Submarine Network Fault Prevention

Switch Fault Pattern Analysis

ELECOMP Capstone Design Project 2022-2023

Sponsoring Company:

Rite-Solutions, Inc.
One Corporate Place
Middletown, RI. 02842
http://www.rite-solutions.com

Company Overview:

Rite-Solutions is an award-winning Veteran-Owned Small Business (VOSB) headquartered in Middletown, Rhode Island, established since April 2000. We have a stable and deep corporate history in Rhode Island, with over 280 employees with core competencies including systems engineering, software development, information technology, and cyber engineering. The stability of our business is demonstrated by both our longevity and our business backlog – well over $200M in recent contract awards that will be executed over the next 5 years. We have achieved both state and national recognition; one of our founders, Joe Marino, is an active member of the RI Science and Technology Advisory Council, along with other trusted Rhode Island business and academic leaders.

Rite-Solutions is known for innovation and dedication to the information and decision support needs of our government and commercial customers, a commitment we have coined as a company slogan, *The Information Advantage™.* We have significant US Navy prime contracts in both warfare systems and business systems development and sustainment, as well as in Information Technology infrastructure support and cyber protection. In our research and development efforts, we demonstrate a proclivity for our own inventive solutions but also for finding and working with non-traditional partners, including academia and small commercial businesses, whose technologies we adapt to national research objectives in practical and meaningful ways that lead to productization.

Our corporate culture is a major part of our drive to both attract and retain people who make a difference. We take pride in our ethos of the F.E.W. (Friends Enjoying Work). Our ownership and management team work to create an open, creative, and stimulating environment. Our founders

embraced the servant-leader concept from company inception, and it continues to guide our management team and approach. We recognize great ideas can and do come from anyone, not just the C-suite. To facilitate ideation in an engaging format, we built and maintain a company-wide platform called Mutual Fun that promotes corporate giving, training, and innovation in technology and our culture in a stock game format. This approach to business nurtures the employer/employee relationship and moves our teams to a model where anyone who wants to can be part of the company's growth and contribute to our success. We have won numerous awards, including a recent award in 2021 as the **Providence Business News "best place to work"** among large organizations (> 150 employees), and received extensive media coverage and academic interest about our culture, including as the subject of case studies by Harvard University and Stanford University, in employee motivation and idea generation techniques. In 2022, we had a national organization – Great Places to Work - conduct a survey where our employees could anonymously rate Rite-Solutions. The results were astounding – **96% of our employees say Rite-Solutions is a great place to work** as compared to 57% at a typical U.S. based company.

Technical Directors:

Name: Matt Pistacchio
Title: Project Manager & Technical Lead
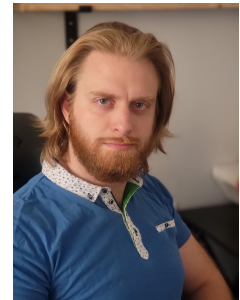Email: mpistacchio@rite-solutions.com
LinkedIn

Name: Matt McAdams
Title: Computer Engineer
Email: mmcadams@rite-solutions.com

Project Motivation:

Our US Navy's Submarine Fleet depends on the accuracy, fidelity, and reliability of its dependent hardware and software communications. These communications are not possible unless we have a resilient network infrastructure to bridge each sub-component together. Contacts, or identified entities in the proximity of a submarine, detected from SONAR need to be communicated to the vessel's Combat System (Tactical Control) to determine the correct course of action. A submarine's network performance is responsible for allowing our warfighters the split-second notice they need to avoid a collision, whether friendly or otherwise.

Our project addresses this responsibility by ultimately focusing on the problem of network fault prevention and analyzing the conditions upon which a network can fail. Applying today's concepts of Data Science and their relationship to Pattern Recognition principles to a publicly available data set, this capstone project is intended to identify and deduce evidence-based conclusions on what contributing factors lead to switch port bandwidth, and thus point-to-point network health, being compromised.

Anticipated Best Outcome:

The goal of this project is to research and analyze curated data from publicly available network datasets to identify system/component anomalies such as anomalous traffic, anomalous network bandwidth, etc. (whatever the assigned team identifies as significant to the overall problem). Network port bandwidth or input and output utilization is a recommended firsdt step. If bandwidth is chosen, it can be assumed for the scope of this project that a port bandwidth failure, measured as a utilization percentage, would be a percentage exceeding 80%. One such applicable public data set is Australia's University of New South Wales (UNSW) Canberra available here, but the assigned team is encouraged to leverage another data set if desired. Tools such as Python, MatLab, and/or R are recommended for performing the requisite analysis. The graphics one can generate from these tools would then be featured throughout a final report or White Paper that introduces, evidences, and concludes the findings of the effort. Ultimately, we want to establish a tool(s) that we can host on our Rite-Solutions network and a process that can be applied/used with tactical or production datasets. Additionally, we will use the data as input to the anomaly detection algorithm developed by Dr Noah Daniels at URI and improved via a grant from the Office of Naval Research.

Project Details:

Overall system concept:

This project can be boiled down to the following 3 stages (deliverables are **bolded**):

1. Exploratory Data Analysis (EDA)
   - It's vital to learn and inquire to the significance of the features isolated by the UNSW data set and how they apply to the problem we're trying to solve.
   - Here's a table isolating and describing some of the more significant parameters or features from "NUSW-NB15_features.csv" in the UNSW provided data set:

No.	Name	Type	"Nice" Name	Description	Comment
1	srcip	nominal		Source IP address	Need to filter on source switch host
2	sport	integer		Source port number	
3	dstip	nominal		Destination IP address	Destination host/IP is important for distinguishing key connections
4	dsport	integer		Destination port number	Destination port is important for distinguishing key connections
8	sbytes	Integer	Bytes Out	Source to destination transaction bytes	May be more useful for utilization calculation than "Packets Out" (more compatible for unit conversion) - need to translate to SNMP's ifOutOctets
9	dbytes	Integer	Bytes In	Destination to source transaction bytes	May be more useful for utilization calculation than "Packets In" (more compatible for unit conversion) - need to translate to SNMP's ifInOctets
12	sloss	Integer	Input Packet Loss	Source packets retransmitted or dropped	Of special interest as a contributing factor
13	dloss	Integer	Output Packet Loss	Destination packets retransmitted or dropped	Of special interest as a contributing factor

16	Dload	Float	Port Speed	Destination bits per second	Can translate to ifSpeed for utilization calculation (needs unit conversion)
17	Spkts	integer	Packets Out	Source to destination packet count	
18	Dpkts	integer	Packets In	Destination to source packet count	
27	Sjit	Float		Source jitter (mSec)	Of special interest as a contributing factor
28	Djit	Float		Destination jitter (mSec)	Of special interest as a contributing factor
48	attack_cat	nominal	Attack Vector	The name of each attack category. In this data set , nine categories e.g. Fuzzers, Analysis, Backdoors, DoS Exploits, Generic, Reconnaissance, Shellcode and Worms	The Attack vector of interest

The above coupled with additional research is intended to help with grooming, filtering, and assessing the UNSW data set. Other features that would be recommended to consider through this or other data sets would be that of switch device temperature, switch device CPU Usage, and port CRC errors.

- o A **"Preliminary Findings" report** is expected as an initial milestone will be used as a checkpoint for mentorship and feedback from your technical directors. The scope of this document is to describe the central features, their significance to the problem at hand, and their relationship with each other.

2. Analysis tool identification & execution
   - o Matlab or Python's matplotlib, numpy, and pandas modules are examples of tools you could consider. Whatever tool is chosen is needed to support the EDA pipeline (e.g., data gathering collection/storage capabilities, data analysis (including statistical analysis), data cleaning (to enhance data quality), and data visualization).
   - o Bandwidth, if used, should be calculated as input and output utilization percentage as described here:

$$\text{Input utilization} = \frac{\Delta ifInOctets \times 8 \times 100}{(\text{number of seconds in } \Delta) \times ifSpeed}$$

$$\text{Output utilization} = \frac{\Delta ifOutOctets \times 8 \times 100}{(\text{number of seconds in } \Delta) \times ifSpeed}$$

   - o The given reference is to the scope of Simple Network Management Protocol (SNMP) query returns from a switch; dimensional analysis or unit conversion is needed to align the formulae with the outputs of the data set you use. See the following references for ifInOctets, ifOutOctets, and ifSpeed.
   - o Bandwidth as it approaches or exceeds the failure threshold (80%) across the ports of a switch, if selected as a feature, should be modeled in plots overtime and cross-plotted against the trends of appropriate related features.
   - o A **PowerPoint brief depicting and describing the models, graphs, and any other visuals created** using the analysis tools and initial conclusions would then serve as a second milestone/deliverable.

3. Finding reporting & demonstration
   - o A **White Paper** more formally articulating the findings of the feature analysis and a **poster demonstration** of the project will be the final milestone.

Composition of Team:

2 Computer or Software Engineers

Skills Required:

Computer Engineering Skills Required:
- Network Engineering basics
- Data Science basics
- Pattern Recognition basics
- MatLab, Python, R, WireShark, or other related analysis tool proficiency

Anticipated Best Outcome's Impact on Company's Business, and Economic Impact

Rite-Solutions is continuing to expand our network monitoring and fault detection capabilities for our DoD/Navy customers. We are currently a prime contractor for the AN/BYG-1 Inter-Subsystem Monitoring Tool (ISMT), a tool which monitors and facilitates the troubleshooting efforts for the hardware and software communications aboard a submarine. R&D efforts such as this enhance the tool and thus the Fleet's capabilities for network threat detection as well as train our potential future workforce on how to solve the problems of this program and beyond.

Broader Implications of the Best Outcome on the Company's Industry:

The US Navy prioritizes Cyber Security and Cyber threat prevention. By upgrading ISMT's capability to prevent network port bandwidth faults, there is also cross-over to preventing denial of service attacks that can have the potential to impede critical submarine communications. This raises the significance of this overall effort and can provide our Submarine Fleet with the solution they need to avoid this worst-case scenario.