live network capture was stopped. Overall however, the objective is considered to have been met as the project produced a working prototype of a system capable of visualising network data, which was used for the research study to investigate its usability.

Objective Two

The second objective was to investigate the usability of different mechanisms for presenting visualised network data to the user. There are many different ways of presenting this information, and for this given use case, this objective was to determine how different techniques might affect the usefulness or usabilityof the system.

Research Question

The original research question for this project, as stated in chapter 1.2 was:

> *Can a high-level, bespoke visualisation system of Air Traffic Control radio networks be usable to engineers without significant expertise in computer networking?*

Chapter 6.1 states that the results of the research study was that the dashboard system is more usable and preferable to both engineers from Park Air Systems, and to unrelated users with only a basic awareness of the system or its purpose. In discussing the system with each participant it became clear that the system was useful and that the company participants felt it would be beneficial to their work.

It is therefore clear that these findings demonstrate that objective two has been met.

6.2.3 Challenges and Limitations

Inevitably when looking back on past work, mistakes tend to present themselves more willingly than they do when initially making them. To highlight and discuss those mistakes, this section will discuss the challenges and limitations of the work done for this thesis.

The nature of the study meant that in order to make use of the most appropriate participants, those from the technical staff at Park Air Systems, there would have to be some compromises to minimise the interference of the research with the daily operations of a business and manufacturing environment. This meant that the time available to conduct the study was limited in multiple ways, firstly the dates on which it would be possible to visit the company and conduct the study were limited due to conflicting schedules of many members of staff. Secondly, it would not have been possible to ask the company to spare a large number of their technical staff for several hours at a time during their daily work, so the length of the study was required to be as short as possible.

A key limitation of the study was the small sample size, only 24 participants with only half from the company. This was a limitation on the amount of potential data that could be collected, however it was not a significant issue for the study as the sample participants were very directly relevant to the project.

In retrospect there are some elements of the study that were not as well considered as they could have been, and were the study to be repeated there are some changes that would be made.

Implementing a mechanism to measure error rates amongst the participants when they are using the system would allow for a much more detailed analysis of the specific performance challenges than the accuracy data that was collected.

6.3 Future Work

There are several potential areas of future research to continue with this project. A simple project would be to conduct a repeat study with similar methodology, but utilise a larger group of participants to gather more data and test for the same results.

One particular area for continued study is the usability of expert level software, given that such software is often focused on performance for the user but can possibly neglect the potential performance improvements of having a system that can be more easily understood and used. One potential study for this would be to conduct a longitudinal study into the perception of usability between beginner and expert level systems. Testing how experience with a system affects the performance and perception of usability with users over time, and if there is a difference between the beginner and expert systems as time progresses and experience builds.

Continuing with studies into how different metrics and usability measurement tools can produce different results from a study of perceived usability can build on the work of [3] and [72] to provide a more solid understanding of the effects of user perception on the usability of a system, and how best to account for and measure this properly with other studies would also be a good potential continuation of this work.

A final area for future study to continue with this work would be to conduct a large scale study of many different expert level systems in a variety of different fields, to gather a database of usability measurement scores which can be used to examine a wide range of different real-world software and how usable they are perceived to be.

References

[1] M. Strohmeier, M. Schäfer, R. Pinheiro, V. Lenders, and I. Martinovic, "On Perception and Reality in Wireless Air Traffic Communication Security", *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 6, pp. 1338–1357, 2017.

[2] J. Hird, M. Hawley, and C. Machin, "Air Traffic Management Security Research in SESAR", in *Availability, Reliability and Security (ARES), 2016 11th International Conference on*, IEEE, 2016, pp. 486–492.

[3] J. Sauro and J. R. Lewis, "Correlations among prototypical usability metrics", *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, p. 1609, 2009. DOI: `10.1145/1518701.1518947`. [Online]. Available: `http://dl.acm.org/citation.cfm?doid=1518701.1518947`.

[4] M Handley, H Schulzrinne, E Schooler, and J Rosenberg, *RFC 2543 - SIP: Session Initiation Protocol*, 1999. [Online]. Available: `https://www.ietf.org/rfc/rfc2543.txt`.

[5] H Schulzrinne, S Casner, R Frederick, and V Jacobson, *RFC 1889 - RTP: A Transport Protocol for Real-Time Applications*, 1996. [Online]. Available: `https://tools.ietf.org/html/rfc1889`.

[6] DARPA, "RFC 791 - Internet Protocol", 1981.

[7] ——, *RFC 793 - Transmission Control Protocol*, 1981. [Online]. Available: `https://tools.ietf.org/html/rfc793#section-3.4`.

[8] A. Ruffi and R. McDonald, *Network Basics Companion Guide*, 1st. Indianapolis, USA: Cisco Press, 2013, ISBN: 1-58713-317-2.

[9] R. A. Alhanani and J. Abouchabaka, "An overview of different techniques and algorithms for network topology discovery", *2014 Second World Conference on Complex Systems (Wccs)*, pp. 530–535, 2014.

[10] J. Jiang, X. Xu, and N. Cao, "Research on Improved Physical Topology Discovery Based on SNMP", *22017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pp. 219–222, 2017. DOI: `10.1109/CSE-EUC.2017.224`. [Online]. Available: `http://ieeexplore.ieee.org/document/8006001/`.

[11] J. Ge and W. Xiao, "Network layer network topology discovery algorithm research", 2013, pp. 542–545.

[12] S. Pandey, M.-J. Choi, S.-J. Lee, and J. Hong, "IP network topology discovery using SNMP", *International Conference on Information Networking, 2009. ICOIN 2009*, pp. 1–5, 2009. [Online]. Available: http://ieeexplore.ieee.org/ielx5/4816117/4897247/04897254.pdf?tp=&arnumber=4897254&isnumber=4897247%5Cnhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4897254.

[13] W. Zhangchao, Z. Yan, Z. Dedong, and W. Hongwei, "An algorithm and implementation of network topology discovery based on SNMP", *2016 First IEEE International Conference on Computer Communication and the Internet (ICCCI)*, pp. 283–286, 2016. DOI: 10.1109/CCI.2016.7778926. [Online]. Available: http://ieeexplore.ieee.org/document/7778926/.

[14] J. Postel, *RFC 768 - User Datagram Protocol*, 1980. [Online]. Available: https://tools.ietf.org/html/rfc768.

[15] J. D. Case, M. S. Fedor, M. L. Schoffstall, and C Davin, *RFC 1157 - A Simple network management protocol (SNMP)*, 1990. [Online]. Available: https://tools.ietf.org/html/rfc1157.

[16] B. Vachon and A. Johnson, *Connecting Networks v6 Companion Guide*, 1st, M. Taub, M. B. Ray, S. Schroeder, C. Christopher, T. Simpson, C. Hutchinson, and R. McDonald, Eds. Indianapolis, USA: Cisco Press, 2017, ISBN: 1-58713-432-2. [Online]. Available: http://www.ciscopress.com/store/connecting-networks-v6-companion-guide-9781587134326.

[17] E. M. Salvador and L. Z. Granville, "Using visualization techniques for SNMP traffic analyses", *Proceedings - IEEE Symposium on Computers and Communications*, pp. 806–811, 2008, ISSN: 15301346. DOI: 10.1109/ISCC.2008.4625672.

[18] O. A. De Lima, V. Fresse, and F. Rousseau, "Evaluation of SNMP-like protocol to manage a NoC emulation platform", *Proceedings of the 2014 International Conference on Field-Programmable Technology, FPT 2014*, no. January 2015, pp. 199–206, 2015. DOI: 10.1109/FPT.2014.7082776.

[19] C. Plummer David, *RFC 826 - An Ethernet Address Resolution Protocol*, 1982. [Online]. Available: https://tools.ietf.org/html/rfc826.

[20] D. Keim, G. Andrienko, J.-D. Fekete, C. Görg, J. Kohlhammer, and G. Melançon, "Visual Analytics: Definition, Process, and Challenges", in *Information Visualization: Human Centered Issues and Perspectives*, A. Kerren, J. Stasko, J.-D. Fekete, and C. North, Eds., 1st ed., Springer-Verlag Berlin Heidelberg, 2008, ch. 7, pp. 154–175, ISBN: 9780123814647. DOI: 10.1109/MCG.2017.25.

[21] S. Kandel, A. Paepcke, J. M. Hellerstein, and J. Heer, "Enterprise data analysis and visualization: An interview study", *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2917–2926, 2012, ISSN: 10772626. DOI: 10.1109/TVCG.2012.219.

[22] M. Sathiyanarayanan and C. Turkay, "Challenges and Opportunities in using Analytics Combined with Visualisation Techniques for Finding Anomalies in Digital Communications", *16th International Conference on Artificial Intelligence and Law*, 2012.

[23] C. O'Neil and R. Schutt, *Doing Data Science: Straight Talk From The Frontline*, 1st ed., M. Loukides and C. Nash, Eds. O'Reilly, 2014, ISBN: 978-1-449-35865-5.

[24] J. C. Roberts, "State of the art: Coordinated & multiple views in exploratory visualization", *Proceedings - Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization, CMV 2007*, pp. 61–71, 2007. DOI: 10.1109/CMV.2007.20.

[25] H. Shiravi, A. Shiravi, and A. A. Ghorbani, "A survey of visualization systems for network security", *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 8, pp. 1313–1329, 2012, ISSN: 10772626. DOI: 10.1109/TVCG.2011.144.

[26] R. Shi, M. Yang, Y. Zhao, F. Zhou, W. Huang, and S. Zhang, "A Matrix-Based Visualization System for Network Traffic Forensics", *IEEE Systems Journal*, vol. 10, no. 4, pp. 1350–1360, 2016, ISSN: 19379234. DOI: 10.1109/JSYST.2014.2358997.

[27] P. Godfrey, J. Gryz, and P. Lasek, "Interactive Visualization of Large Data Sets", *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 8, pp. 2142–2157, 2016, ISSN: 10414347. DOI: 10.1109/TKDE.2016.2557324.

[28] K. Wongsuphasawat and J. Lin, "Using visualizations to monitor changes and harvest insights from a global-scale logging infrastructure at Twitter", in *2014 IEEE Conference on Visual Analytics Science and Technology, VAST 2014 - Proceedings*, Paris, France: IEEE, 2015, pp. 113–122, ISBN: 9781479962273. DOI: 10.1109/VAST.2014.7042487.

[29] S. Tak, A. Toet, and J. Van Erp, "The perception of visual uncertainty representation by non-experts", *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 6, pp. 935–943, 2014, ISSN: 10772626. DOI: 10.1109/TVCG.2013.247.

[30] D. A. Keim, F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler, "Visual Analytics: Scope and Challenges", in *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics*, S. Simoff, M. Böhlen, and A. Mazeika, Eds., Springer, 2008, pp. 76–90, ISBN: 9781509064717. DOI: 10.1109/CCAA.2017.8229987.

[31] S. M. Ali, N. Gupta, G. K. Nayak, and R. K. Lenka, "Big data visualization: Tools and challenges", *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 656–660, 2016. DOI: 10.1109/IC3I.2016.7918044. [Online]. Available: http://ieeexplore.ieee.org/document/7918044/.

[32] M. Celenk, T. Conley, J. Willis, S. Member, and J. Graham, “Predictive Network Anomaly Detection and Visualization”, *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY*, vol. 5, no. 2, pp. 288–299, 2010.

[33] I. Kalamaras, A. Zamichos, A. Salamanis, A. Drosou, D. D. Kehagias, G. Margaritis, S. Papadopoulos, and D. Tzovaras, “An Interactive Visual Analytics Platform for Smart Intelligent Transportation Systems Management”, *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 2, pp. 487–496, 2018, ISSN: 15249050. DOI: 10.1109/TITS.2017.2727143.

[34] Z. M. Fadlullah, F. Tang, B. Mao, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, “State-of-the-Art Deep Learning: Evolving Machine Intelligence Toward Tomorrow’s Intelligent Network Traffic Control Systems”, *IEEE Communications Surveys and Tutorials*, vol. 19, no. 4, pp. 2432–2455, 2017, ISSN: 1553877X. DOI: 10.1109/COMST.2017.2707140.

[35] V. T. Guimaraes, C. M. Dal Sasso Freitas, R Sadre, L. M. Tarouco, and L. Z. Granville, “A Survey on Information Visualization for Network and Service Management”, *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, p. 1, 2015, ISSN: 1553-877X. DOI: 10.1109/COMST.2015.2450538.

[36] J James and K. A. Cook, “A Visual Analytics Agenda”, *IEEE Computer Graphics and Applications*, vol. 26, no. 1, pp. 10–13, 2006.

[37] E. Karapistoli and A. A. Economides, “Wireless sensor network security visualization”, *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, pp. 850–856, 2012, ISSN: 21570221. DOI: 10.1109/ICUMT.2012.6459781.

[38] A. Figueiras, “A typology for data visualization on the web”, *Proceedings of the International Conference on Information Visualisation*, pp. 351–358, 2013, ISSN: 10939547. DOI: 10.1109/IV.2013.45.

[39] N. Elmqvist and J.-d. Fekete, “Hierarchical Aggregation for Information Visualization : Overview , Techniques , and Design Guidelines”, 2012.

[40] B. Lee, P. Isenberg, N. H. Riche, S. Carpendale, B. Lee, P. Isenberg, N. H. Riche, S. Carpendale, B. Mouse, B. Lee, P. Isenberg, N. H. Riche, and S. Carpendale, “Beyond Mouse and Keyboard : Expanding Design Considerations for Information Visualization Interactions”, vol. 18, no. 12, pp. 2689–2698, 2012.

[41] N. Elmqvist, J. Stasko, and P. Tsigas, “DataMeadow: A visual canvas for analysis of large-scale multivariate data”, *Information Visualization*, vol. 7, no. 1, pp. 18 –33, 2008. DOI: 10.1109/VAST.2007.4389013.

[42] D. A. Szafir, “Modeling Color Difference for Visualization Design”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 392–401, 2018, ISSN: 10772626. DOI: 10.1109/TVCG.2017.2744359.

[43] R. S. Amant, “Natural Interaction with Visualization Systems”, *IEEE Internet Computing*, vol. 19, no. 6, pp. 60–64, 2015, ISSN: 10897801. DOI: 10.1109/MIC.2015.129.

[44] E. Wu and A. Nandi, “Towards Perception-aware Interactive Data Visualization Systems”, pp. 1–4, 2015.

[45] E. Zgraggen, Z. Zhao, R. Zeleznik, and T. Kraska, “Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis”, *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pp. 1–12, 2018, ISSN: 0263-774X. DOI: 10.1145/3173574.3174053. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3173574.3174053.

[46] S. Haroz and D. Whitney, “How capacity limits of attention impact visualization”, vol. 18, no. 12, pp. 2402–2410, 2012, ISSN: 1077-2626. DOI: http://dx.doi.org/10.1007/s10811-014-0323-5. [Online]. Available: http://steveharoz.com/research/attention/.

[47] Z. Liu and J. Heer, “The Effects of Interactive Latency on Exploratory Visual Analysis”, *IEEE Transactoins On Visualization and Computer Graphics*, vol. 20, no. 12, 2014.

[48] B. Foucault Welles and W. Xu, “Network visualization and problem-solving support: A cognitive fit study”, *Social Networks*, vol. 54, pp. 162–167, 2018, ISSN: 03788733. DOI: 10.1016/j.socnet.2018.01.005. [Online]. Available: https://doi.org/10.1016/j.socnet.2018.01.005.

[49] B. Shneiderman, “The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations”, *The Craft of Information Visualization*, pp. 364–371, 1996, ISSN: 1049-2615. DOI: 10.1016/B978-155860915-0/50046-9. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/B9781558609150500469.

[50] G. Ellis and A. Dix, “A taxonomy of clutter reduction for information visualisation”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1216–1223, 2007, ISSN: 10772626. DOI: 10.1109/TVCG.2007.70535.

[51] E Duval, “Attention, please! Learning Analytics for Visualization and Recommendation”, in *Proceedings of LAK11: 1st International Conference on Learning Analytics and Knowledge 2011*, 2011. DOI: 10.1109/MPRV.2014.3.

[52] M. Bastian, *Gephi*, 2008. [Online]. Available: https://gephi.org.

[53] Cyram Inc, *NetMiner*, 2001. [Online]. Available: www.netminer.com.

[54] T. M. Orlando and W. D. Sunindyo, “Designing dashboard visualization for heterogeneous stakeholders (case study: ITB central library)”, in *Proceedings of 2017 International Conference on Data and Software Engineering, ICoDSE 2017*, vol. 2018-Janua, IEEE, 2018, pp. 1–6, ISBN: 9781538614495. DOI: 10.1109/ICODSE.2017.8285872.

[55] M. Gleicher, M. Correll, C. Nothelfer, and S. Franconeri, "Perception of average value in multiclass scatterplots", *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2316–2325, 2013, ISSN: 10772626. DOI: 10.1109/TVCG.2013.183.

[56] M. Gleicher, "Considerations for Visualizing Comparison", *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 1, pp. 413–423, 2018, ISSN: 10772626. DOI: 10.1109/TVCG.2017.2744199.

[57] M. Gleicher, D. Albers, R. Walker, I. Jusufi, C. D. Hansen, and J. C. Roberts, "Visual comparison for information visualization", *Information Visualization*, vol. 10, no. 4, pp. 289–309, 2011, ISSN: 14738716. DOI: 10.1177/1473871611416549.

[58] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski, "Visual Methods for Analyzing Time-Oriented Data", *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 1, pp. 47 –60, 2008. DOI: 10.4018/978-1-59904-889-5.ch170.

[59] K. Fatema, V. C. Emeakaroha, P. D. Healy, J. P. Morrison, and T. Lynn, "A survey of Cloud monitoring tools: Taxonomy, capabilities and objectives", *Journal of Parallel and Distributed Computing*, vol. 74, no. 10, pp. 2918–2933, 2014, ISSN: 07437315. DOI: 10.1016/j.jpdc.2014.06.007. [Online]. Available: http://dx.doi.org/10.1016/j.jpdc.2014.06.007.

[60] Tao Zhang, Qi Liao, and Lei Shi, "Bridging the Gap of Network Management and Anomaly Detection through Interactive Visualization", *2014 IEEE Pacific Visualization Symposium*, pp. 253–257, 2014. DOI: 10.1109/PacificVis.2014.22. [Online]. Available: http://ieeexplore.ieee.org/document/6787176/.

[61] M Withall, I Phillips, and D Parish, "Network Visualisation: A Review", *IET Communications*, vol. 1, no. 3, pp. 365 –372, 2007. DOI: 10.1049/iet-com:20060169.

[62] R. Blue, C. Dunne, A. Fuchs, K. King, and A. Schulman, "Visualizing real-time network resource usage", *Visualization for Computer ...*, vol. 5210, no. Chapter 12, pp. 119–135, 2008, ISSN: 978-3-540-85931-4. DOI: 10.1007/978-3-540-85933-8{_}12. [Online]. Available: http://link.springer.com/10.1007/978-3-540-85933-8_12%5Cnpapers3://publication/doi/10.1007/978-3-540-85933-8_12.

[63] *Usability | Definition of Usability in English by Oxford Dictionaries*. [Online]. Available: https://en.oxforddictionaries.com/definition/usability.

[64] International Organization for Standardization, *ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs). Guidance on usability*. ISO, 1998.

[65] ——, *ISO 9241-11:2018 Ergonomics of Human-System Interaction — Part 11: Usability: Definitions and Concepts*, 2nd. Geneva, Switzerland: ISO, 2018, pp. 0–36. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.

[66] J. Nielsen, *Usability Engineering*, 1st. Morgan Kaufmann, 1993, pp. 321–361, ISBN: 9788578110796. DOI: 10.1145/1508044.1508050.

[67] J. Brooke, "A quick and dirty usability scale", *Usability evaluation in industry*, 1996.

[68] J. R. Lewis, "The System Usability Scale: Past, Present, and Future", *International Journal of Human-Computer Interaction*, vol. 34, no. 7, pp. 577–590, 2018, ISSN: 15327590. DOI: 10.1080/10447318.2018.1455307. [Online]. Available: https://doi.org/10.1080/10447318.2018.1455307.

[69] S. Borsci, S. Federici, and M. Lauriola, "On the dimensionality of the System Usability Scale: A test of alternative measurement models", *Cognitive Processing*, vol. 10, no. 3, pp. 193–197, 2009, ISSN: 16124782. DOI: 10.1007/s10339-009-0268-9.

[70] J. Brooke, "SUS : A Retrospective", *Journal of Usability Studies*, vol. 8, no. 2, pp. 29–40, 2013, ISSN: 1931-3357. DOI: 10.1074/jbc.R115.675280. [Online]. Available: http://www.usabilityprofessionals.org/upa_publications/jus/2013february/brooke1.html%5Cnhttp://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html.

[71] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale", *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008, ISSN: 10447318. DOI: 10.1080/10447310802205776.

[72] J. R. Lewis, B. S. Utesch, and D. E. Maher, "Measuring Perceived Usability: The SUS, UMUX-LITE, and AltUsability", *International Journal of Human-Computer Interaction*, vol. 31, no. 8, pp. 496–505, 2015, ISSN: 15327590. DOI: 10.1080/10447318.2015.1064654. [Online]. Available: http://dx.doi.org/10.1080/10447318.2015.1064654.

[73] A. Bangor, P. Kortum, and J. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale", *Journal of usability studies*, vol. 4, no. 3, pp. 114–123, 2009, ISSN: 1931-3357. DOI: 66.39.39.113. [Online]. Available: http://66.39.39.113/upa_publications/jus/2009may/JUS_Bangor_May2009.pdf.

[74] K. Finstad, "The usability metric for user experience", *Interacting with Computers*, vol. 22, no. 4, pp. 327–330, 2010, ISSN: 09535438. DOI: 10.1093/iwc/iwt005. [Online]. Available: http://dx.doi.org/10.1016/j.intcom.2010.04.004.

[75] M. Ilker, B. Msc, and D. Karahoca, "Re-Assessing the Usability Metric for User Experience (UMUX) Scale", *Journal of Usability Studies*, vol. 11, no. 3, pp. 89–109, 2016, ISSN: 1931-3357. DOI: 10.1080/10705511.2014.935842. [Online]. Available: http://www.upassoc.org..

[76] J. R. Lewis, B. S. Utesch, and D. E. Maher, "UMUX-LITE: when there's no time for the SUS", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, p. 2099, 2013. DOI: 10.1145/2470654.2481287. [Online]. Available: http://www.scopus.com/inward/record.url?eid=2-s2.0-84877970092&partnerID=tZOtx3y1.

[77] J. Kirakowski and M. Corbett, "SUMI: the Software Usability Measurement Inventory", *British Journal of Educational Technology*, vol. 24, no. 3, pp. 210–212, 1993, ISSN: 14678535. DOI: 10.1111/j.1467-8535.1993.tb00076.x.

[78] S. Mclellan, A. Muddimer, and S. C. Peres, "The Effect of Experience on System Usability Scale Ratings", *Journal of Usability Studies*, vol. 7, no. 2, pp. 56–67, 2012, ISSN: 1931-3357.

[79] P. Kortum and M. Johnson, "The relationship between levels of user experience with a product and perceived system usability", *Proceedings of the Human Factors and Ergonomics Society*, pp. 197–201, 2013, ISSN: 10711813. DOI: 10.1177/1541931213571044.

[80] O. Bjurling, J. Carlbring, A. Jaber, and J. Weilandt, "Effect of domain knowledge on system usability evaluations", no. 2012, 2016.

[81] J. K. Li, M. Mubarak, R. B. Ross, C. D. Carothers, and K. L. Ma, "Visual Analytics Techniques for Exploring the Design Space of Large-Scale High-Radix Networks", *Proceedings - IEEE International Conference on Cluster Computing, ICCC*, vol. 2017-Septe, pp. 193–203, 2017, ISSN: 15525244. DOI: 10.1109/CLUSTER.2017.26.

[82] J. S. Yi, Y. A. Kang, J. T. Stasko, and J. A. Jacko, "Toward a deeper understanding of the role of interaction in information visualization", *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1224–1231, 2007, ISSN: 10772626. DOI: 10.1109/TVCG.2007.70515.

[83] D. A. Keim, F. Mansmann, and H. Ziegler, "Challenges in Visual Data Analysis", in *Tenth International Conference on Information Visualisation (IV'06)*, London, UK: IEEE, 2006, pp. 9 –16, ISBN: 0769526020. DOI: 10.1109/IV.2006.31.

[84] V. Braun and V. Clarke, "Using thematic analysis in psychology", *Qualitative Research in Psychology*, 2006, ISSN: 14780887. DOI: 10.1191/1478088706qp063oa.

[85] G. Combs, *Wireshark*, 1998. [Online]. Available: https://wireshark.org.

[86] M. Roesch, *Snort*, 1998. [Online]. Available: https://snort.org.

[87] *Python 3.6*, 2019. [Online]. Available: https://www.python.org/.

[88] *MAC Address Vendor Lookup API*, 2019. [Online]. Available: https://macvendors.co/api.

[89] J. Sauro and J. R. Lewis, "Average task times in usability tests", *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, p. 2347, 2010. DOI: 10.1145/1753326.1753679. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1753326.1753679.

[90] K. Orfanou, N. Tselios, and C. Katsanos, "Perceived usability evaluation of learning management systems: Empirical evaluation of the System Usability Scale", *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 2, pp. 227–246, 2015, ISSN: 1492-3831. DOI: 10.19173/irrodl.v16i2.1955. [Online]. Available: http://www.irrodl.org/index.php/irrodl/article/view/1955.

[91] S. Borsci, S. Federici, S. Bacci, M. Gnaldi, and F. Bartolucci, "Assessing User Satisfaction in the Era of User Experience: Comparison of the SUS, UMUX, and UMUX-LITE as a Function of Product Experience", *International Journal of Human-Computer Interaction*, vol. 31, no. 8, pp. 484–495, 2015, ISSN: 15327590. DOI: 10.1080/10447318.2015.1064648. [Online]. Available: http://dx.doi.org/10.1080/10447318.2015.1064648https://doi.org/10.1080/10447318.2015.1064648.

[92] A. Bangor, T. Staff, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean", *Journal of usability studies*, vol. 4, no. 3, pp. 114–123, 2009, ISSN: 1931-3357. DOI: 66.39.39.113. [Online]. Available: http://66.39.39.113/upa_publications/jus/2009may/JUS_Bangor_May2009.pdf.

[93] *RStudio*, Boston, MA, 2018. [Online]. Available: www.rstudio.com.

Appendix A

Study Supplemental Materials

Study Demographic Information Form

Age Range:	18 - 30	31 - 40	41 -50	51 -60	61 +	Prefer not to say
	☐	☐	☐	☐	☐	☐

Gender:	Male	Female	Other	Prefer not to say
	☐	☐	☐	☐

Job Title:

FIGURE A.1: Participant Demographics Form

Research Consent Form

Visualising Radio System Network Traffic Information

Researcher: Adam Walker
Email: 11357886@students.lincoln.ac.uk
Supervisor: Mrs Yvonne James

Please make sure that you have read and understood the provided information sheet before proceeding.

	Please Tick
1. I confirm I have read and understand the information sheet provide and that I have had the opportunity to ask questions.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason.	☐
3. I agree to take part in the above study	☐
4. I agree to the use of anonymised quotes in publications	☐
5. I agree to be voice recorded for a short interview at the end of the study	☐
6. I am 18 years of age, or older.	☐

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

FIGURE A.2: Participant Consent Form

Study Information Sheet

Project Title: Visualising Network Data for Air Traffic Radio Systems

Researchers: Adam Walker, Yvonne James

Project Description: This research is to investigate different techniques for displaying network data visualisations on a bespoke software dashboard, developed for use by Park Air Systems.

Details of the Study: The study will require you to complete two small sets of tasks, each using a different version of the visualisation software. Following the completion of each set of tasks you will be asked to complete a short survey regarding how you felt towards the system.

Throughout the experiment, your performance will be measured in terms of the time taken to complete the tasks and your overall accuracy with each task.

Finally you will be asked to participate in a short (approximately 5-10 minute) interview to discuss the systems you just used.

During the study, we would like to collect the following data:

1. Basic demographic data (your approximate age, gender, etc)
2. Your responses to two short surveys regarding your experience after using each system
3. Task performance metrics (Accuracy of your answers, time taken)
4. A short audio recorded interview at the end of the study regarding your overall experience.

The study typically takes approximately 30-45 minutes to complete.

This study is conducted in accordance with the University of Lincoln Ethical Guidelines for participant studies and has been approved by the Ethics board for the College of Science.

Participants' Rights:

1. You may withdraw from the study at any time without giving reason.
2. You have the right to have your questions about the study answered (except where those answers might interfere with the outcome of the study)

Confidentiality: Any data collected from you during this study will be anonymised to ensure there is no way to link the data to any personal information which may be used to identify you (such as email addresses, names, etc). Following the completion of the study, your anonymised data may be made available to other researchers or possibly used for publication.

Further Information: If you have any questions or require further information you are encouraged to ask the researcher before the study begins.

FIGURE A.3: Participant Information Sheet

Study Introduction Script

Thank you for taking part in this study, before we get started I'm going to give you a quick introduction to these systems so that you understand how to control them.

Version 1

This is the first version that you will be testing today. The system is a simple dashboard style interface as you can see, and it contains different visualisations grouped together in sections.

Hovering the mouse over data points in any visualisation with show the specific value at that point.

Clicking and dragging will produce a selection tool that will change the visualisation to show only what was selected. Double clicking anywhere on the graph will reset the graph. On any graph with multiple sets of data, selecting any data in the legend on the side will toggle the inclusion of that data in the visualisation. This should allow you to focus on specific pieces of data if you wish.

In the top right corner of each visualisation there is a navigation menu with selectable buttons that provide all of these previously mentioned functions, as well as additional ones such as panning through the data or zooming in and out.

Searching or sorting a table can be done easily through the search bar at the top of the table, and each column and row can be selected to sort the data.

Version 2

This is the second version that you will be testing. This version does not have any visualisations displayed for you, instead there is menu bar which can be used to select any visualisation to add to the page. You can add any amount of visualisations in any order, and remove them from the page by deleting them from the menu bar.

The visualisations are all the same as in the first version and can be controlled in the same way.

FIGURE A.4: Study Introduction Script

Participant ID:
Software Version:

System Usability Scale

Instructions: For each of the following statements, please select <u>one</u> box that best describes your feelings towards the system you have just used.

	Strongly Disagree				Strongly Agree
1. I think I would like to use this system frequently.	☐	☐	☐	☐	☐
2. I found this system unnecessarily complex.	☐	☐	☐	☐	☐
3. I thought this system was easy to use.	☐	☐	☐	☐	☐
4. I think that I would need assistance to be able to use this system.	☐	☐	☐	☐	☐
5. I found the various functions in this system were well integrated.	☐	☐	☐	☐	☐
6. I thought there was too much inconsistency in this system.	☐	☐	☐	☐	☐
7. I would imagine that most people would learn to use this system very quickly.	☐	☐	☐	☐	☐
8. I found this system very cumbersome/awkward to use.	☐	☐	☐	☐	☐
9. I felt very confident using this system.	☐	☐	☐	☐	☐
10. I needed to learn a lot of things before I could get going with this system.	☐	☐	☐	☐	☐

Please provide any further comments about this system:

FIGURE A.5: SUS Survey Form

Study Task Sheet 1

Instructions: Please complete each of these tasks and write your answers on the answer sheet provided.

1. Which source device transmitted the most bytes across the network?
2. Which destination device received the second most bytes in total from across the network?
3. At timestamp 13:12:30, how many bytes in total were being transmitted across the network?
4. How many packets were transmitted by TCP at timestamp 13:15:44?
5. What is the timestamp when Websocket packets are first transmitted?
6. Which protocol transmits more packets over the network: ARP or ICMP?
7. What percentage of the total traffic is made up by ARP?
8. Who is the manufacturer of the following device: e0:18:77:ca:77:d6?
9. Which destination device receives the most traffic on the network?
10. What percentage of traffic is transmitted from the source device: 00:15:8b:00:75:3a

FIGURE A.6: Participant Demographics Form

Study Task Sheet 2

Instructions: Please complete each of these tasks and write your answers on the answer sheet provided.

1. Which source device transmitted the least bytes across the network?
2. For the destination device which received the most bytes, how many bytes were received?
3. What is the timestamp at the point of the highest value for total bytes sent across the network?
4. How many packets were transmitted by UDP at timestamp 13:08:14?
5. At timestamp 13:14:28 what is the third most transmitted protocol in terms of packets sent?
6. Which timestamp has more packets transmitted by TCP: 13:06:30 or 13:08:30?
7. How many total packets does HTTP send over the network?
8. What percentage of traffic is made up by TCP?
9. What is the most popular manufacturer of the devices on the network?
10. Which destination device receives a higher percentage of traffic: 00:15:8b:00:6e:cd or 00:15:8b:00:75:3a?

FIGURE A.7: Participant Demographics Form

Participant ID:
Software Version:

Study Task Response Sheet

Instructions: Please complete each of the tasks on the instruction sheet provided, in the order they're written and write your answers in the corresponding spaces on this sheet.

Question 1: ____________________

Question 2: ____________________

Question 3: ____________________

Question 4: ____________________

Question 5: ____________________

Question 6: ____________________

Question 7: ____________________

Question 8: ____________________

Question 9: ____________________

Question 10: ____________________

FIGURE A.8: Participant Demographics Form

Appendix B

Results Supplemental Materials

B.1 Residual Histograms

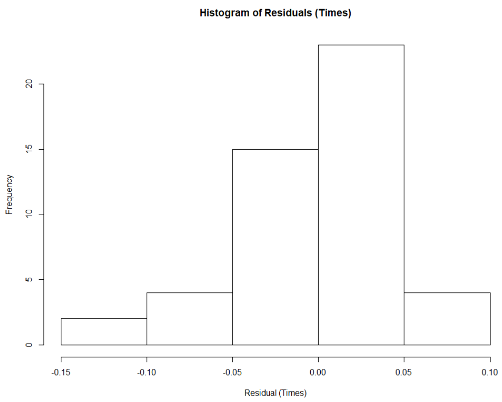

FIGURE B.1: Histogram of Residuals for Time Data

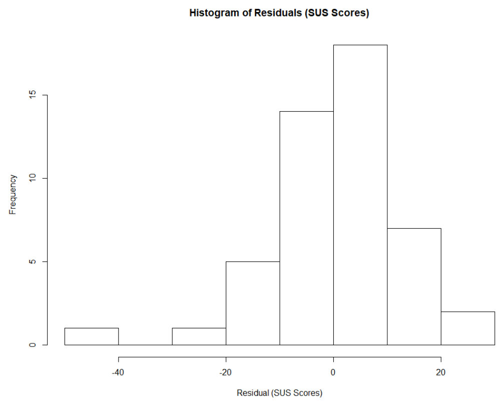

FIGURE B.2: Histogram of Residuals for SUS Data

B.2 Participant Observations

Version 1 Group 1

- Participant gained confidence quickly with system controls
- Some trouble isolating individual traces
- Initial trouble isolated traces
- Searching for values on the wrong chart
- Trouble finding the correct timestamp when adjusting zoom
- Difficulty resetting scale on zoomed graph
- Participant interacted with the system a bit to get the hang of the system

Version 1 Group 2

- Showing all of the graphs in one page makes it easier to see everything
- Sometimes I couldn't find the graph I was looking for because there were too many
- Initial trouble scrolling down the page and not scrolling the graphs
- Finding the correct graph became easier over time

Version 2 Group 1

- Seemed very at ease with the graphs straight away
- Had no trouble using the menu bar to search for graphs
- Accidentally turned panning mode on one graph
- Could not find the correct graph

Version 2 Group 2

- Explored graphs first to find the correct one
- Added only one graph at a time, then removed it before the next one
- Unsure what graph names referred to what graphs
- Gained speed with the system after a few questions
- Opening a graph then closing it before reopening it again

B.3 Participant Comments

Version 1 Group 1

- Protocol Breakdown graph was awkward having double clicked to get one trace, then trying to get back to multiple traces
- I found the software easy to use and contained enough information to allow me to easily diagnose network problems when working on the T6
- Perhaps include a timestamp of the overall collection period
- Very easy to use and understand, good user interface

Version 1 Group 2

- Showing all of the graphs in one page makes it easier to see everything
- Sometimes I couldn't find the graph I was looking for because there were too many

Version 2 Group 1

- Protocol Table gives a count but doesn't make it clear that it's packets
- The implementation aims to be simpler by displaying less on screen however it has a higher learning curve as people may switch to the incorrect table
- I needed to know what to search for which made it take longer to find the answer when I didn't
- Being able to search was useful but a better indication of graph types might help
- I definitely had to think harder with this system. Especially about which information was in which chart

Version 2 Group 2

- Some of the colours used in the protocols graph were difficult to tell apart
- Sometimes I was unsure which graph was related to which task
- Perhaps include a timestamp of the overall collection period
- Very easy to use and understand, good user interface
- I liked being able to pick which graphs I want to compare to each other