
Electronic Thesis and Dissertation Repository

6-24-2020 10:45 AM

Design of Interactive Visualizations for Next-Generation Ultra-Large Communication Networks

Wenjun Chen, *The University of Western Ontario*

Supervisor: Sedig, Kamran, *The University of Western Ontario*

: Haque, Anwar, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in
Computer Science

© Wenjun Chen 2020

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Graphics and Human Computer Interfaces Commons](#), and the [OS and Networks Commons](#)

Recommended Citation

Chen, Wenjun, "Design of Interactive Visualizations for Next-Generation Ultra-Large Communication Networks" (2020). *Electronic Thesis and Dissertation Repository*. 7121.

<https://ir.lib.uwo.ca/etd/7121>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

With the increasing size and complexity of next-generation communication networks, it is critical to utilize interactive information visualization to support the network monitoring, planning, and management. Effectively visualizing large-scale networks has been considered difficult with traditional methods because of the high link density and complicated node relationship. Given the limited screen space, it is essential to explore how to present ultra-large-scale network data efficiently that can assist Internet Service Provider's (ISP) network planning and management activities. This work proposes a design of the real-time interactive visualization system that combines the idea of progressive disclosure and multiple panels to elegantly visualize the large-scale network and avoid the information-overloaded problem. The system also visualizes the configuration of the network elements and provides the network performance information, including the port-level Quality of Service (QoS) metrics. Additionally, the system enables navigation through the port-level connection and provides different modes for multiple purposes.

Keywords

Information Visualization, Network Management, Network Performance, Network Monitoring, Quality of Service (QoS), Human-Information Interaction

Summary for Lay Audience

Conducting network planning and management and monitoring network behaviors have been considered more and more important with the growth in network complexity and volume. The efficient network and service management with the large-scale communication network has been recognized as an essential part of the business goals. The network operators and planners are required to measure and analyze the large-scale network. With the increasing size and complexity of networks, more researchers consider it vital to utilize information interactive visualization tools to encourage data exploration and real-time analysis, and visualization can support timely response and decision making.

Effectively visualizing a moderately large-scale network has been considered difficult with traditional methods because the density of links is high and the relationship between nodes is complicated. Given the limited screen space, it is essential to explore how to manage large quantities of network data efficiently and also contain the necessary network configuration and performance details. Providing performance metrics can benefit the network operators in understanding and monitoring the behavior of the network, and Quality of Service (QoS) management is one of the essential tasks in the network service management. This work proposes a design of the real-time interactive visualization system that combines the idea of progressive disclosure and multiple panels to elegantly visualize the large-scale network according to its hierarchy and avoid the information-overloaded problem. The system also visualizes the configuration and the inside structure of the basic network elements. In addition, the visualization system provides the network performance information including the port-level QoS metrics information, and the system also enables navigation through the port level connection. The above are expected to benefit the network planners / network managers in conducting their network health analysis in a more effective way.

Acknowledgments

First, I wish to express my deepest gratitude to my supervisors Dr. Kamran Sedig and Dr. Anwar Haque, who convincingly guided and steadfastly encouraged me in all the time of research. Thank you so much for always being supportive and patient, and for giving me lots of guidance both for research and life. This accomplishment would not have been possible without you.

I would like to thank my friends TDM and Jeff for providing me warmly support and encouragement throughout the research.

I owe my very profound gratitude to my family, especially my grandparents, for their unfailing love and care throughout my years of study and writing this thesis.

Finally, my appreciations go to the people serving on the frontlines of COVID-19 in essential services. Thank you so much for your dedication and sacrifices.

Table of Contents

Abstract.....	ii
Summary for Lay Audience.....	iii
Acknowledgments.....	iv
Table of Contents.....	v
List of Tables.....	vii
List of Figures.....	viii
Chapter 1.....	1
1 Introduction.....	1
1.1 Motivation.....	1
1.2 Problem Background.....	2
1.3 Problem Statement.....	3
1.4 Thesis Outline.....	4
Chapter 2.....	5
2 Background.....	5
2.1 Network Basics.....	5
2.1.1 Network Topology.....	5
2.1.2 Router Architecture and Hierarchical Network Design.....	6
2.1.3 Quality of Service (QoS).....	10
2.2 Information Visualization Background.....	12
2.2.1 Information Space and Task Space.....	14
2.2.2 Patterns and Blendings.....	14
2.2.3 Visualization Techniques.....	15
2.2.4 Concrete Encoding and Interaction.....	18
Chapter 3.....	21

3	Related Work	21
3.1	Survey of past network traffic monitoring visualization tools.....	21
3.2	Contemporary Network traffic monitoring visualization tool	23
Chapter 4.....		28
4	Methodology and System Design	28
4.1	System Structure Design.....	28
4.2	Network Data	30
4.3	Visualization Tool Design	34
4.3.1	Method	34
4.3.2	Overview.....	35
4.3.3	The Network Panel	36
4.3.4	The Inside Panel.....	38
4.3.5	Interaction	39
4.3.6	Filtering Function and Alert Mode	45
4.4	Case Study	48
Chapter 5.....		53
5	Conclusion	53
5.1	Summary.....	53
5.2	Limitation and Future Work	54
References.....		56

List of Tables

Table 1: The data range for the simulated configuration information of the access router	31
Table 2: The data range for the simulated configuration information of the edge router	32
Table 3: The data range for the simulated configuration information of the core router	32

List of Figures

Figure 1: The topology layout of the full-mesh network and partial mesh network	6
Figure 2: A high-level view of a generic router architecture [20]	7
Figure 3: The structure of the hierarchical network [21].....	9
Figure 4: The reference model for visualization proposed by Card [8].....	13
Figure 5: The visualization design process proposed by Sedig and Parsons [26]	13
Figure 6: a- The arc diagram (up) [30]; b- The adjacency matrix layout (bottom left) [31]; c- The hive plots (bottom right) [32]	16
Figure 7: The force-directed layout [33].....	17
Figure 8: The Voronoi diagram [35] and The pseudo-code to implement the Voronoi diagram [26].....	18
Figure 9: An overview of SeeNet visualization [58]	22
Figure 10: The screenshots of the DNS attack in country zoom level and host zoom level in SURFmap [63]	24
Figure 11: A top-layer overview of a complex network and a prototype GUI with Satellite View, Palette Window and Properties Window highlighted [10].....	24
Figure 12: a- The GPS mode (3D) in MANET [64]; b- The example of displaying physical connections [5].....	25
Figure 13: A cluster view (different colors represent different clusters) using the Walktrap algorithm. The cluster visualization helps in understanding the communities [6].	26
Figure 14: The system design configuration architecture.....	28
Figure 15: The network data example	33

Figure 16: The generated filtered data file as the QoS report.....	34
Figure 17: The overview of the visualization tool	36
Figure 18: The network panels	37
Figure 19: The B3 inside panel.....	39
Figure 20: Hovering the nodes or the links.....	40
Figure 21: Clicking one of the core routers in the A1 panel.....	41
Figure 22: a- Clicking the edge router in the A1 panel (Up); b- Clicking the edge router in the A2 panel (Middle); c- Clicking the access router in the A2 panel (Down)	42
Figure 23: Clicking the user in the A3 panel	43
Figure 24: When mouse hovering on the port in the B3 panel	44
Figure 25: When clicking the port, it will navigate to show the inside data of the router where the linked port located and the information of the linked port.	45
Figure 26: Filtering the routers according to the router utilization.....	46
Figure 27: Filtering the ports according to the port delay	47
Figure 28: The comparison between the normal mode and the alert mode	47
Figure 29: a- The normal mode (left); b- The alert mode (right).....	49
Figure 30: The selected router with its linked routers and the link between them are also highlighted. The different levels of connections are on different panels.....	50
Figure 31: Filter the ports, and the visualization tool only shows the ports whose delay is higher than 200ms.....	51
Figure 32: Filter the routers with specific utilization range.....	52

Chapter 1

1 Introduction

1.1 Motivation

There has been a rapid growth in both the variety and volume of network traffic, and at the same time, it is becoming ever more important for network analysts to understand network behaviors and conduct network management [1]. Network management is the act of initializing, monitoring, and modifying the operation of the primary network functions that directly support user requirements. Network management also refers to the methods, activities, procedures, and tools that pertain to the operation, administration, maintenance, and provisioning of networked systems [2]. Network operators and engineers teams should be enabled to monitor, measure, and analyze the network ecosystem to keep it healthy as long as possible [3]. Network monitoring and measurement provide valuable information to network administrators, ISPs, and content providers [4]. The concerns about an efficient network and service management have been considered essential for business goals since it is one of the pillars to ensure competitive advantages in the market [3].

In a large-scale network with various equipment, it is vital to understand the physical and logical connections, the hierarchy of equipment, geographical information, and temporal variance of the network [5]. Some previous research works enable network administrators to retrieve management information from remote devices, to analyze the collected data, and to take decisions to fix or optimize the network by reconfiguring not well-tuned devices. With the complexity of the context data, more and more researchers consider it essential to allow administrators to shift through massive traffic logs in a visually appealing and interactive manner that encourages data exploration rather than hindering it [6]. In the network management domain, however, the gap between the daily network monitoring and the high level of decision making is currently not well bridged by the current interactive visualization tools [7]. Information Visualization as a tool to help human operators in management tasks has always been highlighted as a research challenge in the area of network management [8].

Real-time analysis and visualization are worth more attention because of the support of timely response and decision making [9]. It has been stated that visualization in near real-time may improve network administrators' diagnoses and reactions [3]. Pras et al. [2] have claimed that the existing information visualization techniques and tools do not satisfy the network operators' needs. With the increasing size and complexity of networks, the traditional topological views do not scale well, and there is a growing need for online or real-time visualizations to reduce detection and reaction time.

1.2 Problem Background

Effectively visualizing a moderately large-scale network has been considered difficult with traditional methods. The density of links is high, and the relationship between nodes is complicated. Given the limited screen space, it is essential to explore how to manage large quantities of management data efficiently and how to view this data at a user-defined level of detail. Understanding the network as a whole by visualizing objects position, their semantic and relationships in near real-time is a long-term dream of the network operators. The layout and view of topology and the organizing of topology play an important role in network visualization [10]. Early visualizations were aimed at understanding overall network loads and topology, and recent works focus more on helping network operators and network planners understand the network at different scales [1]. In this case, the interactive features with enhanced usability are mandatory, such as zooming, filtering, and details on demand. Various equipment, such as routers, is used in a large-scale network, and the network structure is based on the hierarchical information of the equipment. There remains a research problem of how to arrange the layout for the large-scale hierarchical data to avoid the information overload problem and also display the necessary details and information entirely. Also, visualizing the network elements only according to its hierarchical data may lose other essential information, for example, it is necessary to explore how to easily determine both the geographic information and the hierarchy information of the network elements [5]. In addition, using inappropriate ways of producing visual representations always result in increased development effort and doubtful results from the user perspective. Besides the layout and interaction, the design of

the visual representations also needs careful consideration to better visualize the elements and traffic in the network, and therefore it can enhance the monitoring and planning ability.

Moreover, it is essential to provide enough details in the network visualization to help the administrators for their decision making. Quality of Service (QoS) management is one of the essential tasks in the network service management [11]. QoS monitoring tends to allow network administrators to track the ongoing QoS, compare monitored QoS against expected performance, detect possible QoS degradation, and then tune network resources accordingly to sustain the guaranteed QoS [12]. Besides, with detailed information, the visualization tool may not only be utilized for network management but also be beneficial for network planning. Network planning is concerned with the cost-effective deployment of a communication infrastructure to provide adequate throughput, coverage and quality for the end-user services [13]. The network planners could use interactive information visualization tools to aid in capacity planning since the network congestion points can be easily identified.

1.3 Problem Statement

Given an ultra-large-scale network topology details, real-time network configuration and QoS parameters details, the research goal is to develop an interactive visualization tool for planning and monitoring such large-scale communication network that would support the following features:

1. visualization of ultra-large-scale communication networks and solve the information-overloaded problem;
2. visualization of the configuration and the inside structure of routers, and port level QoS metrics;
3. navigation between the linked nodes and ports in the communication network;
4. different modes to identify the capacity, anomaly and the quality degradation in the network;
5. visual elements (data) filtering and generating the data managed file for future use;

6. real-time visualization of the communication network for monitoring purposes.

1.4 Thesis Outline

The thesis is organized as follows: **Chapter 2** reviews the background of both the network basics and the visualization design methodologies and interaction techniques. **Chapter 3** is the exploration and analysis of previous related network monitoring visualization tools. **Chapter 4** describes the approach and methodology utilized in designing and developing this system and illustrates the design of the visualization in detail. In addition, it also introduces the case study to further illustrate how the system solves the task. In **Chapter 5**, summary and future work are proposed for further examination and future studies.

Chapter 2

2 Background

In this chapter, various related network concepts, as well as information visualization terms and techniques, are discussed.

2.1 Network Basics

2.1.1 Network Topology

In graph theory, a graph is made up of nodes that are connected by edges, and the layout of computer networks is an application of graph theory [14]. Network nodes are the points of connection of the transmission medium to transmitters and receivers of the electrical, optical, or radio signals carries in the medium. The links refer to the transmission media used to link devices to form a computer network.

Understanding the network as a whole by visualizing devices position, their semantic and relationships is a long-term dream of network operators and planners [3] as it allows them to find bottlenecks in the network and to figure out if current hardware is configured correctly [15]. Network topology is a representation of the interconnection between directly connected peers in the network, and it is the arrangement of the elements (links and nodes) of a communication network. Topologies can be physical or logical. The physical topology is the way the network physically laid out, and the logical topology is how the information flows on the network [16], [17]. There are mainly five basic topologies: bus, star, ring, mesh and hybrid. In bus topology, there is a main cable and all the devices are connected to this main cable. In star networks, all the devices are connected to a single hub through a cable. In ring topology, each device is connected with the two devices on either side of it. A combination of two or more topology is known as hybrid topology. In mesh networks, all nodes cooperate to distribute data amongst each other. Each network device is interconnected with one another, allowing for most transmissions to be distributed even if one of the connections goes down. Mesh topology can be a full mesh topology or a partially connected mesh topology (Figure 1). In a full mesh topology, every node in the network has a connection to each of the other devices in the network.

The number of connections in this network can be calculated using the following formula: $n(n-1)/2$. The advantages of a mesh topology are as follows: a mesh topology could manage a high amount of traffic; failure of one device does not break the network or transmission of data; adding additional devices does not disrupt data transmission between other devices [18]. Nevertheless, the full mesh topology is difficult and time-consuming, and there may be redundant connections, which lead to reduced efficiency. In a partially connected mesh topology, at least two of the computers in the network have connections to multiple other computers in that network. The partial mesh is less expensive to implement but yields less redundancy than full mesh topology.

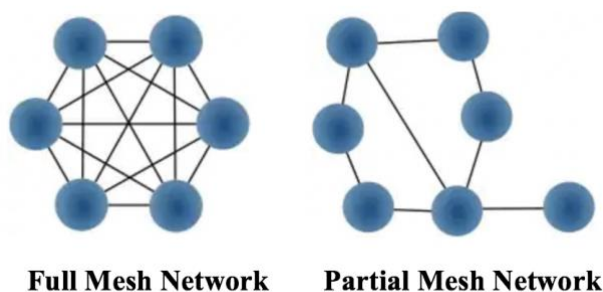


Figure 1: The topology layout of the full-mesh network and partial mesh network

2.1.2 Router Architecture and Hierarchical Network Design

Configuration management is one of the essential areas for network management. It allows a network operator to track which devices are on the managed network and the hardware and software configurations of these devices [19]. The router is one of the network-core devices. It has always been considered as the node in the network, and the connection between the routers is considered as link. A router is a networking device that forwards data packets between computer networks. A packet is forwarded from one router to another through the networks until it reaches its destination node. Forwarding and routing are two major functions of routers [19]. Routers receive the packets from its input ports, checks its header and checksum, and then look up to the routing table to find the appropriate output port to dump the packets onto. Routing is the process that routers determine what is the best path for the packet to reach the destination. One of the aims of our visualization tool

is to monitor the condition and configuration of the inside components of routers. The high-level view of a generic router architecture is represented in Figure 2.

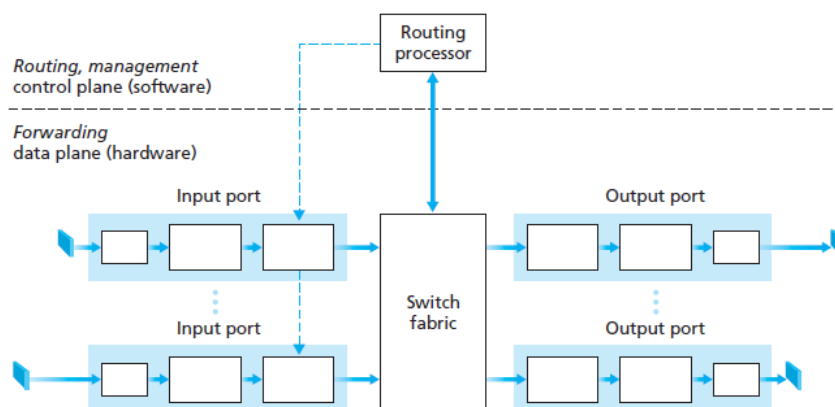


Figure 2: A high-level view of a generic router architecture [20]

The basic components of routers are as following:

- **Central Processing Unit**

The central processing unit in a router performs the primary computing of the device. It is the main component that processes the instructions necessary for the router to operate. It executes the instructions of the operating system and also performs functions such as system initialization and network interface control.

- **Power Supply**

The power supply for the router can be internal or external. Some routers have multiple power supplies for redundancy.

- **Chassis**

A chassis is a frame for mounting the circuit components for routers, and the frame also protects all the vital internal equipment from dust, moisture and tampering.

- **Line card**

The line cards provide interfaces to the network, and it is a circuit board that provides transmitting or receiving ports.

- **Port**

Multiple ports are often gathered together on a single line card within a router [19]. There are both input ports and output ports in different line cards in a router. The input port terminates an incoming physical link to a router, and control packets are forwarded from the input port to the routing processor. The output ports perform the reverse data link and physical layer functionality as the input port. The output ports store the datagrams that have been forwarded to it through the switching fabric and transmits the datagrams on the outgoing link.

- **Flash Memory**

Flash Memory is an electronically erasable and reprogrammable memory chip. The operating system images are held in Flash memory. The content in Flash Memory will be retained when router restarts or powers down.

- **Random Access Memory (RAM)**

Random Access Memory is responsible for handling all the short-term storage requirements for the router. Random Access Memory stored processes and utilities such as a routing table and configuration files. Routing tables have vital information such as the source and destination IP addresses of data packets and metrics that are used to determine the most efficient route when sending data. The content in the RAM will be lost when a router restarts or powers down.

- **Read-Only Memory (ROM)**

The Read-Only Memory is read-only memory available on a router's processor board. The Read-Only Memory usually stored initial system bootstrap software, operating system software and power-on diagnostic test programs (the POST).

- **Non-Volatile RAM (NVRAM)**

The Non-Volatile RAM is where the router holds its startup configuration, and it will not be erased when the router is reloaded or switched off.

The cisco hierarchical internetworking model (Figure 3) has been widely adopted in the industry for designing a flexible, scalable and reliable network [21]. Compared with flat network design, broadcasts can be easily controlled, and undesirable traffic can be efficiently filtered with the hierarchical design. A typical hierarchical network design includes the following three layers: access layer, distribution layer and the core layer. The access layer connects client nodes to the network and access routers are usually located at the customer side like home and branch. The distribution layer provides policy-based connectivity and controls the boundary between the access and core layers. The core layer takes the responsibility to transport data across a network fast and reliably. Also, each ISP has a number of border / gateway routers that connect a specific network (i.e., ISP) to another network or ISP.

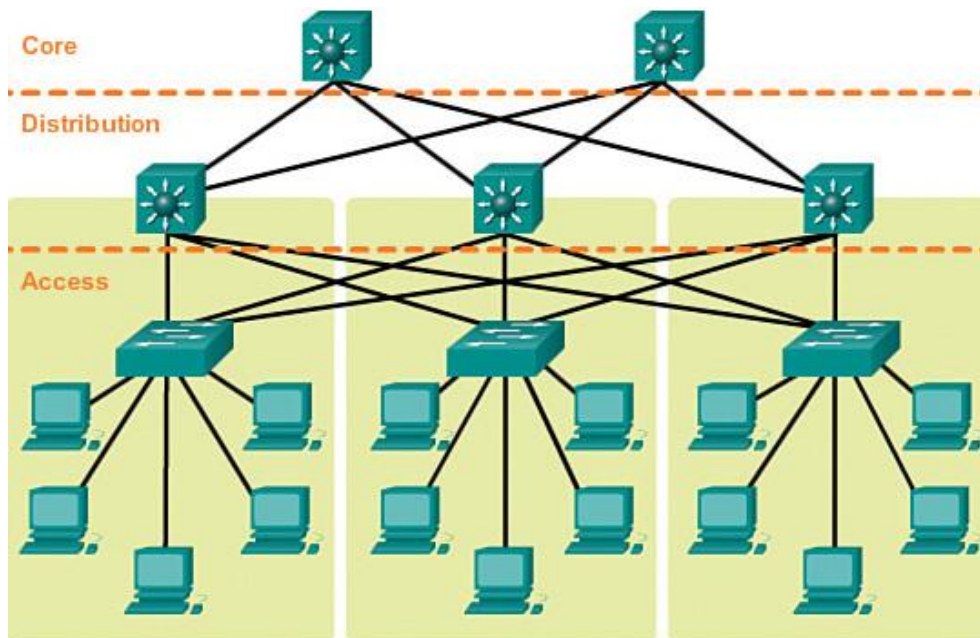


Figure 3: The structure of the hierarchical network [21]

According to the hierarchical model, the routers that belong to each layer have different responsibilities. The access routers are located at home and customer sites. The edge (distribution) routers aggregate traffic from multiple access routers. Core router may provide a collapsed backbone interconnecting the distribution router from various buildings of a campus or significant enterprise locations. Besides the three types of routers, the border router always connects the enterprise networks and the service provider networks where the ISP connections terminate. The border routers are responsible for routing packets outside the area [19]. Therefore, the border router has also been considered and visualized in the visualization system proposed in this project.

2.1.3 Quality of Service (QoS)

The goal of performance management is to quantify, measure, report, analyze, and control the performance of different network components [19]. Providing Quality of Service guarantee is an essential requirement for network performance management [12]. Quality of service is the measurement of the overall performance of a web service, particularly the performance seen by the users of the network. QoS metrics indicate objective, system-related characteristics that provide insights into the delivery service's performance at the transmission level [22]. QoS metrics offer a way to mark traffic prioritization over networks in an attempt to guarantee a certain level of performance. Quality of Service can be affected both by human factors and technical factors. Human factors consist of stability of service quality, availability of service, waiting times and user information [23]. The stability of operation of the system and the fault clearance times may affect the QoS. Technical factors include reliability, scalability, effectiveness, maintainability and network congestion[23]. The QoS may also be formed by the transmission quality factor.

QoS can be measured quantitatively by following main parameters:

- **Packet loss**

Packet loss occurs when the network fails to deliver some packets due to network congestion. Packet loss can be caused by the following common issues: network congestion, software bugs, problems with network hardware and security threats. It

is normally calculated as the percentage of packets that are not received or received malformed, and it is also recognized as a frame loss ratio. The formula for packet loss is as follows:

$$\text{Packet Loss} = \frac{\text{Packets transmitted} - \text{Packets received}}{\text{Packets transmitted}} \times 100$$

- **Latency/delay**

Network latency/delay for a data packet is the time it takes for the packet to travel from its source to its destination. Network latency/delay is the sum of a number of individual latency/delay components which include queuing delay, processing delay, transmission delay, and propagation delay. Queuing delay refers to the delay experienced by the data packet inside a router queue. Processing delay refers to the delay as a result of various operations performed on the data packet by the router's processing engine. Transmission delay refers to the time it takes for the transmission port to transmit the packet over the channel/link. Propagation delay is the ratio between the network link length and the propagation speed over a specific media (such as copper, fiber, etc.).

- **Jitter**

Jitter is the variation in the delay of received packets. Jitter can be caused by a number of factors including network congestion, improper queuing, or configuration errors.

- **Bandwidth**

Bandwidth is the capacity of a network communications link to transmit the maximum amount of data from one point to another in a given amount of time. It is the maximum throughput of a logical or physical communication path in a digital communication system.

- **Reliability**

In a given time window, network reliability is the ratio between the total uptime (i.e., total time – total downtime) and the total time. Network reliability is expressed as a percentage value. Reliability can be calculated as follows:

$$Reliability = \frac{(Total\ time - Total\ downtime)}{Total\ time} \times 100$$

- **Utilization**

Network traffic refers to the data moving across a network at any given time [24]. Network link utilization is the ratio between the amount of traffic on the network link and the total capacity of that link. It is generally specified as a percentage. The formula for utilization is:

$$Utilization = \frac{Link\ traffic}{Link\ capacity} \times 100$$

2.2 Information Visualization Background

Data visualization can incorporate human capabilities into an intuitive visual interface, and it has been widely used in the network field. There are three major fields in data visualization: scientific visualization, information visualization and visual analytics. While scientific visualization deals with scientific data, information visualization deals with abstract data and visual analytics combines automated analysis techniques. In this thesis, we focus on information visualization since the network management datasets are aligned with information visualization assumptions. Information visualization has been recognized by Card et al. [8] as the use of computer-supported, interactive, visual representations of abstract data to amplify cognition. Information visualization focuses on the representation of abstract, unstructured, and high-dimensional data, including business, social network and textual data [25]. Network traffic datasets are generally large scale and high-dimensional, and therefore visualizing network traffic data mostly employs information visualization [9].

A classical reference model of visualization (Figure 4) has been proposed by Card et al. [8] to outline the main components present either in the use of a technique or the development of a new technique. After collecting and processing the raw data to generate the desired information, the information is processed and transformed into some visual representations. Then the end-user interacts with the visual representation in a view. In our cases, users (network administrators/planners) interact with the visualization to perform tasks and to browse the network information.

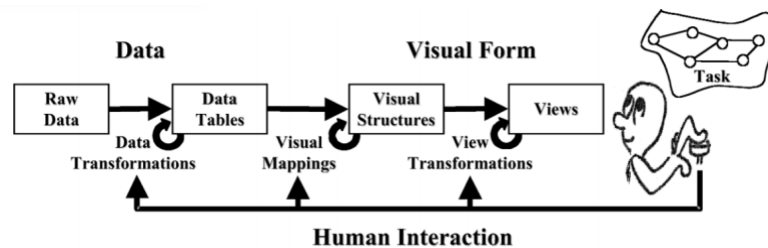


Figure 4: The reference model for visualization proposed by Card [8]

A framework for designing visualization for human-information interaction has been proposed by Sedig and Parsons [26] and the process is made of four major stages: information space and task space, patterns and blendings, visualization techniques and concrete encoding and interaction (Figure 5). The four stages will be discussed in more detail in the following subsections.

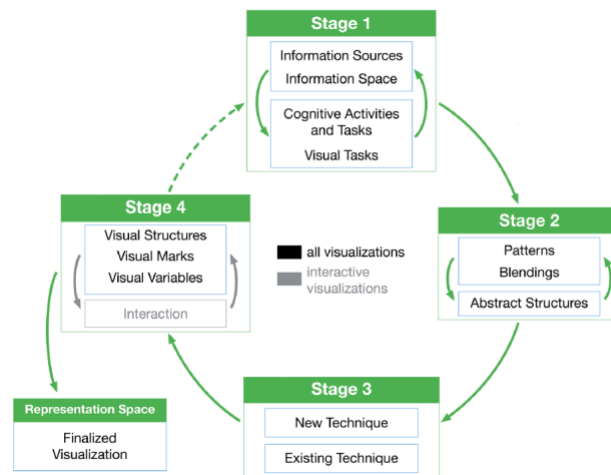


Figure 5: The visualization design process proposed by Sedig and Parsons [27]

2.2.1 Information Space and Task Space

An information space is a body of information that is conceptualized as having spatial characteristics. Within any given information space, information can be combined and blended from multiple sources and environments. It has been claimed that the concept of information space is more powerful than just data in terms of supporting design thinking because spatial metaphor is vital in human thinking, and data does not conceptually capture all of the possibilities for the underlying content of visualization [26].

A deep understanding of human cognition is necessary for research in human-centered visualization and the development of the visualization tool [28]. Cognitive activities involve high-level cognitive processes, for example, sensemaking, decision-making, learning, planning, problem-solving and analytical reasoning. A cognitive activity is made of tasks, which can be cognitive, interactive and visual. Different tasks require different distribution of cognitive, interactive and visual processing. Interactive tasks can be viewed at the level of interaction patterns, techniques, or events. Visual tasks can be divided as low-level and high-level tasks. Low-level visual tasks are performed pre-attentively and/or performed in a concise period of time. High-level tasks and activities emerge from a combination of repeated low-level interactions.

2.2.2 Patterns and Blendings

The patterns have been defined as regularity in some dimension, and they are derived empirically from observations rather than from first principles [29]. Patterns present abstractions at a meaning level of granularity. These patterns are not concrete structures and need to be instantiated as visual structures. The patterns can be instantiated using many different structures and this flexibility benefits creativity. The patterns were divided into three groups: (1)primary, (2) substrate and (3)relational [26]. The following are the descriptions of some patterns and their instantiation in our cases. The Token pattern belongs to the first category, and it acts as primary building blocks for designing visualization. The Token pattern can be utilized to map network nodes from an information space onto a single unitized visual representation. This pattern is useful when a designer wants to represent an item, in our case, the router or the port. The simplest instance of

Token can be seen when a designer maps an item onto the atomic visual mark. The Area pattern can be utilized to map items from an information space onto visualizations such that their boundary, shape and region, and the Area pattern belongs to the second group, substrate. The Group pattern is used to organize information items in the representation space by congregating them close to each other. The nodes can be grouped in terms of their location, and the boundaries are displayed using the Voronoi algorithm. In order to show the network connection, the Link pattern can be used to connect nodes using lines. Both the Group and Link pattern are in the relational group. These patterns can be used across domains to help create novel visualizations.

These patterns could be blended in numerous ways, using different combinations. Different patterns can be blended to create sophisticated visualizations that are useful for showing different aspects and features of the data [30]. The pattern language employs a syntax to represent the blending of different patterns. The symbol “•” denotes a blending, where blended patterns appear together in square brackets “[]” and the symbol “ \in ” to denote that a visualization or representational structure “is derived from,” “is based on,” “instantiates,” or “is an instance of” a blending. The blending can promote flexibility and novelty and result in different instantiations as well.

2.2.3 Visualization Techniques

This stage involves either the creation of a novel visualization technique or the selection of an existing technique from the space of techniques. The patterns and blending decided in the last stage can inspire the designer to generate a new visualization or narrow down the space of design possibilities in this stage. The network presents the possibility of many-to-many connections, and many visualization techniques have been put forward for representing networks: adjacency matrix, arc diagrams, trees and force-directed network diagrams. In each case, the graphical representation will be very different. In this subsection, these techniques will be discussed.

In the adjacency matrix (Figure 6b), nodes will be encoded as rows and columns, and the connection between nodes will be presented as a grid. It usually uses colors or other highlighting techniques to emphasize the weight of each connection. Figure 6b utilizes the

adjacency matrix to show the co-appearance between the characters in Les Miserables. Though this layout seems to use the space very efficiently, it is hard to browse with extensive data and to show the relations. In arc diagrams (Figure 6a), nodes will be positioned along a line and draw the links as arcs above or below that line. It is easy to tell nodes that are isolated from nodes that have much connection. Arc diagram is suitable to visualize a small number of entities and relations, but when the number of objects and links increases, it is hard to read. The hive plots (Figure 6c) are arc diagrams arranged on radially distributed linear axes, and it is more suitable at managing the visual complexity arising from a large number of links. Both the adjacency matrix and arc diagram benefit from the linear manner and the control of the order and position of the nodes.

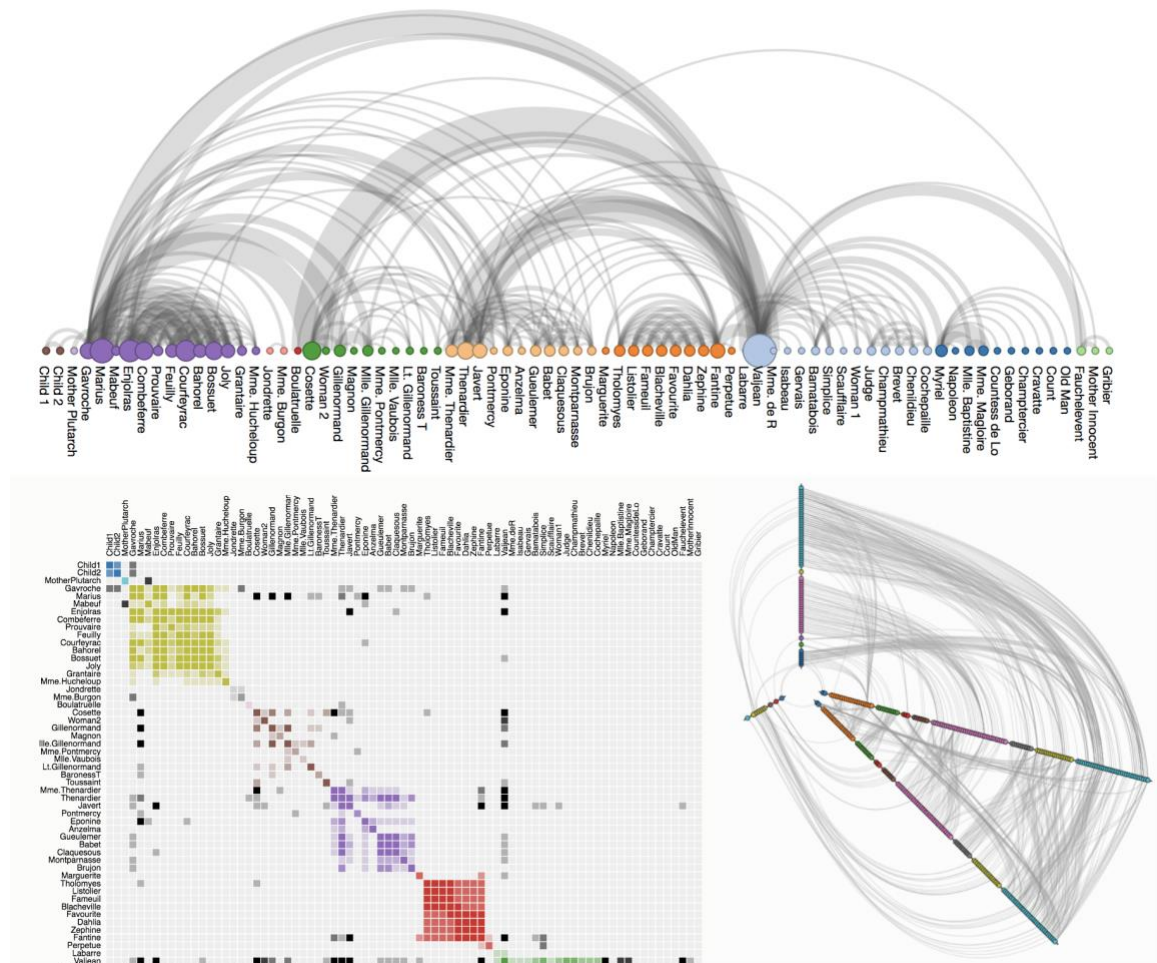


Figure 6: a- The arc diagram (up) [31]; b- The adjacency matrix layout (bottom left) [32]; c- The hive plots (bottom right) [33]

The layouts that arrange the networks can be combined with other techniques to better achieve the tasks. In our visualization tool, the force-directed layout is integrated with the Voronoi diagram (Figure 8). The Voronoi diagram partitions the space based on the minimal distance to each site. Let $P = p_1, \dots, p_n$ be a set of n distinct points in the plane, called sites, and $d(p, q)$ the Euclidean distance between points p and q . The first order Voronoi diagram of P is a subdivision of the plane in n cells, one for each site in P , such that a point q lies in the cell of a site p_i if and only if $d(q, p_i) < d(q, p_j)$ for all sites $p_j \in P$ with $j \neq i$ [35]. The cell boundaries lie on the perpendicular bisectors of the line segments $p_i p_j$ [35].

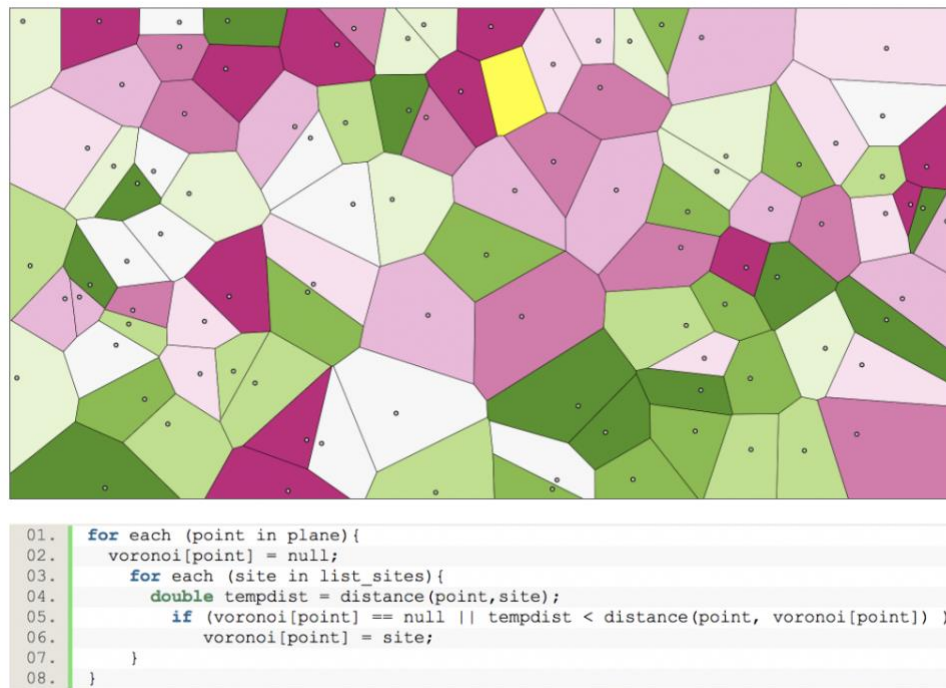


Figure 8: The Voronoi diagram [36] and The pseudo-code to implement the Voronoi diagram [27]

2.2.4 Concrete Encoding and Interaction

This stage is concerned with concrete encoding of the information items, and the designer needs to determine what specific visual structures and marks should be used and to select visual variables that will support the user's low-level visual tasks. In addition, for

interactive visualization, the designer needs to select low-level epistemic interactions to support the tasks and activities further.

Representation space is the space in which information items are encoded into perceptible visual forms. Representation space contains visualization not only of information items from an information space, but also of navigational elements, controls, and other information [26]. All complex visual representations can be broken down into basic visual marks. Visual marks are the objects of a mapping from an information space to a representation space and they are the primitive building blocks of visualization. The common visual marks are as follows: dots, dingbats, letters, digits, lines, circles, polygons, arrows, etc. Visual variables are a number of visual properties that can meaningfully encode information [26]. The common ones are as follows: planar dimension, color, size, angle/orientation, texture, shape, curvature, and motion. In most design situations, multiple variables are utilized together to encode information.

In the context of visualization design, making visualizations interactive is imperative for supporting complex activities. One way to effectively reduce the mental load is to make visualizations interactive. The interaction between a user and the visualization system can be categorized at multiple levels of granularity. Sedig et al. [27] categorize the interaction into four main levels: activities, tasks, actions and events. Activities always occur at a high level and are often open-ended and complex; tasks are goal-oriented behaviors that occur at a lower level during the performance of activities; interactions take place at the lower level and involve specific actions that are performed upon an interface; events occur at the lowest level of physical interaction with a visualization tool. How action is operationalized determines the quality of users' discourse with data and thus is an essential consideration for designers [37]. According to the framework developed by Sedig and Parsons [27] for interaction design, several action patterns were chosen to facilitate mental information-processing needs in this thesis. The following are the descriptions of some of the interaction patterns.

- The drilling action is utilized to bring out, make available and display interior, in-depth information.

- The filtering action displays a subset of their elements according to certain criteria.
- The navigating action enables moving on, through, and/or around visualizations.
- The searching action enables seeking out the existence of or location position of specific items, structures or relationships.
- The selecting action enables the user to focus on or choose the visualization, either as an individual or as a group.
- The transforming action changes the visualization's geometric form.

Similar to the visualization pattern, these interaction patterns can be recognized as the conceptual basics of interaction design in the context of complex cognitive activities. These patterns can be instantiated using different interaction techniques; for example, the drilling pattern can be instantiated using the following techniques: mouse-over, click, semantic zooming, gestures and spatial proximity.

Some interaction techniques that are selected to be utilized in this visualization system will be discussed as follows. Brushing is a change in the encoding of one or more items essentially immediately following and in response to an interaction with another item [38]. The intention of brushing is to highlight brushed items in different views of a visualization. The details-on-demand technique is also selected to facilitate the exploration of the network. The details-on-demand technique permits interactively selecting parts of data to be visualized more detailed while providing an overview of the whole informational concept. Semantic zooming is a form of details-on-demand that lets the user see different amounts of details in a view by zooming in and out. Semantic zooming is usually a combination of two epistemic actions: drilling and transforming. It could provide the user with the ability to act upon the visualization to bring more information to the surface and to manipulate the semantic and geometric properties of the visualization. Semantic zooming adjusts the contents and density of information that is shown instead of only changing visual detail and scale.

Chapter 3

3 Related Work

There is a growing need to efficiently visualize large network topologies in the communication network. Communication network management is required to visualize relationships between a large number of network elements, such as the routing relationships between the internet routers. In practice, network administrators need to monitor the network in real-time so that problems due to unusual traffic patterns, equipment failures or misconfiguration can be detected [39]. This raises the problem of how to visualize large network topologies in real-time and provide plenty of useful information for monitoring.

There has been a vast amount of work to design different visualization solutions to display the network, e.g. node-link representation [40]–[42], fisheye tree [43], [44], treemap [45]–[49], matrix display [50]–[52], circular layout [53], [54] and parallel coordinate [55]–[59]. It has been widely acknowledged that node and link displays are the most intuitive and effective layouts to visualize the small sparse network for a most purpose [60]. However, there are some problems when visualizing larger networks with the node-link layout. The displays are easily overwhelmed and become cluttered and visually confused when displaying too much information. Also, the running time and processing load of this layout are high. Besides, the interpretation of the layout is dependent on the node positioning, and the different node positioning algorithms lead to quite different interpretations of the data [61]. Many research works have been conducted in order to overcome the limitation of the node-link layout in a large-scale network.

3.1 Survey of past network traffic monitoring visualization tools

Information visualization has its beginnings as an area between the end of the '80s and the beginning of the '90s [3]. The significant evolution with the internet happened years later. The '90s were characterized by relevant contributions from the AT&T Bell Laboratories team. Many research projects have been conducted in the 1990s. SeeNet tool was developed by Becker et al. [62] with three complementary graphical techniques to display network data: link map (that represents network connections), node map (that displays

nodes by showing a glyph or symbol with characteristics such as size, shape, and color) and traffic matrix (Figure 9). Several interactive techniques were integrated into this tool, such as manipulating line length between links, symbol size and animation speed, zooming and brushing. The analyst manipulates the display parameters interactively while watching the display change. SeeNet tried to solve the information overload problem through parameter focusing and dynamic interaction. However, this system still cannot handle the large-scale network efficiently.

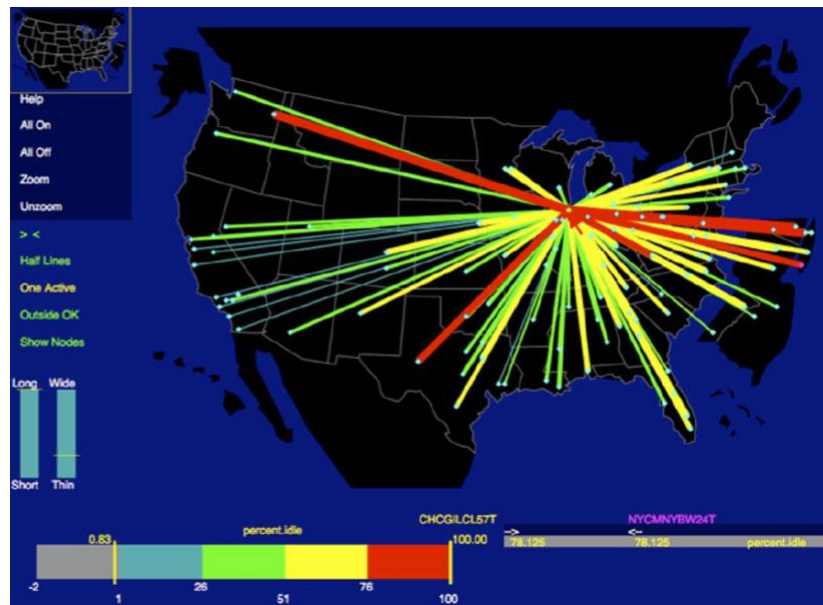


Figure 9: An overview of SeeNet visualization [62]

Some researchers also started to focus on how to develop large scale network visualization systems. An approach for large scale network visualization has been described by Abello et al. [63] from AT&T lab, and it was an experiment with a novel graphical representation of networks using a hierarchy of 3D surfaces that forms a hierarchical graph surface. The 3D layout can solve the link intersecting problem, and therefore, the display is perceived to be less cluttered. The difficulty with the 3D network display is that they are often tricky and confusing to navigate, which may cause the user to lose the sense of the overall context.

Some architectures and designs were introduced addressing real-time visualization in 1990s for further guidance. Parulkar et al. [64] introduced an architecture of a visualization system that is configurable and can support multiple simultaneous real-time displays of a

network. They presented a design for the visualization system called View Choreographer, which filtered the events to present and perform user-specified mappings of network events. A real-time system on Android terminals for visualizing network topology and packet flow in the network has been designed by Tsutsui et al. [65] to conduct efficient MANET field experiments. Nakai et al. [66] presented an approach to visualize networks in real-time with a bifocal display technique. Though the visualization is simultaneously capable of displaying both context and local details within a single window without the loss of information, it is still not suitable for ultra-large-scale network data.

3.2 Contemporary Network traffic monitoring visualization tool

In order to overcome the limitation of node-link layout, some recent research works continue to explore the layout of visualizing complex networks and try to use visualization techniques and some other techniques to increase efficiency in presenting computer network data.

Some systems have explored using multiple layers or multiple views to represent and visualize the network. Hofstede and Fioreze [67] created a monitoring tool SURFmap to provide traffic information at a geographical dimension by using the Google Maps API (Figure 10). Interactive zooming and details-on-demand techniques were applied in this tool. Though SURFmap added a geographical dimension into network data, it lost the other network information, like the network performance information, was not integrated into the representation in the visualization. It only showed limited information, and most of the information in the tool was described using text. In addition, though it provided four zoom levels including country, region, city, and host levels to arrange the network data, the context of data in higher hierarchical zoom level and the hierarchy information of the network will be lost. Wang [10] described a visualization prototype system for large-scale and complex networks. Organizing the large topology into multiple layers has been suggested. The system provides a visual representation of network topology, device information, connection information, and monitoring messages. Though the visualization system also provides a satellite view for users to find which part of the whole topology map is now in their view (Figure 11), the visualization will lose the context when switching

between different views. In addition, the system has not integrated with the network performance monitoring, and the visualization is not real-time. Though this visualization still does not perfectly visualize the large-scale network for monitoring purposes, it provides us with the idea to organize a large and complex topology into multiple layouts. The top layer only reveals the most crucial structure of the topology, and the lower layers contain more details of the network.



Figure 10: The screenshots of the DNS attack in country zoom level and host zoom level in SURFmap [67]

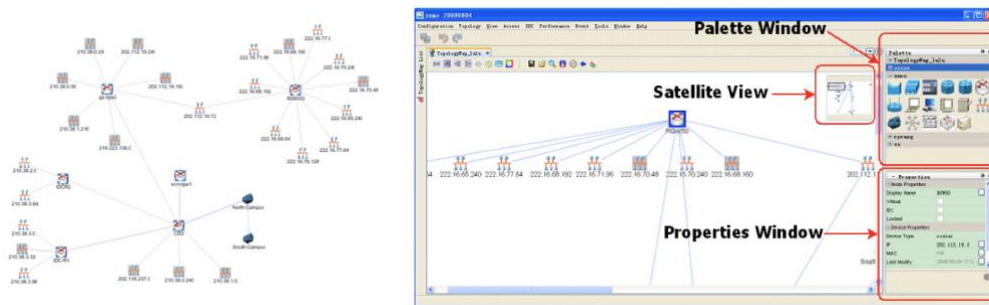


Figure 11: A top-layer overview of a complex network and a prototype GUI with Satellite View, Palette Window and Properties Window highlighted [10]

The 3D embeddings have also been utilized to solve the information-overloaded problem in node-link layout in recent works. Inoue et al. [68] proposed a 3D visualization system to display network topology, node information, link quality, and packet flow. For visualization of topology, GPS mode (Figure 12a) and manual mode were implemented,

and the latitude information of the node was provided in the 3D canvas. When dealing with large-scale data, it is not easy for administrators to navigate through the 3D layout. Besides, the interaction techniques are not utilized to display more information. A network visualization system MANET using 3-D space and time axes for determining failure points has been proposed by Tateishi et al. [5]. Instead of displaying geographical information and hierarchical information of topology separately, a visualization method with 3D axes was proposed. Geographical information was displayed on an X-Y plane while the hierarchical of network equipment was displayed on Z axis (Figure 12b). The viewpoint and direction could be freely selected in a 3D view. Though this visualization enables administrators to determine failure points and affected areas at a glance, when it involves the large-scale data, the visualization will still be confusing with human's cognitive workload. In addition, it does not visualize various types of data, like network performance information and router information, in a complicated way.

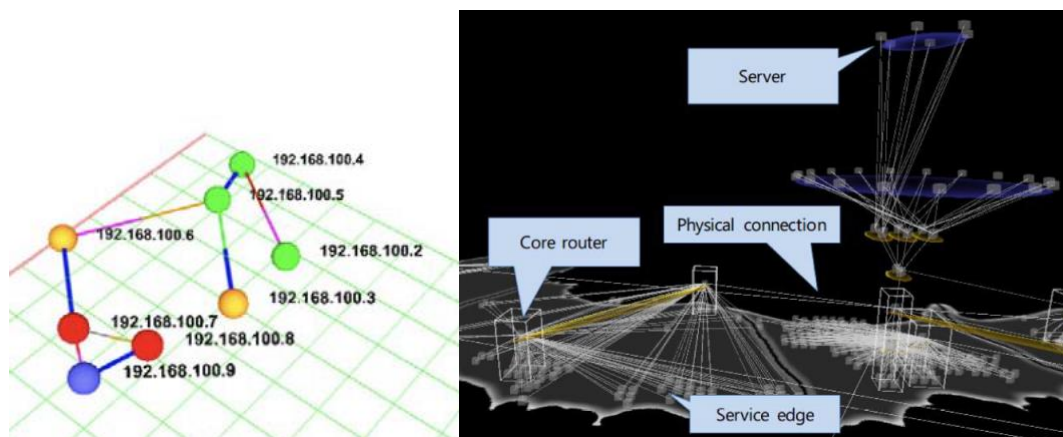


Figure 12: a- The GPS mode (3D) in MANET [68]; b- The example of displaying physical connections [5].

Some novel techniques, including the machine learning algorithms, have been explored to augment the domain knowledge of human investigators and therefore reduce the workload of the visualization tool. Liao et al.[6] developed a tool called ENAVis, which offers visual analysis features including connectivity graphs of combinations of hosts, users and applications. The visualization tool was implemented through the use of dynamic correlation and novel graph modeling. As it has been claimed by Liao et al. [6] that it is

vital to determine the set of variants and invariants of the network graphs, the maximum common subgraph and minimum common supergraph are visualized. The Walktrap algorithm is also selected as it takes the approach of random walk for graph cluster visualization, as represented in Figure 13. The system is also integrated with the graph mining process to maximize the knowledge/insight acquisition. All these algorithms and methods can guide the visual exploration process to only important things.

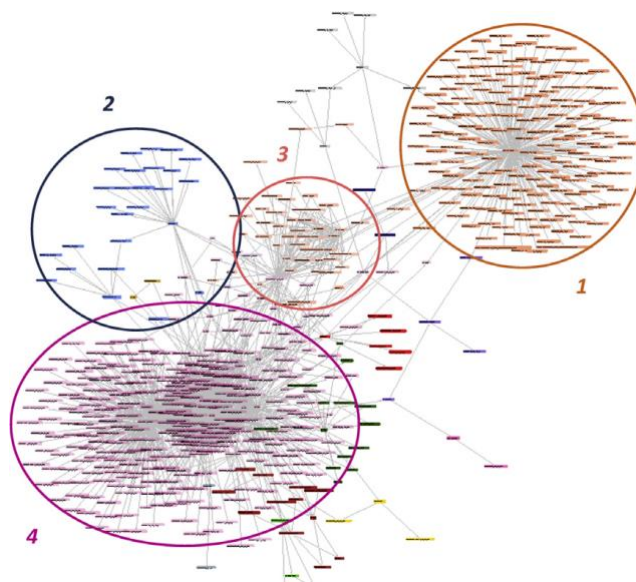


Figure 13: A cluster view (different colors represent different clusters) using the Walktrap algorithm. The cluster visualization helps in understanding the communities [6].

The above research tried to overcome the limitation of the node-link layout in a large-scale network with adopting parameter focusing, zooming techniques, 3D layout, or novel graph modeling. However, there still exist some problems with these methods. Some visualization systems may lose the context or be confusing to navigate. Others may not be able to contain enough details or visualize the large-scale network data efficiently. Therefore, these previous works have motivated us to explore how to better visualize the large-scale network data and also provide enough low-level details and thus enables the network administrator to plan and monitor the next-generation communication network efficiently. This thesis proposes a design of the real-time interactive visualization system that combines the idea of progressive disclosure with the semantic zooming techniques and

multiple panels to elegantly visualize the large-scale network according to its hierarchy and avoid the information-overloaded problem. In this way, the whole network will be partially visualized in different panels, and the users still do not lose the context during exploration. The system also visualizes the configuration and the inside structure of the network elements and provides the network performance information, including the port-level Quality of Service (QoS) metrics. The low-level information could benefit the network planners / network managers in conducting their network health analysis in a more effective way. In addition, the system enables navigation through the port-level connection and provides different modes for multiple purposes. More details will be further illustrated in **Chapter 4**.

Chapter 4

4 Methodology and System Design

4.1 System Structure Design

The system design configuration is illustrated in Figure 14. The client-server architecture has been adopted, and there are three layers in the system architecture (from left to right in Figure 14): the storage layer, back-end layer and the visualization layer (front-end layer). The information stored in the storage layer is discussed in detail in **Section 4.2**. The server is implemented to retrieve the data from the storage layer and continuously generate and publish new network data to the visualization layer. The users interact with the visualization tool at the visualization layer to explore and monitor the network. The users can also input some numbers and select the attributes to display the user-defined range of data elements. The visualization layer will send the query to the server, and after the server receiving the request, the server will filter the data according to the query range and then save the filtered data files for management purposes.

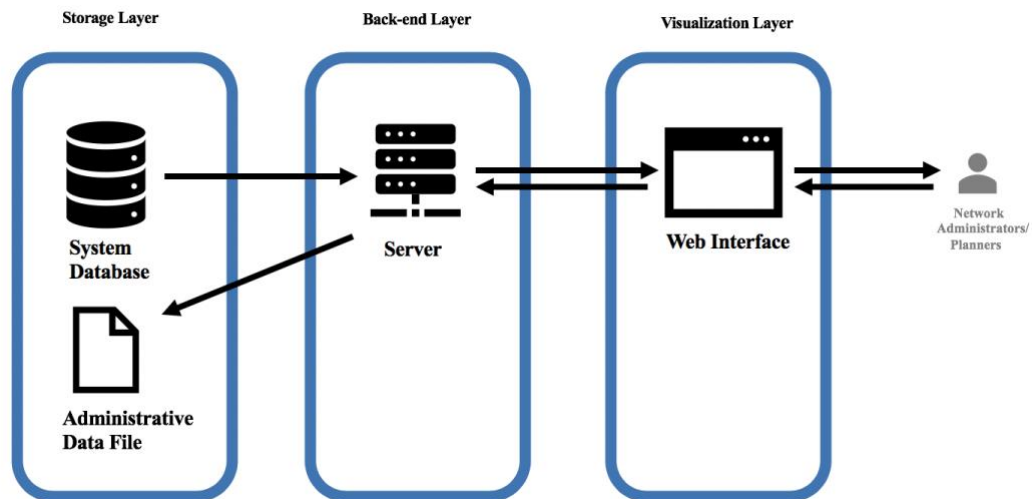


Figure 14: The system design configuration architecture

Our visualization tool is required to visualize the real-time network data, and the users can also interact with the tool and send queries to the server. Therefore, besides continuously sending the data to the visualization layer, continuous and full-duplex communication

between the server and the visualization layer is required. Due to this reason, WebSocket has been applied in this real-time visualization tool. WebSocket is a communications protocol that enables interactive, event-driven communication, facilitating real-time data transfer between client and server over a single TCP connection. The WebSocket protocol supports transferring messages between a web client and a web server at any time in any direction. It offers a standardized way for the server to send messages to the client without being first requested by the client, and allowing messages to be passed back and forth while keeping the connection open [69]. It is designed to be implemented in web browsers and web servers, but it can be used by any client or server applications. The server side runs on a 2.3 GHz Intel Core i5 macOS machine with 16 GB 2133 MHz LPDDR3 memory that can handle a large amount of processing in a quick time. It is implemented using Python and the Tornado library is utilized. Tornado is a python web framework and asynchronous networking library [70]. By using non-blocking network I/O, Tornado can scale to tens of thousands of open connections.

In order to provide easy access to network management, the network visualization system would be developed as an interactive web tool. The creation of a browser-based tool can be utilized as a generalized presentation layer for network management and planning applications. The client application in our system is an SVG enabled web browser that connects to the server and retrieves information from the server. Scalable Vector Graphics (SVG) is an Extensible Markup Language (XML)-based vector image format for 2D graphics with support for interactivity and animation [71]. The visualization layer is built with D3.js and Bootstrap. D3.js is an immensely popular JavaScript library that provides useful building blocks for making data visualizations [72]. It facilitates mapping the processed data into visual representations. D3 allows users to bind arbitrary data to a Document Object Model (DOM) and then apply data-driven transformations to the elements. The Bootstrap library can help to style the elements of the page consistently and appealingly. Bootstrap is an open-source toolkit for developing with HTML, CSS, and JS. It contains CSS and JavaScript-based design templates for forms, navigation, buttons and other interface components.

4.2 Network Data

Instead of collecting, structuring and transferring data, simulated network data is applied to the visualization system at this stage. The simulated structured data is generated through the following process: selecting a structure of the data; using the random number generator to generate a sample from the assumed structure; formatting the simulated data. There are mainly two collections of data to be visualized in this project: the network topology information and the router detail information. The network topology information includes the link configuration that displays the communication in the network and the link utilization information. The router detail information stored the router configuration information including the condition of power, chassis, processor, etc., and also the slot information and port information. The router detail data also includes the utilization information of each router and the QoS metrics of each port. The detail of the structure and the data range for each attribute will be described as follows.

This project tends to simulate the topology of routers' connection in the real-life situation. The percentages of the quantity of each type of routers are structured to simulate the real-life situation, and the structure and number of the connections between different hierarchical routers are also designed according to the industry standards. The whole topology of the network is a partial mesh, which means routers are much more tightly coupled than in any of the basic topologies but are not fully interconnected. A partial mesh can reduce the startup and operational expenses by not interconnecting low-traffic segments. Therefore, the partial mesh is more scalable and affordable than a full mesh topology.

We referred to some databases [73]–[75] that provides router utilization and QoS metrics data, and explored the information of the routers that are in daily use in industrial application. According to these data examples, we settled some standards for each type of routers and defined the range for the slot utilization, port utilization and port bandwidth. They are illustrated in Table 1, Table 2 and Table 3. To simulate the QoS performance data of the ports, the range for each QoS metric is defined. The range for the delay is 1 to 300 milliseconds (ms) and the range for jitter is from 0 to 300 ms. The packet loss is within 0 to 25 percent. The range for reliability is from 0 to 100 percent. In addition, the range for

the speed of the port is according to the line card's type where the port located. The data range is determined for generating the simulated test data to examine the visualization tool. It is only used for facilitating the simulated network data generating process and the real-world data could be within different data ranges. The data is generated using the Random module in Python and the random integers are from the discrete uniform distribution. The discrete uniform distribution with parameters (a, b) constructs a random variable with an equal probability of being any one of the integers in the half-open range [a, b]. If a is not given, it is assumed to be zero and the only parameter is b. With the simulated network data, we could operate our proposed visualization tool experimentally and confirm the effectiveness of our network visualization system.

The data is represented using JavaScript Object Notation (JSON). It can be noticed that in Figure 15, in the router detail collection, the port data is embedded into the slot object, and the slot object is embedded into the router node. The delay, packet loss, reliability, jitter and bandwidth information of each port are stored. Another example in Figure 15 is the data in the network topology collection, and the data includes the source and target information and the link utilization.

Router Type: Access	Config	Card Type	Total Ports	Slot Utilization (%)	Port Utilization (%)	Port Bandwidth (%)
Chassis	Standard					
Memory	Standard				30 - 90	
Processor	Dual				50 - 90	
Power	Dual					
Customer Facing Slots	15			50-90		
		10Mbps	10			
		100Mbps	10			
		GE	20			
		10GE	20			
Edge Facing Slots	5			70-90		
		GE	20		50 - 80	50-90
		10GE	20		50 - 80	50-90

Table 1: The data range for the simulated configuration information of the access router

Router Type: Edge	Config	Card Type	Total Ports	Slot Utilization (%)	Port Utilization (%)	Port Bandwidth (%)
Chassis	Standard					
Memory	Standard				30 - 90	
Processor	Dual				50 - 90	
Power	Dual					
Access Facing Slots	10			50-90		
		10Mbps	10		50 - 80	50-90
		100Mbps	10		50 - 80	50-90
		GE	20		50 - 80	50-90
		10GE	20		50 - 80	30-90
		40GE	10		50 - 80	30-90
Core Facing Slots	10			50-90		
		GE	10		50-80	50-90
		10GE	10		50-80	30-90
		40GE	6		30 - 80	30-90
		100GE	4		30 - 80	30-90

Table 2: The data range for the simulated configuration information of the edge router

Router Type: Core	Config	Card Type	Total Ports	Slot Utilization (%)	Port Utilization (%)	Port Bandwidth (%)
Chassis	Standard					
Memory	Standard				30 - 90	
Processor	Dual				50 - 90	
Power	Dual					
Edge Facing Slots	10			50-90		
		GE	10		50-80	50-90
		10GE	10		50-80	50-90
		40GE	6		30 - 80	50-90
		100GE	4		30 - 80	30-90
Border Facing Slots	8			50-90		
		40GE	6		30 - 80	50-90
		100GE	4		30 - 80	30-90

Table 3: The data range for the simulated configuration information of the core router

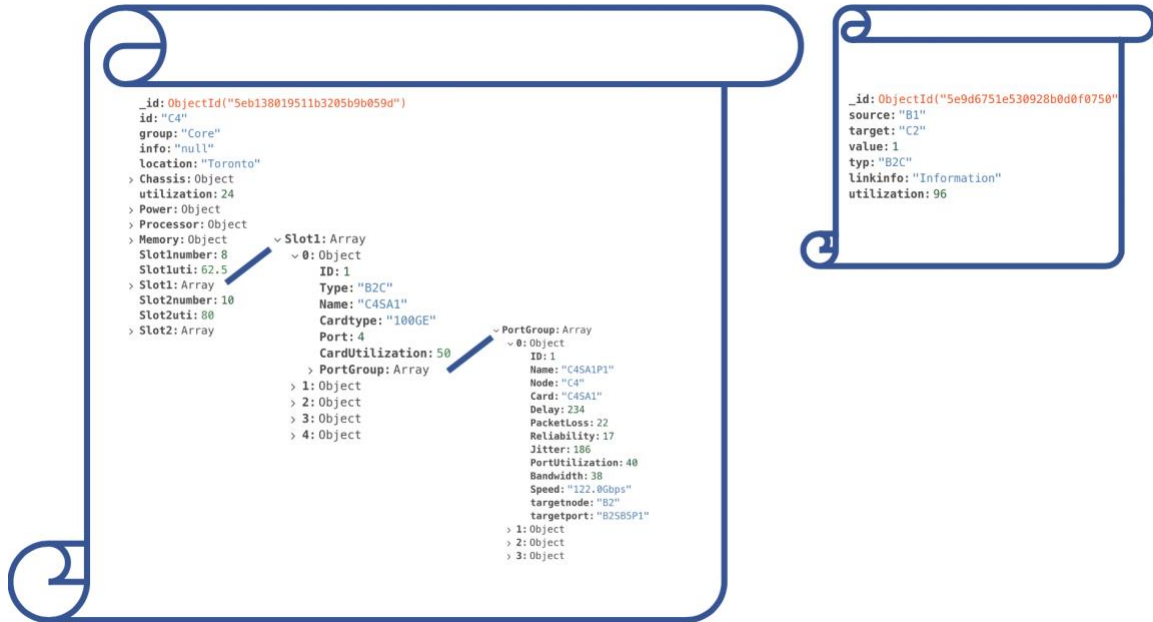


Figure 15: The network data example

When the network administrators and planners interacting with the visualization tool, they can input some numbers and select the attributes to display the user-defined range of data elements. The visualization layer will send the query to the server, and the server will filter the data according to the query range and then save the filtered data files as the QoS report for management and planning purposes. The filtered data are saved as the excel files, and an example of the part of the report is shown in Figure 16. The report represented in Figure 16 is generated to list the ports whose delay is higher than 280 milliseconds. The port information which includes the port ID, the ID of the router the port located in, the line card ID, the QoS information of the port, and the ID of the target router and target port is recorded in the report to support the management and future planning. The ID of the port is defined by combining the id of the line card and its position in the line card. The ID of the line card is composed of the ID of the router and its position in the line card. The 'SA' means the line card slots are dealing with the incoming packets and the 'SB' means the line cards perform the output function.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1		Name	Node	Card	brtUtilizati	Delay	PacketLoss	Reliability	Jitter	Bandwidth	Speed	targetnode	targetport
2	0	C1SA6P4	C1	C1SA6	57	291	18	34	159	65	168.0Gbps	B1	B1SB1P1
3	1	C1SA7P3	C1	C1SA7	55	298	7	13	224	72	156.0Gbps	B1	B1SB5P3
4	2	C2SA6P4	C2	C2SA6	67	300	23	91	166	36	149.0Gbps	B1	B1SB4P2
5	3	C4SA2P3	C4	C4SA2	53	286	20	23	54	44	161.0Gbps	B2	B2SB2P2
6	4	C4SA3P1	C4	C4SA3	79	300	17	23	0	59	43.0Gbps	B2	B2SB5P5
7	5	C5SA1P1	C5	C5SA1	71	295	15	77	74	51	42.0Gbps	B3	B3SB1P1
8	6	C6SA3P3	C6	C6SA3	67	291	4	6	141	39	106.0Gbps	B3	B3SB1P3
9	7	C9SA2P3	C9	C9SA2	69	289	19	36	29	84	42.0Gbps	B5	B5SB3P1
10	8	C10SA2P1	C10	C10SA2	61	298	16	78	33	80	40.0Gbps	B5	B5SB1P1
11	9	C10SA3P1	C10	C10SA3	45	286	4	14	159	87	43.0Gbps	B5	B5SB1P1
12	10	C10SA7P3	C10	C10SA7	62	287	22	89	121	57	40.0Gbps	B5	B5SB1P4
13	11	C11SA2P4	C11	C11SA2	63	299	25	16	273	71	45.0Gbps	B6	B6SB3P1
14	12	C11SA2P5	C11	C11SA2	37	287	22	4	160	90	47.0Gbps	B6	B6SB1P5
15	13	C11SA3P3	C11	C11SA3	74	299	3	51	248	81	41.0Gbps	B6	B6SB1P3
16	14	C12SA6P5	C12	C12SA6	57	281	22	51	277	53	40.0Gbps	B6	B6SB5P1
17	15	C14SA6P2	C14	C14SA6	51	300	15	13	133	79	46.0Gbps	B7	B7SB7P6
18	16	C15SA6P5	C15	C15SA6	75	299	16	100	281	56	42.0Gbps	B9	B9SB7P1
19	17	C15SA7P1	C15	C15SA7	67	290	6	47	80	32	140.0Gbps	B9	B9SB4P1
20	18	C16SA3P4	C16	C16SA3	75	299	5	58	197	69	46.0Gbps	B9	B9SB1P1
21	19	C16SA6P2	C16	C16SA6	63	298	23	59	104	61	151.0Gbps	B9	B9SB4P3
22	20	C16SA6P3	C16	C16SA6	39	283	10	72	292	64	196.0Gbps	B9	B9SB4P4
23	21	C17SA2P6	C17	C17SA2	77	294	13	75	36	50	41.0Gbps	B10	B10SB2P3
24	22	C17SA6P2	C17	C17SA6	52	292	14	44	101	38	151.0Gbps	B10	B10SB5P4
25	23	C18SA1P3	C18	C18SA1	49	296	23	8	9	53	41.0Gbps	B11	B11SB3P1
26	24	C19SA2P1	C19	C19SA2	33	291	23	67	59	57	172.0Gbps	B12	B12SB3P2
27	25	C19SA6P1	C19	C19SA6	52	296	1	82	54	86	178.0Gbps	B12	B12SB5P2
28	26	C20SA1P3	C20	C20SA1	67	288	17	42	85	57	166.0Gbps	B8	B8SB4P3

Figure 16: The generated filtered data file as the QoS report

4.3 Visualization Tool Design

4.3.1 Method

When visualizing large scale network data for the monitoring and planning purposes, the complexity of network architecture and large-scale data are the major issues. Analyzing the overall network properties might be too coarse to be useful for network management [76]. Identifying and locating faults naturally becomes more difficult with large and more diverse networks. Presenting the complex network data into simple visualization is a significant challenge. Instead of using parameter focusing and 3D embeddings to solve the information overload problem in large-scale network visualization, the visualization in this thesis is designed to combine the idea of progressive disclosure and using multiple panels to cleverly position nodes to show the network structure.

Progressive disclosure is an interaction design technique to help maintain the focus of a user's attention by reducing the cognitive workload. It improves usability by presenting only the minimum data required for the task at hand. Progressive disclosure can limit what

is shown on a screen and only show users what they need when they need it. A hybrid to progressive disclosure called staged disclosure had been introduced by Nielsen [77], characterized by the back-next(wizard) interaction technique. It is a step-by-step process that allows users to input information in a prescribed order and in which subsequent steps may depend on information entered in the previous one. However, the wizard method is not gracefully interruptible, and if the users quit the process midway, they might not only lose their work but also may need to click again through the preceding steps. Even though sometimes the wizard support saving of the state and resuming later, it is not easy for users to recover the context and the mental model of the process. Therefore, our visualization is designed with multiple panels to show each state of the visualization. In this way, the visualization tool provides the context and more information than considering each component independently.

In order to conquer the problem of the overload information problem, the design of our visualization divides the network according to its hierarchy and visualize it separately in different panels. Combined with the interaction techniques, the users can interact with the tool and only explore the part of the network that they are interested in or the elements where some anomaly activities may exist. Clicking the selected node and its lower level linked nodes will appear in the next panel. In this way, the large-scale data can be divided and only be shown separately with some interaction. If every panel could support to visualize N nodes in a delectable layout and manner, and the visualization has q panels, the order of magnitude of the data the tool can support is N^q . Thus, large-scale data can be managed to be visualized elegantly and also not lose the context. Details of the design of the visualization tool will be described in the following sections.

4.3.2 Overview

The whole visualization can be divided into two types of panels: the network panel (A1, A2 and A3 in Figure 17) and the inside panel (B3 in Figure 17 and B1 and B2 are hidden). The network panels are arranged in the upper position, and the inside panels are in the lower position. There are both three network panels and three inside panels. As shown in Figure 17, the A1 panel visualizes the whole border routers, core routers and edge routers in the network. The A2 panel only visualizes the selected edge routers and its linked access

routers. The A3 panel visualizes the selected access routers and its linked users. The overall network information and condition are visualized in the network panel. Each inside panel visualizes the inside router configuration information and the port-level QoS metrics of the selected router. Some buttons, drop-down menus and filtering form are at the top of the tool. The legend which indicates the color encoding of this visualization tool is placed on the bottom of the page for better supporting users' exploration. More details of these two types of panels will be illustrated and discussed in **Section 4.3.3** and **Section 4.3.4**.

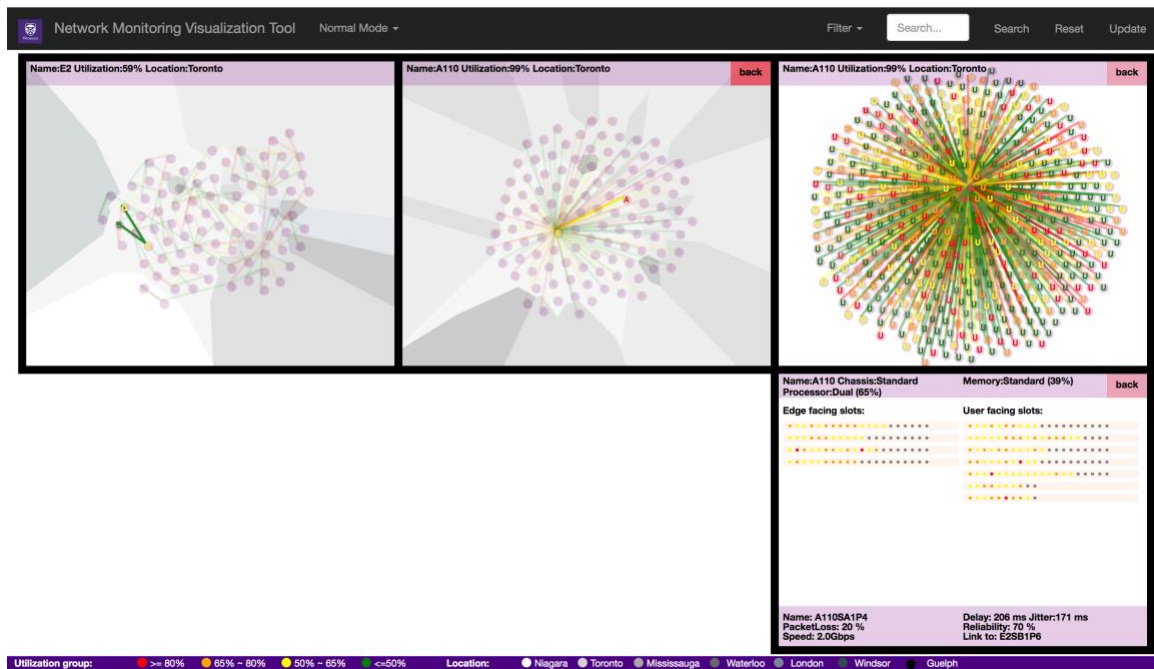


Figure 17: The overview of the visualization tool

4.3.3 The Network Panel

The elements of visualization in the network panels are as follows. The glyph is made up of a circle and a letter to represent each router. The letter comes from the initials of the type of routers. The router type that each letter stands for are as follows: Border routers are represented as B; Core routers are represented as C; Edge routers are represented as E; Access routers are represented as A and User routers are represented as U. The nodes are connected with links and links are represented as lines in our visualization. The color of the nodes or links indicates the current utilization of the nodes or the links, and they are with the same color encoding. The network administrators/planners only concern about the

range that the utilization falls to and not the specific value. Therefore, in our visualization tool, the quantitative utilization data are categorized into different levels, and the different utilization groups are encoded with different colors. Due to the monitoring purpose, the green color indicates the ‘healthy’ nodes or lines of which the utilization is below 50%, and the red color alerts the users that the nodes or links are under a heavy workload so that some actions may need to be adopted to fix this issue. The specific color encoding design for the nodes and links are as follows: the green color shows that the utilization is below 50%; the yellow color shows that the utilization of this node or link is larger or equal to 50% and smaller than 65%; the orange color shows that the utilization of this node or link is larger or equal to 65% and smaller than 80%; the red color shows that the utilization of this node or link is larger or equal to 80%. Such color encoding could aid network administrators and planners to easily identify where the utilization achieves the high level and consequently worsens the network performance. The legend which indicates the color encoding of the utilization group is also provided in the visualization tool to aid in monitoring and planning. The three network panels are demonstrated in Figure 18.

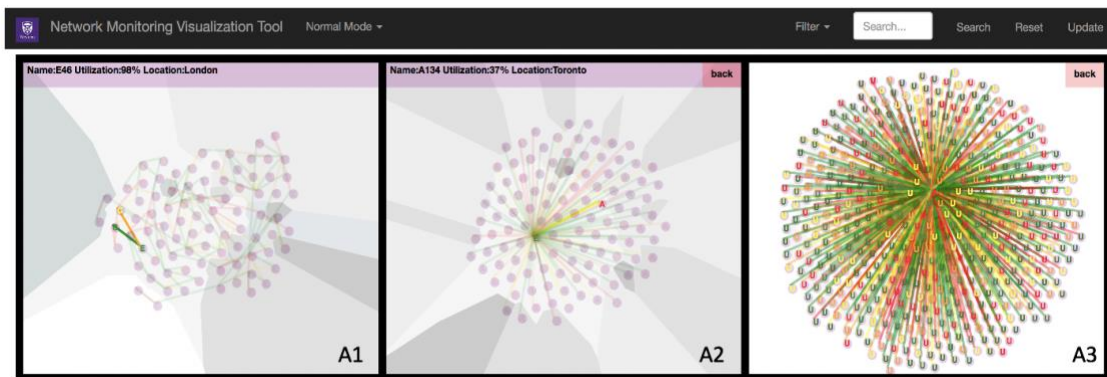


Figure 18: The network panels

Instead of using other layouts, for example, matrices or tree layout, the force-directed graph is selected to display the layout since it is more direct and intuitive. The nodes and links are organized in the force-directed graph layout using position Verlet integration to allow simple constraints. Verlet integration is a numerical method used to integrate Newton’s equations of motion, and it is generally utilized to calculate trajectories of particles in

molecular dynamics simulations [78]. This implementation also utilizes a quadtree to accelerate charge interaction using the Barnes–Hut approximation [79]. Both the repulsive charge force and a pseudo-gravity force are applied among the nodes. The pseudo-gravity force keeps nodes centered in the visible area and avoids expulsion of disconnected subgraphs, while links are fixed-distance geometric constraints. In addition, the force is customized to place the nodes that are in the same location together in the visualization and set apart nodes that are geographically away.

The background areas of the nodes indicate their location. The Voronoi diagram is utilized to partition the plane into regions close to each of a given set of objects. Visualizing the diagram is straightforwardly done by drawing the set of disjoint, adjacent planar polygons that represent the diagram [80]. The different location areas are represented using different levels of grey color. As shown in Figure 17, the legend that indicates the color encoding of the nodes' location is displayed at the bottom right of the page to directly support the exploration of the visualization. In this way, it is intuitive to determine both the hierarchical layer and the geographical position of the routers.

4.3.4 The Inside Panel

The inside panel visualizes the inside information of the router and it is demonstrated in Figure 19. The rectangles represent the line cards, and as the inside structure of routers, the line cards are divided into two slots. Each slot faces different levels of routers. There are ports on the line cards in the routers. As the hierarchy of the routers, border routers' ports only connect with core routers, and the user routers only connect with the access routers. Other routers each have two slots and the ports in these two slots link to an upper and lower hierarchical level of routers respectively. In our visualization, the ports are represented as circles and they are placed on the rectangles that represent line cards. The color of the ports also encodes the utilization of the port and they have the same encoding with the nodes and links' utilization. The grey ports depict that the ports are not in utilization. The concrete information of the router has been shown on the top of the panel which includes the name, the chassis type, the memory type, the memory's utilization, the processor type and processors' utilization.

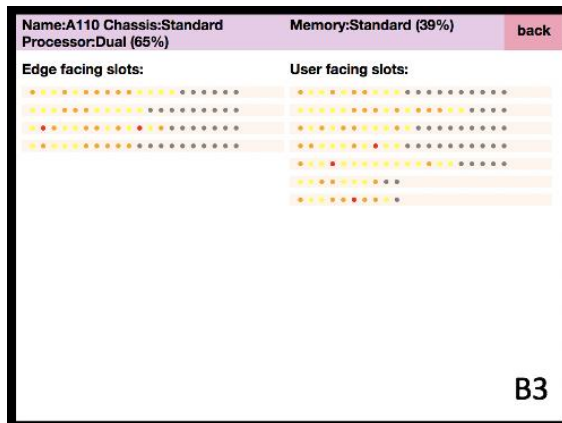


Figure 19: The B3 inside panel

4.3.5 Interaction

It is widely recognized that a large part of the usefulness of a network management tool lies in its interactivity. The element of interactivity is basically defined as the ability to respond to actions and dynamically change graphical representations. Interactivity is the very element that generally makes the difference between a real usable application and a mere depiction [60]. The design and development difficulty is that the implementation of interactivity incorporates many unnecessary details and is widely cumbersome. Designing suitable highly interactive visual tools will contribute to the network management. In our visualization tool, the dynamic interaction is very vital and can help to solve the large-scale information-overloaded problem and benefit the exploration of network data and the planning and monitoring procedures.

After clicking the update button on the upper right corner of the page in the navigation bar, the simulated data sent from the server will be visualized and displayed in the web browser. The A1 panel which contains all the main routers in the network including border routers, core routers and edge routers will be displayed in the original view of the visualization tool. As demonstrated in Figure 20, when hovering the mouse on a particular node in the network panels, the overall information of this node including the node name, exact utilization number and location information will appear in the purple box on the top of the panel. When the mouse leaves the node, the information will still be shown, and it will only be changed when the mouse is hovering other nodes. When the mouse is hovering the link,

the link information which contains the link source, the link target, and link utilization appears in the lower line in the purple box. When the mouse leaves the link, the link information will be disappeared.

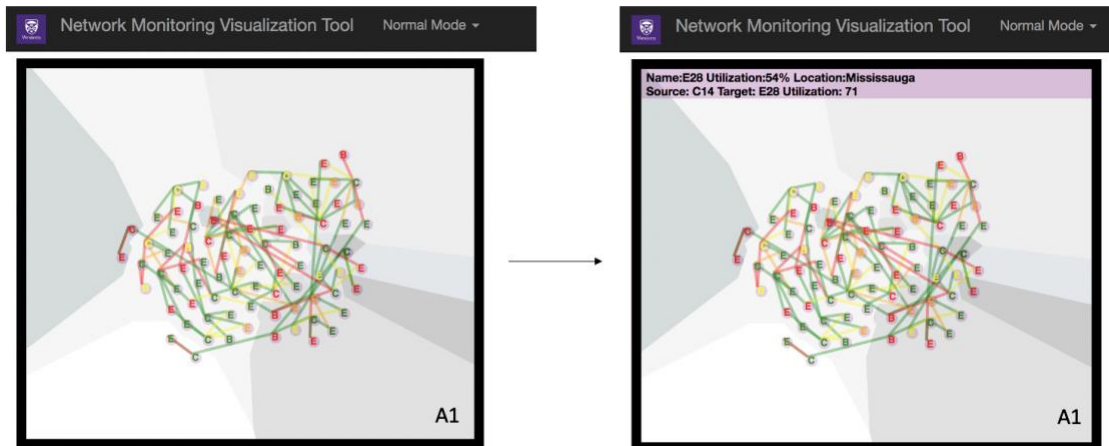


Figure 20: Hovering the nodes or the links

In the A1 panel, when clicking the border router or the core routers, the typology of how the selected nodes are connected to other nodes comes into view. As shown in Figure 21, the symbol of the node and its connected nodes are still displayed, while other nodes are blurred and their symbols are hidden. Only the circles of other nodes are shown to indicate their existence and positions. Only the links that link the clicked node to its connected nodes are vividly shown in the visualization, and the opacity of other links is decreased and turned to gloomy. Thus, the way how the nodes are connected are visualized clearly but also not lose the context. The users can easily catch the connection information of the node they are interested in while still keep the idea of how other nodes are connected. At the same time, the B1 panel will come into view. The inside structure of the selected nodes will be visualized. Users can explore the inside information of the nodes. When clicking the back button on the upper right side of the B1 panel, the B1 panel will disappear, and the A1 panel will return to its original state. No node is selected, and the whole topology of the border routers, core routers and edge routers are displayed totally.

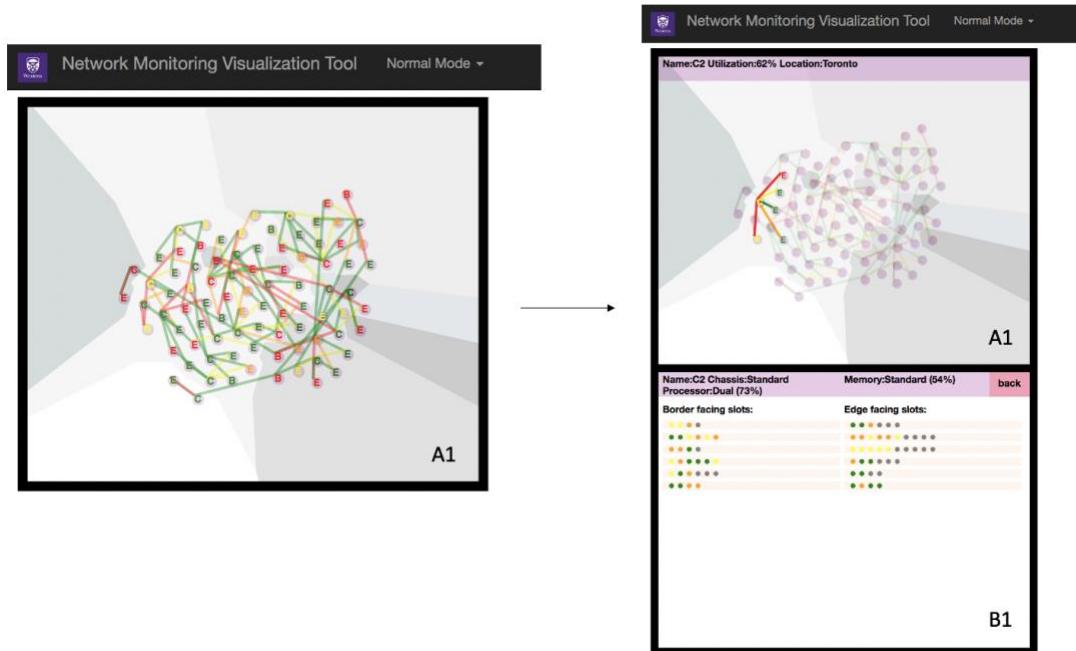


Figure 21: Clicking one of the core routers in the A1 panel

In the A1 panel, if clicking the edge router, the topology of how the edge router is connected to the core router also be highlighted (Figure 22). Meanwhile, the A2 panel comes into view. When clicking the back button on the upper left side of the A2 panel, the A2 panel will disappear, and the A1 panel will return to its original state. No node is selected, and the whole topology of the border routers, core routers and edge routers are displayed totally in the A1 panel.

The selected edge routers and its linked access routers are visualized in the A2 panel. When clicking the edge router in the A2 panel, the B2 panel which indicates the inside structure of the edge router is shown. When clicking the back button on the upper right side of the B2 panel, the B2 panel will disappear. When clicking the access router in the A2 panel, the topology of how the access router is connected to the edge router is highlighted. Only the symbols of the selected access node and its linked edge node are shown in the A2 view, and the symbols of other nodes are hidden. Meanwhile, the A3 panel which visualizes the topology of selected access nodes and its linked user nodes comes into view. The interaction process has been represented in Figure 22.

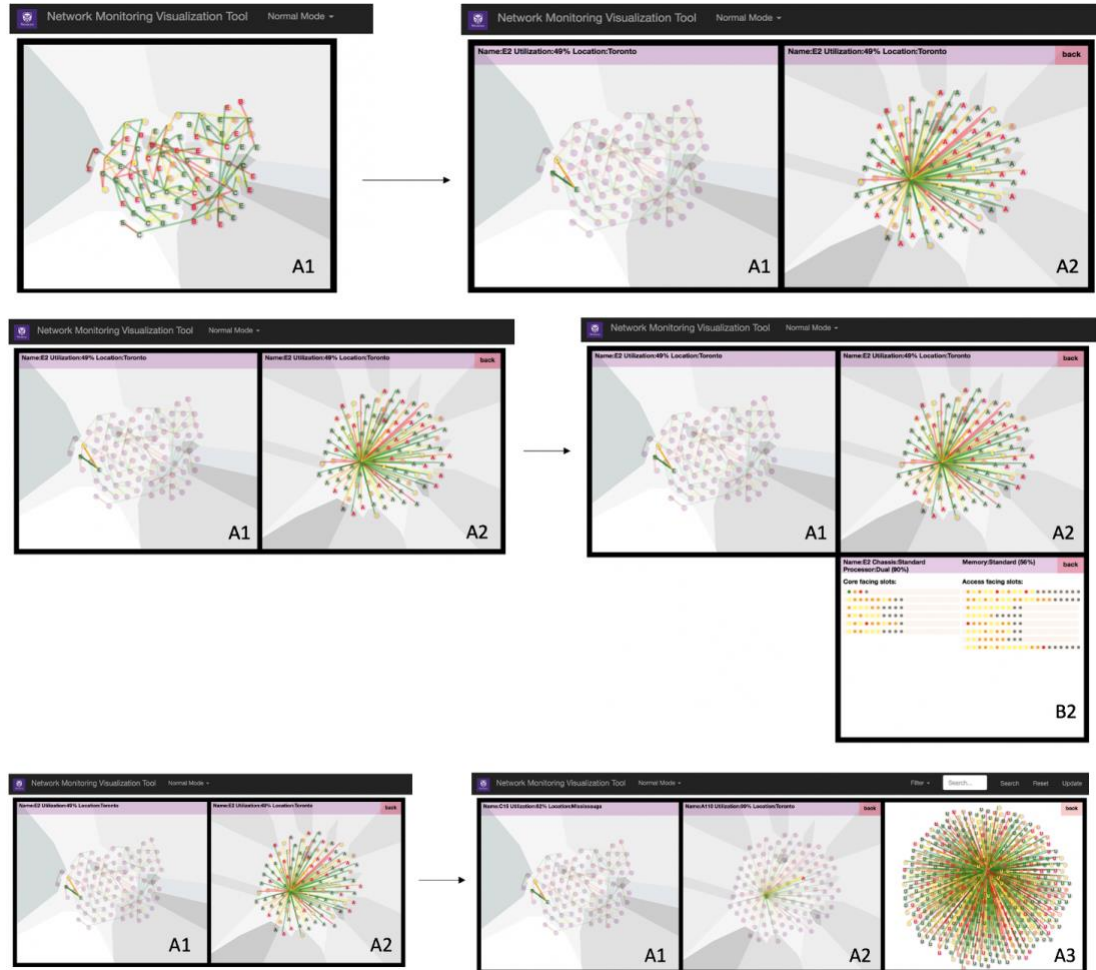


Figure 22: a- Clicking the edge router in the A1 panel (Up); b- Clicking the edge router in the A2 panel (Middle); c- Clicking the access router in the A2 panel (Down)

When clicking the access node in the A3 view, the B3 panel will be shown. If the back button on the B3 panel is pressed, the B3 panel will be folded. If clicking the user node in the A3 view, the user and its linked access node will be highlighted in the A3 panel, and the B3 panel will be shown. When pressing the back button on the B3 panel, the B3 panel will be folded, and the A3 will return to its original state.

The overall interaction pattern of the node clicking action in the network panel can be summarized as follows. When clicking the node in the network panels, the typology of how the selected node is connected to other nodes comes into view. The symbol of the node and

its connected nodes are still displayed, while other nodes are blurred and their symbols are hidden. Only the circles of other nodes are shown to indicate their existence and positions. Only the links that link the clicked node to its connected nodes are vividly shown in the visualization, and the opacity of other links is decreased and turned to gloomy. At the same time, when the node being clicked is in its last panel, the inside panel of this node will be displayed. If the node being clicked is not in its last panel, the next panel will be shown, and both the clicked node and the lower hierarchical level of nodes that are linked to the clicked node will be displayed in the next panel. In the visualization tool, the border routers, core routers and edge routers are visualized in the A1 panel. The selected edge routers and its linked access routers are shown in the A2 panel. The A3 panel displays the selected access router and its connected user routers.

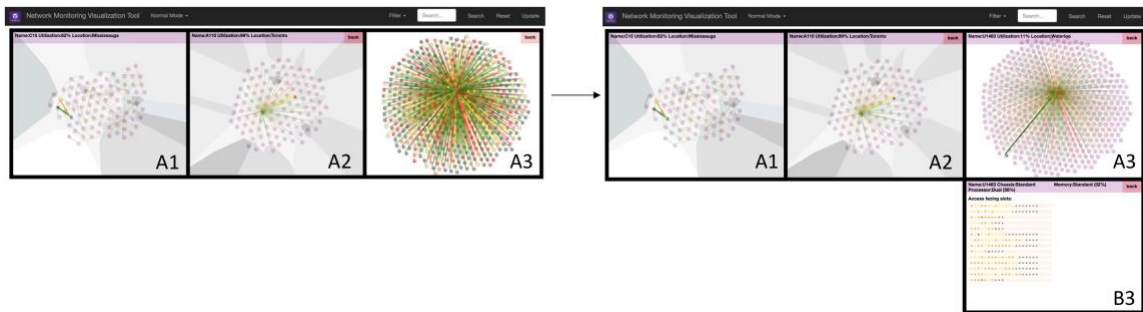


Figure 23: Clicking the user in the A3 panel

In the inside (B1, B2 and B3) panel, when the mouse over the ports, the port name, the port it links to, the speed and the QoS information will be displayed as text in the purple box at the bottom of the panel (as represented in Figure 24). The QoS information, including the delay, jitter, packet loss, reliability, will be displayed.

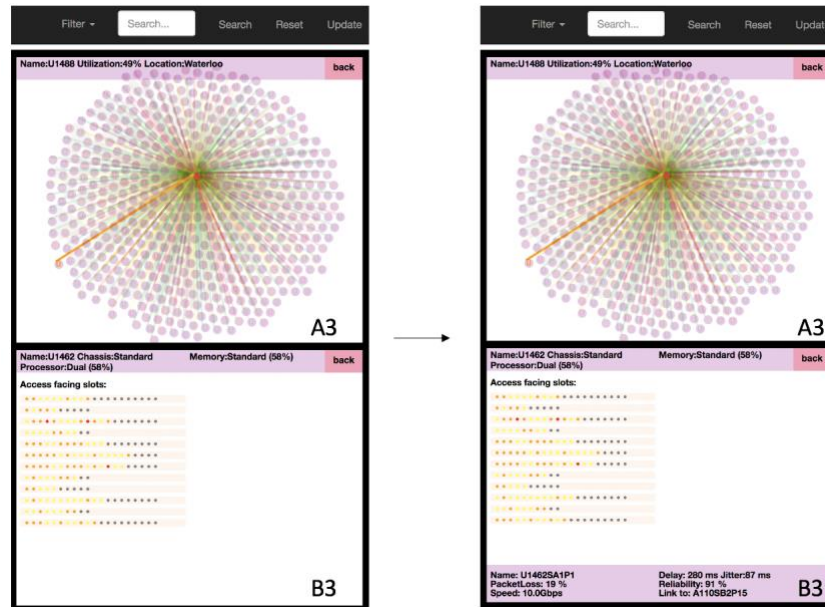


Figure 24: When mouse hovering on the port in the B3 panel

When clicking any of the port, it will turn to show the information of its linked port and the inside information of the router the linked port is in. The inside panel of the current node will be folded, and the inside panel of the node which the linked port is in will be displayed. The visualization will also navigate to show the information of the router where the linked port is located, and the network panel will turn to display the connection information of this node. It will only show the information of the current hierarchy of the node where the linked port locates and for example, if the node that the port links to is in the A2 panel, the A3 panel will be disappeared. Also, the node and its linked nodes will be highlighted. It will have the same display as if the node is clicked. As in Figure 25, it can be seen that the ID of the port being hovered is A110SA4P1, and it lies in the line cards in the edge facing slot. This port links to the port E2SB6P7, and the port E2SB6P7 belongs to the edge router E2. When clicking the port A110SA4P1, the visualization will navigate to show the information of the router E2. The A3 and B3 panel will be hidden. The B2 panel will be displayed to show all the inside information of the edge router E2 and also the information of the port E2SB6P7 including its QoS metrics.

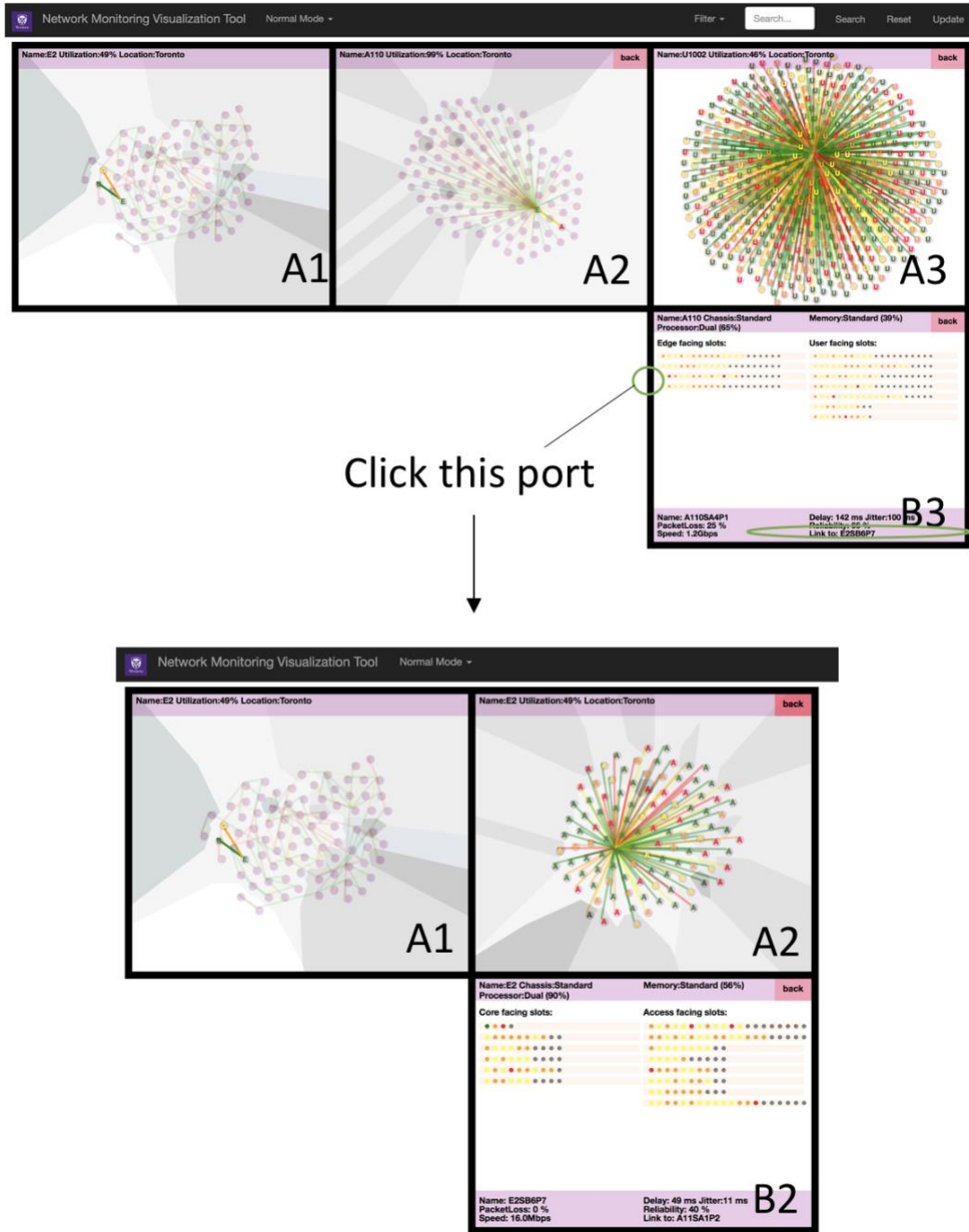


Figure 25: When clicking the port, it will navigate to show the inside data of the router where the linked port located and the information of the linked port.

4.3.6 Filtering Function and Alert Mode

When the user tends to filter the nodes or the ports according to node utilization, port utilization, port delay, or port packet loss, they could type the number in the form and select

the attributes on the drop-down menu. The tool will only show the nodes or ports whose selected attribute is higher than the number. After filtering the nodes, the users can also select the port attribute on the drop-down menu and input the number. The visualization tool can display both part of the nodes and part of the ports whose utilization, delay or packet loss is higher than the input number. The high utilization, delay, or packet loss indicates that there may be some anomalous changes or activities, and it will lead to degraded services and utilization. Therefore, showing the routers or the ports that may have potential problems is beneficial for troubleshooting and performance monitoring in network management and network planning. The views after filtering the routers according to their utilizations are demonstrated in Figure 26, and Figure 27 illustrates the change in the views when filtering the ports according to the port delay.



Figure 26: Filtering the routers according to the router utilization

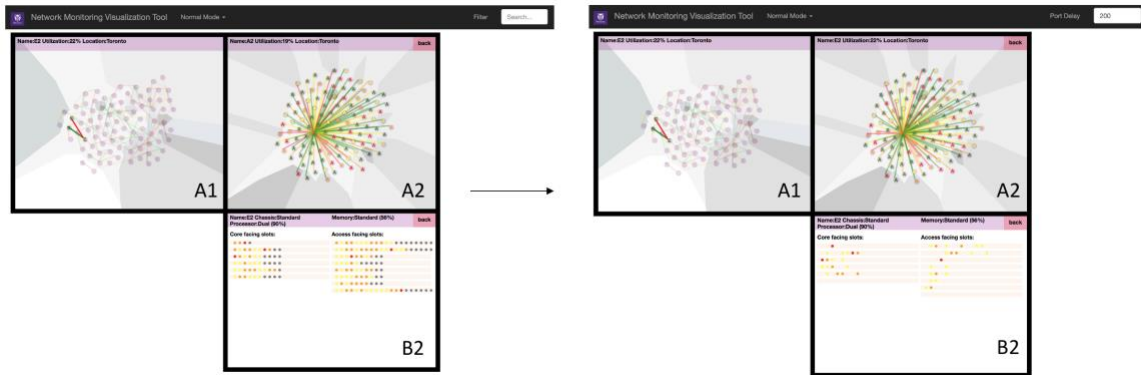


Figure 27: Filtering the ports according to the port delay

Besides the normal mode described above, the visualization tool also provides an alert mode for network administrators to directly catch the anomaly so they could handle it efficiently. The comparison between the normal mode view and the alert mode view is demonstrated in Figure 28. When choosing the alert mode in the drop-down menu on the top bar in the tool, the visualization will only display the red nodes, links and ports. The red color means the utilization is higher than 80%. In this way, the workload of monitoring could be reduced, and the network administrators could quickly identify the meaningful anomalous activities.



Figure 28: The comparison between the normal mode and the alert mode

4.4 Case Study

The intention of this network monitoring visualization tool is to help network administrators to efficiently detect possible degradation and easily figure out the possible anomalous devices so that it will be easy to tune the network resources accordingly. Also, it should benefit the network planners in conducting their network health analysis and aid in capacity planning since the network congestion points can be easily identified. According to the requirements, several direct case studies have been proposed to illustrate and examine this visualization tool.

A: Can the network administrators/planners efficiently detect the anomalous routers, links or ports?

B: Can the network administrators/planners easily identify the node and its linked node?

C: Can the network administrators/planners easily figure out the configuration of the routers?

D: Can the network administrators/planners monitor the QoS and detect possible QoS degradation?

E: Can the network administrators identify the connection between each port? Can the network administrators/planners navigate through the network in the level of ports?

F: Can the network administrators/planners interact with the network visualization tool and filter the data to better monitor the network and derive the filtered data report for future purposes?

A: In the normal mode (Figure 29a), through the color of each node, link and port, it is easy to figure out which range their utilization is in. When mouse over the nodes or links, their exact utilization number will be displayed in the window on the top of the panel. In the alert mode (Figure 29b), it will only display all the nodes, links and ports with extremely high utilization, which means all the elements displayed on the screen needs particular attention. Instead of being driven by request and query, the alert mode provides more direct ways to alert the network administrators. Network administrators can interact with these

elements in the same way as in the normal mode. Network administrators can easily convert between these two modes without losing the exploration stage and increasing the workload. In this way, the anomalous nodes, links, and ports can be easily detected.

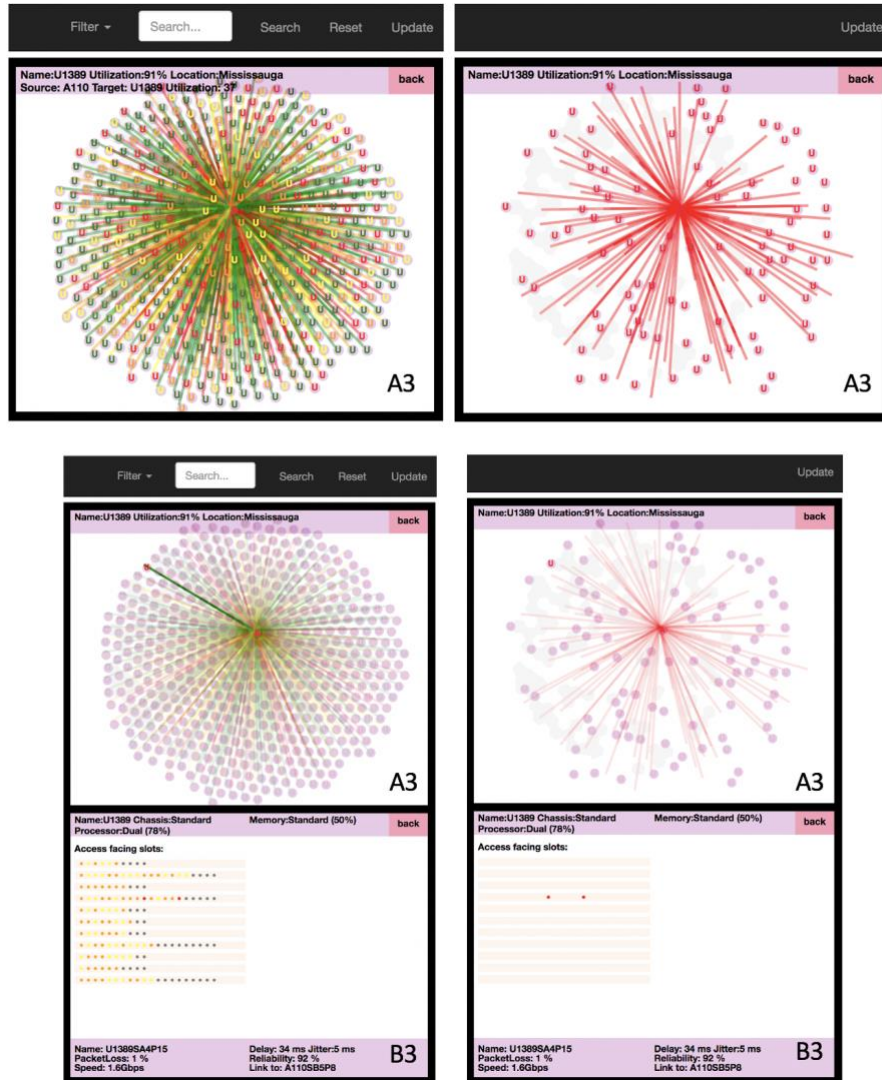


Figure 29: a- The normal mode (left); b- The alert mode (right)

B: As illustrated in Figure 30, when clicking one of the nodes, only the symbols of the selected routers and its linked routers are shown, and the symbols of other routers will be hidden. In this way, only the selected router, its linked routers and the links between them are highlighted, and other routers and links are faded. In addition, the three network panels

make it clear to identify the hierarchy of each connection and the number of hops between two nodes.

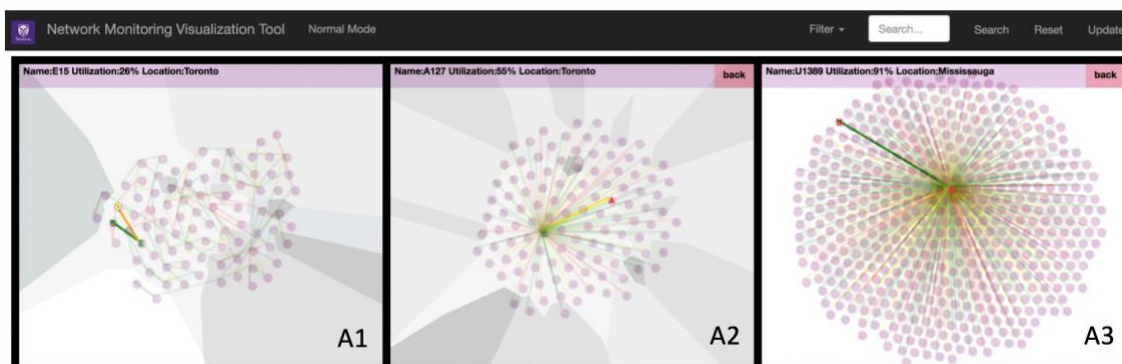


Figure 30: The selected router with its linked routers and the link between them are also highlighted. The different levels of connections are on different panels.

C: When mouse hovering around the router, its ID, its utilization and its location will be displayed as text in the window at the top of the panel. The background color of the area where the router is placed indicates the location of the router as well. When the inside panel of the selected router is displayed, it is easy to identify its facing slot(s) and its line cards. The rectangular represents the line card inside the router. It is easy to identify the quantity of line cards in each slot in the router and the condition of the ports located in each card. Also, the concrete information of the selected router will be shown on the top of the inside panel, which includes the name of the router, the chassis type, the memory type, the memory's utilization, the processor type, and the processors' utilization.

D: In the inside panel, when mouse over the ports, the specific QoS information of the port will be shown as text in the window at the bottom of the panel. In addition, the network administrators can filter the ports according to their port utilization, port delay or port packet loss. Only the ports whose port utilization, port delay or port packet loss is beyond the standard will be displayed, and an example is demonstrated in Figure 31. Users can interact with the visualization and explore the network with these ports. In this way, it is very efficient to detect possible QoS degradation, conduct health analysis, and still can obtain enough information of the network.

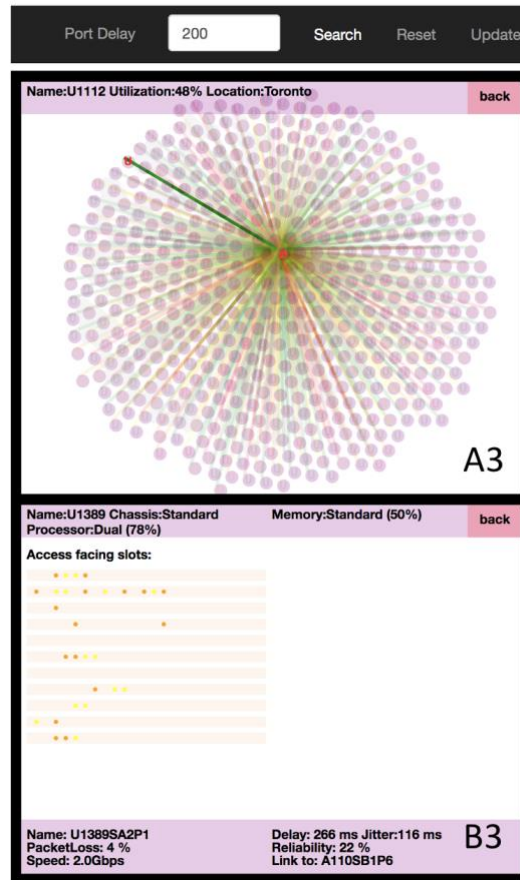


Figure 31: Filter the ports, and the visualization tool only shows the ports whose delay is higher than 200ms.

E: In the inside panel, when clicking the port, it will navigate to the port it links to. The inside panel of the current router will be folded, and the inside panel of the router where the linked port is located will be shown. The information of the linked port including the speed and QoS metrics will be displayed as text at the bottom of the inside panel. Meanwhile, the router in the network panel will be highlighted. If some port has a degraded service, it is very convenient to identify and locate which port it links to and check the condition of the linked port and its routers. It is beneficial for the network administrators or planners to identify the anomalous element and efficiently check the condition of elements that may be affected.

F: As mentioned in the D case, the visualization tool enables users to filter the port to better detect possible QoS degradation. Besides this, as shown in Figure 32, the nodes can also be filtered according to the router utilization. Only the suspect routers are displayed in the panel. In addition, the data information of the filtered ports or the filtered routers will be stored in the additional data file as the performance report for future use.

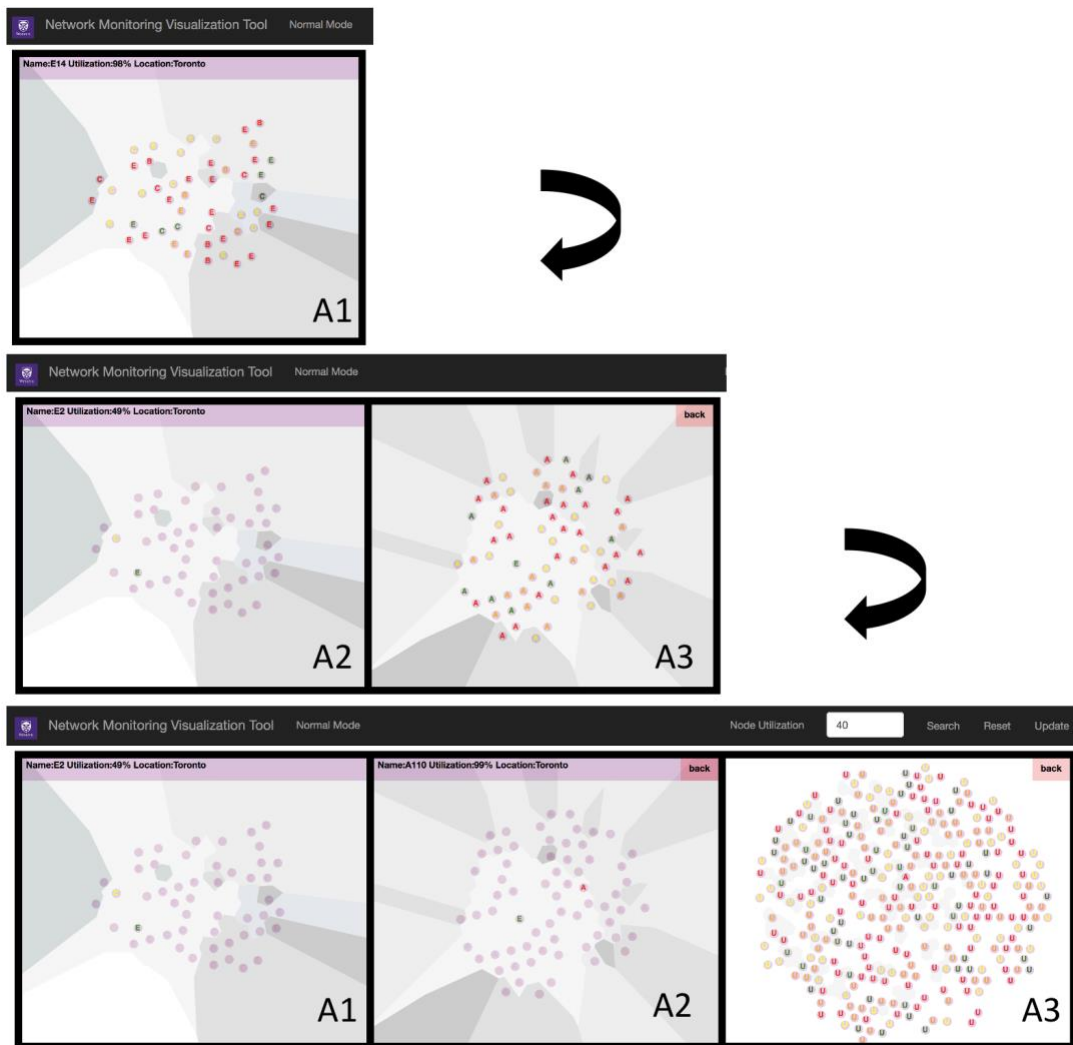


Figure 32: Filter the routers with specific utilization range

Chapter 5

5 Conclusion

5.1 Summary

With the increasing size and complexity of next-generation communication networks, it is critical to utilize interactive information visualization to support the network monitoring, planning, and management. Effectively visualizing large-scale networks has been considered difficult with traditional methods because the density of link is high, and the relationship between nodes is complicated. Given the limited screen space, it is essential to explore how to present ultra-large-scale network data efficiently that can assist Internet Service Provider's (ISP) network planning and management activities. To overcome the limitation of the node-link layout in a large-scale network, many previous works have explored this problem and proposed the following methods: adopting parameter focusing, zooming techniques, 3D layout, and novel graph modeling. However, there still exist some problems with these methods. Some visualization systems may lose the context or be confusing to navigate. Others may not be able to contain enough details or visualize the ultra-large-scale network data efficiently.

This thesis is motivated to explore how to better visualize the large-scale network data and provide enough low-level details and thus enables the network planners or administrators to plan and monitor the next-generation communication network efficiently. The system proposed in this thesis introduces a way of elegantly visualizing the topology of a large-scale computer network and visualizes the relevant information concerning it, including the router configuration and port level QoS metrics. The design of the visualization system combines the idea of progressive disclosure and multiple panels, and it divides the routers according to its hierarchy and visualizes them in different panels with the semantic zooming and brushing interaction techniques for users to monitor and explore the network. Representing additional information has often been considered cumbersome due to inefficient visual network topology abstraction and overloaded layout. Compared with the previous works, this system can clearly visualize and display the information of the inside configuration of the routers and the port level QoS metrics, with the designed multiple

panels layout. Also, network administrators and network planners will not lose the context of the connection and condition information during the exploration, and it is easy to navigate through both the router-level connection and port-level connection. It could improve their response time and speed up identifying the trouble to better solve anomaly problems and conduct health analysis. Additionally, the system provides the data filtering feature for the users to better plan and monitor the communication network. The different modes, including the normal mode and alert mode, can satisfy the diverse needs and benefit the monitoring process. Moreover, the visualization system proposed in this thesis is real-time, and it can improve diagnoses and support timely response and decision making.

5.2 Limitation and Future Work

There are mainly two limitations in this work at this stage. The first one is that currently the visualization tool is fed with the simulated data. In the future, the system could involve the real-world large-scale data collection, salient features extraction and data analysis to better structure the real-world data for visualization. Secondly, the current visualization is designed to allow updating the visualization only when clicking the button. The real-time data coming from the server can only be visualized when the ‘update’ button is clicked. In the future, more options may be provided for users to choose whether updating the visualization automatically or manually.

Besides all the limitations discussed above, other advanced features and techniques could be added or combined into this visualization system in future work. Some machine learning techniques could be combined to predict the anomaly and the fluctuation of the QoS. Some other machine learning techniques may be applied to cluster the routers, the links and the nodes. The router nodes could be divided according to these clusters and visualized in different panels, besides according to the hierarchy of the routers. In addition, the QoS metrics information could be analyzed, and some summary statistic charts could be provided. Also, a timeline with long-term historical data over the same view may be provided to aid in capacity planning. In this way, overloaded elements of the network could be easily identified and analyzed. Moreover, the traffic flow monitoring and packet-level capture can be visualized upon this visualization tool in the future, and some animation of

events could be provided to recognize the unusual behavior and benefit the resource allocation and planning.

References

- [1] L. Xiao, J. Gerth, and P. Hanrahan, “Enhancing Visual Analysis of Network Traffic Using a Knowledge Representation,” *2006 IEEE Symp. Vis. Anal. Sci. Technol.*, pp. 107–114, 2006.
- [2] W. Dargie and C. Poellabauer, *Fundamentals of wireless sensor networks*. A John Wiley and Sons, Ltd., Publication.
- [3] V. T. Guimarães *et al.*, “A Survey on Information Visualization for Network and Service Management,” *IEEE Commun. Surv. Tutorials*, vol. 18, no. 1, pp. 285–323, 2016.
- [4] B. Li, J. Springer, G. Bebis, and M. Hadi, “A survey of network flow applications,” *J. Netw. Comput. Appl.*, vol. 36, no. 2, pp. 567–581, 2013.
- [5] N. Tateishi, M. Tahara, N. Tanji, and H. Seshake, “Method for visualizing information from large-scale carrier networks,” in *2013 15th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, 2013, pp. 1–6.
- [6] Q. Liao, A. Blaich, D. Vanbruggen, and A. Striegel, “Managing networks through context : Graph visualization and exploration,” *Comput. Networks*, vol. 54, no. 16, pp. 2809–2824, 2010.
- [7] D. Lalanne, E. Bertini, P. Hertzog, and P. Bados, *Visual Analysis of Corporate Network Intelligence : Abstracting and Reasoning on Yesterdays for Acting Today*. Springer, Berlin, Heidelberg.
- [8] S. T. Kard, J. D. Mackinlay, and B. Scheiderman, *Information Visualization: Using Vision To Think*. Morgan Kaufmann, 1993.
- [9] T. Zhang, X. Wang, Z. Li, F. Guo, Y. Ma, and W. Chen, “A survey of network anomaly visualization,” *Sci. China Inf. Sci.*, vol. 60, no. 12, pp. 1–17, 2017.
- [10] H. Wang, “From A Mess to Graphic Maps : Visualization of Large-scale Heterogeneous Networks,” in *2010 Second International Conference on Computer Modeling and Simulation*, 2010, vol. 1, pp. 531–535.
- [11] L. Atzori, A. Iera, and G. Morabito, “The Internet of Things : A survey,” *Comput. Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [12] B. C. Tham, Y. Jiang, and C. Ko, “Monitoring QoS distribution in multimedia networks,” *Int. J. Netw. Manag.*, vol. 10, no. 2, pp. 75–90, 2000.

- [13] Y. Wu, S. Member, P. A. Chou, Q. Zhang, S. Member, and K. Jain, "Network Planning in Wireless Ad Hoc Networks : A Cross-Layer Approach," vol. 23, no. 1, pp. 136–150, 2005.
- [14] F. Harary, "Graph Theory In Network Analysis," *Soc. Networks*, vol. 5, pp. 235–244, 1983.
- [15] R. Siamwalla, R. Sharma, and S. Keshav, "Discovering Internet Topology." [Online]. Available: <http://www.cs.cornell.edu/skeshav/%0Apapers.html>. [Accessed: 20-May-2019].
- [16] Y. Xin, G. N. Rouskas, S. Member, H. G. Perros, and S. Member, "On the Physical and Logical Topology Design of Large-Scale Optical Networks," *J. Light. Technol.*, vol. 21, no. 4, pp. 904–915, 2003.
- [17] R. Ramaswami, S. Member, and K. N. Sivarajan, "Design of Logical Topologies for Wavelength-Routed Optical Networks," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 840–851, 1996.
- [18] R. D. T. Valle, D. Passos, C. Albuquerque, and D. C. M. Saade, "Mesh Topology Viewer (MTV): an SVG-Based Interactive Mesh Network Topology Visualization Tool," in *IEEE Symposium on Computers and Communications*, 2008.
- [19] J. F. Kurose and K. W. Ross, *Computer Networking A Top-Down Approach*. Pearson.
- [20] J. Garud, "Router Architecture - Electronics Post." [Online]. Available: <https://electronicspost.com/router-architecture>. [Accessed: 20-May-2019].
- [21] Cisco Academy, "Hierarchical Network Design Overview." [Online]. Available: <https://www.ciscopress.com/articles/article.asp?p=2202410&seqNum=4>. [Accessed: 20-May-2019].
- [22] H. Truong, R. Samborski, and T. Fahringer, "Towards a Framework for Monitoring and Analyzing QoS Metrics of Grid," in *IEEE International Conference on e-Science and Grid Computing*, 2006.
- [23] M. Peuhkuri, "IP Quality of Service," *Helsinki University of Technology, Laboratory of Telecommunications Technology*, 1999. [Online]. Available: <http://www.netlab.tkk.fi/u/puhuri/htyo/Tik-110.551/iwork.pdfv>. [Accessed: 20-

- May-2019].
- [24] D. Mistry, P. Modi, K. Deokule, A. Patel, H. Patki, and O. Abuzagheh, “Network Traffic Measurement And Analysis,” *IEEE Long Isl. Syst. Appl. Technol. Conf.*, no. 1, pp. 1–7, 2016.
- [25] D. A. Keim, “Information Visualization and Visual Data Mining,” *IEEE Trans. Vis. Comput. Graph.*, vol. 8, no. 1, pp. 100–107, 2002.
- [26] S. Kamran and P. Parsons, “Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach,” *Trans. Human-Computer Interact.*, vol. 5, no. 2, pp. 84–133, 2013.
- [27] K. Sedig and P. Parsons, *Design of Visualizations for Human-Information Interaction A Pattern-Based Framework*. Morgan & Claypool Publishers, 2016.
- [28] M. Hegarty, “The Cognitive Science of Visual-Spatial Displays : Implications for Design,” *Top. Cogn. Sci.*, vol. 3, pp. 446–474, 2011.
- [29] N. A. Salingaros, *The Structure of Pattern Languages*. Cambridge University Press, 2005.
- [30] O. Ola and K. Sedig, “The Challenge of Big Data in Public Health : An Opportunity for Visual Analytics,” *Online J. Public Health Inform.*, vol. 5, no. 3, pp. 1–21, 2014.
- [31] “Les Misérables Matrix.” [Online]. Available: <https://bost.ocks.org/mike/miserables/>. [Accessed: 20-May-2019].
- [32] “Knuth: The Stanford GraphBase.” [Online]. Available: <https://www-cs-faculty.stanford.edu/~knuth/sgb.html>. [Accessed: 20-May-2019].
- [33] M. Krzywinski, “Hive Plots - Linear Layout for Network Visualization - Visually Interpreting Network Structure and Content Made Possible.” [Online]. Available: <http://egweb.bcgsc.ca/>. [Accessed: 20-May-2019].
- [34] “Force-Directed Graph.” [Online]. Available: <https://observablehq.com/@d3/force-directed-graph>. [Accessed: 20-May-2019].
- [35] F. Preparata and M. Shamos, *Computational Geometry An Introduction*. Springer Science & Business Media, 2012.
- [36] “Voronoi Diagram.” [Online]. Available: <https://strongriley.github.io/d3/ex/voronoi.html>. [Accessed: 20-May-2019].

- [37] K. Sedig, P. Parsons, H. Liang, and J. Morey, “Supporting Sensemaking of Complex Objects with Visualizations : Visibility and Complementarity of Interactions,” *Informatics*, vol. 3, no. 4, p. 20, 2016.
- [38] R. Spence, *Information Visualization*. Springer.
- [39] S. C. Au, C. Leckie, A. Parhar, and G. Wong, “Efficient Visualisation of Large Routing Topologies,” *Int. J. Netw. Manag.*, vol. 14, no. 2, pp. 105–118, 2004.
- [40] H. Burch, B. Cheswick, and B. Labs, “Mapping and Visualizing the Internet,” in *USENIX Annual Technical Conference*, 2000.
- [41] B. Cheswick and H. Burch, “Mapping and Visualizing the Internet,” *Computer (Long. Beach. Calif.)*, vol. 32, no. 4, pp. 97–98, 1999.
- [42] T. Qiu *et al.*, “What Happened in my Network ? Mining Network Events from Router Syslogs Categories and Subject Descriptors,” in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 2010, pp. 472–484.
- [43] J. Abello, S. G. Kobourov, and R. Yusufov, “Visualizing Large Graphs with Compound-Fisheye Views and Treemaps,” in *International Symposium on Graph Drawing*, 2004, pp. 431–441.
- [44] E. Gansner, S. C. North, and L. Anonymization, “Topological Fisheye Views for Visualizing Large Graphs,” *IEEE Trans. Vis. Comput. Graph.*, vol. 11, no. 4, pp. 457–468, 2004.
- [45] F. Mansmann *et al.*, “Visual Support for Analyzing Network Traffic and Intrusion Detection Events using TreeMap and Graph Representations,” *Proc. Symp. Comput. Hum. Interact. Manag. Inf. Technol.*, pp. 19–28, 2009.
- [46] Z. Kan, C. Hu, Z. Wang, G. Wang, and X. Huang, “NetVis : A Network Security Management Visualization Tool Based On Treemap,” in *2nd International Conference on Advanced Computer Control*, 2010, pp. 1–4.
- [47] F. Fischer, F. Mansmann, D. A. Keim, and S. Pietzko, “Large-scale Network Monitoring for Visual Analysis of Attacks,” *Int. Work. Vis. Comput. Secur.*, pp. 111–118, 2008.
- [48] L. Williams, R. Lippmann, and K. Ingols, *An Interactive Attack Graph Cascade and Reachability Display*. Springer-Verlag Berlin Heidelberg.
- [49] L. Harrison, M. Iannacone, E. Downing, and J. R. Goodall, “NV : Nessus

- Vulnerability Visualization for the Web Categories and Subject Descriptors,” in *Proceedings of the Ninth International Symposium on Visualization for Cyber Security*, 2012, pp. 25–32.
- [50] W. Xu and M. Shehab, “Visualization Based Policy Analysis : Case Study in SELinux,” in *Proceedings of the 13th ACM symposium on Access control models and technologies*, pp. 165–174.
- [51] A. Lu and W. Wang, “Sybil attack detection through global topology pattern visualization,” *Inf. Vis.*, vol. 10, no. May 2010, pp. 32–46, 2011.
- [52] J. R. Goodall and W. G. Lutters, “Preserving the Big Picture : Visual Network Traffic Analysis with TNV,” *IEEE Work. Vis. Comput. Secur.*, pp. 47–54, 2005.
- [53] M. Daniel, S. Bohn, A. Wynne, and A. William, “Real-Time Visualization of Network Behaviors for Situational Awareness,” in *Proceedings of the Seventh International Symposium on Visualization for Cyber Security*, 2010, pp. 79–90.
- [54] A. Gubin, W. Yurcik, and L. Brumbaugh, “Network management visualization with PingTV Network Management Visualization With PingTV,” in *Annual IEEE Conference on Local Computer Network*, 2014, no. February 2001.
- [55] H. Choi, H. Lee, and H. Kim, “Fast detection and visualization of network attacks on parallel coordinates,” *Comput. Secur.*, vol. 28, no. 5, pp. 276–288, 2009.
- [56] S. Tricaud, K. Nance, and P. Saadé, “Visualizing Network Activity using Parallel Coordinates,” in *44th Hawaii International Conference on System Sciences*, 2011, pp. 1–8.
- [57] J. Cho, H. Lee, K. Choi, S. Nam, and J. Moon, “Visualization of Abnormal Behavior Detection using Parallel Coordinate and Correspondence Analysis,” *Int. J. Inf.*, pp. 1–15.
- [58] S. Tricaud and P. Saadé, “Applied parallel coordinates for logs and network traffic attack analysis,” *J. Comput. Virol.*, pp. 1–29, 2010.
- [59] T. Nunnally, P. Chi, K. Abdullah, A. S. Uluagac, J. A. Copeland, and R. Beyah, “P3D : A Parallel 3D Coordinate Visualization for Advanced Network Scans,” pp. 2052–2057, 2013.
- [60] A. Douitsis and D. Kalogeras, “Interactive Network Management Visualization with SVG and AJAX,” in *Large Installation System Administration Conference*,

- pp. 233–246.
- [61] S. G. Eick, “Aspects of Network Visualization,” *IEEE Comput. Graph. Appl.*, vol. 16, no. 2, 1996.
 - [62] A. W. R. Becker, S. Eick, E. Miller, “Dynamic graphics for network visualization,” in *Proc. 1st IEEE Conf. Vis.*, 1990, pp. 93–96.
 - [63] E. R. Gansner, S. C. North, and F. Park, “Large-Scale Network Visualization,” *SIGGRAPH Newsl.*, pp. 1–5, 1999.
 - [64] G. P. ; D. S. ; E. K. ; J. T. ; A. Kantawala, “An architecture for monitoring visualization and control.pdf,” *IEEE Netw.*, vol. 11, no. 5, pp. 34–43, 1997.
 - [65] Y. Tsutsui, Y. Kitaura, K. Kikuchi, T. Ohta, and Y. Kakuda, “Development of a Real-time System for Visualizing Network Topology and Packet Flow in MANET Field Experiments,” in *2014 10th International Conference on Mobile Ad-hoc and Sensor Networks*, 2014, pp. 151–157.
 - [66] H. Fuji, S. Nakai, H. Matoba, and H. Takano, “Real-time Bifocal Network-Visualiization,” in *IEEE Network Operations and Management Symposium*, 1994.
 - [67] R. H. and T. Fioreze, “SURFmap: A network monitoring tool based on the Google Maps API,” in *Proc. IFIP/IEEE Int. Symp. IM Netw.*, pp. 676–690.
 - [68] A. Inoue, Y. Takahashi, and A. Koyama, “MANET Viewer III : 3D Visualization System for Mobile Ad-hoc Networks,” in *2013 16th International Conference on Network-Based Information Systems*, 2013, pp. 178–185.
 - [69] “WebSocket.” [Online]. Available: <https://en.wikipedia.org/wiki/WebSocket>. [Accessed: 20-May-2019].
 - [70] Tornadoweb.org, “Tornado Web Server — Tornado 6.0.4 documentation.” [Online]. Available: <https://www.tornadoweb.org/en/stable/>. [Accessed: 22-Oct-2019].
 - [71] “Scalable Vector Graphics.” [Online]. Available: https://en.wikipedia.org/wiki/Scalable_Vector_Graphics. [Accessed: 20-May-2019].
 - [72] M. Bostock, “D3.js - Data-Driven Documents.” [Online]. Available: <https://d3js.org/>. [Accessed: 20-May-2019].
 - [73] “wsdream.” [Online]. Available: <https://github.com/wsdream/wsdream-datasetv>.

- [Accessed: 20-May-2019].
- [74] “QWS Dataset.” [Online]. Available: <https://qwsdata.github.io/>. [Accessed: 20-May-2019].
- [75] “Simulated Cloud Service QoS Dataset.” [Online]. Available: <https://data.mendeley.com/datasets/5vffs75j85/1>. [Accessed: 20-May-2019].
- [76] A. Striegel, “Intelligent Network Management Using Graph Differential Anomaly Visualization,” in *IEEE Network Operations and Management Symposium*, 2012.
- [77] N. Nielsen, “Progressive Disclosure.” [Online]. Available: <https://www.nngroup.com/articles/progressive-disclosure/>. [Accessed: 20-May-2019].
- [78] “Verlet integration.” [Online]. Available: <https://resources.saylor.org/wwwresources/archived/site/wp-content/uploads/2011/06/MA221-6.1.pdf>. [Accessed: 20-Sep-2019].
- [79] “d3/d3-force.” [Online]. Available: <https://github.com/d3/d3-force>. [Accessed: 05-Feb-2019].
- [80] A. Telea and J. J. Van Wijk, “Visualization of Generalized Voronoi Diagrams,” *Eurographics*, pp. 165–174.

Curriculum Vitae

Name: Wenjun Chen

Post-secondary Education and Degrees: The University of Western Ontario
London, Ontario, Canada
2018-2020 M.Sc

University of Liverpool
Liverpool, United Kingdom
2016-2018 B.S.

Xi'an Jiaotong-Liverpool University
Suzhou, China
2014-2016 B.S.

Honours and Awards: Western Graduate Research Scholarship
2018-2019

Related Work Experience Teaching Assistant
The University of Western Ontario
2018-2020