Theses and Dissertations 1. Thesis and Dissertation Collection, all items

2020-09

SYNTHETIC NETWORK GENERATION AND VULNERABILITY ANALYSIS OF INTERNET INFRASTRUCTURE SYSTEMS IN THE U.S. VIRGIN ISLANDS

Moeller, Brian T.

Monterey, CA; Naval Postgraduate School

https://hdl.handle.net/10945/66112

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

SYNTHETIC NETWORK GENERATION AND VULNERABILITY ANALYSIS OF INTERNET INFRASTRUCTURE SYSTEMS IN THE U.S. VIRGIN ISLANDS

by

Brian T. Moeller

September 2020

Thesis Advisor:	Daniel Eisenberg
Co-Advisor:	Justin P. Rohrer
Second Reader:	David L. Alderson Jr.

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE				*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.				
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** September 2020	**3. REPORT TYPE AND DATES COVERED** Master's thesis		
4. TITLE AND SUBTITLE SYNTHETIC NETWORK GENERATION AND VULNERABILITY ANALYSIS OF INTERNET INFRASTRUCTURE SYSTEMS IN THE U.S. VIRGIN ISLANDS				**5. FUNDING NUMBERS**
6. AUTHOR(S) Brian T. Moeller				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000				**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A				**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.				**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

In September 2017, Category-5 Hurricanes Irma and Maria struck the U.S. Virgin Islands and caused significant damage to all critical infrastructure within the territory. This thesis assesses the vulnerability of the fiber optic telecommunications system on the island of St. Croix to determine ways to ensure Internet access during future disasters. Specifically, we use publicly available information to (1) develop a geospatial data set for St. Croix Internet infrastructure, (2) generate a synthetic network model that approximates Internet demands and traffic by St. Croix communities and critical facilities, and (3) identifies network vulnerabilities to recommend disaster hardening. Results show that the synthetic model is vulnerable to fiber cuts that can disconnect all households and critical facilities from the Internet. Recommendations for system hardening include the need to ensure redundant physical fiber paths off-island and switching locations between internet service providers.

14. SUBJECT TERMS U.S. Virgin Islands, synthetic network generation, network analysis, telecommunications, optimization, critical infrastructure, resilience			**15. NUMBER OF PAGES** 103
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

SYNTHETIC NETWORK GENERATION AND VULNERABILITY ANALYSIS OF INTERNET INFRASTRUCTURE SYSTEMS IN THE U.S. VIRGIN ISLANDS

Brian T. Moeller
Major, United States Marine Corps
BS, U.S. Naval Academy, 2009

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2020

Approved by: Daniel Eisenberg
Advisor

Justin P. Rohrer
Co-Advisor

David L. Alderson Jr.
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In September 2017, Category-5 Hurricanes Irma and Maria struck the U.S. Virgin Islands and caused significant damage to all critical infrastructure within the territory. This thesis assesses the vulnerability of the fiber optic telecommunications system on the island of St. Croix to determine ways to ensure Internet access during future disasters. Specifically, we use publicly available information to (1) develop a geospatial data set for St. Croix Internet infrastructure, (2) generate a synthetic network model that approximates Internet demands and traffic by St. Croix communities and critical facilities, and (3) identifies network vulnerabilities to recommend disaster hardening. Results show that the synthetic model is vulnerable to fiber cuts that can disconnect all households and critical facilities from the Internet. Recommendations for system hardening include the need to ensure redundant physical fiber paths off-island and switching locations between internet service providers.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

1 Introduction **1**

1.1 Overview of Critical Infrastructure Systems in the U.S. Virgin Islands . . . 1

1.2 Internet Infrastructure Systems in the USVI 4

1.3 Vulnerabilities in USVI Internet Infrastructure 10

1.4 Thesis Goals 12

2 Literature Review **13**

2.1 Structure of the Internet 13

2.2 Internet Systems and Resilience 14

2.3 Synthetic Internet Topology Generation 15

2.4 Modeling Internet Traffic Demands 17

2.5 Modeling Internet Behavior 19

2.6 Our Contribution 20

3 Data Curation and Model Development **21**

3.1 Data Curation. 21

3.2 A Network Model of the Internet 28

3.3 STX Internet System Model 46

3.4 Summary 53

4 Analysis **55**

4.1 How St. Croix Internet Customers Connect to viNGN Middle Mile Infrastructure 55

4.2 Synthetic St. Croix Internet Traffic and Network Operations 62

4.3 Synthetic Internet Traffic and Network Operations for the USVI Territory . . 74

4.4 Summary 75

5 Conclusion **77**

5.1 Summary 77

5.2 Model Limitations 78

5.3 Future Work 78

List of References **79**

Initial Distribution List **83**

List of Figures

Figure 1.1	Population Distribution in the USVI.	2
Figure 1.2	viNGN Facility Access Point on St. Croix	5
Figure 1.3	Telecommunication Companies Operating in the USVI	6
Figure 1.4	Submarine Cables in the Leeward Islands	7
Figure 1.5	Map of Community Anchor Institutions in the USVI	8
Figure 1.6	Proposed viNGN Middle Mile Network Diagram	9
Figure 1.7	St. Andrews Public Computing Center in Charlotte Amalie	10
Figure 1.8	Hurricane Maria Damage on St. Croix	11
Figure 2.1	The Open Systems Interconnection Model of Internet Layers.	14
Figure 3.1	St. Croix Private Internet Customers	22
Figure 3.2	St. Croix Public Internet Customers	23
Figure 3.3	St. Croix Last Mile Infrastructure	24
Figure 3.4	St. Croix Facility Access Point Locations	25
Figure 3.5	St. Croix Terrestrial Middle Mile Fiber Optic Cables	26
Figure 3.6	USVI Internet Exchange points, Submarine Middle Mile, and Global Connectivity.	27
Figure 3.7	Simple Model of the USVI Internet System	30
Figure 3.8	Node Types for Simple Model	33
Figure 3.9	Simple Model: Layer-1 Edges	37
Figure 3.10	Simple Model: Layer 2 Edges	38
Figure 3.11	Simple Model: Layer-3 Edges	39

Figure 3.12 Simple Model: Layer-4 Edges 39
Figure 3.13 Layer 1 Traffic Assignment in Simple Model, All Edges Available 44
Figure 3.14 Layer-1 Traffic Assignment in Simple Model, One Edge Destroyed 45
Figure 3.15 Nodes in the STX Internet System Model 47
Figure 3.16 Layer-1 Edges in the STX Internet System Model 49
Figure 3.17 Primary and Alternate Layer-2 Links in STX Internet Model . . . 50
Figure 3.18 Layer-2 Edges in the STX Internet System Model 51
Figure 3.19 Layer-3 Edges in STX Internet System Model 52

Figure 4.1 Customers Served by STX-02 and STX-03 Fiber Access Points . 56
Figure 4.2 St. Croix Internet Customers Colored by Primary Facility Access Point 57
Figure 4.3 St. Croix Estates Colored by Primary Facility Access Point . . . 58
Figure 4.4 Population Serviced by Each Facility Access Point 59
Figure 4.5 Community Anchor Institutions Serviced By Each Facility Access Point. 60
Figure 4.6 Average and Peak Demand per Fiber Access Point 61
Figure 4.7 St. Croix Middle Mile Traffic under Scenario 1 63
Figure 4.8 St. Croix Layer-2 Connectivity to Customers Under Scenario 1. . 64
Figure 4.9 St. Croix Middle Mile Traffic under Scenario 2 66
Figure 4.10 St. Croix Layer-2 Connectivity to Customers Under Scenario 2 . 67
Figure 4.11 St. Croix Middle Mile Traffic under Scenario 3 68
Figure 4.12 St. Croix Layer-2 Connectivity to Customers Under Scenario 3 . 69
Figure 4.13 St. Croix Layer-2 Connectivity to Customers Under Scenario 4 . 70
Figure 4.14 St. Croix Layer-2 Connectivity to Customers Under Scenario 5 . 71

List of Tables

Table 3.1 Community Anchor Institution Categories 23

Table 3.2 Internet Node Attributes 32

Table 3.3 Simple Model Node Type and Layer Attributes 34

Table 3.4 Simple Model Population and CAI Node Demand Attributes . . . 34

Table 3.5 Simple Model Node Uplink Attributes 35

Table 3.6 Internet Edge Attributes 36

Table 3.7 Simple Model Traffic Matrix, Average Demand 42

Table 3.8 Simple Model Traffic Matrix, Peak Demand 42

Table 3.9 Internet Traffic Metrics for the Simple Model 46

Table 4.1 Layer 2 Aggregate Traffic Between Fiber Access Points (FAPs) under Scenario 1. 63

Table 4.2 Layer 2 Aggregate Traffic Between FAPs under Scenario 2. 66

Table 4.3 Layer 2 Aggregate Traffic Between FAPs under Scenario 3. 68

Table 4.4 Layer 2 Aggregate Traffic Between FAPs under Peak Demand . . 72

Table 4.5 Scenario Comparison 72

Table 4.6 Layer 2 Aggregate Traffic Between Islands and the Continental United States. 75

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

A-D	Attacker-Defender
BTOP	Broadband Technology Opportunities Program
BIT	Bureau of Information Technology
CAI	Community Anchor Institution
CI	Critical Infrastructure
CONUS	Continental U.S.
DHS	U.S. Department of Homeland Security
DPW	Department of Public Works
FAP	Fiber Access Point
FEMA	Federal Emergency Management Agency
Gbps	Gigabits Per Second
GPS	Global Positioning System
ISP	Internet Service Provider
IT	Information Technology
IXP	Internet Exchange Point
Mbit	Megabit
Mbps	Megabits Per Second
NAP	Network Access Point
NDA	Non-Disclosure Agreement

NOFA	Notice of Funding Availability
NPS	Naval Postgraduate School
NTIA	National Telecommunication and Information Administration
OSI	Open Systems Interconnection
PCC	Public Computing Center
QGIS	Quantum Geographic Information System
SCADA	Supervisory Control and Data Acquisition
STJ	St. John
STT	St. Thomas
STX	St. Croix
U.S.	United States
USVI	U.S. Virgin Islands
UVI	University of the Virgin Islands
viNGN	Virgin Islands Next Generation Network
VIPFA	Virgin Islands Public Finance Authority
VITEMA	Virgin Islands Territorial Emergency Management Agency

Executive Summary

In September 2017, Category-5 Hurricanes Irma and Maria struck the U.S. Virgin Islands within a two-week period. This event devastated territorial infrastructure systems, including power, water, transportation, and telecommunications. The loss of hardline fiber optic infrastructure that provides Internet access and services impacted emergency response operations and restricted coordination with the public. To support the territory and its communities in future disasters, it is important to assess and reduce vulnerabilities of this infrastructure system.

This thesis focuses on producing a representative model of the St. Croix Internet system that enables vulnerability analysis. In particular, we:

1. gather publicly available information about the U.S. Virgin Islands Internet infrastructure to produce geospatial data sets for visualization and modeling;
2. develop network analysis techniques to generate a synthetic St. Croix Internet system model that is realistic to inform its operations; and,
3. analyze the structure and characteristics of the St. Croix Internet system model to measure how it performs under different demands and disruptions.

Our geospatial data set and model indicates that the U.S. Virgin Islands Internet includes public and private assets across physical and virtual networks. The Virgin Islands Next Generation Network (viNGN), a semi-autonomous government agency who provides fee-based broadband service, owns and operates fiber optic rings across all major islands that enable Internet connectivity within the territory and to the mainland U.S. Internet customers, including public agencies, community organizations, schools, medical facilities, emergency responders, and private households, purchase Internet service from commercial internet service providers that connect to the viNGN network. Internet traffic between public and private Internet customers is exchanged at few locations in the territory before being sent to the mainland U.S. via submarine cables.

Synthetic Internet network generation results suggest that Internet customers are unevenly distributed among viNGN middle mile infrastructure. Customers connect into the Internet via Facility Access Points (FAPs). Some FAPs provide connectivity for a large amount of

Internet customers while some connect relatively few. Furthermore, some FAPs connect a larger number of medical facilities and public safety organizations than others. Ensuring Internet access in future disasters requires hardening these critical FAPs.

St. Croix Internet demand and traffic analysis also suggests that the network is vulnerable to fiber optic cable cuts and spikes in Internet demand. The distribution of customers across FAPs means that the loss of key viNGN infrastructure will disconnect large portions of public and private Internet customers. We also identify possible single points of failure, where a single fiber cut may disconnect all Internet customers on St. Croix. Moreover, the rated bandwidth on network bottlenecks may limit customer access during demand spikes.

Despite our model identifying significant vulnerabilities in the St. Croix Internet network, our work is based on publicly available information which may be out of date or incomplete. Future work should focus on validating data and modeling assumptions. Moreover, the Internet model we assess is based on pre-hurricane infrastructure networks. Future work should build on the data, models, and analysis presented herein to guide protection of the post-storm system.

This analysis has considered only the system on St. Croix (STX), and as such it potentially misses system properties that arise from the interconnections St. Thomas (STT) and St. John (STJ). A complete analysis of the Internet infrastructure across the entire Territory is needed to understand the resilience of the aggregate system. Moreover, because of the distinct differences between these islands, a comparable analysis of the system on each island also deserves attention.

Acknowledgments

To my co-advisors, Dr. Daniel Eisenberg and Dr. Justin Rohrer, thank you for your guidance and patience over the past six months. Dr. Rohrer, your technical expertise and oversight were instrumental in developing a model that captures the complicated operation of the Internet. Dr. Eisenberg, your commitment to me was inspiring. I appreciate all of the hard work and late nights that you contributed to this project.

To my second reader, Dr. David Alderson, and the rest of the Virgin Islands team at the Naval Postgraduate School, thank you for your support throughout this process. I trust that our collective efforts will improve operation and resilience of critical infrastructure in the Virgin Islands.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 1:
Introduction

In September 2017, Category-5 Hurricanes Irma and Maria struck the United States (U.S.) Virgin Islands (USVI) within a two-week period and collectively devastated homes, businesses, and infrastructure throughout the territory. In particular, the loss of hardline fiber optic telecommunications infrastructure during the storms significantly impacted emergency response and recovery operations (USVI Hurricane Recovery and Resilience Task Force 2018). As of April 2019, approximately 7% of the aerial fiber optic lines that enable Internet traffic and data sharing were not restored (Virgin Islands Next Generation Network 2019). The delayed recovery of fiber optic Internet systems impacted the operations and recovery of all other lifeline infrastructure systems (e.g., electric power, water, and transportation among others).

This thesis is in support of Federal Emergency Management Agency (FEMA) response and recovery activities and part of a broader FEMA-funded effort by the Naval Postgraduate School (NPS) to assess and improve the resilience of interdependent USVI lifeline infrastructure systems (Alderson et al. 2018). This thesis additionally supports several other complementary efforts with the University of the Virgin Islands (UVI) to develop a next-generation Hazard Mitigation and Resilience Plan for the Territory.

1.1 Overview of Critical Infrastructure Systems in the U.S. Virgin Islands

1.1.1 The U.S. Virgin Islands Territory

The USVI is a territory of the United States located in the Leeward Islands of the Lesser Antilles. It is approximately 40 miles east of Puerto Rico and over 1,100 miles east of Miami, Florida. The territory is comprised of three main islands—St. Croix (STX), St. Thomas (STT), and St. John (STJ)—and several smaller surrounding islands. STX is the southernmost island in the territory and geographically the largest at 82 square miles. STT is the second largest of the islands at 32 square miles and is located 40 miles north of STX.

The territory's capital, Charlotte Amalie, is located on STT. STJ is the smallest of the three islands and is located two miles east of STT. Each island is divided into community estates with a total population of approximately 106,000 people as of the 2010 census; 50,601 across 212 estates in STX, 51,634 across 73 estates in STT, and 4,170 across 51 estates in STJ (U.S. Census Bureau 2010a); see Figure 1.1.

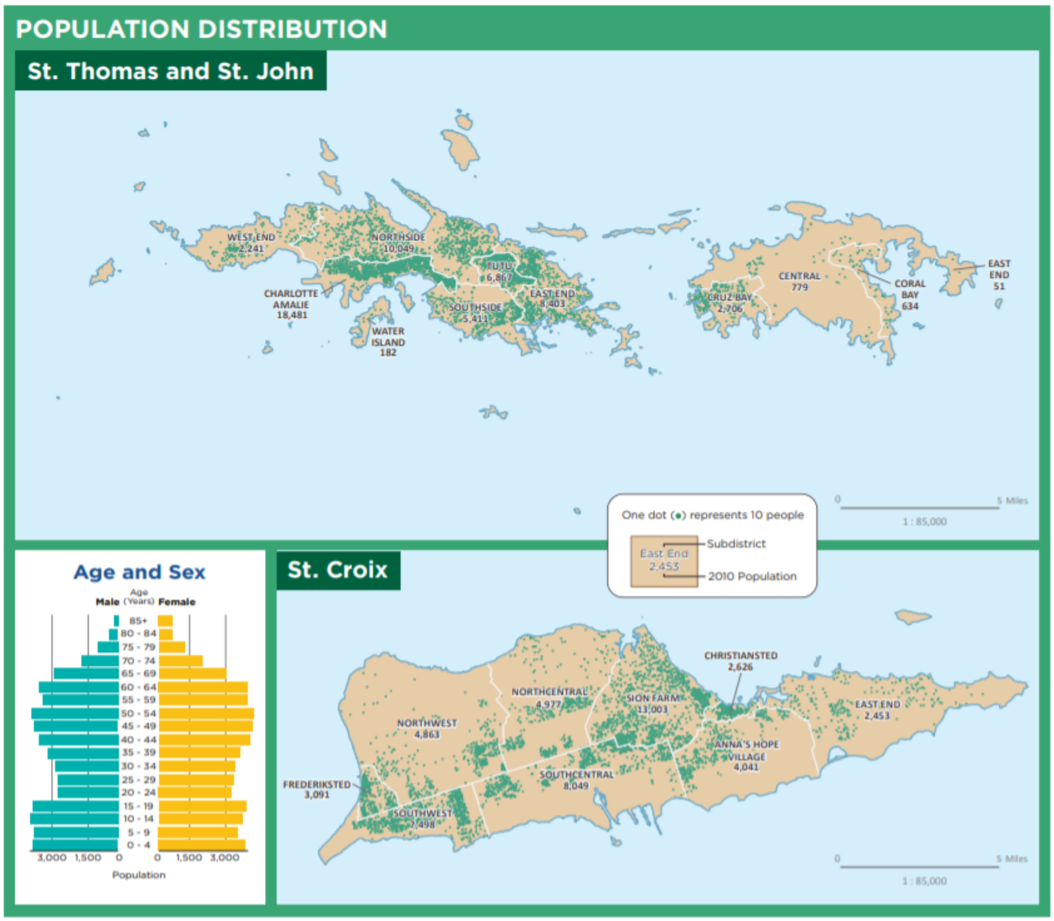

Figure 1.1. Population Distribution in the USVI. Source: U.S. Census Bureau (2010b).

The territory has limited natural resources, domestic manufacturing, and agriculture. As a result, it must import most of its energy, consumer goods, and food which creates a high cost of living on the island. STT is the primary center for tourism, government, finance, trade, and commerce in the territory. Its economy is largely dependent on tourism and the cruise ship industry. The economy of STX also relies on tourism, although less so than STT. STX is the center of other industries in the Territory including rum production at the

Cruzan and Diageo distilleries and oil production at the Limetree Bay Terminal. Limetree Bay (previously Hovensa refinery) was a major employer for the island prior to its closure in 2012. Recently, it has begun the process of re-opening (Austin 2018).

1.1.2 Critical Infrastructure in the USVI

The 2001 USA PATRIOT ACT defines Critical Infrastructure (CI) as, "systems and assets, whether physical or virtual, so vital to the United States that the incapacity or destruction of such systems and assets would have a debilitating impact on security, national economic security, national public health or safety, or any combination of those matters." CI systems span 16 sectors defined by the U.S. Department of Homeland Security to include lifeline systems like electric power, transportation, water, telecommunications, financial services, and healthcare among others (Cybersecurity and Infrastructure Security Agency 2020). CI systems in the USVI are shaped by the remoteness, natural geography, and history of each island.

This thesis focuses on assessing the operational resilience of fiber optic telecommunications networks that provide Internet services to the USVI. For this reason, the remainder of the thesis discusses only these systems in detail. It is important to note, however, that most CI systems are interdependent. For example, communication systems require other CI systems like electric power and transportation to function, and vice versa. We refer readers interested in learning more about other CI systems in the USVI to the following studies:

- *An Operational Model of Interdependent Water and Power Distribution Infrastructure Systems* (Bunn 2018);
- *Simulation Optimization for Operational Resilience of Interdependent Water-Power Systems in the US Virgin Islands* (Wille 2019);
- *An Operational Model of Food Supply Chain and Transportation Systems in the U.S. Virgin Islands* (Good 2019);
- *Analyzing Cell Phone Network Resilience in the US Virgin Islands* (Wine 2020); and
- *Measuring and Modeling Potable Water Demand in the United States Virgin Islands* (Borgdorff 2020).

1.2 Internet Infrastructure Systems in the USVI

The ownership, operation, and management of the USVI telecommunication network is spread out over a mixture of public and private stakeholders. These stakeholders often protect or obscure information about their networks in order to ensure network security and maintain economic competitiveness (Alderson et al. 2018). As a result, it is difficult to establish an operational understanding of how the various elements of the USVI telecommunications landscape interact with each other to form a single network.

1.2.1 Public Communications

There are two public telecommunications providers in the USVI; the services and functions of the Bureau of Information Technology (BIT) and the infrastructure and services of the Virgin Islands Next Generation Network (viNGN).

Bureau of Information Technology

BIT is the government agency that responsible for the territory's emergency communication infrastructure for first responders and for establishing, maintaining, and improving the Information Technology (IT) environment for the Government of the USVI (USVI Hurricane Recovery and Resilience Task Force 2018). Although BIT officially serves as the information technology arm of the USVI government, other government agencies are not required to use them and there are "no restrictions that prevent government agencies from acquiring their own IT capabilities, technicians, application, or contractors" (Alderson et al. 2018). As a result, USVI "has been unable to establish a common acquisitions or cybersecurity plan, create IT situational awareness across the government, or create IT commonality and interoperability between agencies" (Alderson et al. 2018). In practice, BIT approves contracts for government agencies IT requirements, runs a microwave radio network that some government agencies use for Internet access, provides data storage services at its STX offices, and operates the public safety public safety radio network (USVI Hurricane Recovery and Resilience Task Force 2018).

Virgin Islands Next Generation Network

viNGN is a semi-autonomous government agency that is a wholly owned subsidiary of the Virgin Islands Public Finance Authority (VIPFA). The agency provides a "territory-wide, fee-based broadband 'middle-mile' fiber network to private and Internet Service Providers (ISPs), which enables the ISPs to provide broadband Internet connection services to their customers" (USVI Hurricane Recovery and Resilience Task Force 2018). viNGN's middle mile network is composed of over 220 miles of terrestrial and undersea fiber optic cables and 23 Fiber Access Points (FAPs) which are buildings that connect fiber optic cable segments and house network equipment which facilitate the routing of Internet traffic (Figure 1.2 shows the exterior of a FAP). viNGN leases capacity on commercial transcontinental submarine cables that connect the USVI with Continental U.S. (CONUS). Additionally, viNGN provides Internet service to all of the territory's Community Anchor Institutions (CAIs) which include schools, libraries, medical and healthcare providers, public safety entities, community colleges and other institutes of higher learning, and other community support organizations and entities (Virgin Islands Next Generation Network December 2012).

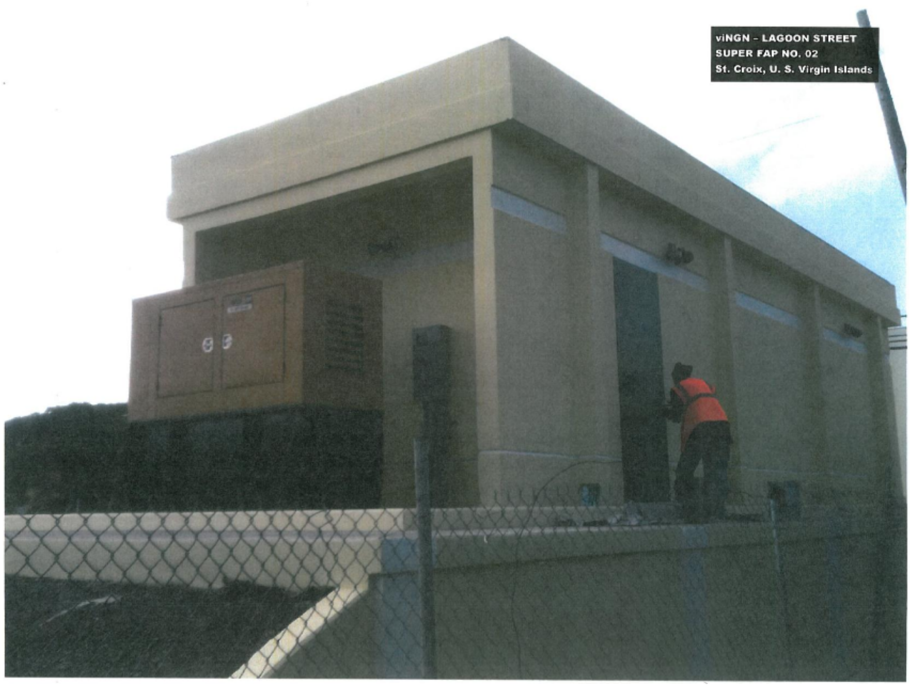

Figure 1.2. viNGN FAP on STX. Source: Virgin Islands Next Generation Network (December 2012).

1.2.2 Private Communications

There are three primary elements of the private telecommunications system in the USVI: ISPs, wireless service providers, and transcontinental submarine cables.

Internet Service Providers (ISPs)

An ISP is a private company that provides Internet service to customers via wired connections. ISPs in the USVI are comprised of three types of connections: "*backbone*, which is the high-capacity fiber (or trunk line) that carries massive amounts of data for local or regional exchange, *middle mile*, which connect the backbone to the ISPs' or telecommunications providers' core network or telecommunications exchange, and *last mile*, which delivers the data connection to customers' homes and businesses" (USVI Hurricane Recovery and Resilience Task Force 2018, emphasis added). ISPs, in general, are secretive about their networks, and it is difficult to determine the physical topology (i.e., where cables and exchanges are physically located) or logical topology (i.e., how ISPs route their Internet traffic). In the USVI, ISPs are encouraged to lease capacity on viNGN network for their backbone and middle mile connection, though they are not required to do so (Alderson et al. 2018). As a result, ISPs may own and operate redundant infrastructure. There are multiple ISPs operating in the USVI including Viya and Broadband VI.

Company	Landline telephone	Wireless/ cellular network	Cable television	Broadcast TV and radio	Internet service providers	Cable landing stations
AT&T		X				X
PR Wireless		X				
Verizon		Contracted				
Viya	X	X	X		X	
CenturyLink						X
T-Mobile		Contracted				
SmartNet					X	
BBVI					X	

Figure 1.3. Telecommunication Companies Operating in the USVI. Source: USVI Hurricane Recovery and Resilience Task Force (2018).

Wireless Service Providers

A wireless provider is a private company who provides cellular or mobile Internet service to customers via wireless connections. The wireless network consists of communication towers that wireless devices connect to, a microwave or fiber optic backhaul that connect communication towers to switching stations, and switching stations that allow wireless devices to access the Internet (USVI Hurricane Recovery and Resilience Task Force 2018). At the time of this report, five wireless providers operate in the USVI: AT&T, PR Wireless, Viya, Verizon, and T-Mobile (USVI Hurricane Recovery and Resilience Task Force 2018).

Like with ISPs, the lack of transparency in their operations makes it is difficult to determine the physical and logical topology of their networks.

Transcontinental Submarine Cables

The USVI depends on transcontinental submarine cables for sending “99 percent of all data traffic, including Internet, phone calls, and text messages, from the USVI to the rest of the world” (USVI Hurricane Recovery and Resilience Task Force 2018). The USVI’s location in the Leeward Islands makes a strategic connection point for 11 submarine cables that service the Caribbean and Americas (Telegeography 2020). ISPs and viNGN must lease capacity on these cables in order to use them.

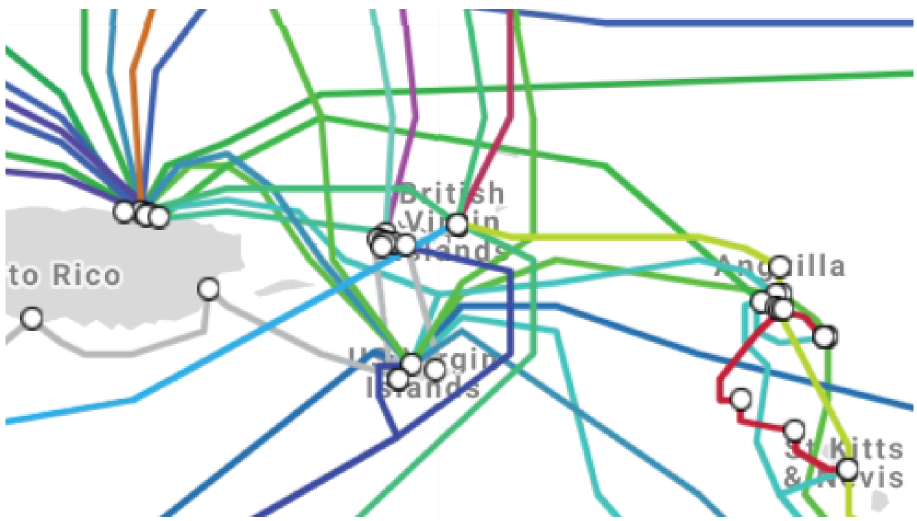

Figure 1.4. Submarine Cables in the Leeward Islands. Source: Telegeography (2020).

1.2.3 The Origins of USVI’s Internet Infrastructure

In 2009, the VIPFA submitted for and received five grants from the National Telecommunication and Information Administration (NTIA) under the Broadband Technology Opportunities Program (BTOP) and State Broadband Initiative programs to improve telecommunication infrastructure and services on the island. The most significant of these grants were the $1.29M *State Broadband Data and Development Program* (Virgin Islands Public Finance Authority 2009c), the $58.89M *viNGN Comprehensive Community Infrastructure*

Program (Virgin Islands Public Finance Authority 2009a), and the $3.02M *viNGN Public Computer Center Program* (Virgin Islands Public Finance Authority 2009b).

The *State Broadband Data and Development Program* focused on gathering, verifying, and disseminating information about broadband services in the USVI (Virgin Islands Public Finance Authority 2009c). This information was used for evaluating broadband needs, for expansion and capacity planning, and as a cost analysis tool for building the infrastructure throughout the territory. Although much of this data is kept confidential under Notice of Funding Availability (NOFA) requirements and Non-Disclosure Agreement (NDA), several geospatial files presumably produced under this grant are available on the NTIA website. Of particular use is a geospatial file of 316 CAIs and their maximum upload and download bandwidth (Virgin Islands Public Finance Authority 2009c).

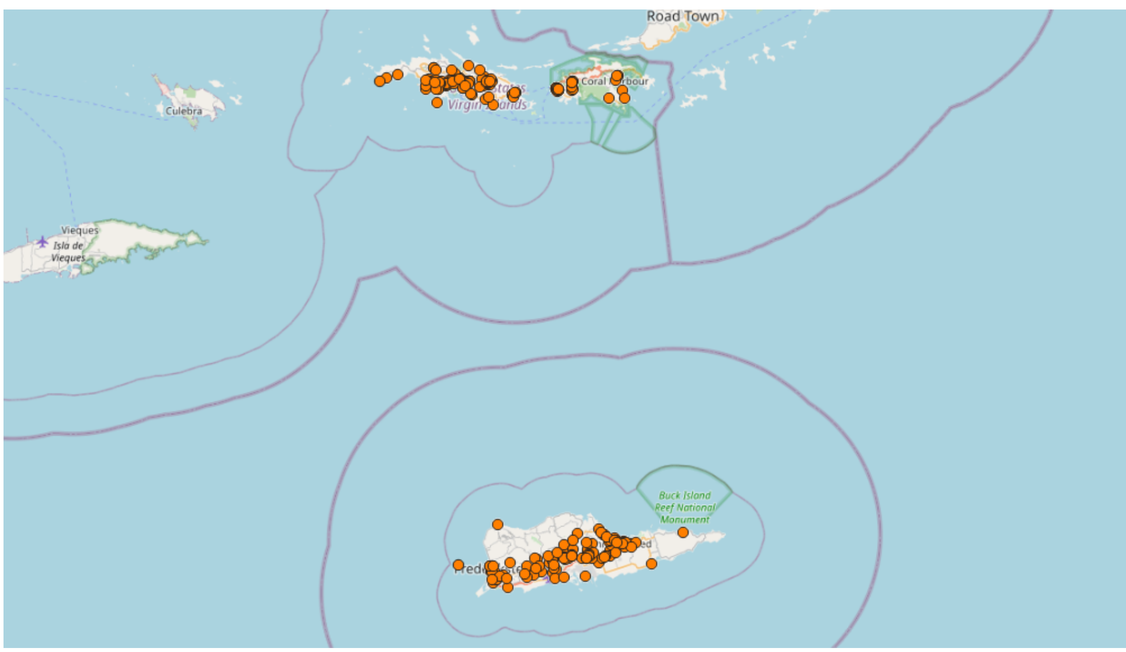

Figure 1.5. Map of CAIs in USVI. Source: Virgin Islands Public Finance Authority (2009c).

The *Comprehensive Community Infrastructure Program* proposed to deploy a viNGN-owned and operated middle mile fiber optic infrastructure throughout the territory (Virgin Islands Public Finance Authority 2009a). The proposed middle mile network consists of 3,000 miles of existing fiber optic cable and 244 miles of new construction to offer broadband speeds between 10 Megabits Per Second (Mbps) and 10 Gigabits Per Second (Gbps) to CAIs and local ISPs. The project application, quarterly and annual progress reports, and fact sheet provide insight into the physical and logical topology of the USVI Internet including cable landing sites, Internet exchanges, existing and proposed fiber optic infrastructure, viNGN FAPs, and the names and locations of 325 proposed CAIs. This project was completed in the second quarter of 2014.

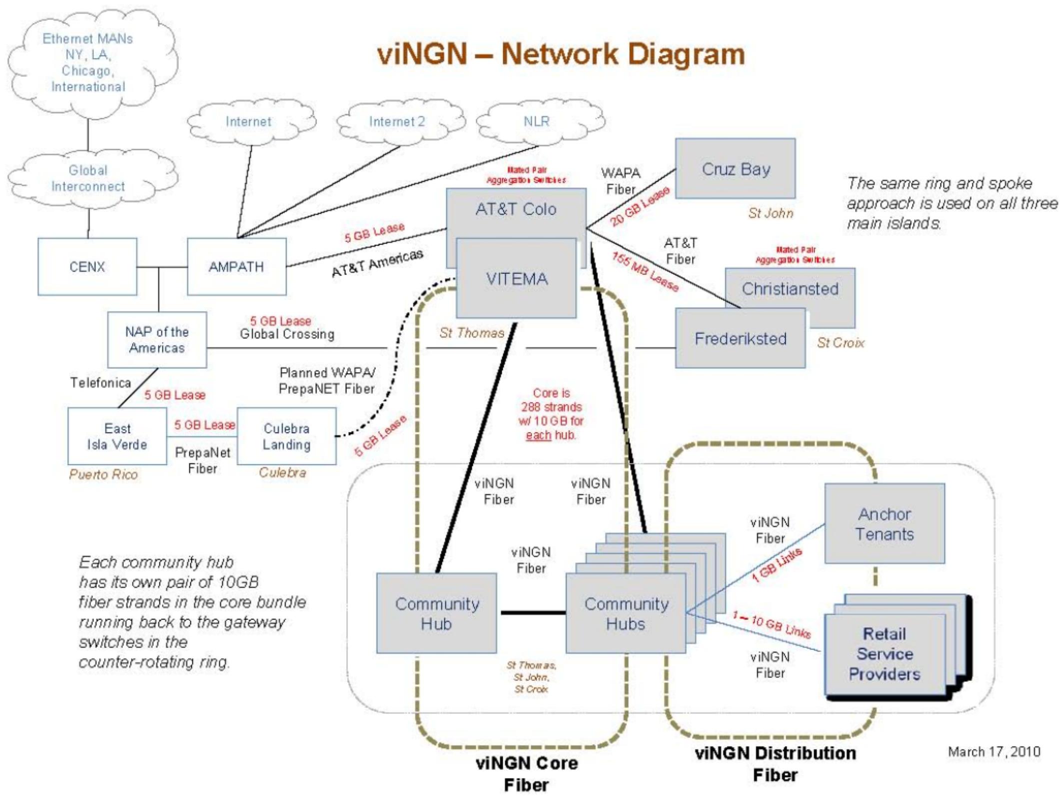

Figure 1.6. Proposed viNGN Middle Mile Network Diagram. Source: Virgin Islands Public Finance Authority (2009a).

The *viNGN Public Computer Centers Program* proposed to "deploy as many as 740 new workstations in over 40 new and five upgraded computer centers, increasing public access to computer centers by up to 1,500 hours per week and accommodating approximately 13,000 additional users each week" (Virgin Islands Public Finance Authority 2009b). A comparison between the list of Public Computing Center (PCC) and CIAs reveals that PCCs are included in the CIA list. This project was completed in the second quarter of 2014.

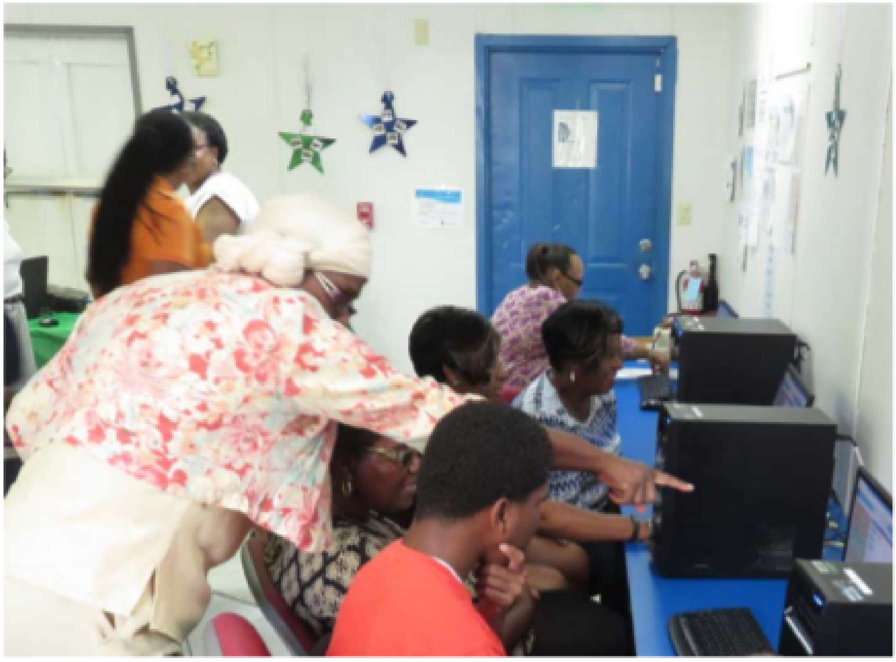

Figure 1.7. St. Andrews PCC in Charlotte Amalie. Source: USVI Hurricane Recovery and Resilience Task Force (2018).

1.3 Vulnerabilities in USVI Internet Infrastructure

The USVI's Internet system has proven to be vulnerable to both natural and human events. Hurricanes Irma and Maria collectively devastated the territory's telecommunications infrastructure which produced Internet outages lasting several months. More recently, Internet service has been disrupted in the USVI by the accidental cutting of fiber optic cables.

1.3.1 Hurricanes Irma and Maria

Both public and private telecommunication networks "sustained significant damage in Hurricanes Irma and Maria which, in many cases, resulted in systems being out of service for several months" (USVI Hurricane Recovery and Resilience Task Force 2018). Ninety percent of viNGN's aerial cable infrastructure was destroyed or rendered unusable and two FAPs suffered structural and water damage. ISPs also suffered significant damage to their last mile aerial cables which, combined with the losses to viNGN's network, prevented customers on both wired and wireless devices from accessing the Internet (USVI Hurricane Recovery and Resilience Task Force 2018). There was approximately a 60%-to-70% drop off in Internet traffic on the territory immediately following the storms. Both viNGN's buried cable infrastructure and the territory's submarine cables were largely unaffected by the storms directly but were impacted by recovery operations.

Figure 1.8. Hurricane Maria Damage on STX. Source: Reuters (2017).

Damage sustained by the territory's hardline telecommunication infrastructure significantly impacted emergency response and recovery operations. Government agencies were severely restricted in their ability to coordinate with each other, access the Internet to provide updates and receive information, or pass critical information to citizens. The lack of Internet also impacted the public's ability to contact emergency response services or provide updates to friends and family (USVI Hurricane Recovery and Resilience Task Force 2018).

1.3.2 Recent Internet Disruptions

Internet service in the USVI continues to be disrupted due to vulnerabilities in the Territory's Internet infrastructure. In September 2019, Viya, a ISP in the Territory, reported an Internet service interruption on STT when a USVI Department of Public Works (DPW) contractor severed an underground fiber optic cable (Viya 2019). In February 2020, AT&T accidentally cut viNGN cables on STT interrupting service to STJ (Lee 2020).

Incidents such as these highlight two important vulnerabilities in the USVI Internet Infrastructure. First, although viNGN provides a middle mile network for ISPs to connect through, they are not required to use it. As a result, there are multiple public and private service providers that overlap throughout the territory. Stephens Adams, the Chief Executive Officer of viNGN, noted that "on St. Thomas the overlap is greater than 70 percent and on St. Croix it more than 80 percent" (Lee 2020). Second, "Call Before You Dig" legislation, which was passed in 2015 and is designed to prevent accidental cable cuts, has never been fully implemented (Lee 2020).

1.4 Thesis Goals

The goal of this study is to produce a representative model of the USVI fiber optic infrastructure that will allow us to conduct analysis at the operational level. In order to establish an operational view of the USVI telecommunications infrastructure, we first collect and curate data on the known parts of the network. This includes gathering information on infrastructure that can visibly be observed using tools developed by the NPS Computer Science Department (Woodman 2016), information that is available in the public domain, and information that public and private entities share with us. Data is curated into geospatial data sets for modeling and analysis using the open-source Quantum Geographic Information System (QGIS) (QGIS Development Team 2019).

Once portions of the telecommunications infrastructure are determined, we fill in the gaps through synthetic network generation techniques to infer a network that represents actual USVI Internet infrastructure. Then, we identify vulnerabilities, critical points in the network, opportunities for improvement, and efforts to prioritize.

Ultimately, this study makes several contributions.

- A new, curated geospatial data set is created that aggregates publicly available and proprietary telecommunications infrastructure data into a single repository for modeling and analysis.
- A working synthetic network generation model is created in Pyomo that serves USVI communities and generates Internet network maps that are realistic, but not real.
- Network and vulnerability analysis methods are implemented and presented in a way that serves USVI stakeholder needs.

We begin with a review of relevant past work.

CHAPTER 2:
Literature Review

Significant research has aimed to improve our understanding of the Internet's complex structure and the associated Internet infrastructure system found in the USVI. In general, researchers develop frameworks for analyzing and improving Internet architectures to be more resilient (i.e., continue to function before, during, and after large-scale and surprising disruptions). More specifically, a number of technical methods developed in the literature inform the data curation, modeling, and analysis completed in this thesis. These methods include:

- Synthetic network topology generation to produce realistic Internet infrastructure systems for robustness and failure assessment;
- Internet traffic models that consider the structure, function, and relationships across layers to deliver traffic in a realistic way; and,
- Internet behavior models that capture how network routing and congestion change as infrastructure fails and disaster challenge system operations.

We review the most relevant of each in turn.

2.1 Structure of the Internet

An important starting point is to understand what the Internet is and how it works. The Internet is a global network of interconnected computer networks that communicate with each other through layered protocols. The Open Systems Interconnection (OSI) model of Internet layers is "widely accepted as a basis for the understanding of how a network protocol stack should operate and as a reference tool for comparing network stack implementation" (Murhammer et al. 1998).

In the OSI model, "each layer provides a set of functions to the layer above and, in turn, relies on the functions provided by the layer below" (Murhammer et al. 1998). Figure 2.1 provides a simple depiction of the OSI model where two Internet devices, called *hosts*, exchange information through layers. The lowest layer is the physical layer which provides physical

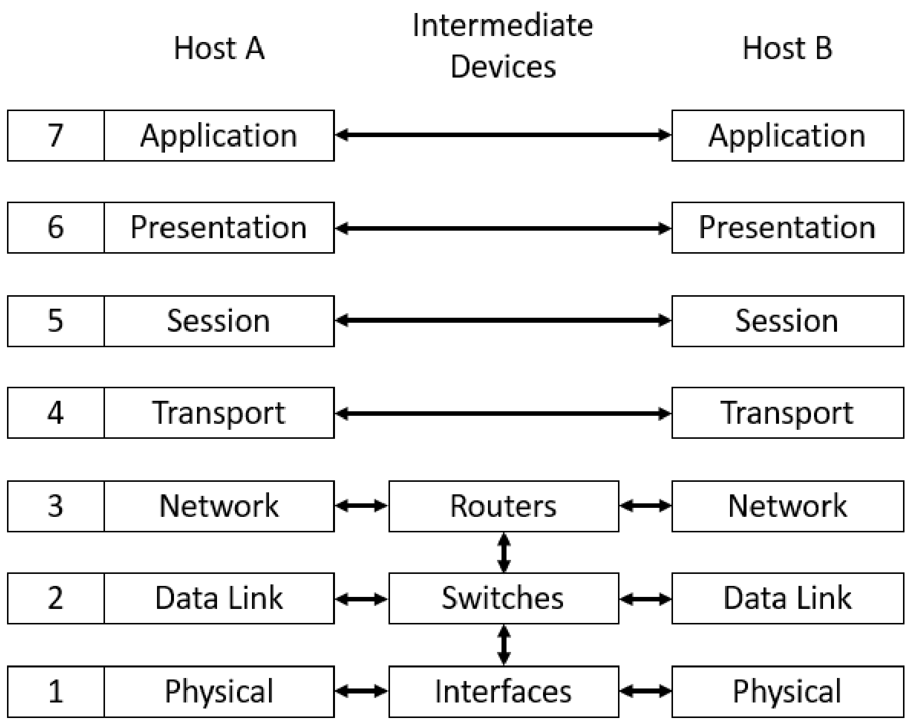

Figure 2.1. The OSI Model of Internet Layers. The OSI model provides a framework for understanding how devices, called hosts, communicate with each other through the layered Internet connections. Adapted from Murhammer et al. (1998).

pathways, such as network interfaces, cables, and hardware devices; over which raw bits of information travel. The second layer is the data link layer. This layer creates *frames* of data and provides error checking. Most switches, hardware devices which forward Internet traffic from one host to another, operate at the data link layer. The third layer is the network layer. This layer forms *packets* of information and decides which physical pathway those packets will take through routers. The fourth layer is the transport layer which establishes host-to-host connections. Layers 5, 6, and 7 are primarily concerned with how those hosts communicate with each other.

2.2 Internet Systems and Resilience

The Internet is itself a critical component of modern society that governments, businesses, and private citizens rely on to communicate. Additionally, many other CI sectors—such as energy, transportation, and potable water—rely on Internet-connected Supervisory Control and Data Acquisition (SCADA) systems for monitoring and control. Disruptions to Internet service have potentially catastrophic public safety, economic, and social consequences. It is critical, then, that Internet systems are designed to be survivable and resilient (Sterbenz et al. 2013). Sterbenz et al. (2013) define resilience as “the ability of the network to provide desired service even when challenged by attacks, large-scale disasters, and other failures.”

The principles of redundancy and diversity facilitate resilient Internet design and engineering (Sterbenz et al. 2013; Çetinkaya et al. 2012; Rohrer et al. 2014). Redundancy is the the replication of elements within a network to provide fault tolerance (Sterbenz et al. 2013). For example, a resilient network may have multiple cables connecting two Internet nodes or multiple servers available in case of disruption. If a single element is compromised, there are other elements available to ensure reliable operation. Diversity is providing alternative means to provide reliable function (Sterbenz et al. 2013). Geographic diversity is a critical component to ensure a resilient physical network; separating key network elements by distance decrease their chances of all being affected by the same event (Çetinkaya et al. 2012). Rohrer et al. (2014) demonstrates that path diversification can accurately predict the survivability of Internet topologies.

2.3 Synthetic Internet Topology Generation

There are at least two classes of methods researchers use to generate Internet network topology to estimate the structure and function of the built infrastructure system:

1. Complex Networks: Network topology generated using complex networks focus on using the statistical and graph properties of Internet systems to produce typologies that share similar characteristics.
2. Optimization: Network topology generated with optimization uses more engineering and real-world constraints on physical network structure to generate realistic networks that mimic the decisions made by owners and operators during initial infrastructure construction.

2.3.1 Synthetic Internet Topologies based on Complex Networks

Significant research has been conducted into generating Internet topologies by modeling the graphical properties of a network. Traditionally, these models were based on network nodes randomly linked together, which ignores economic and policy considerations that guide ISPs when they are designing their networks (Jabbar et al. 2008). As a result, networks produced by random graphs do not adequately represent the actual system.

Jabbar et al. (2008) proposed to incorporate location and cost parameters into existing topology generators which allow for more realistic analysis. Pure random models are the most basic form. In these models, pairs of nodes are connected with an independent probability. The locality model and Waxman model are more complex and connect pairs of nodes based on the distance between them (Jabbar et al. 2008). In the locality model, pairs of nodes are connected with a probability based on whether the distance between the nodes is above or below a certain threshold. In the Waxman model, modes are connected with a probability that takes into account the distance between the nodes, the maximum distance between any pair of nodes, and the ratio between the shortest and longest link. The Waxman model produces realistic topologies that are more representative of the actual systems than random graphs (Jabbar et al. 2008).

2.3.2 Synthetic Internet Topologies based on Optimization

Alderson et al. (2006) propose that "Internet topology at the router-level can be understood in terms of the tradeoffs between network performance and the technological and economic factors constraining design" through optimization. Some of the principle divers of network topology are the following.

- Link costs. The installation, operation, and maintenance of communication links are extremely expensive and cost tends to increase with the length of the link. As a result, network designers are economically motivated minimize the number of long distance links by aggregating traffic.
- Router technology. Routers are networking devices that forward packets of data between parts of an Internet network. Routers have a maximum number of link connections and maximum bandwidth which creates an " 'efficient frontier' of possible bandwidth-degree combinations available for each router" (Alderson et al. 2006).

Network designers are economically motivated to employ routers as close to the 'efficient frontier' as possible in order to gain the most efficiency out of their network.

- Customer constraints. Internet customers themselves drive many features of Internet networks. The concentration of Internet customers and the bandwidth that they are willing to pay for will inform the type of Internet infrastructure that network designers will install at certain locations.
- Service requirements. In order to be commercially viable, Internet networks must deliver quality network performance which this paper defines as the maximum proportional throughout of a network under heavy traffic. In other words, an Internet topology must support its customer's requirements.

Derosier (2008) utilized optimization-based reverse engineering to produce router-level topologies that capture the technical capabilities, economic constraints, operational requirements, and performance objectives faced by real ISPs (Derosier 2008). First, he analyzed existing router-level topology for Tier-1 ISPs and reversed engineered their key design principles. These principles include the types of routers that are typically located at Internet nodes and the types of links between Internet nodes.

Next, Derosier (2008) forward-engineered a network topology generation process based on his observed design principles. This process takes the population and the penetration of Internet users in the market to determine the number of Internet customers at a location and the number of access routers required to support them. He then uses that information to determine the number of backbone routers required and the number and location of backbone links between nodes.

Derosier (2008) provides a methodology to generate 'realistic, yet fictitious' network topologies that adhere to basic technological and economic constraints faced by real network designers. Perhaps most importantly, these networks are generated from readily available data such as population and geographic information.

2.4 Modeling Internet Traffic Demands

Crain (2012) and Martin (2014) both used gravity models to estimate Internet traffic demand between geographically distinct locations in order to model telecommunication networks. Gravity models are based on Newton's law of gravitation which states that objects are attracted to each other with a force proportional to their mass and inversely proportional to the distances between them squared. Gravity models can be applied to telecommunication networks by using population, number of Internet hosts, or Internet traffic entering or exiting a node, or the total traffic in a network as surrogates for mass and distance. The output of a gravity model is a traffic matrix which shows the amount of Internet traffic entering a node and the amount of Internet traffic exiting the node destine for other nodes in the network.

Here, we repeat the formulation of Crain (2012), who used a gravity model to approximate Internet traffic between countries through submarine cables. The traffic matrix is built using the estimated total traffic across the total network T^{tot} and the number of Internet hosts H_n in each country.

Index use

$n \in N$	Countries (alias s, t)

Data [units]

T^{tot}	Total traffic across the network [Gbps]
T_s	Total traffic associated with country s [Gbps]
T_{st}	Total traffic exchanged between country s and country t [Gbps]
b_s^t	Traffic from country s to country t [Gbps]
H_n	Number of Internet hosts in country n [scalar]

Formulation

$$T_s = \frac{H_s}{\sum_{n \in N} H_n} T^{tot} \qquad \forall s \in N$$

$$T_{st} = \frac{H_t}{\sum_{n \in N} H_n} T_s \qquad \forall s, t \in N$$

$$b_s^t = \frac{H_s}{H_s + H_t} T_{st} \qquad \forall s, t \in N, s \neq t$$

$$b_t^t = -\sum_{s:s \neq t} b_s^t \qquad \forall t \in N$$

Martin (2014) used a gravity model to develop an Internet traffic matrix for the fictional country of Dystopia. Dystopia is fictional country used for modeling disaster situations and does not have a T^{tot} or T_s built in. Martin (2014) estimated the population around Internet nodes, the penetration of Internet into that node, and the estimated daily use of Internet by use which is formulated as

$$\text{Daily\_Traffic}(n) = \text{Population}(n) \times \text{Penetration}(n) \times \text{Daily\_Use}(n).$$

Once the daily traffic at each node was determined, Martin (2014) multiplied the daily traffic for each node by every other node's traffic demand to produce a traffic matrix.

2.5 Modeling Internet Behavior

Alderson et al. (2015) describe a method for assessing the operational resilience of CI systems using Attacker-Defender Models (Alderson et al. 2015). The system is first viewed from the operator's perspective. The operator may be a human, automated system, or a combination of the two that makes decisions about the behavior of the system in order maximize objectives while minimizing expenses subject to constraints. In this context, the system's objectives are what the operator wants the system to do (e.g., minimize the cost of moving a commodity through a system). The systems constraints are what it can do (e.g., each arc can only move a certain amount of the commodity). Alderson et al. (2015) formulate the *Operator Model* as a constrained optimization model:

$$\min_{y \in Y(\hat{x})} = f(\hat{x}, \mathbf{y})$$

where $\hat{x}$ is a vector that represents the state of the system, the set $Y(\hat{x})$ represents the possible actions the operator can take given the state $\hat{x}$, and $f(\hat{x}, \mathbf{y})$ is a function that measures the performance of action $\mathbf{y}$ (Alderson et al. 2015). This produces a baseline solution to the Operator Model.

The system is then viewed from the attacker's perspective. Attacker in this sense represents any source of disruption to the system whether it be natural (i.e., hurricanes) or man-made (i.e., terrorism). Alderson et al. present an example where an attacker has the ability to target a single link in the system which will maximize the costs incurred by the operator. He formulates the *Attacker Model* as:

$$\max_{x \in X} \min_{y \in Y(x)} = f(\mathbf{x}, \mathbf{y})$$

where $\mathbf{x}$ now represents a decision variable belonging to the attacker and X represents the set of possible actions the attacker can take. The operator still faces the same minimization function as before but the state parameter $\hat{x}$ has become a decision variable for the attackers. This form of optimization is an *Attacker-Defender Model*. Brown et al. (2006) and Alderson et al. (2013) apply *Attacker-Defender Model* to defending critical infrastructure defense.

2.6 Our Contribution

This thesis produces a synthetic network topology of the USVI's Internet infrastructure that embodies the functions of the actual system. This is accomplished by curating publicly available data about the USVI's Internet infrastructure into geospatial data sets in QGIS to produce a graph of the network. We apply graph theory topological generation methods found in Sterbenz et al. (2013) and Rohrer et al. (2014) to produce sets of node pairs that describe how the Internet functions and utilize optimization methods identified by Alderson et al. (2006) and Derosier (2008) to analyze its performance. Finally, we conduct 'what if' style analysis on the network to study how the network performs against disruptions.

CHAPTER 3: Data Curation and Model Development

We generate a synthetic model of the USVI Internet infrastructure with methods for data creation and development, network generation, and multilayer network analysis. Together, these methods enable network traffic assignment and vulnerability assessment.

3.1 Data Curation

We curate publicly available information about the USVI telecommunications infrastructure and develop geospatial data sets that capture the physical location and other key attributes of Internet infrastructure. The USVI Internet is comprised of several related network component types which are organized into layers to transport data to and from network customers. Traffic is generated by Internet customers who purchase Internet access from commercial ISPs to upload or download data. Last mile Internet connections connect Internet customers to middle mile fiber optic infrastructure. The middle mile network achieves global connectivity through backbone connections to Network Access Points (NAPs) in the continental United States. We develop geospatial data sets for each node and edge that comprise these interrelated layers.

3.1.1 Internet Customers

Internet customers in the USVI are divided into two categories; private Internet customers and public Internet customers.

Private Internet Customers

Private Internet customers are households that purchase Internet access from an ISP. Information for private Internet customers within the USVI was derived from previous work on supply chain modeling in the USVI (Good 2019). The USVI tracks census information by estates which are smaller than traditional census sub-districts but serve the same purpose. Good (2019) created population nodes for STX by dividing the island's 212 estates into a geospatial file containing 233 nodes located along surface streets. Figure 3.1 provides a visual depiction of these nodes representing private Internet customers on STX.

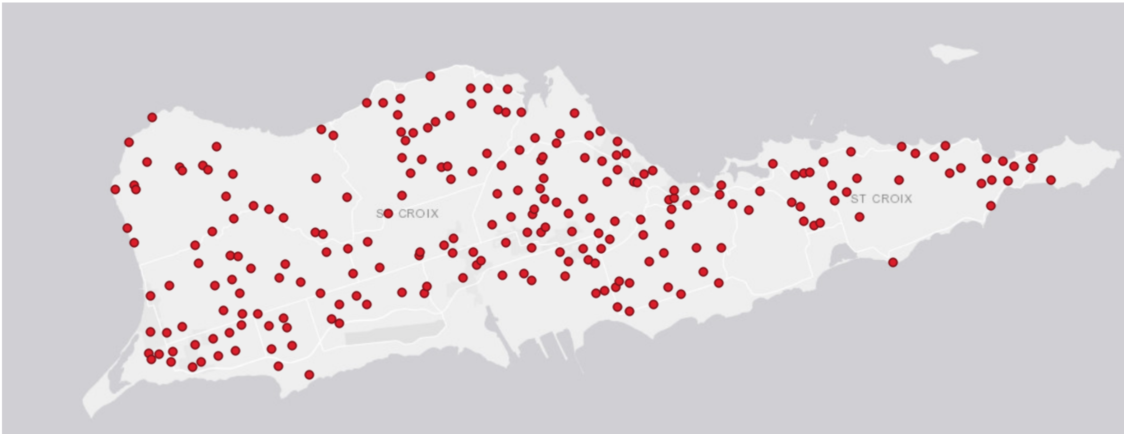

Figure 3.1. St. Croix Private Internet Customers. There are 233 nodes on STX, depicted as red circles, representing population centers in estates. Each population node represents a group of private Internet customers and is annotated with attribute information such as the population and number of households.

Public Internet Customers

Public Internet customers include both CAIs and PCCs. The viNGN website maintains a list of 25 PCCs which are currently in operation, and VIPFA's final progress report for the viNGN *Comprehensive Community Infrastructure Program* identifies 316 CAIs that successfully connected to the network and (Virgin Islands Public Finance Authority 2009a). A comparison between the PCC and CAI lists determined that all PCCs are located at CAIs.

A geospatial file containing all territorial CAIs was obtained from the *State Broadband Data and Development Program* documents (Virgin Islands Public Finance Authority 2009c). The file coded CAIs with one of seven category codes relating the facility to its purpose to the USVI government (see Table 3.1). The file was imported into QGIS, split by zip code, and then re-compiled into three separate layers, one for each of the USVI main islands. Several of the CAIs contained inaccurate zip codes and Global Positioning System (GPS) coordinates which required updating. Figure 3.2 provides a visual depiction of CAI nodes on STX.

Table 3.1. Community Anchor Institution Categories

CAI Category	Type/Services
1	School K-12
2	Library
3	Medical/Healthcare
4	Public Safety
5	University, College, other Post Secondary School
6	Other Community Support— Governmental
7	Other Community Support— Non-Governmental

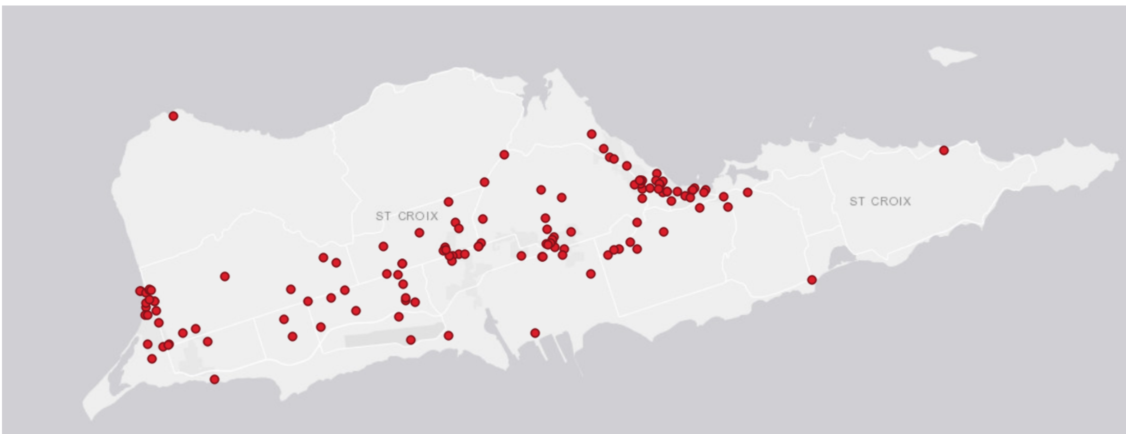

Figure 3.2. St. Croix Public Internet customers. STX has 133 CAIs, depicted as red circles. We annotate each CAI node with attribute information including type and maximum upload and download data rate.

3.1.2 Last Mile Infrastructure

Last mile infrastructure consists of ISP or viNGN fiber optic cables which connect public and private Internet customers to FAPs on the middle mile network. Fiber optic infrastructure tends to follow roads because they provide an inexpensive right-of-way where cables can be hung on utility poles or buried (Çetinkaya et al. 2015). The transportation network developed in Good (2019) was adopted to represent possible last mile Internet infrastructure segments. First, we found the closest node on the transportation network to each CAI using a distance matrix in QGIS and then generated arcs connecting those CAIs to their closest transportation node. Next, FAPs were connected to the transportation network using the same process. Finally, the start and end point of each segment were added to it's attribute table. This method provided a geospatial representation of possible last mile infrastructure segments with their source and and destination. Figure 3.3 provides a visual depiction of STX's last mile infrastructure.

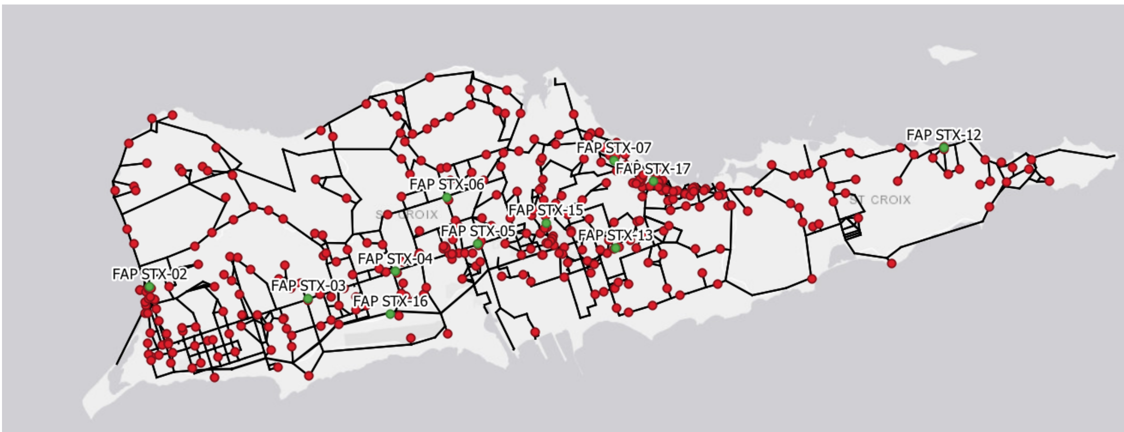

Figure 3.3. St. Croix Last Mile Infrastructure. Black lines represent fiber optic cables, red circles represent public and private Internet customers, and green circles represent FAPs.

3.1.3 Middle Mile Infrastructure

The middle mile is the core of the viNGN open service network. It consists of FAPs interconnected by rings of fiber optic cable owned and operated by viNGN. These cables include aerial and buried conduit on each island and submarine cables which connect STX, STT, and STJ.

Facility Access Points

Available information suggests that 23 FAPs were constructed at various locations along the middle mile (Virgin Islands Public Finance Authority 2009a; Virgin Islands Next Generation Network December 2012). The locations and presence of these FAPs have all been verified by faculty, staff, or student site visits. Two FAPs, STX-02 in Frederiksted on STX and STT-01 at the Virgin Islands Territorial Emergency Management Agency (VITEMA) headquarters on STT, are designated as “Super FAPs” which contain the infrastructure to power the viNGN submarine cables between islands. Figure 3.4 provides a visual depiction of each FAP on STX.

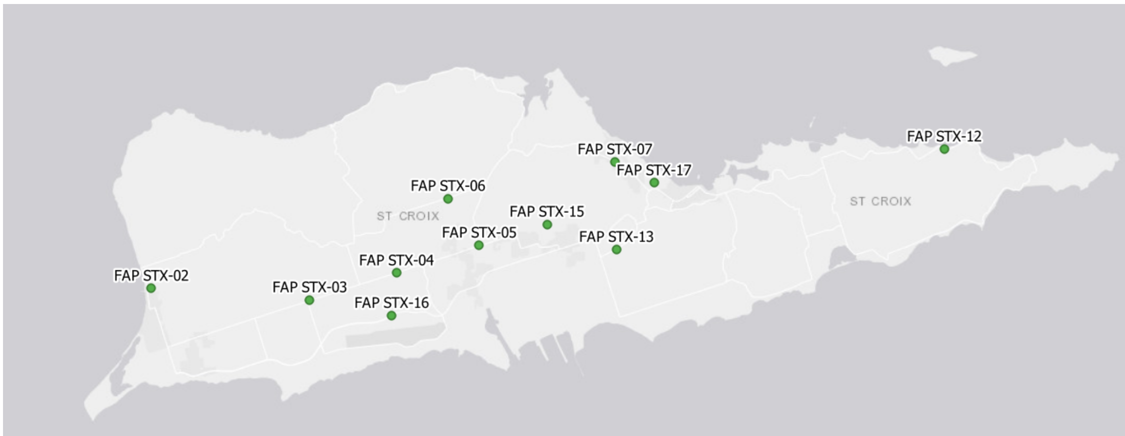

Figure 3.4. St. Croix Facility Access Point Locations. There are 11 FAPs on STX, represented as green circles, that connect Internet customers to middle mile fiber optic cables. FAP STX-02, located on the west coast of STX, is a Super FAP that contains the infrastructure to power the viNGN submarine cables between the islands.

Terrestrial Fiber Optic Cables

viNGN’s terrestrial fiber optic cables physically connect FAPs to each other. Available evidence suggests that there is over 3,000 miles of fiber optic cable in the viNGN middle mile (Virgin Islands Public Finance Authority 2009a; Virgin Islands Next Generation Network December 2012). Fiber optic cables are either buried, mounted on wood poles, or mounted on composite poles. Different segments of fiber optic cable connect to each other through junctions. Figure 3.5 provides a visual depiction of the terrestrial fiber optic cables on STX colored by type. A similar ring exists on STT connecting to STJ.

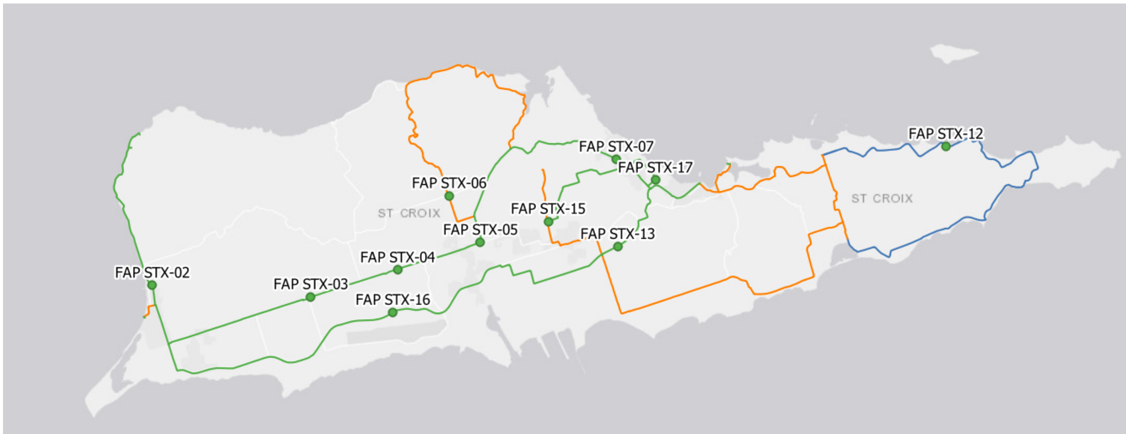

Figure 3.5. St. Croix Terrestrial Middle Mile Fiber Optic Cables. Middle mile terrestrial fiber optic cables connect FAPs across STX. Buried cables are represented as green lines, wood pole cables as orange lines, composite poles as blue lines, and FAPs as green circles. Cable junctions are omitted for clarity.

Submarine Fiber Optic Cables

viNGN's submarine fiber optic cables physically connect the middle mile fiber optic rings on STX, STT, STJ to each other. Applications for the submarine cable landing permits filed in 2012 show the locations of the cable landing sites on Western and Eastern sides of STX and STT with dual connections between STT and STJ.

Figure 3.6 depicts the viNGN submarine cable system as planned in original public documents. However, conversations with USVI telecommunication stakeholders suggest only portions of this cable system were installed and/or are currently in use (Federal Communications Commission 2012). Specifically, stakeholders suggest that the cables connecting western and eastern STX and eastern STX to the eastern STT are either not installed or are not in use.

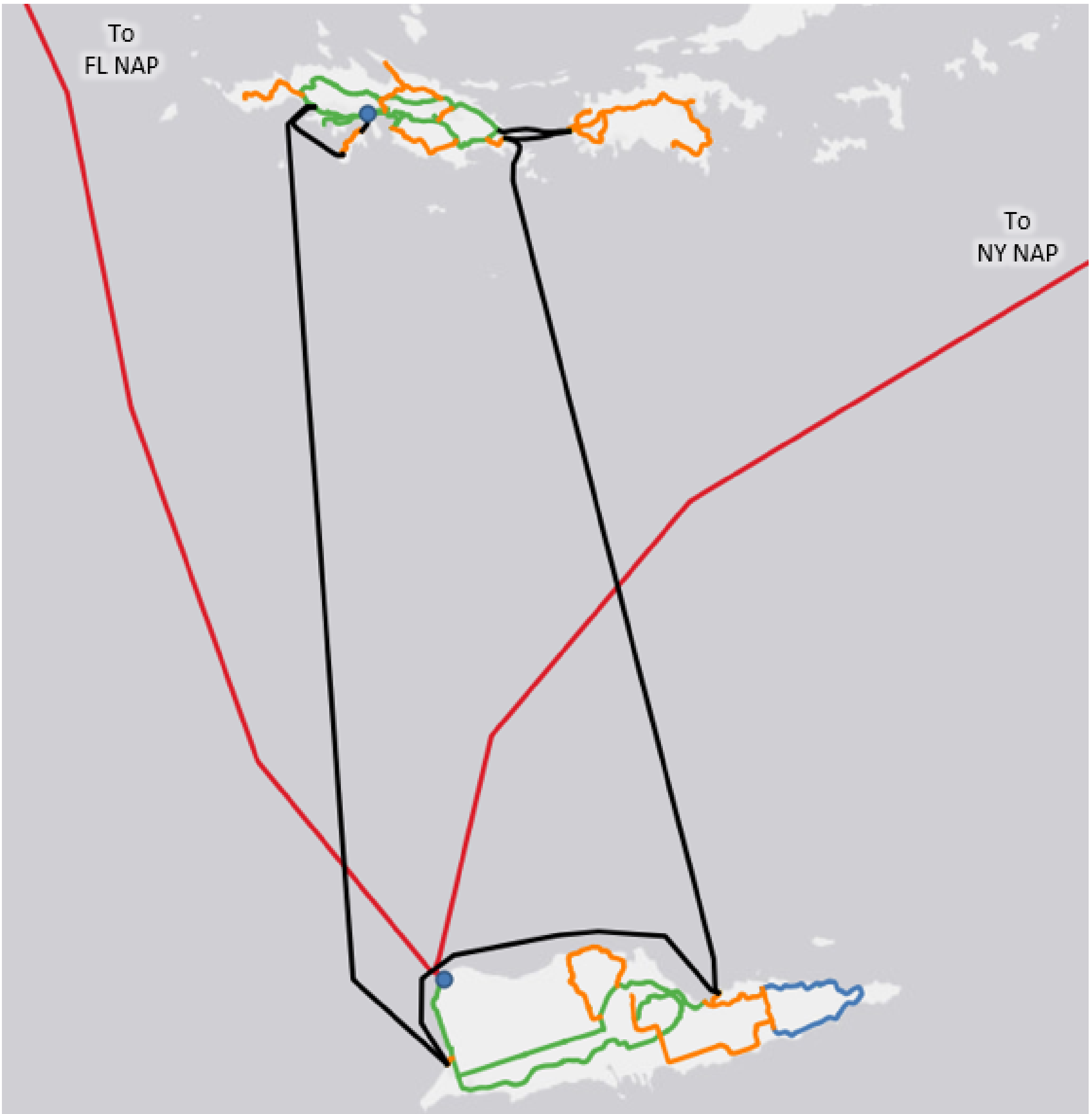

Figure 3.6. USVI Internet Exchange Points, Submarine Middle Mile, and Global Connectivity. Internet traffic is sent by middle mile cables to Internet Exchange Points (IXPs) on STX and STT, then underwater to NAPs for global connectivity. viNGN submarine middle mile submarine cables are represented as black lines, underground cables as green, aerial cables on wooden poles as orange, and aerial cables on composite poles are blue. IXPs are represented as blue circles. Backbone connections to the continental U.S. are represented as red lines, with Florida and New York NAPs located off the map to the west.

Local Internet Exchange Points (IXPs)

The final component of the viNGN middle mile network are IXPs. IXPs are co-location facilities where intra-territory Internet traffic is exchanged between different ISPs. They receive Internet traffic from Super FAPs through the territory's middle mile infrastructure. There are two IXPs in the USVI: one at the Global Crossing Center on STX located near the Frederiksted Super FAP and at the AT&T co-location center on STT near the VITEMA Super FAP (Virgin Islands Public Finance Authority 2009a). Unless the source and destination are both within the same on-island ISP, all Internet traffic in the USVI must run through an IXP before it reaches its intended recipient.

3.1.4 Backbone Connection and Global Connectivity

The USVI telecommunication backbone connects to the rest of the world via submarine cables (Virgin Islands Public Finance Authority 2009a). viNGN leases dark fibers on the cable system operated by Mid-Atlantic Crossing (formerly Global Crossing); this system connects STX to Florida and to New York via two separate physical submarine conduits at 10 Gbps each. On the mainland, viNGN leases space at NAPs in Florida and New York (Virgin Islands Public Finance Authority 2009a). Like IXPs, NAPs are facilities where ISP exchange Internet traffic. USVI Internet customers access the global Internet through the Florida and New York NAPs.

3.2 A Network Model of the Internet

We develop methods to represent the Internet as a multi-layer network where individual network layers provide distinct networking services by creating physical and virtual connections. Throughout this work, we model the Internet as a multilayer directed graph $G(N, E)$, where $n \in N$ are the nodes in the Internet and $(s, d, l) \in E$ are edge-tuples representing the source-destination pair (s, d) of nodes and the associated network layer (l). Internet nodes exist at all layers, but may not operate at all layers. In contrast, edges only exist at a single layer. Edges at all layers are assumed to be bidirectional and full-duplex, but are modelled as directed arcs with upload and download traffic non-interfering.

Our methods focus on the first four layers of the OSI Internet model (i.e., $l \in \{1, 2, 3, 4\}$). Each Internet node and network layer provides distinct functionality for system operation.

3.2.1 Model Primitives for OSI Layers 1-4

Layer 1 ($l = 1$) is comprised of fiber *runs* connected at network *junctions*. A fiber run refers to the real-world fiber optic cable that comprises the physical backbone of the Internet. When a fiber run is installed, it is typically in semi-rigid tubing called *conduit* containing large bundles (e.g., 288 fibers in a typical underground conduit in the USVI). The fibers in a conduit may "share fate," in the sense that if dug up by a backhoe one expects them all to be severed. Long runs may require optical repeaters, which use electricity to amplify the optical signal. Changing the curvature of a fiber (e.g., when an aerial cable swings in the wind) will change the ability of the fiber to transmit light, possibly resulting in lost data.

Layer 2 ($l = 2$) is comprised of *links* that connect network *switches*. A link is comprised of one or more layer-1 fiber runs. A link is available if and only if all of its underlying fiber runs and junctions are available. The link transmits frames which include error detection.

Layer 3 ($l = 3$) is comprised of *segments* that connect network *routers* and/or *hosts*. Hosts are the physical endpoints of the Internet (e.g., computers, servers) that are used to run applications. A segment at layer 3 is comprised of one or more layer-2 links. The segment is fully available as long as the layer-2 graph is connected. Alternatively, the segment is fully available as long as the layer-2 graph is not partitioned. A segment on the Internet is reachable from the rest of the Internet via advertised routes. The global Internet advertises approximately 0.5 million routed segments.

Layer 4 ($l = 4$) is comprised of traffic *flows* between *hosts*. Applications (running on hosts) are the logical endpoints of the Internet, which generate and consume network traffic. Flows at layer 4 traverse one or more segments at layer 3. Hosts connect using flows to upload and download data. Uploading and downloading occurs at different rates.

A Simple Model of the USVI Internet

Figure 3.7 provides a depiction of a simplified multi-layer Internet network consistent with the system in the USVI. In Figure 3.7, nodes are represented as grey boxes and given names of real locations and infrastructure types found in the USVI. Colored circles represent the layers of the OSI model at which the nodes operate.

Colored lines in Figure 3.7 represent edges between nodes at different OSI layers. The black

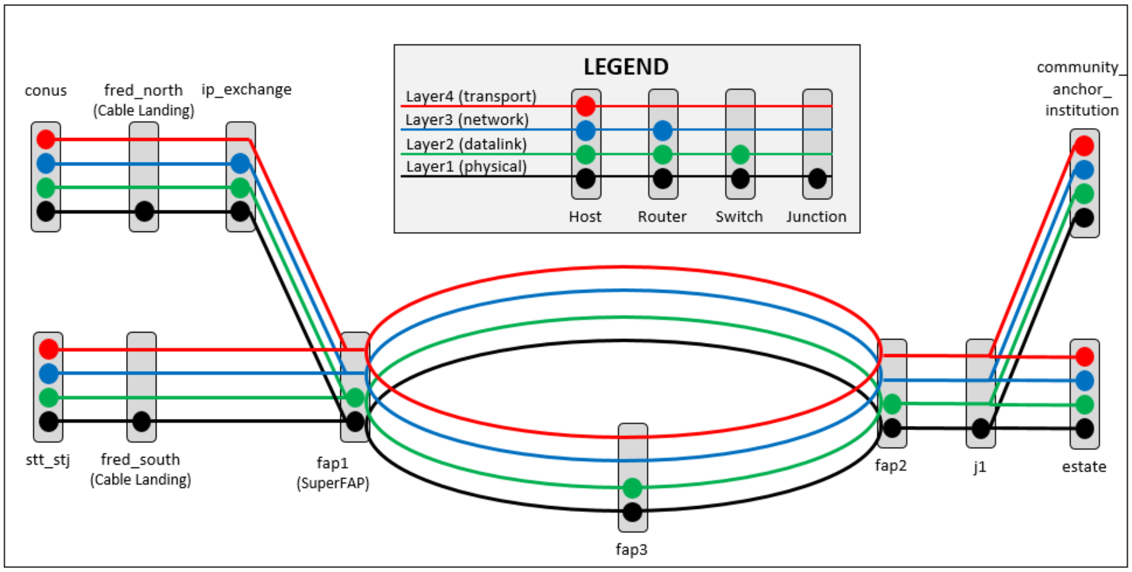

Figure 3.7. Simple Model of the USVI Internet System. Grey rectangles represent nodes that connect physical infrastructure together. There are four node types: junction, switch, router, and host; these are indicated by the colored circles representing the layer at which each node operates. Colored lines represent the connectivity (network edges) between layers of the OSI model. There are four edge types: layer-1 runs (black), layer-2 links (green), layer-3 segments (blue), and layer-4 flows (red). All nodes operate in physical layer (black). However, each node does not take part in all OSI Internet layers. Nodes that do not operate at higher layers in the stack act as pass-through nodes and are ignored during analysis of that layer. For example, CAI and FAP 2 nodes operate at layer 2 and may have a link connecting them. All nodes between them that do not operate at layer 2 (i.e., 'j1') act as pass-through for this link.

edges are layer-1 *runs* between nodes. Every node in the simple model has a physical run to at least one other node represented in the real world as a fiber optic cable. Every node exists in layer 1 and may operate at higher layers within the Internet. Every node operates at all layers under its highest level (e.g., a node that operates at layer 4 must also operate at layer 3, layer 2, and layer 1).

The green edges are layer-2 *links* between nodes. Some junctions do not provide switching operations and are pass-through nodes in layer 2. For example, consider the layer-2 link between 'stt_stj' and 'fap1' on the bottom left of Figure 3.7. The link runs through an intermediate junction, 'fred_south', even though 'fred_south' does not operate at layer 2.

This is because the layer-2 link *physically* runs over 'fred_south' but the link *logically* does not recognize the node.

Blue edges are layer-3 *segments* between routers and red edges are layer-4 *flows* between *hosts*. Unlike layer 2, only a single link-path must be available to connect routers along a layer-3 segment. Similarly, only a single segment-path must be available to connect layer-4 hosts.

3.2.2 Internet Node Attributes

Table 3.2 defines node attributes that capture the function of actual Internet infrastructure and enable network modeling. In general, all nodes in the network graph G are given a unique ID labeled with the OSI layers in which it operates (as a Boolean). Junctions and routers provide no additional service to the network and do not require specialized attributes. We assign additional attributes depending on whether the node is a switch or a host.

Every switch node has an *uplinks* attribute which provides a prioritized layer-2 link that the switch will use for transmitting Internet traffic. In general, a network switch only uses a single uplink at a time, but has multiple uplinks it can choose from for transmitting traffic.

Network traffic results from a combination of flows moving between hosts in both the upload and download direction. We estimate the aggregate traffic flows using average and peak demand measures. We also add additional attributes used for demand calculation and future analyses. These include population at a node, whether the node requires electricity to operate, and whether the node is available for network operation (e.g., during a disaster).

Table 3.2. Internet Node Attributes

Attribute	Description
Node ID (string)	A unique identifier for each node.
Node Type (integer)	Value (1-8) assigned based on the physical services the node provides. Each node is only one type.
is_layer1 (boolean)	'True' if node operates in layer 1 (junction), 'False' otherwise.
is_layer2 (boolean)	'True' value if node operates in layer 2 (switch), 'False' otherwise.
is_layer3 (boolean)	'True' value if node operates in layer 3 (router), 'False' otherwise.
is_layer4 (boolean)	'True' value if node operates in layer 4 (host), 'False' otherwise.
Population (integer)	The number of people that reside at that node, 0 if not population center.
Average Demand (float)	If host node, calculated average Internet demand in Gbps.
Peak Demand (float)	If host node, calculated peak Internet demand in Gbps.
Requires Power (boolean)	'True' if node requires power to operate, 'False' otherwise.
Backup Power (integer)	If node requires power, the number of hours the node can operate without electricity, 0 otherwise.
Uplinks (list of dictionaries)	If node is a switch, a prioritized list of links the node can connect with in OSI layer 2, *Null* otherwise
Available? (boolean)	'True' if node is installed and operating, 'False' otherwise. False values apply to proposed equipment never installed and non-functional installed equipment.
Geometry (string)	A formatted geometry string for geospatial representation.

Simple Model Node Attributes

We develop methods to assign node attributes for the USVI Internet model based on Table 3.2. We demonstrate these methods for the simple network in Figure 3.7. Figure 3.8 visually presents these nodes. Each node is assigned 14 attributes based on realistic function and characteristics that junctions, switches, routers, and hosts have in the USVI.

Several node attributes are assigned and calculated based on USVI data. First, each nodes is assigned one of the following node *types* based on the primary functions completed by the node when operating in the USVI Internet: (Type 1) Population Node, (Type 2) CAI Node, (Type 3) FAP Node, (Type 4) Super FAP Node, (Type 5) Junction Node, (Type 6) IXP Node, (Type 7) Submarine Cable Landing Site, (Type 8) NAP node.

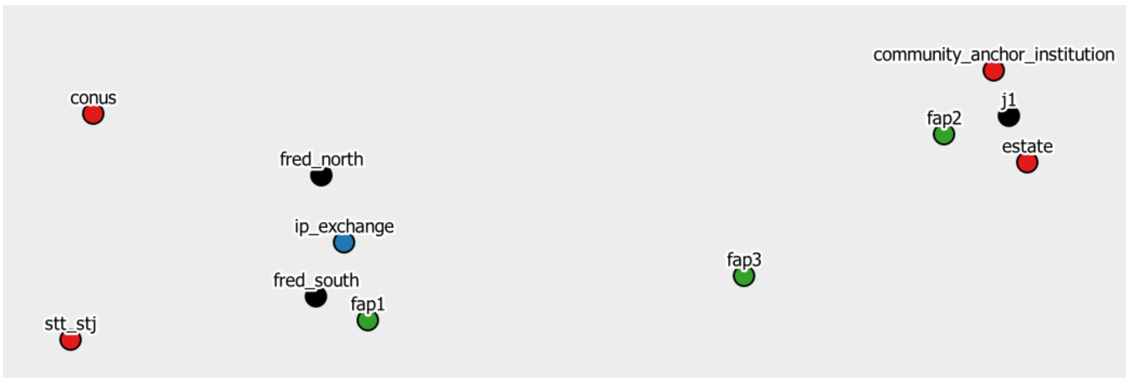

Figure 3.8. Node Types for Simple Model. The color of each node corresponds to the highest level of the OSI model at which each node operates. Black nodes are junction nodes, green nodes are switch nodes, blue nodes are router nodes, and red nodes are host nodes.

Then, we calculate demands for host nodes that drive Internet traffic. Specifically, we calculate demand at population nodes (Type 1) as:

$$\text{avg\_demand}_p = \text{Population}_p \times \text{avg\_use}_p \quad \forall p \in N \tag{3.1}$$

$$\text{peak\_demand}_p = \text{Population}_p \times \text{peak\_use}_p \quad \forall p \in N \tag{3.2}$$

where $p \in N$ is a Type-1 population node. We assume $\text{avg\_use}_p = 0.001$ Gbps and $\text{peak\_use}_p = 0.005$ Gbps for STX communities.

We calculate demands for CAI nodes (Type 2) as:

$$\text{avg\_demand}_c = 1 \text{ Gbps} \tag{3.3}$$

$$\text{peak\_demand}_c = 10 \text{ Gbps} \tag{3.4}$$

where $c \in N$ is a Type-2 CAI node.

Finally, we assign uplinks for layer-2 nodes to determine prioritized layer-2 links that the node will use when sending Internet traffic. Uplinks are manually added to layer-2 nodes. Each uplink is given a priority, where hosts are assigned a single, primary uplink and FAPs and IXPs are given primary and alternate uplinks. This method captures the redundant routing options within the middle mile infrastructure and the operational fact that uplink routes are generally assigned by hand for FAPs. Table 3.3 provides the node type and layer attributes, Table 3.4 provides the average and peak demand of population and CAI nodes, and Table 3.5 provides the uplinks.

Table 3.3. Simple Model Node Type and Layer Attributes

Node ID	node_type	is_layer1	is_layer2	is_layer3	is_layer4
fap1	4	TRUE	TRUE	FALSE	FALSE
fap3	3	TRUE	TRUE	FALSE	FALSE
fap2	3	TRUE	TRUE	FALSE	FALSE
ip_exchange	6	TRUE	TRUE	TRUE	FALSE
j1	5	TRUE	FALSE	FALSE	FALSE
fred_north	7	TRUE	FALSE	FALSE	FALSE
fred_south	7	TRUE	FALSE	FALSE	FALSE
estate	1	TRUE	TRUE	TRUE	TRUE
stt_stj	1	TRUE	TRUE	TRUE	TRUE
conus	1	TRUE	TRUE	TRUE	TRUE
community_anchor_institution	2	TRUE	TRUE	TRUE	TRUE

Table 3.4. Simple Model Population and CAI Node Demand Attributes

Node ID	average_demand (Gbps)	peak_demand (Gbps)
estate	0.0005	0.5
stt_stj	0.002	2
conus	0.05	50
community_anchor_institution	1.0	10.0

Table 3.5. Simple Model Node Uplink Attributes

Node ID	primary_uplink	alternate_uplink
fap1	(‘fap1’,‘ip_exchange’)	None
fap3	(‘fap3’,‘fap1’)	(‘fap3’,‘fap2’)
fap2	(‘fap2’,‘fap1’)	(‘fap2’,‘fap3’)
ip_exchange	(‘ip_exchange’,‘conus’)	None
j1	None	None
fred_north	None	None
fred_south	None	None
estate	(‘estate’,‘fap2’)	None
stt_stj	(‘stt_stj’,‘fap1’)	None
conus	(‘conus’,‘ip_exchange’)	None
community_anchor_institution	(‘community_anchor_institution’,‘fap2’)	None

3.2.3 Internet Edge Attributes

Once the nodes have been established, we build edges for layer 1, layer 2, layer 3, and layer 4 as four separate files in QGIS. We assign each edge nine attributes to integrate each edge into a single master edge list. Edge attributes are presented in Table 3.6.

We assign several edge attributes for network analysis based on edge type. All edges share some similar attributes. Whereas nodes can operate at multiple OSI layers simultaneously, each edge operates only at a single layer. We assign an integer value representing the layer (1-4) for the edge. We also assign a boolean variable representing whether the edge is available or unavailable and a geometry string for geospatial representation.

Some attributes are layer-specific. For layer-1 edges, we assign the medium and type, primarily for future analysis that considers edge materials and physical orientation.

For layer-2 edges, we assign the bandwidth provided for network traffic. We calculate provided bandwidth differently depending on the source and destination of the edge. We

Table 3.6. Internet Edge Attributes

Parameter	Description
Source (string)	ID of source node.
Destination (string)	ID of destination node.
Layer (integer)	The OSI layer of the edge.
Component Arcs (list)	If layer 2 edge, list of layer 1 runs associated with the link. *Null* otherwise.
Medium (integer)	Transmission medium (e.g., 1: fiber optic, 2: coaxial, etc.)
Type (integer)	If layer 1 edge, physical orientation (e.g., 1: wood pole, 2: composite pole, etc.)
Provided Bandwidth (integer)	If layer 2 edge, bandwidth available for Internet traffic.
Available? (boolean)	'True' if edge installed and operating, False, otherwise.
Geometry (string)	Formatted string for geospatial representation.

calculate provided bandwidth for edges from population nodes to FAP as:

$$\text{provided\_bandwidth}_{p,f} = \text{Population}_p \times \text{host\_upload}_p \quad \forall p \in N \tag{3.5}$$

$$\text{provided\_bandwidth}_{f,p} = \text{Population}_p \times \text{host\_download}_p \quad \forall p \in N \tag{3.6}$$

where $p \in N$ are population nodes and $f \in N$ are FAP nodes. Population_p is the population at p, host_upload_p is the advertised upload speed of a residential Internet connection (1 Gbps) at p, and host_download_p is the advertised download speed of a residential Internet connection (5 Gbps) at node p.

The provided bandwidth for edges from CAI nodes to FAP nodes are calculated as:

$$\text{provided\_bandwidth}_{c,f} = \text{cai\_upload}_c \quad \forall c \in N \tag{3.7}$$

$$\text{provided\_bandwidth}_{f,c} = \text{cai\_download}_c \quad \forall c \in N \tag{3.8}$$

where $c \in N$ are CAI nodes and $f \in N$ are FAP nodes. Also, cai_upload_c and cai_download_c are the published upload and download speeds for CAI nodes (10 Gbps).

Simple Model: Layer-1 Edges. Figure 3.9 depicts the layer-1 edges in the simple model. Layer-1 edges represent physical cable runs between nodes. The simple model layer-1 edges mimic the design of a portion of the actual USVI Internet infrastructure where population and CAI nodes connect to a nearby FAP, and FAPs are joined together by a ring. Node 'fap1' is the super FAP and serves as a gateway for all FAP traffic and connects to the IXP 'ip_exchange' which provides a physical path to the rest of the Internet depicted as 'conus'.

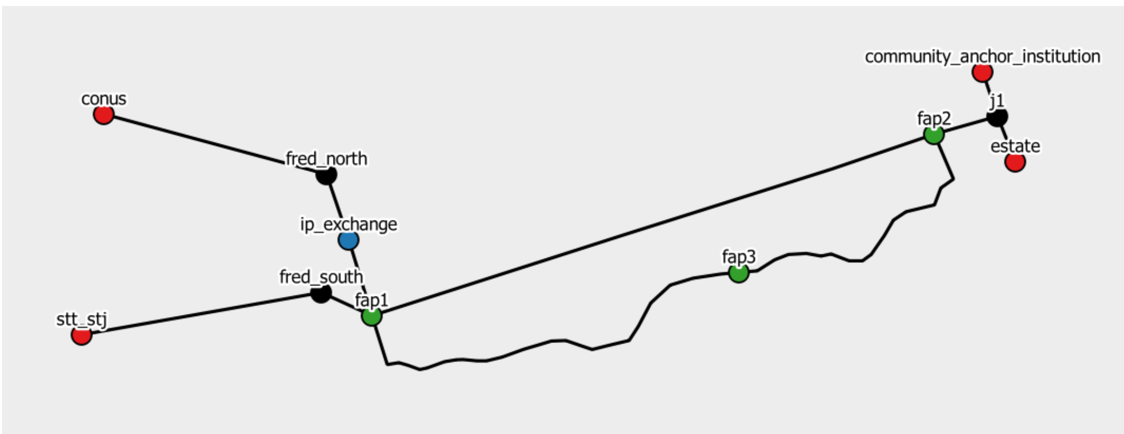

Figure 3.9. Simple Model: Layer-1 Edges. Layer 1 edges, represented as black lines, physically connect all nodes in the simple model.

Simple Model: Layer-2 Edges. Layer-2 edges represent the links between nodes that perform switching functions. Although layer-2 links logically ignore nodes that only perform layer-1 functions, we represent layer-2 links in Figure 3.10 as following the layer-1 paths that they operate through. Nodes that perform layer-2 functions are assigned a primary layer-2 and sometimes an alternate layer-2 uplink to adjacent nodes. For example, 'fap1' only has a primary uplink through layer-2 edge ('fap1', 'ip_exchange'). On the other hand, 'fap2' has a primary uplink on layer-2 edge ('fap2', 'fap1') and an alternate uplink on layer-2 edge ('fap2', 'fap3'). Each layer-2 link has a provided bandwidth attribute which is calculated later.

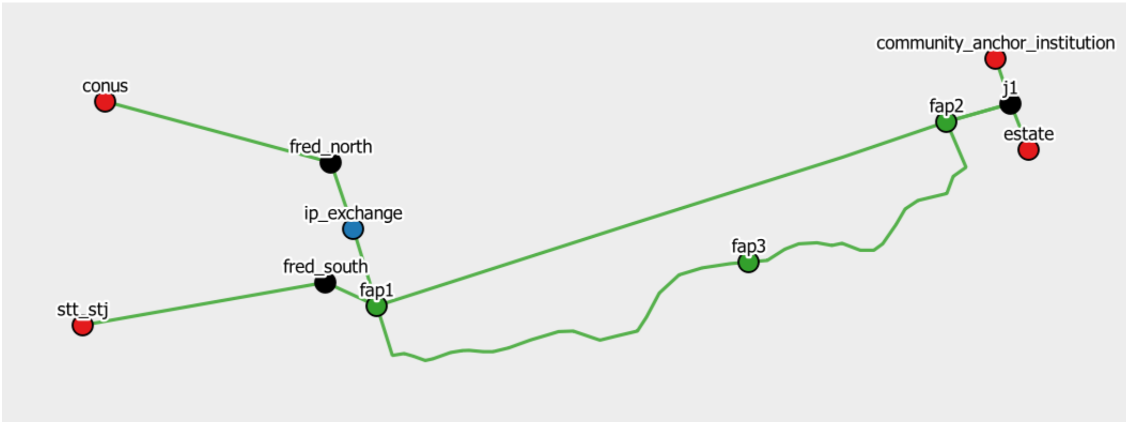

Figure 3.10. Simple Model: Layer-2 Edges. Layer-2 edges, represented as green lines, provide links between nodes that perform layer-2 functions. Although layer-2 links logically ignore nodes that only perform layer-1 functions, for simplicity this image depicts layer-2 links along the layer-1 paths they follow.

Simple Model: Layer-3 Edges. Layer-3 edges represent the segments between routers (i.e., nodes that perform routing functions). Since there are multiple possible layer-1 paths and layer-2 links between routers, it is not possible to describe the physical locations of these segments. As a result, we display layer-3 segments as point-to-point lines between routers, ignoring the possible intermediate nodes. Figure 3.11 depicts the layer-3 edges in our simple model. All population nodes and CAI nodes are connected to the IXP which then routes traffic to its final destination.

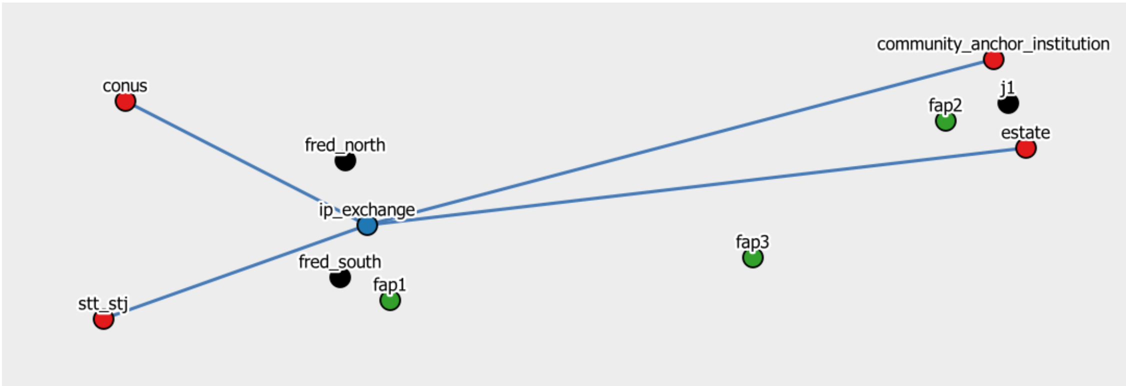

Figure 3.11. Simple Model: Layer-3 Edges. Layer-3 edges, represented as blue lines, provide routes between nodes that perform layer-3 functions. Layer-3 edges are shown as logical end-to-end routes between nodes because there are multiple possible layer-2 links and layer-1 paths between them.

Simple Model: Layer-4 Edges. Layer-4 edges represent the connections between applications that generate and consume Internet traffic. Like layer-3 routes, there are multiple possible layer-1 paths, layer-2 links, and layer-3 segments that provide service to a layer-4 flows so it is not possible to describe the physical location of these connections. The layer-4 host-to-host connections are depicted in Figure 3.12.

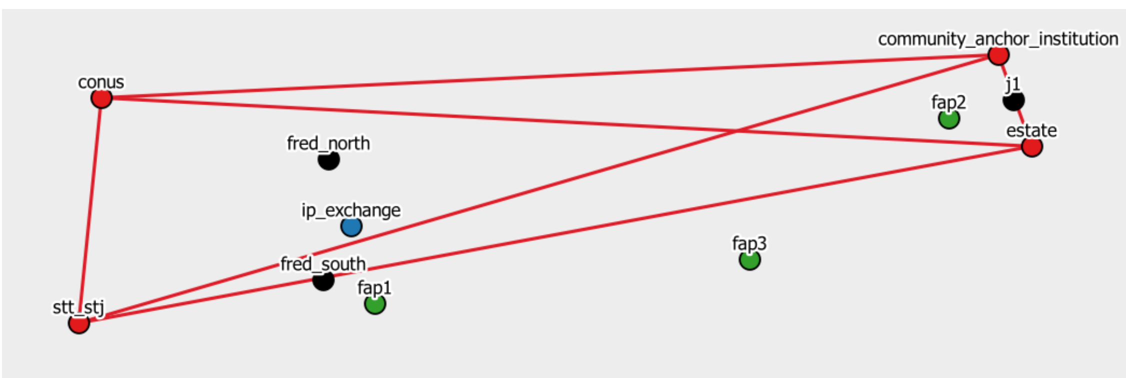

Figure 3.12. Simple Model: Layer-4 Edges. Layer-4 edges, represented as red lines, provide connections between hosts. Layer-4 edges are shown as logical host-to-host connections between nodes because there are multiple possible layer-3 segments, layer-2 links, and layer-1 paths between them.

3.2.4 Synthetic Network Generation

To generate a synthetic network for our simple model, we import data for the nodes and edges (layers 1-4) and then combine them into a single GeoDataFrame object for processing (GeoPandas Development Team 2019). Because information flows both ways in the Internet, we create two directed arcs for each undirected edge.

Layer-1 Network Construction

All nodes and layer-1 edges initially have their `Available?` attribute set to `TRUE`. We developed a function in Python that takes layer-1 nodes or edges we wish to disable as an input and changes their `Available?` attribute to `FALSE`. If a node is input as disabled, the node and every layer-1 edge adjacent to the node is set to `FALSE`. If a layer-1 edge is input as disabled, only that layer-1 edge is set to `FALSE`. The function is cable of handling multiple node and layer-1 edges as inputs. The output of this function is an updated Internet graph G where all nodes and layer-1 edges have the appropriate `Available?` attribute.

Layer-2 Network Construction

Next we create the layer-2 network. To do this, we developed a function in Python that takes the updated layer-1 nodes and edges with the appropriate `available?` attribute as an input. First, the function examines each layer-2 edge and checks to see if any of its layer-1 component arcs are `Available?` = `FALSE`. If any component arc is `FALSE`, the associated layer-2 edge is flagged as `FALSE`. If all component arcs are available, the layer-2 edge is flagged as `TRUE`. Next, the function examines each layer-2 node and selects the appropriate uplink based on `TRUE` layer-2 edges. Each layer-2 node has at least one primary uplink. If the primary uplink is available, the primary uplink is assigned `Available? = True`. If the primary uplink is `FALSE`, but the layer-2 node has an available alternate uplink, the alternate is assigned `TRUE`. If neither the primary nor the alternate are `TRUE`, then no uplink is assigned `TRUE`. This process is completed once for each layer-2 node. All layer-2 edges with `Available?` = `True` are added to G.

Then, the function checks to see if the network requires self-healing. In practice, if the layer-2 network is partitioned into multiple connected components, it will attempt to self heal and use new layer-1 runs that can transmit traffic. To simulate this process, we check to see if the layer-2 network is connected. If multiple network components exist, we test whether adding a secondary uplink for all nodes will reduce the number of components. If adding in an additional secondary link reestablishes layer-2 connectivity, it is added to the network. If it does not change the number of connected components, it is not added to the network. This routine continues until either the network is full connected, or once all secondary uplinks are tried and the network remains partitioned. Taken together, the final output of the network construction and self-healing routines is a synthetic layer-2 network that represents a viable routing for USVI Internet traffic.

Layer-3 Network Construction

The layer-3 network for the USVI is a star network, as all traffic transiting from one ISP to another needs to pass through an IXP, and there is only a single IXP on STX. For this reason, we generate the network by connecting all routers to the network IXP (Node Type 6).

Layer-4 Network Construction

Layer 4 is a fully connected network $|flows| = hosts(hosts - 1)$ weighted based on upload and download demands. The amount of traffic passed between hosts is weighted based on the "size" of the host. We use population and CAI upload and download rates as a surrogate weight for host nodes. We then determine host-host traffic flows by modifying the gravity model presented in Crain (2012) to consider population and CAI Internet demands.

Index use

$n \in N$	Nodes (alias s, t)

Data [units]

D^{tot}	Total Internet demand across the network [Gbps]
D_s	Demand associated with node s [Gbps]
D_t	Demand associated with node t [Gbps]
T_{st}	Total traffic exchanged between node s and node t [Gbps]
b_s^t	Traffic from node s to node t [Gbps]

Formulation

$T_{st} = \frac{D_t}{D^{Tot}} D_s \qquad \forall s,t \in N$

$b_s^t = \frac{D_s}{D_s + D_t} T_{st} \qquad \forall s,t \in N, s \neq t$

$b_t^t = -D_t \qquad \forall t \in N$

Using the above equations, we developed a Python function that takes nodes and the type of demand we want to analyze (average or peak) and outputs a traffic matrix that shows the portion of traffic passed between each host node. The simple model average and peak traffic matrices are depicted in Tables 3.7 and 3.8.

Table 3.7. Simple Model Traffic Matrix, Average Demand

	cai	estate	stt_stj	conus
cai	-1.0	0.0005	0.0019	0.045
estate	0	-0.0005	0	0
stt_stj	0	0	-0.0020	0
conus	0.0023	0	0	-0.050

Note: 'community_anchor_institution' is abbreviated as 'cai'. Rows represent source nodes and columns represent destination nodes. Diagonal values are negative the source node average demand.

Table 3.8. Simple Model Traffic Matrix, Peak Demand

	cai	estate	stt_stj	conus
cai	-10	0.076	0.27	1.3
estate	0.0038	-0.50	0.0032	0.0040
stt_stj	0.053	0.013	-2.0	0.062
conus	6.7	0.40	1.5	-50

Note: 'community_anchor_institution' is abbreviated as 'cai'. Rows represent source nodes and columns represent destination nodes. Diagonal values are negative the source node peak demand.

Traffic Assignment

The final step in generating our synthetic network is to assign layer-4 traffic over layer-3 routes, layer-2 links, and layer-1 runs. To do this, we developed a function in Python that takes the available Internet nodes, edges for all layers, and the layer-4 traffic matrix as input. The function first creates a network for layers 4, 3, and 2 in the NetworkX software library (Hagberg et al. 2008). It then iterates over all host-host pairs ($(s,t,4) \in E$) and finds the shortest path between (s,t) among layer-3 edges using the build-in function `dijkstra_path`. This determines the shortest segment-path between hosts. We continue this procedure on layer 2 to determine the shortest link-path between hosts (if one exists). If a link-path between hosts exists, we then determine the associated run-path between hosts using layer-2 component arcs. Once (s,t) host-to-host routing is established on layers 1, 2, 3, and 4, we assign upload and download traffic to layer-2 links and layer-1 runs. Additionally, we measure the amount of traffic each host node receives to measure how much traffic was demanded and served for each host. The output of this function are GeoDataFrames for layer-1 edges, layer-2 edges, and Internet nodes with traffic assignments. These GeoDataFrames can be exported and visualized in QGIS to see the concentration of traffic moving through the network.

We measure Internet traffic using three metrics. First, we measure amount of demand disconnected from the Internet, $D_{disconnected}$. When network paths exist at all layers, $D_{disconnected} = 0$. Otherwise, it is equal to the total demand that could not flow across the Internet in Gbps.

Second, we measure network congestion, C_{st}, for each link in layer 2 ($(s,t,2) \in E$):

$$C_{st} = \begin{cases} 1 - \frac{\text{Bandwidth}_{st} - \text{Traffic}_{st}}{\text{Bandwidth}_{st}}, & \text{if Traffic}_{st} \leq \text{Bandwidth}_{st}; \\ 1, & \text{otherwise.} \end{cases} \tag{3.9}$$

For a given Internet traffic scenario, we report network congestion as the layer-2 link closest to its maximum bandwidth, C_{st}^{max}. In general, Internet networks are capacity-constrained at the layer-2 link level (layer-1 runs have theoretically infinite capacity). The closer a layer-2 link is to its maximum bandwidth, the closer the STX Internet is to having unserved consumer demand. $C_{st}^{max} > 0$ if the network is operating within normal bandwidth limits and $C_{st}^{max} = 1$ when there is greater demand than layer 2 can support.

Finally, we measure the total unserved demand, U. We measure this as the amount of traffic beyond the capacity of layer-2 links. U is zero if all layer-2 links are under capacity (i.e., $C_{st}^{max} < 1$). Otherwise, U is the amount of traffic that exceeds the network flow capacities in Gbps.

3.2.5 Network Analysis for Simple Model

Consider a situation where all layer-1 path segments in the simple model are available and all hosts are using their average demand. Analyzing the layer-1 edges, which capture the aggregate traffic flow over each layer-1 edge segment, provides valuable insight into the behavior of the simple model.

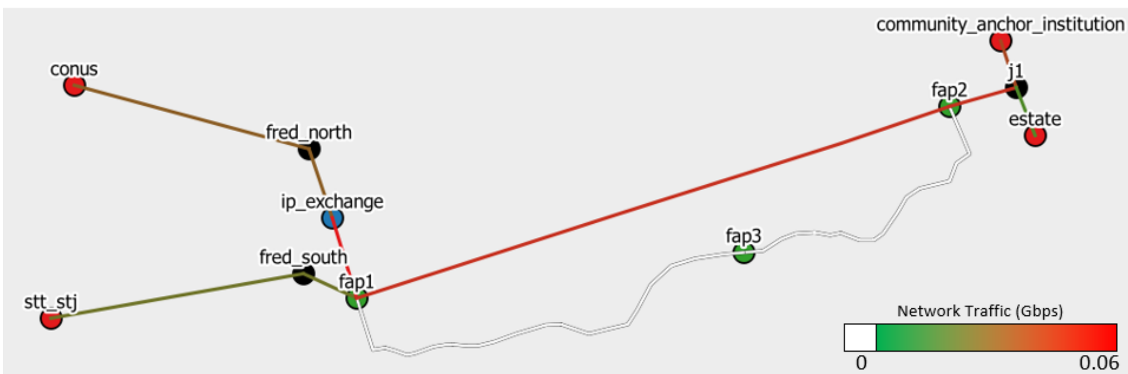

Figure 3.13. Layer 1 Traffic Assignment in Simple Model, All Edges Available. Aggregate traffic over layer 1 edges with all layer 1 edges available. White lines indicate no traffic flow. Traffic flow over edges increases from green to red as the amount of traffic flow over the layer 1 edge increases.

Figure 3.13 depicts all layer-1 edges with their traffic flows. Edge segment are colored based on the amount of traffic flowing over each segment. White segments indicate no traffic flow. Light yellow indicates a small amount of traffic flow which gradually transitions to dark red to indicate a large amount of traffic flow. In Figure 3.13, we see that there is no flow on the ('fap2', 'fap3') and ('fap3', 'fap1') edges. This is because 'fap2' is sending all of its traffic flow over its primary uplink, to 'fap1', and 'fap3' has no hosts demanding traffic through it. The layer-1 edge ('estate', 'j1') has the least traffic flow, indicated by its light yellow color, because 'estate' has the least demand. The layer-1 edge ('ip_exchange','fap1') has the most traffic flow, indicated by its dark red color, because all of the demand from 'stt_stj', the 'community_anchor_institution', and the flow through it from 'estate' to reach the 'ip_exchange'.

Consider another situation where the layer-1 edge segment ('fap1', 'fap2') has been destroyed and all hosts are using their average demand.

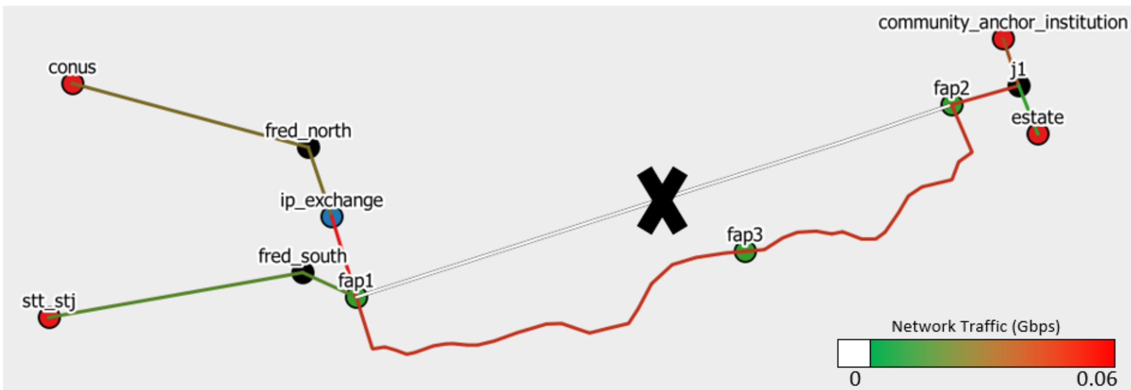

Figure 3.14. Layer-1 Traffic Assignment in Simple Model, One Edge Destroyed. Aggregate traffic over layer-1 edges with edge ('fap1', 'fap2') destroyed. The black 'X' represents a fiber optic cable cut. White lines indicate no traffic flow. Color on edges increases from green to red as the amount of traffic flow over the layer-1 edge increases. Traffic flow has been re-routed over an alternate layer-2 route.

Figure 3.14 depicts all new layer-1 edges and the same color scale applies. Because edge ('fap1', 'fap2') was removed from the network it no longer has any flow over it. Instead, 'fap2' is using its alternate uplink to 'fap3' and 'fap3' is using its primary uplink to send that traffic to 'fap1'. The other layer-1 edge segments see no change in the flow over them.

We assess operational impacts of traffic re-routing using $D_{disconnected}$, C_{st}^{max}, and U (Table 3.9). In general, network re-routing will lead to some layer-2 links having increased traffic, congestion, and unserved demand. Moreover, depending on the location of the destroyed infrastructure, customers may be disconnected. In this simple scenario, however, we note that neither customers are disconnected, congestion increases, nor demand becomes unserved. This is due to the equivalent operations of both network paths, even though re-routing over the alternate layer-2 route requires more hops and is less efficient with respect to physical and logical distance.

Table 3.9. Internet Traffic Metrics for the Simple Model

Scenario	$D_{disconnected}$ (Gbps)	C_{st}^{max}	U (Gbps)
All Edges Available	0	0.005	0
One Edge Destroyed	0	0.005	0

3.3 STX Internet System Model

The multi-layer Internet network model captures the types of infrastructure, functions, and relationships of the USVI Internet system. We use the format of the model and simple example to build a descriptive model of the USVI Internet system, focusing on STX. Building on the data previously curated into geospatial data files in QGIS during the data curation phase of our work, we conduct additional processing to have the correct attributes and data formats.

3.3.1 Nodes in the STX Internet System Model

Our STX Internet system model contains a total of 1,068 nodes. Like in the simple model, each node is assigned 14 attributes based on the function and characteristics of that node. These nodes are visually depicted in Figure 3.15. The color of each node corresponds with the highest level of the OSI model at which the node operates.

Population Node Calculations

Population nodes operate at all four layers of the OSI model. The population nodes for the STX Internet system model include the 233 private Internet customers previously curated as well as synthetic population nodes representing a combined STT/STJ and CONUS. For purposes of defining traffic demand, the STT/STJ synthetic node represents the entire population of STT and STJ as a single node, and the CONUS synthetic node represents the entire population of the US. The average and peak demand for STX population nodes were calculated using the formula:

$$\text{avg\_demand}_p = \text{households}_p \times \text{penetration}_p \times \text{avg\_scrb\_dmd}_p \tag{3.10}$$

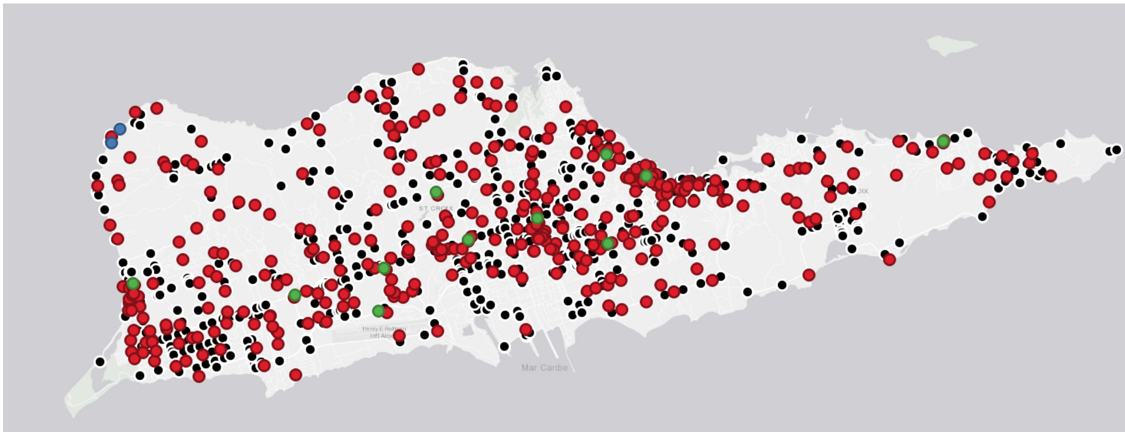

Figure 3.15. Nodes in the STX Internet System Model. Node colors align with the highest level of the OSI model at which the node operates: black nodes are junction nodes (fiber junctions and cable landing sites), green nodes are switch nodes (FAPs and Super FAP), blue nodes are router nodes (IXPs and NAPs), and red nodes are host nodes (population centers and CAIs). Nodes representing locations on STT/STJ and CONUS exist but are not shown here.

$$\text{peak\_demand}_p = \text{households}_p \times \text{penetration}_p \times \text{peak\_scrb\_dmd}_p \tag{3.11}$$

where p is a population node on STX, households_p is the number of households at the node, penetration_p is the percentage of the households that have Internet service, and avg_scrb_dmd_p) and peak_scrb_dmd_p are the average and peak subscriber Internet demand. The penetration of Internet in the USVI is 54.8% (Central Intelligence Agency 2020), the average Internet demand is assumed to be 1 Mbps, and the peak Internet demand is assumed to be 5 Mbps.

The average and peak demand for the STT/STJ and CONUS population nodes were calculated using the formula:

$$\text{avg\_demand}_p = \frac{\text{population}_p}{\text{avg\_household\_size}_p} \times \text{penetration}_p \times \text{avg\_scrb\_dmd}_p \tag{3.12}$$

$$\text{peak\_demand}_p = \frac{\text{population}_p}{\text{avg\_household\_size}_p} \times \text{penetration}_p \times \text{peak\_scrb\_dmd}_p \tag{3.13}$$

where p is a population node in STT/STJ or CONUS, population(p) is population of the node, avg_household_size(p) is the average household size in the US, penetration(p) is the percentage of the households that have Internet service, and avg_scrb_dmd(p) and peak_scrb_dmd(p) are the average and peak household Internet demand. The average household size in the US is 2.52 people, the penetration of Internet at STT/STJ is 54.8%, the penetration of Internet in CONUS node is 73.00%, the average Internet demand is assumed to be 1 Mbps, and the peak Internet demand is assumed to be 5 Mbps (Central Intelligence Agency 2020; Pew Research Center 2019; U.S. Census Bureau 2019).

CAI Node Calculations

CAI nodes also operate at all four layers of the OSI model. The CAI nodes for the STX Internet system model include the 133 CAIs previously curated. Although CAIs do not have a population assigned to them, they do have a maximum upload and download speed for each node which was used to calculate the average and peak Internet demand for each CAI using the formula:

$$\text{avg\_demand}_c = \text{subscrb\_dwn}_c \times \text{avg\_cai\_use}_c \tag{3.14}$$

$$\text{peak\_demand}_c = \text{subscrb\_dwn}_c \times \text{peak\_cai\_use}_c \tag{3.15}$$

where c is a CAI node, subscrb_dwn(c) is the given CAI maximum download rate, avg_cai_use(c) is the average amount of the maximum download rate that is used, and peak_cai_use(c) is the peak amount of the maximum download rate that is used. avg_cai_use(c) is assumed to be 10% and peak_cai_use(c) is assumed to be 100%.

3.3.2 Layer-1 Edges in the STX Internet System Model

We compile edges for the STX Internet system model from known elements of the last mile, middle mile, and backbone infrastructure. Additionally, synthetic layer-1 edges were created to connect the STT/STJ synthetic population node to the middle mile infrastructure and the CONUS synthetic population node to the two CONUS NAPs. In total, there are 2,478 individual layer-1 edge segments representing the entire physical fiber optic cable infrastructure. The 2,478 STX Internet system model layer-1 edges are visually depicted in Figure 3.16. The middle mile and backbone edges are black, and last mile edges are grey. The destination of edges that provide paths to nodes off the map are annotated.

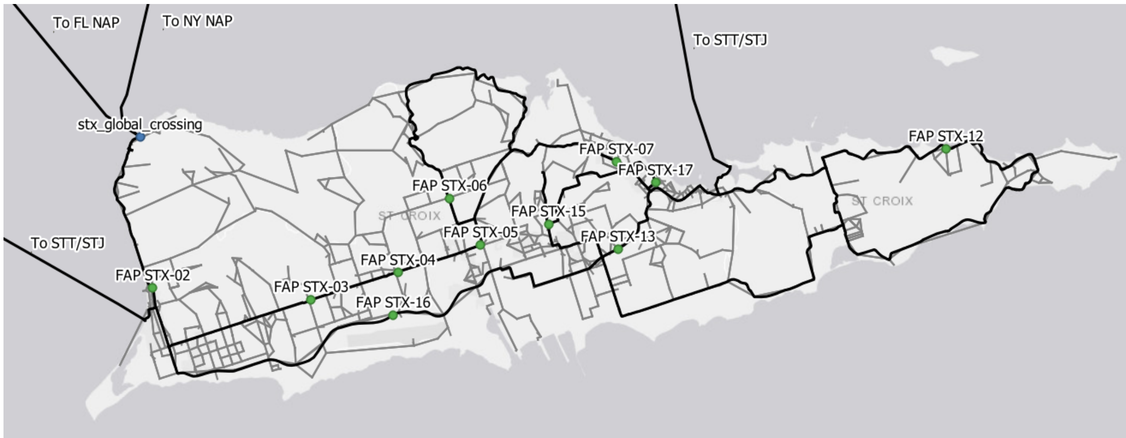

Figure 3.16. Layer-1 Edges in the STX Internet System Model. Middle mile and backbone edges are black and last mile edges are grey. The destination of edges that provide paths to nodes off the map are annotated. FAP, Super FAP, and IXP nodes are identified in green and blue, respectively.

3.3.3 Layer-2 Edges in the STX Internet System Model

The layer-2 edges in this model were developed differently for different source destination pairs. For layer-2 edges between customers on STX and FAPs, a Dijkstra's shortest path algorithm was used to calculate the shortest path (in meters) from each customer node to every FAP. The closest FAP was selected and a layer-2 edge was then created between that customer and its nearest FAP. The STT/STJ synthetic node was provided a layer-2 link to FAP STX-02 through the viNGN submarine cables.

The remaining layer-2 edges were manually decided. For FAP-to-FAP layer-2 edges, a primary and alternate link was chosen from each FAP to adjacent FAPs. Because the middle-mile infrastructure on STX generally forms rings, we chose the primary link destination and layer-1 runs it follows based on the shorter direction around the ring towards the Super FAP. Alternate layer-2 links were chosen to provide layer-1 path diversity (i.e., different layer-1 paths than the primary path) to a FAP in the other direction around the ring (i.e., away from the Super FAP). Figure 3.17 provides an example of pair of primary and alternate FAP layer-2 links.

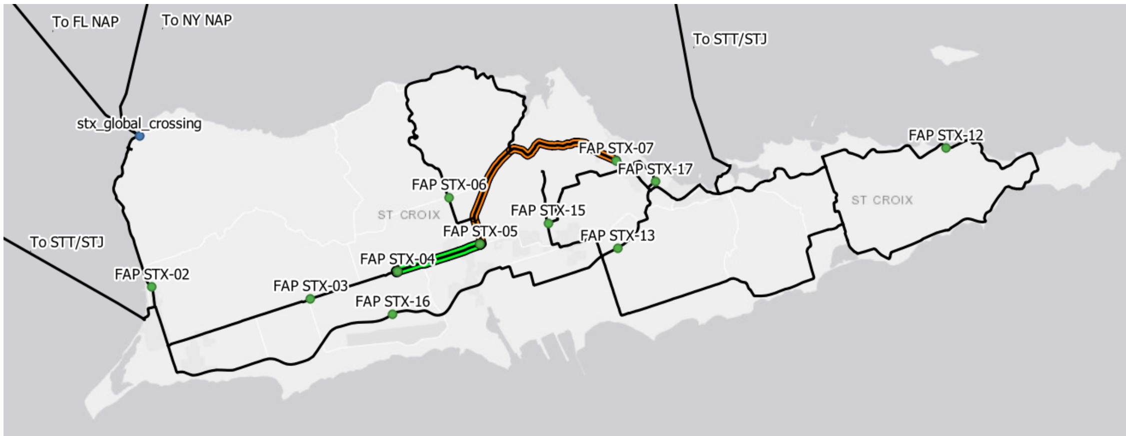

Figure 3.17. Primary and Alternate Layer-2 Links in STX Internet Model. The primary layer-2 link from FAP STX-06, highlighted in bright green, is counter-clockwise through the middle mile ring to FAP STX-04. The primary link allows FAP STX-06 to move layer-2 traffic closer to FAP STX-02, the Super FAP. The alternate layer-2 link, highlighted in orange, is clockwise through the middle mile ring to FAP STX-07. The alternate link provides a backup layer-1 path for moving Internet traffic through the middle mile.

FAP STX-02, the Super FAP was provided a primary layer-2 link to the Global Crossing NAP. FAP STX-02 does not have an alternate layer-2 link, since there is no alternate set of layer-1 runs that would connect it the the IXP. The Global Crossing NAP is assigned a primary layer-2 link to the Florida NAP and an alternate to the New York NAP. The CONUS synthetic node was manually assigned a primary layer 2 link to the Florida NAP and an alternate to the New York NAP. The Florida NAP and New York NAP were each assigned layer-2 links to the Global Crossing NAP through unique submarine cable routes. The 392 edges in the USVI Internet system layer-2 are visually depicted in Figure 3.18. Some layer-1 paths, such as the eastern submarine cable between STX and STT, are not used to provide layer-2 links because no nodes uplink traffic through them.

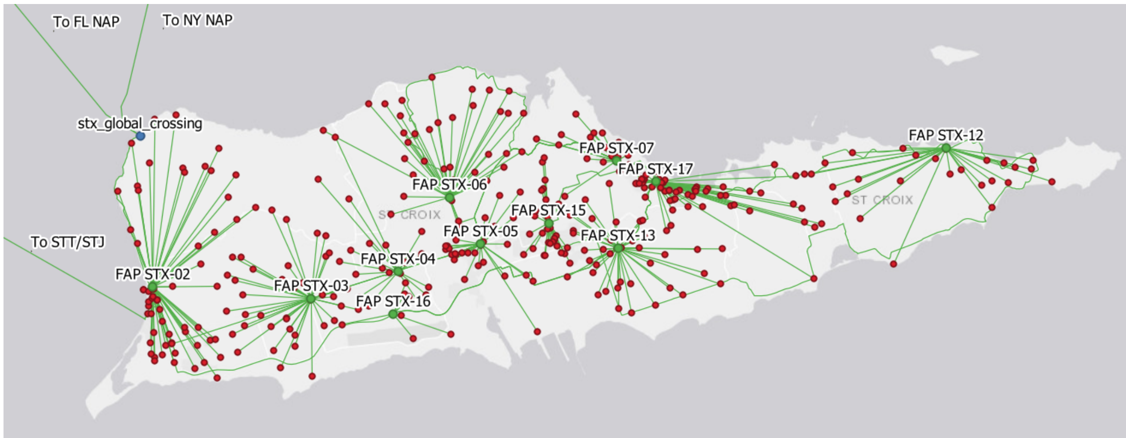

Figure 3.18. Layer-2 Edges in the STX Internet System Model. Red nodes represent customer nodes, green nodes represent FAP nodes, and blue nodes represent IXP nodes. Layer-2 edges are represented by green lines. Layer-2 edges between customer nodes and FAPs are represented as straight green lines and do not follow layer-1 paths. All other layer-2 links, such as those between FAPs, do follow layer-1 paths.

3.3.4 Layer-3 Edges in the STX Internet System Model

In our STX Internet system model, layer-3 edges are routes between all customer nodes and the Global Crossing IXP on STX. Since there are multiple possible layer-1 paths and layer-2 links between customer nodes and the IXP, it is not possible to describe the physical locations of these routes. The point-to-point layer-3 routes between all customer nodes and the IXP are visually depicted in QGIS in Figure 3.19.

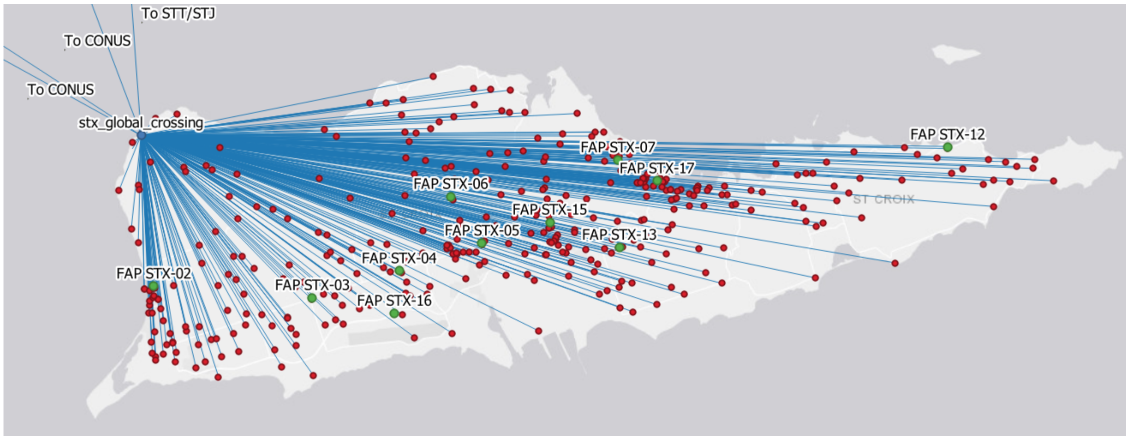

Figure 3.19. Layer-3 Edges in STX Internet System Model. Red nodes represent customer nodes and blue nodes represent IXP nodes. Layer-3 edges between all customer nodes and the IXP are represented by blue lines.

3.3.5 Layer-4 Edges in the STX Internet System Model

In our STX Internet system model, layer-4 edges are connections between all customer nodes. Like layer-3 routes, there are multiple possible layer-1 paths, layer-2 links, and layer-3 routes that form a layer-4 connection so it is not possible to describe the physical location of these connections. There are 67,528 layer-4 host-to-host connections included in the STX Internet system model to connect all host-host pairs.

3.3.6 Synthetic Network Generation in STX Internet System Model

In our STX Internet system model, synthetic network generation follows the same process as the simple model. Working with real data for STX, however, requires a few additional attributes for population and CAI nodes which allows us to make more realistic calculations for provided bandwidth on layer-2 edges. Provided bandwidth for edges from population nodes to FAP nodes are calculated as:

$$\text{provided\_bandwidth}_{(p,f)} = \frac{\text{population}_p}{\text{avg\_household\_size}_p} \times \text{penetration}_p \times \text{isp\_upload}_p \tag{3.16}$$

$$\text{provided\_bandwidth}_{(f,p)}) = \frac{\text{population}_p}{\text{avg\_household\_size}_p} \times \text{penetration}_p \times \text{isp\_download}_p \tag{3.17}$$

where p are population nodes and f are FAP nodes. Here, population_p is the population at p, $\text{avg\_household\_size}_p$ is the average household size at p, penetration_p is the Internet penetration at p, and isp_download(p) is the average ISP download speed. The average household size in the US is 2.52 people, the penetration of Internet in the USVI is 54.8%, the ISP upload from Viya is 5 Mbps, and the ISP download from Viya is 25 Mbps (Central Intelligence Agency 2020; Pew Research Center 2019; Viya 2020).

Provided bandwidth for edges from CAI nodes to FAP nodes are calculated as:

$$\text{provided\_bandwidth}_{(c,f)} = \text{peak\_demand}_c. \tag{3.18}$$

$$\text{provided\_bandwidth}_{(f,c)} = \text{peak\_demand}_c. \tag{3.19}$$

where c are CAI nodes and f are FAP nodes. Here, peak_demand_c is the peak demand previously calculated for each CAI node.

The provided bandwidth on layer-2 edges between all other nodes is assumed to be 10 Gbps.

3.4 Summary

The Internet is a complicated system, and understanding its structure and behavior requires careful attention to the attributes of each of its layers. Even a system of modest size, like the system in STX, has thousands of components and connections. The methods presented here serve as a rigorous means to capture essential features. Understanding the implications of these features in terms of the system traffic and vulnerability is the subject of the next chapter.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 4:
Analysis

Using the methods presented in Chapter 3, we analyze the operation of the Internet on STX. We generate a model of the STX Internet and assess the relationship between Internet infrastructure and customers. We estimate the Internet demand on the viNGN middle mile and measure approximated traffic flows. Results provide an estimate of viNGN infrastructure criticality and vulnerability to support recovery and mitigation operations for future disasters. Results also show how synthetic network generation methods may produce a reasonable model of the STX Internet alongside modeling issues that need to be resolved to capture real network operations.

4.1 How St. Croix Internet Customers Connect to viNGN Middle Mile Infrastructure

We assess how dependent STX communities, households, and CAIs are served by viNGN middle mile infrastructure using synthetic network generation techniques. We use these results to estimate average and peak Internet demand for viNGN FAPs and middle mile infrastructure.

4.1.1 Customer Dependence on FAPs

Synthetic network generation methods for layer 1 provide a means to estimate which populations and geographic regions are served by each FAP. We implement synthetic network generation techniques found in the literature by assuming each population center connects to its nearest FAP based on the STX road network. In general, this is a reasonable approximation as the majority of Internet fiber runs lie underneath roads or are adjacent to roads on above ground power poles. Moreover, ISPs would often construct their layer-1 network with as few fiber runs as possible to be cost effective.

Relating Middle Mile Infrastructure to STX Estates

We associate STX Internet customers to FAPs to determine how geographic regions are served by the viNGN middle mile. Figure 4.1 presents the synthetic network generation results for FAPs STX-02 and STX-03. Results indicate that population connectivity to the viNGN middle mile does not necessarily match the local geographic and jurisdictional boundaries associated with estates. For example, Figure 4.1 shows estates near Fredericksted can have multiple population centers that connect to either STX-02 or STX-03. This result is possible in the 'real' network due to neighborhood growth and ISPs expanding their service regions. Additionally, some estates do not appear the be serviced by any FAPs due to a lack of population.

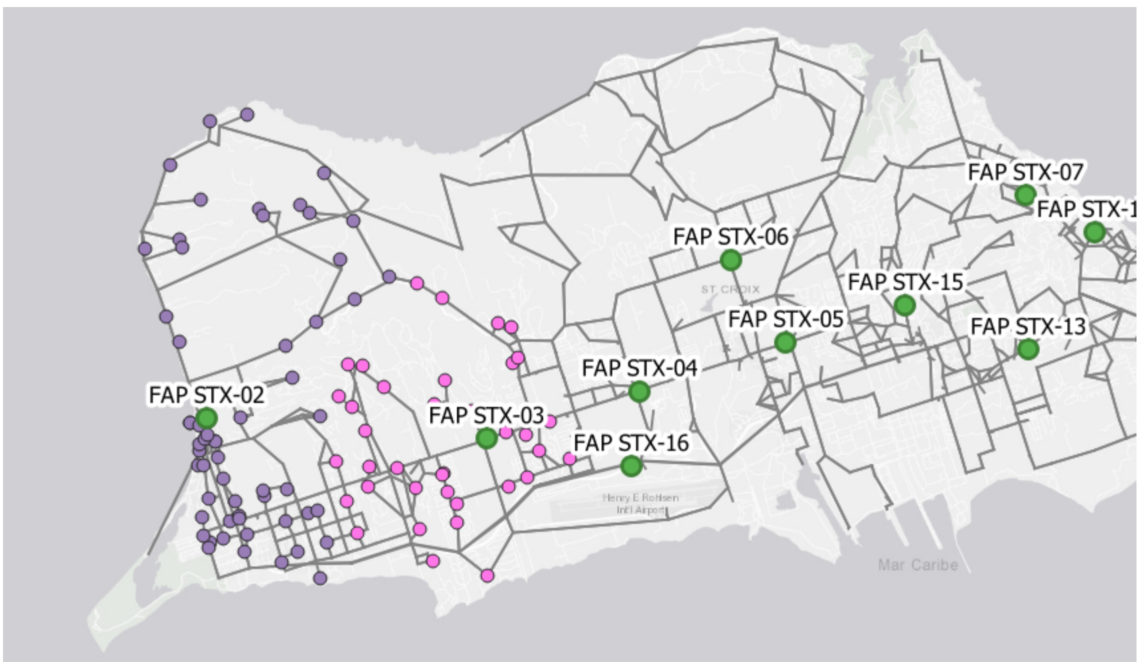

Figure 4.1. Customers Served by STX-02 and STX-03 Fiber Access Points. The customers (i.e., populations and CAIs) that connect to FAP STX-02 are represented in purple and FAP STX-03 represented in pink. We find regions where population centers near each other connect to different FAPs. We also find customers far from each FAP connecting in to the middle mile over potentially indirect routes. FAPs are represented in green and layer-1 last mile infrastructure represented by grey lines. Other FAP customers are omitted for visual clarity.

To better relate the STX Internet system model to real-world demands, we estimate which

FAP is the *primary* connection point for each estate through manual inspection. We present the relationship between Internet customers and estates in Figure 4.2 and FAPs in Figure 4.3. When synthetic network generation produces estates that are wholly occupied by customers of a single FAP, manual assignment is trivial. In contrast, instances where an estate has no known residential customers and/or are served by multiple FAPs require additional methods. Specifically, if an estate has no customers, the estate was assigned to a FAP based on the customers in neighboring estates and geographic proximity. If the estate has customers assigned to more than one FAP, that estate is assigned based on the population majority and geographic proximity.

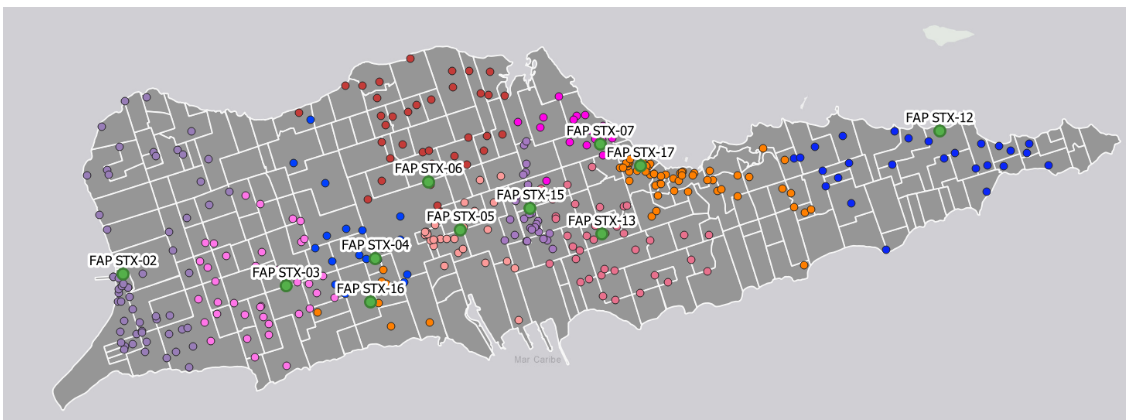

Figure 4.2. St. Croix Internet Customers Colored by Assigned Facility Access Point. Due to overlapping regions, STX estates are assigned to a FAP by manual inspection. FAPs and their assigned customers overlaid on STX estate boarders (white lines). FAPs are presented in green. Customer colors correspond to estate colors in Figure 4.3.

Both synthetic network generation and manual inspection indicate that the location of viNGN middle mile infrastructure does not necessarily correspond to existing population centers and estates. This may be due to the early focus of viNGN to ensure connectivity of CAIs, rather than communities. This result may also be due to rules of right of way, leasing land for laying fiber runs, or changes in infrastructure systems that layer-1 network structure is interdependent with (e.g., the road network).

Overall, this result suggests that there may be idiosyncrasies of the STX Internet not captured in common synthetic network generation methods. For example, Figure 4.3 shows that our synthetic network generation methods would connect FAP STX-04 to customers or regions

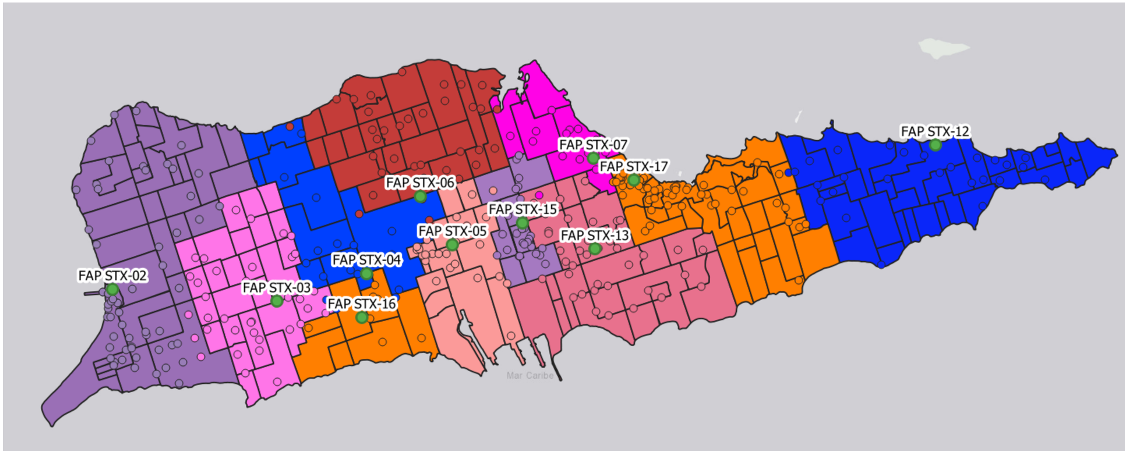

Figure 4.3. St. Croix Estates Colored by Primary Facility Access Point. We assign each estate to a FAP by manual inspection. Estate color correspond to customer nodes presented in Figure 4.2. FAPs are presented in green.

without service on the northwest coast. This result is non-intuitive, and implausible based on the distance between the viNGN middle mile and the northwest coast. In addition, methods generate instances where FAPs do not service estates directly adjacent to them. While this takes into account the last mile infrastructure and road network, the distinction between which customers are served by nearby FAPs STX-07 and STX-17, for example, is potentially unrealistic.

Relating Populations and CAIs to FAPs

We use the results of synthetic network generation to estimate the number of people and critical facilities served by each FAP. Figure 4.4 presents the total population connecting into the viNGN middle mile at each FAPs. Results show the differences in importance for each FAP for STX communities. In particular, FAP STX-02, STX-03, STX-13, and STX-17 service more private Internet customers, with each FAP providing connectivity for over 6,000 people. These FAPs also correspond with the largest population centers on the island: STX-02 and STX-03 are located near Frederiksted on the west end of the island, and STX-13 and STX-17 are located near Christiansted on the east end of the island.

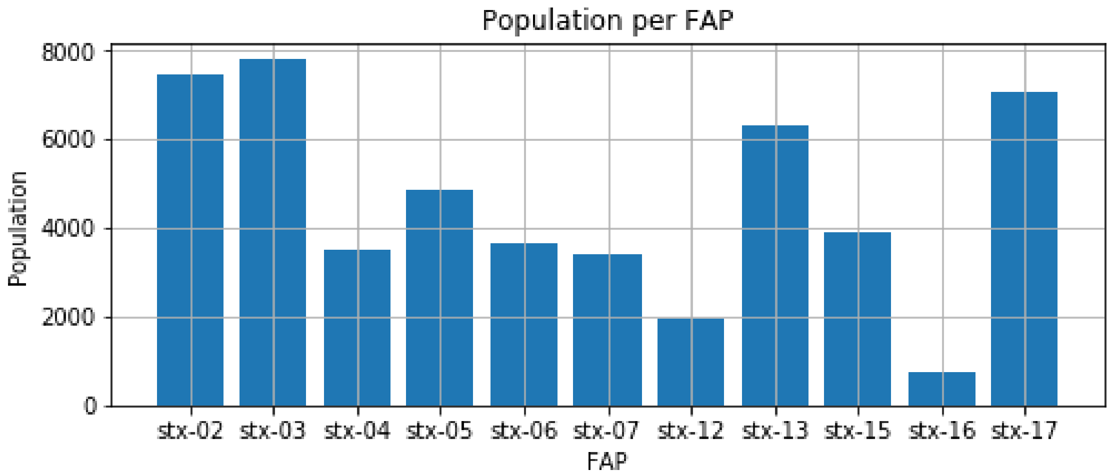

Figure 4.4. Population Serviced by Each Facility Access Point. The population serviced by each FAP. FAPs located in more densely populated areas service larger populations.

In addition to determining FAP importance based on population, we also measure importance based on CAIs. Figure 4.5 presents the number of CAIs connected to each FAP. In general, we find FAPs that service large populations also service a large number of CAIs, such as STX-02 and STX-17. This result is intuitive because government buildings, schools, and churches are located in population centers. A notable departure from this trend is FAP STX-03 which services 15.38% of the population, the most out of any FAP, but only services 7.52% of the CAIs.

More specifically, the importance of a given FAP differs for different CAI types. For example, medical facilities, public safety agencies, and government buildings may be prioritized over libraries or schools during or immediately after a disaster. We find that certain FAPs may not provide Internet connectivity to these critical facilities. Specifically, FAP STX-06 and STX-07 do not serve any public safety facilities (CAI Type 4). Moreover, only FAPs STX-02, STX-13, STX-15, and STX-17 provide Internet to medical facilities (CAI Type 3).

Taken together, results provide an rough estimate of the importance of viNGN FAPs. In particular, STX-02 and STX-17 may be the two most important FAPs for both population and CAI Internet access. This estimate of importance, combined with a future trafficdependency analysis will enable prioritization of FAPs for hardening against hurricanes and other disasters.

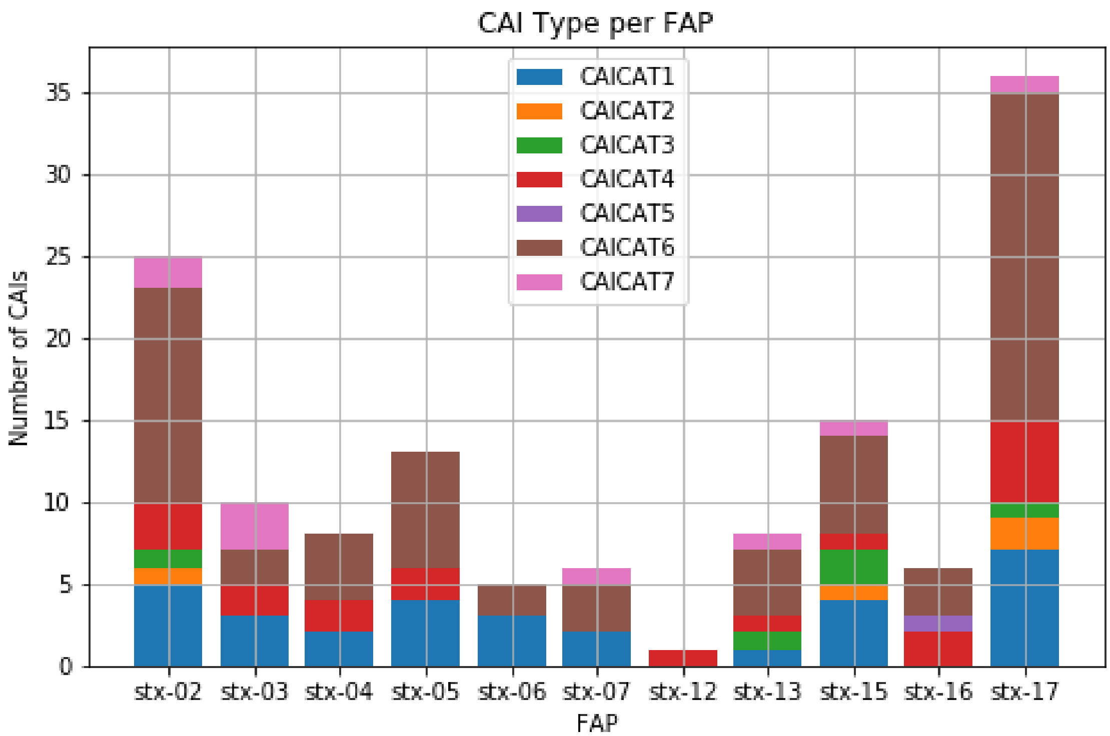

Figure 4.5. Community Anchor Institutions Serviced By Each Facility Access Point. Each FAP provides Internet connectivity for a different combination of CAIs on STX. For example, FAP stx-17 and stx-02 serve the greatest number of CAIs. CAI categories include: (1) School K-12, (2) Library, (3) Medical/healthcare, (4) Public Safety, (5) University College or Post-secondary School, (6) Other community support — governmental, and, (7) Other community support — non-governmental.

4.1.2 FAP Bandwidth Requirements to Serve Customers

We use the above results to estimate the bandwidth requirements for each FAP on STX. Figure 4.6 presents the estimates of average and peak demand by FAP. All download and upload Internet traffic must enter the viNGN middle mile through FAPs. Based on population and CAI requirements, we estimate that under average Internet demand conditions, there is a total of 12.54 Gbps of customer demand in the network with a mean demand of 1.15 Gbps per FAP. FAP STX-17 has the highest demand with 2.00 Gbps followed closely by STX-02 and STX-03 with 1.88 Gbps and 1.80 Gbps respectively. FAP STX-16 has the smallest demand with 0.23 Gbps. Under peak conditions, there is a total of 69.69 Gbps of customer demand in the network with a mean demand of 6.34 Gbps per FAP. FAP STX-17 has the highest demand with 12.73 Gbps followed by STX-02 and STX-03 with 10.61 Gbps and 9.46 Gbps respectively. FAP STX-16 has the smallest demand with 1.43 Gbps.

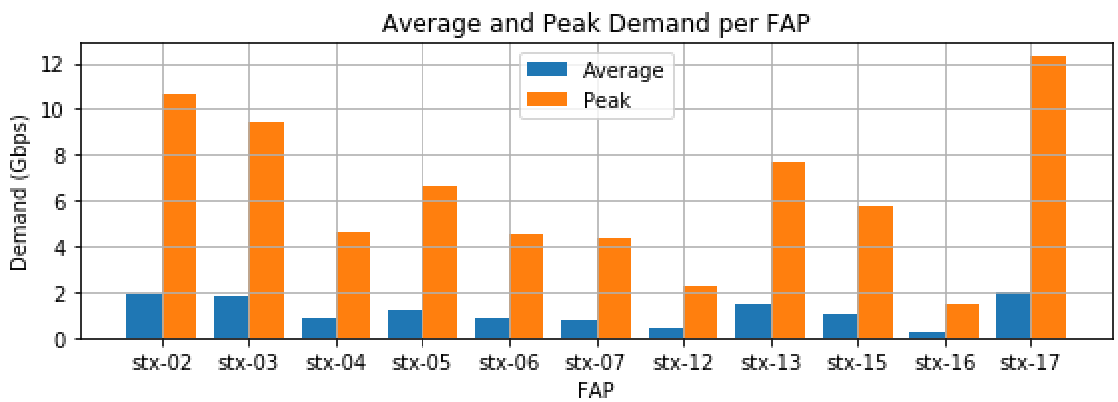

Figure 4.6. Average and Peak Demand per Fiber Access Point. Average Internet demand measure the estimated traffic generated at each FAP during a normal day. Peak Internet demand represents the possible maximum demand that could be generated from Internet customers based on max upload and download rates.

It is possible that synthetic network generation overestimates Internet demand on STX. We know that the bandwidth on each layer-2 link in the viNGN middle mile is 10 Gbps. However, synthetic network generation produces average demand on the island that exceeds this limit (12.54 Gbps). If this result is accurate, it would mean daily Internet use on STX is rate limited, and network protocols would be reducing total traffic from Internet customers. Moreover, the total peak demand is much greater than bandwidth limits (69.69 Gbps). If these were accurate, it would mean that the existing network is operating a severely congested state, which is not typical of wholesale middle-mile providers.

Still, the distribution of demand across each FAP may be accurate, where FAPs serving large populations also have the greatest total demand. This means that the results presented in Figure 4.6 are useful to determine relative volumes of Internet traffic flows, but the absolute value of unserved demands may be overestimated.

4.2 Synthetic St. Croix Internet Traffic and Network Operations

We use the customer and FAP results to approximate network traffic on the STX Internet system model and assess the implications of infrastructure failures for Internet users.

4.2.1 Traffic Assignment Scenarios for the STX Internet System Model

We generate flow maps and aggregate traffic to show how the STX Internet system model functions in normal and failure scenarios. We consider traffic in five fictitious scenarios that help reveal the vulnerability of Internet access to layer-1 cable cuts and demand spikes:

- **Scenario 1:** Average Internet demand, all infrastructure operational;
- **Scenario 2:** Average Internet demand, single fiber run cut along the middle mile ring;
- **Scenario 3:** Average Internet demand, two fiber run cuts along the middle mile ring;
- **Scenario 4:** Average Internet demand, single fiber run cut between the SuperFAP and the IXP;
- **Scenario 5:** Peak Internet demand, all infrastructure operational.

Scenario 1: Average Demand, All Infrastructure Operational

We measure flows on the STX Internet system model with average demands and all infrastructure available to assess Internet traffic during normal operations. Figures 4.7 and 4.8 presents Internet traffic from FAPs into the viNGN middle mile and then off island via submarine cables to North America. Table 4.1 presents the total traffic over each layer-2 link within the middle mile.

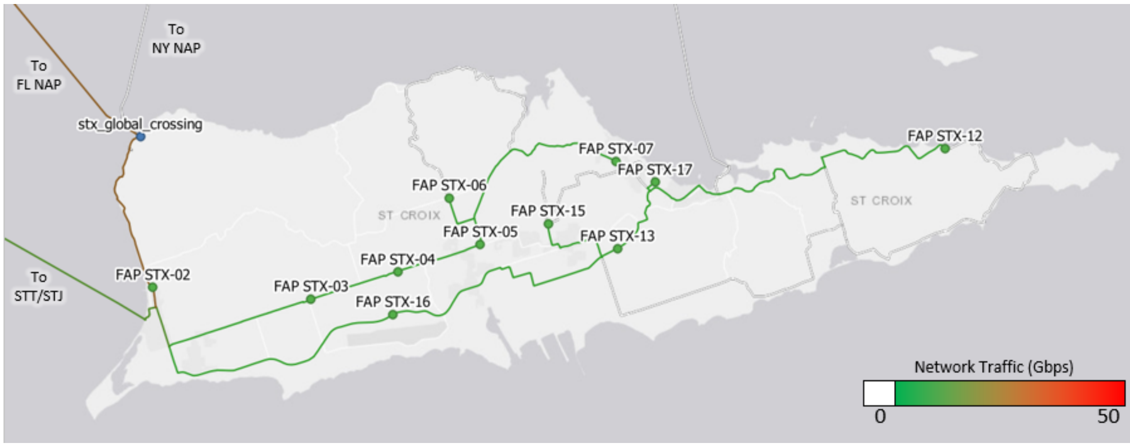

Figure 4.7. St. Croix Middle Mile Traffic under Scenario 1. The amount of Internet traffic flowing on the STX middle mile is represented by shaded lines. Green lines represent light traffic and red lines represent high traffic. Unused layer-2 links are represented as grey lines.

Table 4.1. Layer 2 Aggregate Traffic Between FAPs under Scenario 1.

Source	Destination	Traffic (Gbps)
STX-02	STX-03	5.5
STX-02	STX-16	5.1
STX-16	STX-13	4.9
STX-03	STX-04	3.7
STX-04	STX-05	2.9
STX-13	STX-17	2.5
STX-13	STX-15	1.0
STX-05	STX-06	0.9
STX-05	STX-07	0.8
STX-17	STX-12	0.4

Note: All FAP-to-FAP layer-2 links are bidirectional (upload/download), but traffic flows are dominated by download rates. We assume each FAP-FAP link has a maximum bandwidth of 10 Gbps. Links that do not flow traffic or are disconnected from the network are omitted.

Figure 4.7 provides a baseline for understanding how Internet traffic flows on STX. Results indicate that not all layer-2 links or layer-1 fiber runs within the viNGN middle mile transport

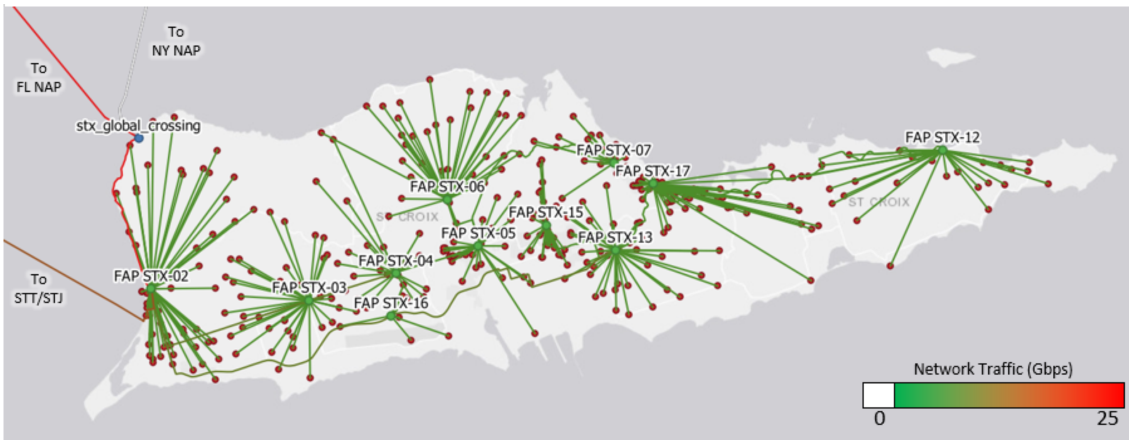

Figure 4.8. St. Croix Layer-2 Connectivity to Customers Under Scenario 1. Population centers with Internet subscribers are represented by red circles with total traffic (upload and download) represented by shaded lines. Green lines represent light network traffic and red lines represent heavy traffic. Unused layer-2 links are represented as grey lines.

Internet traffic during normal operations. Specifically, Figure 4.7 shows that viNGN middle mile forms a tree network, where SuperFAP STX-02 serves as the central node and two branching link-paths form along central and southern STX. The link-path through central STX starts at Christiansted with FAP STX-07 and transports traffic along Centerline Road. The southern link-path starts at FAP STX-12 on the eastern end of the island and transports traffic along the Melvin H. Evans Highway. FAP STX-15 forms a second branch off the southern link-path. All paths terminate at the Super FAP which then forwards all traffic to the IXP for exchange between ISPs and North America. Notably, the links STX-07–STX-17, and STX-15–STX-17 are not utilized during normal operations.

Figure 4.8 shows how the layer-2 traffic from Figure 4.7 relates to population centers and CAIs. Under this scenario, all customers are connected to the Internet.

Table 4.1 shows which layer-2 links flow the greatest Internet traffic during normal operations. In general, no layer-2 link within the viNGN middle mile ring is near its maximum bandwidth. However, as traffic is aggregated at the Super FAP, STX-02, and then forwarded to the IXP and North America, it becomes bandwidth limited. Here, all 12.54 Gbps converge on to a single link between STX-02 and the IXP. This value exceeds the assumed bandwidth constraint of 10 Gbps. viNGN leases dark fiber between the IXP and Miami, and the IXP to New York, each operating with 10 Gbps capacity.

Internet traffic during normal network operations indicates possible vulnerabilities within the STX Internet system model. While layer-2 determines final traffic flow paths, failures in layer-1 (physical) may lead to traffic re-routing, increased network congestion, unserved demands, and disconnected customers. As Layer-2 links along Centerline Road and the Melvin H. Evans highway form the only connections between middle mile FAPs and the Super FAP, cable cuts in layer 1 on either or both of these roads may lead to customers losing Internet access. Equivalently, the link between the Super FAP and the IXP appears to be a possible "most-vital arc" (Alderson et al. 2013). A cable cut in layer-1 along this link may disconnect the entire island. Moreover, even if layer-1 is fully available, we expect to see bandwidth limitations between the Super FAP and the IXP during situations of peak demand. We assess these scenarios to understand the vulnerability of the STX Internet system model.

Scenario 2: Average Demand, Fiber Cut on Melvin H. Evans Highway

We study traffic re-routing and network congestion for a cable cut on the Melvin H. Evans Highway. Figures 4.9 and 4.10 present traffic routing and connectivity. Table 4.2 presents the total traffic flows over the viNGN middle mile.

Scenario 1 revealed that a fiber run cut along the highway could force traffic to re-route over Centerline Road (i.e., the central STX link-path). Figure 4.9 shows how the layer-2 topology can self-heal to circumvent the fiber cut and to ensure all Internet customers remain connected to the network. In particular, the previously un-used link between FAP STX-07 and STX-17 is utilized to ensure all FAPs can connect to the Super FAP and IXP. Accordingly, all Internet traffic is routed over Centerline Road. Figure 4.10 shows that all Internet customers on STX remain connected to the network after the cable cut.

Despite customers maintaining connectivity, Table 4.2 shows that re-routing may lead to higher network congestion and unserved Internet demand. In Scenario 1, all traffic on STX converged at the Super FAP, such that the link between STX-02 and the IXP was the only congestion constrained portion of the network. In Scenario 2, re-routing leads to the link between STX-02 and STX-03 also being congestion constrained with a total flow of 10.7 Gbps. This flow exceeds the assumed link bandwidth of 10 Gbps.

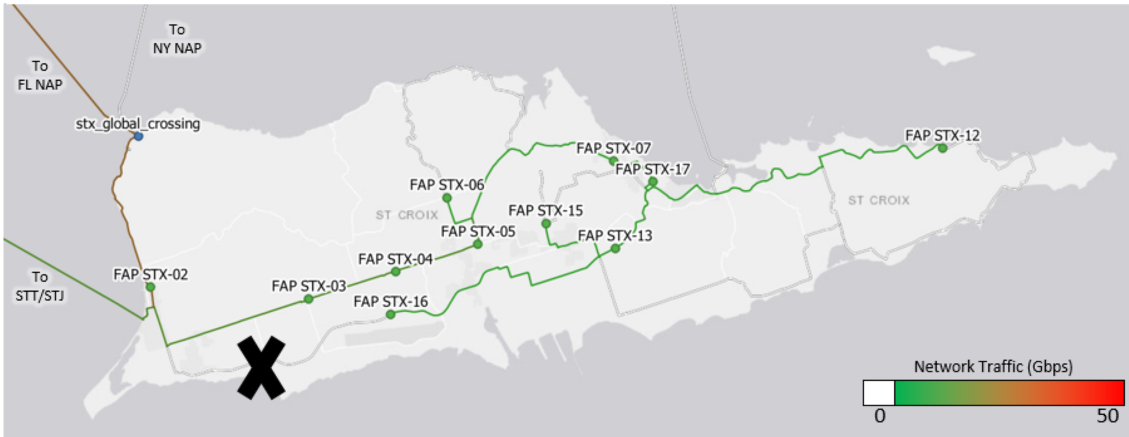

Figure 4.9. St. Croix Middle Mile Traffic under Scenario 2. A fiber cut along the Melvin H. Evans Highway leads to network self-healing and traffic re-routing. All middle mile Internet traffic flows along Centerline Road. to reach the Super FAP (STX-02). The amount of Internet traffic flowing on the STX middle mile is represented by shaded lines. Green lines represent light traffic and red lines represent high traffic. Unused layer-2 links are represented as grey lines. The location of the fiber cut is represented by a black 'X'.

Table 4.2. Layer 2 Aggregate Traffic Between FAPs under Scenario 2.

Source	Destination	Traffic (Gbps)
STX-03	STX-02	10.7
STX-04	STX-03	8.9
STX-05	STX-04	8.0
STX-07	STX-05	6.0
STX-07	STX-17	5.1
STX-17	STX-13	2.7
STX-15	STX-13	1.0
STX-06	STX-05	0.9
STX-12	STX-17	0.4

Note: All FAP-to-FAP layer-2 links are bidirectional (upload/download), but traffic flows are dominated by download rates. We assume each FAP-FAP link has a maximum bandwidth of 10 Gbps. Links that do not flow traffic or are disconnected from the network are omitted.

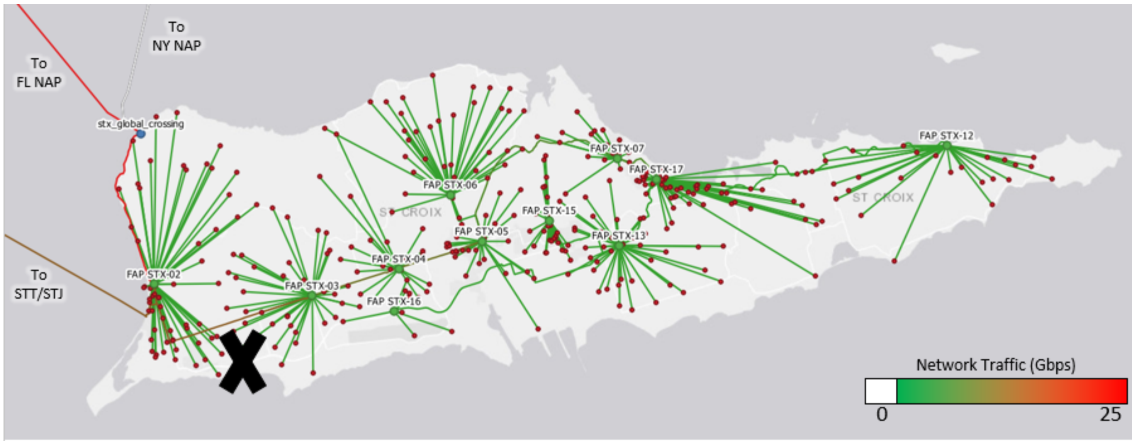

Figure 4.10. St. Croix Layer-2 Connectivity to Customers Under Scenario 2. Even with a fiber cut, the middle mile self-heals and no Internet customers lose connectivity. Customers that with Internet connectivity are represented by red circles with total traffic (upload and download) represented by shaded lines. Green lines represent light network traffic and red lines represent heavy traffic. Unused layer-2 links are represented as grey lines.The location of the fiber cut is represented by a black 'X'.

Scenario 3: Average Demand, Fiber Cuts on Melvin H. Evans Highway and Centerline Road

We study traffic re-routing and network congestion for a cable cut on the Melvin H. Evans Highway and Centerline Road — both the southern and central branches of the viNGN middle mile ring. Figures 4.11 and 4.12 present traffic routing and connectivity. Table 4.3 presents the total traffic flows over the viNGN middle mile.

Figure 4.11 presents the Internet traffic across the viNGN middle STX with two cable cuts. In this scenario, the middle mile ring is unable to self heal, and the majority of STX customers are unable to connect to the Internet. Figure 4.12 shows that only four FAPs — STX-02, STX-03, STX-04, and STX-16 — have connectivity to the IXP and North America. A total of 7.8 Gbps from the central and eastern portions of STX is disconnected from the Internet. Due to customer disconnection, there is a corresponding drop in traffic along layer-2 links. This reduces Internet congestion, as the total flow of Internet demand from STX is only 4.74 Gbps (Table 4.3).

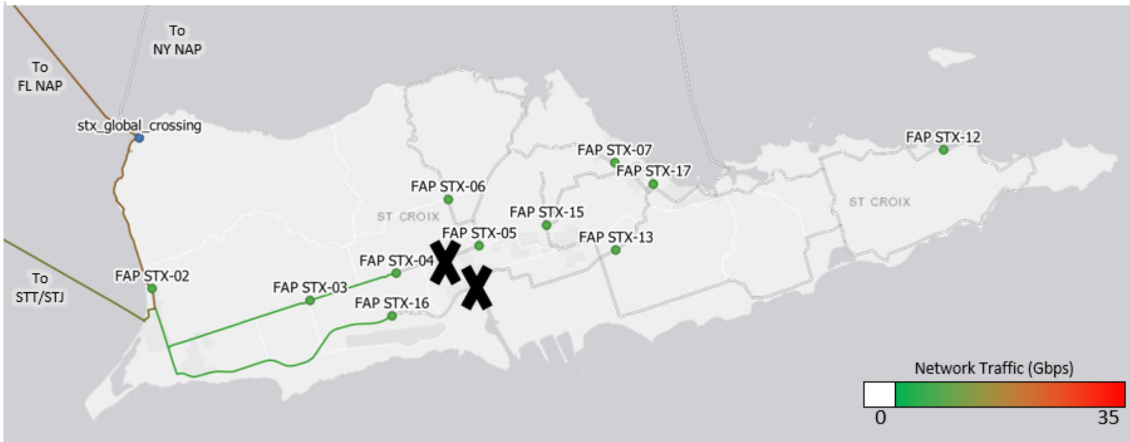

Figure 4.11. St. Croix Middle Mile Traffic under Scenario 3. A fiber cut along the Melvin H. Evans Highway and Centerline Road disconnects a large portion of the viNGN middle mile from the Internet. The amount of Internet traffic flowing on the STX middle mile is represented by shaded lines. Green lines represent light traffic and red lines represent high traffic. Unused layer-2 links are represented as grey lines. The location of the fiber cut is represented by a black 'X'.

Table 4.3. Layer 2 Aggregate Traffic Between FAPs under Scenario 3.

Source	Destination	Traffic (Gbps)
STX-03	STX-02	2.7
STX-04	STX-03	0.9
STX-16	STX-02	0.2

Note: All FAP-to-FAP layer-2 links are bidirectional (upload/download), but traffic flows are dominated by download rates. We assume each FAP-FAP link has a maximum bandwidth of 10 Gbps. Links that do not transport traffic or are disconnected from the network are omitted.

Scenario 4: Average Demand, Fiber Cut Between the Super FAP and IXP

In our previous scenarios, we observed a high traffic concentration between SuperFAP STX-02 and the IXP, suggesting a likely single point of failure in the network. We study a fiber cut between these two critical network points.

Figure 4.13 presents Internet traffic and customer connectivity to the viNGN middle mile

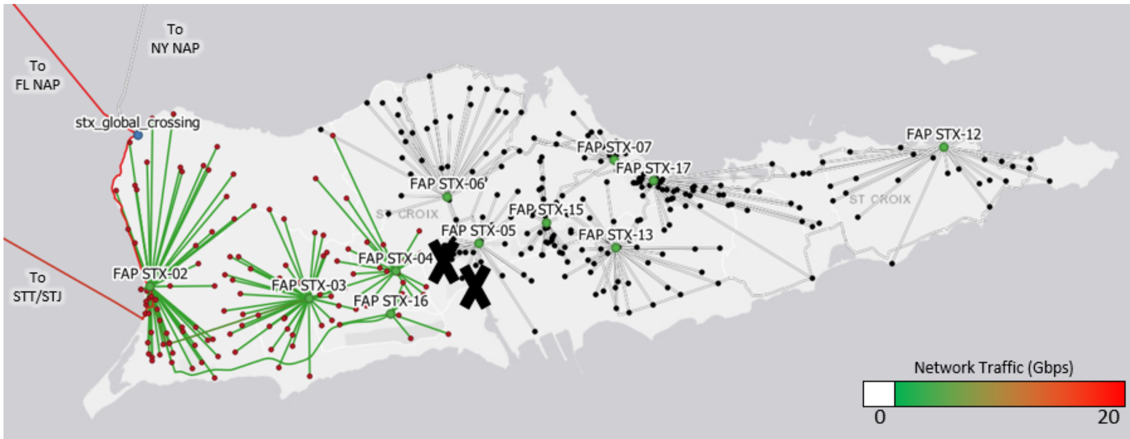

Figure 4.12. St. Croix Layer-2 Connectivity to Customers Under Scenario 3. With two fiber cuts, a large number of customers are disconnected from the Internet (black circles). Customers that with Internet connectivity are represented by red circles with total traffic (upload and download) represented by shaded lines. Green lines represent light network traffic and red lines represent heavy traffic. Unused layer-2 links are represented as grey lines. The location of fiber cuts is represented by a black 'X'.

with a single fiber optic cable cut between SuperFAP STX-02 and the IXP. This cut prevents a layer-2 link from being established between Super FAP STX-02 and IXP and disconnects the entire territory from the Internet. As a result, there is no traffic at any layer of the OSI model across STX. Due to the structure and function of the STX Internet system model, we expect similar results if any fiber run between the Super FAP and IXP is cut, the Super FAP and/or IXP nodes become unavailable, or if a cable cut occurs at the Internet landing site north of the IXP.

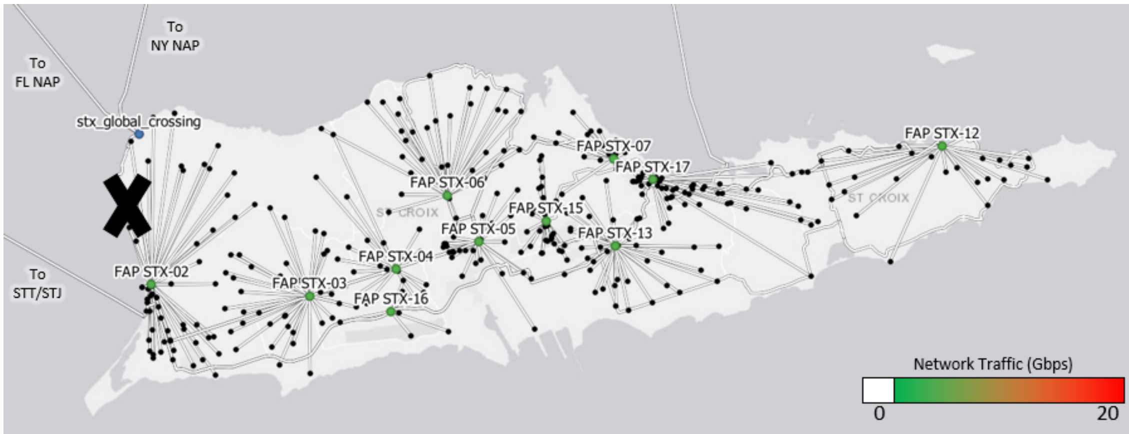

Figure 4.13. St. Croix Layer-2 Connectivity to Customers Under Scenario 4. A single fiber cut between the Super FAP and IXP leads to all customers on STX losing Internet connectivity (black circles). There is no Internet traffic within the STX middle mile. Unused layer-2 links are represented as grey lines. The location of the fiber cut is represented by a black 'X'.

Scenario 5: Peak Demand, All Infrastructure Available

Finally, we study the impacts of a demand spike on Internet connectivity and congestion. Here, there are no fiber cuts, but rather all customers on STX demand Internet content at their provider maximum bandwidth. In this scenario, traffic re-routing does not occur and follows similar paths to Scenario 1 (see Figure 4.7). However, network congestion does occur due to significant bandwidth requirements.

Figure 4.14 presents of aggregate (upload and download) traffic on the viNGN middle mile under peak demand with all layer-1 edges available. Under this scenario, all customers remain connected to the Internet. The model experiences heavy aggregate layer-2 traffic over links between CONUS and the SuperFAP (130.3 Gbps) and between the SuperFAP and STT/STJ (60.6 Gbps).

Table 4.4 shows which FAP-FAP connections are most congested during peak demand. Because Traffic routing follows the same hub and spoke configuration as average demand scenarios, congestion occurs over layer-2 links that flow the highest traffic in normal situations. For example, the layer-2 link between the SuperFAP and the IXP dominate network congestion. In addition, six more FAP-FAP connections are also expected to be congested due to their network traffic exceeding 10 Gbps: STX-02-STX-03, STX-02-STX-16, STX-16-STX-13, STX-03-STX-04, STX-04-STX-05, and STX-13-STX-17. Overall, we expect this scenario to lead to high network congestion and a significant amount of unserved demand.

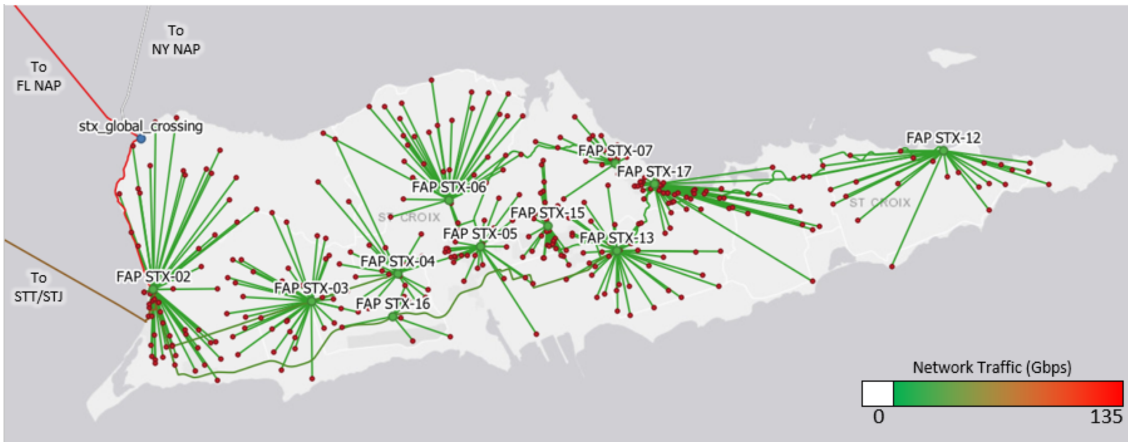

Figure 4.14. St. Croix Layer-2 Connectivity to Customers Under Scenario 5. All Internet customers remain connected to the viNGN middle mile and routing match Scenario 1 (see Figure 4.7). However, peak demand will lead to significant network congestion. Customers that with Internet connectivity are represented by red circles with total traffic (upload and download) represented by shaded lines. Green lines represent light network traffic and red lines represent heavy traffic. Unused layer-2 links are represented as grey lines.

Table 4.4. Layer 2 Aggregate Traffic Between FAPs under Peak Demand

Source	Destination	Traffic (Gbps)
STX-02	STX-03	29.6
STX-02	STX-16	29.5
STX-16	STX-13	28.0
STX-03	STX-04	20.2
STX-04	STX-05	15.5
STX-13	STX-17	14.6
STX-13	STX-15	5.8
STX-05	STX-06	4.5
STX-05	STX-07	4.4
STX-17	STX-12	2.3

Note: All FAP to FAP layer-2 links are bidirectional (upload/download), but traffic flows are dominated by download rates. We assume each FAP-FAP link has a maximum bandwidth of 10 Gbps. Links that do not flow traffic or are disconnected from the network are omitted.

4.2.2 Estimating STX Internet Operations and Vulnerability

We estimate the impact of re-routing and increased demands on the STX Internet system model using the three metrics introduced in Section 3.2.4. $D_{disconnected}$ is the amount of customer demand in Gbps disconnected from the Internet. C_{st}^{max} is the network congestion, estimated by the most congested layer-2 FAP to FAP link. And U is the amount of layer-2 traffic that exceeds link bandwidth on FAP to FAP links. Table 4.5 presents these three metrics for all five scenarios.

Table 4.5. Scenario Comparison

Scenario	$D_{disconnected}$ (Gbps)	C_{st}^{max}	U (Gbps)
1: Average Demand	0.0	0.55	0
2: Average Demand, One Cut	0.0	1	0.7
3: Average Demand, Two Cuts	7.8	0.27	0
4: Average Demand, Catastrophic Cut	12.5	0.0	0
5: Peak Demand	0.0	1	77.4

During normal network operations (Scenario 1), all customers are connected to the Internet ($D_{disconnected} = 0$) and the network is not congested, $C_{st}^{max} = 0.45$. These results suggest that the viNGN middle mile is capable of supporting our estimated average demand when all layer-1 fiber runs are operational.

Scenario 2 shows that network re-routing leads to minor network congestion. When a cable is cut along the Melvin H. Evans Highway, all customers remain connected to the Internet $D_{disconnected} = 0$. However, layer-2 traffic is transported over Centerline Road causing demand to exceed layer-2 bandwidth limits ($C_{st}^{max} = 1$). Still, total unserved demand is relatively small at $U = 0.7$ Gbps. These results suggest that the viNGN network is relatively robust to a single fiber optic cable cut along either the central or southern link-path.

Scenario 3 is the first scenario where customers are disconnected from the Internet. Specifically, a cable cut along the Highway and Centerline Road disconnects Internet customers in central and eastern STX, resulting in $D_{disconnected} = 7.8$ Gbps. Disconnecting customers reduces network demands, such that network congestion is also reduced to $C_{st}^{max} = 0.27$. These results indicate that, while multiple cable cuts may disconnect a significant number of customers, those that remain connected to the network will not experience any network degradation. In fact, customers on the western side of STX may experience better Internet connectivity due to reduced congestion.

Scenario 4 demonstrates the vulnerability of the STX Internet system model to single points of failure. With a single fiber cut between the SuperFAP and the IXP, we see that all customers are disconnected from the network ($D_{disconnected} = 24.7$ Gbps). This value includes traffic from STT/STJ that traverse inter-island submarine cables towards the IXP and North American connectivity. Since there is no traffic flowing through the network, $C_{st}^{max} = 0$ and $U = 0$.

Finally, the peak demand scenario (Scenario 5), provides high estimate of network congestion. Here, all customers are connected to the Internet $D_{disconnected} = 0$, but $C_{st}^{max} = 1$, indicating that a layer-2 link has more traffic flowing over it than it has bandwidth to support. Also, we observe $U = 77.4$ Gbps, indicating that there are there is a total of 77.4 Gbps more layer-2 traffic between FAPs than their bandwidth can support. Such a high value for U would significantly degrade Internet service in the USVI. These results suggest that the viNGN network is unable to support peak Internet demands or that our model severely over-estimates peak demands.

4.3 Synthetic Internet Traffic and Network Operations for the USVI Territory

In addition to assessing traffic on STX, we consider synthetic Internet flows and operations between islands and the mainland US. Table 4.6 presents flows between STT and STX, the links that carry combined island flow on STX (i.e., STX Super FAP and IXP), and between the USVI territory and CONUS (Miami and New York NAPs). While the viNGN middle mile network serves traffic flows on STX with average demands, when combining flows across the territory, the network becomes bandwidth limited with average demands.

All links are assumed to be 10 Gbps, yet Table 4.6 shows traffic demand greater than 10 Gbps on all inter-island links. Our synthetic Internet model assigns traffic without considering bandwidth constraints to estimate unserved demands. Thus, results where traffic exceeds 10 Gbps indicate possible bottlenecks within the USVI Internet system. In particular, the link between the STX Super FAP and the IXP carries all Internet traffic for the territory and is bandwidth limited with average Internet traffic, with unserved demand $U = 14.7$ Gbps. Unserved demand is especially large during peak demand, with $U = 120.4$ Gbps, given our assumptions about demand for viNGNs Internet services in the territory.

The STX Super FAP - IXP link has such large unserved demands because all Internet traffic across the entire territory is routed over this single link. This limitation on the network further emphasizes the vulnerability of the USVI Internet to a single fiber cut. Whereas Scenario 4 above shows the entire island of STX loses Internet connectivity if a fiber run along the STX Super FAP - IXP link is cut, Table 4.6 indicates that all customers within the territory may lose Internet connectivity from this catastrophic scenario.

Table 4.6. Layer 2 Aggregate Traffic Between Islands and the Continental United States.

Source	Destination	Traffic Average (Gbps)	Traffic Peak (Gbps)
STT Super FAP	STX Super FAP	12.1	60.7
STX Super FAP	IXP	24.7	130.4
IXP	Miami FAP	12.3	65.2
IXP	New York FAP	12.3	65.2

Note: All layer-2 links are bidirectional (upload/download), but traffic flows are dominated by download rates. We assume each link has a maximum bandwidth of 10 Gbps, and we do not expect traffic across any link to exceed this value. All demands in excess of 10 Gbps are assumed to be unserved.

4.4 Summary

Taken together, we find the STX Internet system model to be more vulnerable to infrastructure failures and demand spikes than expected. The distribution of customers across FAPs means that the loss of key middle mile infrastructure and disconnect a large portion of the STX population and CAIs. While the viNGN network is designed to have redundant layer-2 links and self-healing capabilities, there are also single points of failure along the layer-1 physical network that can disconnect customers. Routing of traffic between islands indicates that submarine links and the link between the STX Super FAP and IXP are bandwidth limited during average demand scenarios. Moreover, the available bandwidth at layer-2 on STX cannot support surges in demand with or without fiber cuts.

THIS PAGE INTENTIONALLY LEFT BLANK

CHAPTER 5: Conclusion

We conclude with a summary of our synthetic network generation results, limitations of our model, and recommendations for future work.

5.1 Summary

This thesis completes the following data, modeling, and analysis tasks:

1. Gathered publicly available information about the USVI Internet system into geospatial data sets for visualization and modeling;
2. Developed network generation techniques for producing synthetic Internet models;
3. Produced a synthetic model of the STX Internet system; and,
4. Analyzed the structure and characteristics of the STX Internet system model to measure how it performs under different demands and disruptions.

Analysis of the STX Internet system model suggests that private Internet customers, CAIs, and Internet demand are not evenly distributed between FAPs. Some FAPs appear to be more essential than others based on the size of the population they support and the types of CAIs they support. Further analysis of traffic dependencies is needed to arrive at a final priority ranking of FAPs for infrastructure hardening and restoration after disruptions.

However, our model suggests that the viNGN network on STX has several vulnerabilities. The most pressing are several single points of failure that, if cut, completely disconnect all customers on STX from the Internet. Specifically, if the submarine cable landing site at the IXP, the Super FAP, or any fiber optic cables in between these nodes are disrupted, the network will suffer a catastrophic failure.

In the absence of a catastrophic cut, our model suggests that the viNGN network is fairly robust to a single fiber optic cable cut. The network can re-direct Internet traffic along alternate routes to ensure that all Internet customers remain connected to the Internet and only experience minimal unserved demand with average traffic loads.

5.2 Model Limitations

Development of the STX Internet system model was conducted using only publicly available information. Much of this information is over a decade old, generally incomplete, and occasionally contradictory. Although certain features—such as the location of residential populations and main right-of-ways along the road system—are not believed to have change, much of the data that we used has not been validated. Moreover, the Internet model we assess is based on pre-hurricane infrastructure networks. Future work should build on the data, models, and analysis presented herein to guide protection of the post-storm system.

Moreover, our analysis required a number of assumptions based on Internet use in the mainland US that may not be appropriate for the island context. For example, average and peak demands were based on general market trends in the US. We chose to make this assumption, because infrastructure systems in the USVI look similar to those in the mainland. However, systems in the USVI also have unique features that require local knowledge to capture. We have tried to the best of our ability to capture these details through site visits and frequent discussions. But some of this knowledge, such as network configuration decisions, is known only by the local operators of these systems. For this reason, we believe that the results here should be taken only as "approximately right" and the model inputs adjusted as more data becomes available in the future.

5.3 Future Work

This analysis has considered only the system on STX, and as such it potentially misses system properties that arise from the interconnections STT and STJ. A complete analysis of the Internet infrastructure across the entire Territory is needed to understand the resilience of the aggregate system. Moreover, because of the distinct differences between these islands, a comparable analysis of the system on each island also deserves attention.

The function and performance of this model can be significantly improved through collaboration with local stakeholders who can validate our data and modeling assumptions. We look forward to the opportunity to work with local stakeholders in the USVI to improve the results presented here.

List of References

Alderson D, Bunn B, Eisenberg D, Howard A, Nussbaum D, Templeton J (2018) Interdependent Infrastructure Resilience in the US Virgin Islands: Preliminary Assessment. Technical Report NPS-OR-18-005, Naval Postgraduate School, Monterey, CA.

Alderson D, Willinger W, Li L, Doyle J (2006) An optimization-based approach to modeling internet topology. *Telecommunications Planning: Innovations in Pricing, Network Design and Management*, 101–136 (Springer).

Alderson DL, Brown GG, Carlyle WM (2015) Operational models of infrastructure resilience. *Risk Analysis* 35(4):562–586.

Alderson DL, Brown GG, Carlyle WM, Cox Jr LA (2013) Sometimes there is no "most-vital" arc: Assessing and improving the operational resilience of systems. *Military Operations Research* 18(1):21–37.

Austin DA (2018) Economic and Fiscal Conditions in the US Virgin Islands. Congressional Research Service Report R45235, Washington, D.C.

Borgdorff A (2020) *A Demand and Distribution Model for Potable Water in the USVI*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Brown G, Carlyle M, Salmerón J, Wood K (2006) Defending critical infrastructure. *Interfaces* 36(6):530–544.

Bunn B (2018) *An Operational Model of Interdependent Water and Power Distribution Infrastructure Systems*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Çetinkaya EK, Alenazi MJ, Peck AM, Rohrer JP, Sterbenz JP (2015) Multilevel resilience analysis of transportation and communication networks. *Telecommunication Systems* 60(4):515–537.

Central Intelligence Agency (2020) The World Factbook. Washington, DC.

Çetinkaya EK, Rohrer JP, Sterbenz JP (2012) Resilience of backbone provider networks .

Crain JK (2012) *Assessing resilience in the global undersea cable infrastructure*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Cybersecurity and Infrastructure Security Agency (2020) Critical infrastructure sectors. https://www.cisa.gov/critical-infrastructure-sectors.

Derosier JA (2008) *Internet topology generation based on reverse-engineered design principles: Performance tradeoffs between heuristic and optimization-based approaches*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Federal Communications Commission (2012) Application for domestic submarine cable landing license and request for streamlined treatment. Retrieved 31 May 2020, http://licensing.fcc.gov/myibfs/download.do?attachment_key=979131.

GeoPandas Development Team (2019) GeoPandas Open Source Project. Available online http://geopandas.org.

Good J (2019) *An Operational Model of Food Supply Chain and Transportation Systems in the U.S. Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Hagberg A, Swart P, S Chult D (2008) Exploring network structure, dynamics, and function using NetworkX. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

Jabbar A, Shi Q, Çetinkaya EK, Sterbenz JP (2008) KU-LocGEN: Location and cost-constrained network topology generator. Technical Report ITTC-FY2009-TR-45030-01, The University of Kansas, Lawrence, KS.

Lee B (2020) At&t leaves stj without internet, creates $100,000 mistake. *The St. Thomas Source* Retrieved 31 May 2020, https://stthomassource.com/content/2020/03/12/att-leaves-stj-without-internet-creates-100000-mistake/.

Martin KM (2014) *A geographic and functional network flow analysis tool*. Master's thesis, Department of Computer Science, Naval Postgraduate School, Monterey, CA.

Murhammer MW, Atakan O, Bretz S, Pugh LR, Suzuki K, Wood DH (1998) *TCP/IP tutorial and technical overview* (Prentice Hall Upper Saddle River, NJ), accessed May 5, 2020, https://download.itadmins.net/Linux/IBM%20Redbook%20-%20TCP%20IP%20tutorial.pdf.

Pew Research Center (2019) Internet/broadband fact sheet. Retrieved 31 May 2020, https://www.pewresearch.org/internet/fact-sheet/internet-broadband/.

QGIS Development Team (2019) QGIS Geographic Information System. Open Source Geospatial Foundation Project. Available online http://qgis.org.

Reuters (2017) Cars make their way under a partially-collapsed utility pole one week after hurricane maria raked the island, in frederiksted, st. croix, u.s. virgin islands. *Reuters* Accessed 26 Sept 2017, https://news.trust.org/item/20170926230056-y1xtb.

Rohrer JP, Jabbar A, Sterbenz JP (2014) Path diversification for future internet end-to-end resilience and survivability. *Telecommunication Systems* 56(1):49–67.

Sterbenz JP, Çetinkaya EK, Hameed MA, Jabbar A, Qian S, Rohrer JP (2013) Evaluation of network resilience, survivability, and disruption tolerance: analysis, topology generation, simulation, and experimentation. *Telecommunication systems* 52(2):705–736.

Telegeography (2020) Submarine cable map. Accessed May 5, 2020, https://www.submarinecablemap.com/.

US Census Bureau (2010a) Population counts for estates: 2010. Retrieved from https://www.census.gov/population/www/cen2010/island_area/usvi.html.

US Census Bureau (2010b) Understanding the population of the U.S. Virgin Islands. Retrieved from https://www.census.gov/content/dam/Census/programs-surveys/sis/resources/2020/sis_2020map_usvi_k-12.pdf.

US Census Bureau (2019) Historical household tables. Retrieved from https://www.census.gov/data/tables/time-series/demo/families/households.html.

USVI Hurricane Recovery and Resilience Task Force (2018) Final report. Technical report, St. Thomas, USVI, available electronically from https://www.usvihurricanetaskforce.org/.

Virgin Islands Next Generation Network (2019) Network restoration status. Accessed April 16, 2019, https://vingn.com/network-restoration-status/.

Virgin Islands Next Generation Network (December 2012) Broadband business & technology meeting. Technical report.

Virgin Islands Public Finance Authority (2009a) Comprehensive community infrastructure program. Technical Report NT10BIX5570096.

Virgin Islands Public Finance Authority (2009b) Public computer centers program. Technical Report 78-42-B10568.

Virgin Islands Public Finance Authority (2009c) State broadband data and development grant program. Technical Report 78-50-M09055.

Viya (2019) Viya network update. *Facebook* Retrieved 31 May 2020, https://www.facebook.com/ViyaUSVI/posts/10157700261291385.

Viya (2020) Viya residential internet plans. Accessed 31 May 2020, https://viyavi.com/home/internet/plans/.

Wille D (2019) *Simulation Optimization for Operational Resilience of Interdependent Water-Power Systems in the US Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Wine W (2020) *Building a Resilient Cell Phone Network in the US Virgin Islands*. Master's thesis, Department of Operations Research, Naval Postgraduate School, Monterey, CA.

Woodman DG (2016) *Crowdsourcing physical network topology mapping with net.Tagger*. Master's thesis, Department of Computer Science, Naval Postgraduate School, Monterey, CA.

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California