

The Center for Cyber Defenders

Expanding computer security knowledge

Warmachine

Analyzing Transport Layer Security (TLS) Traffic with Machine Learning

Trevor Sorrells, Georgia Institute of Technology



Project Mentors: Joe Ingram, Org. 9365 and Jereme Lamps, Org. 5828



Overview

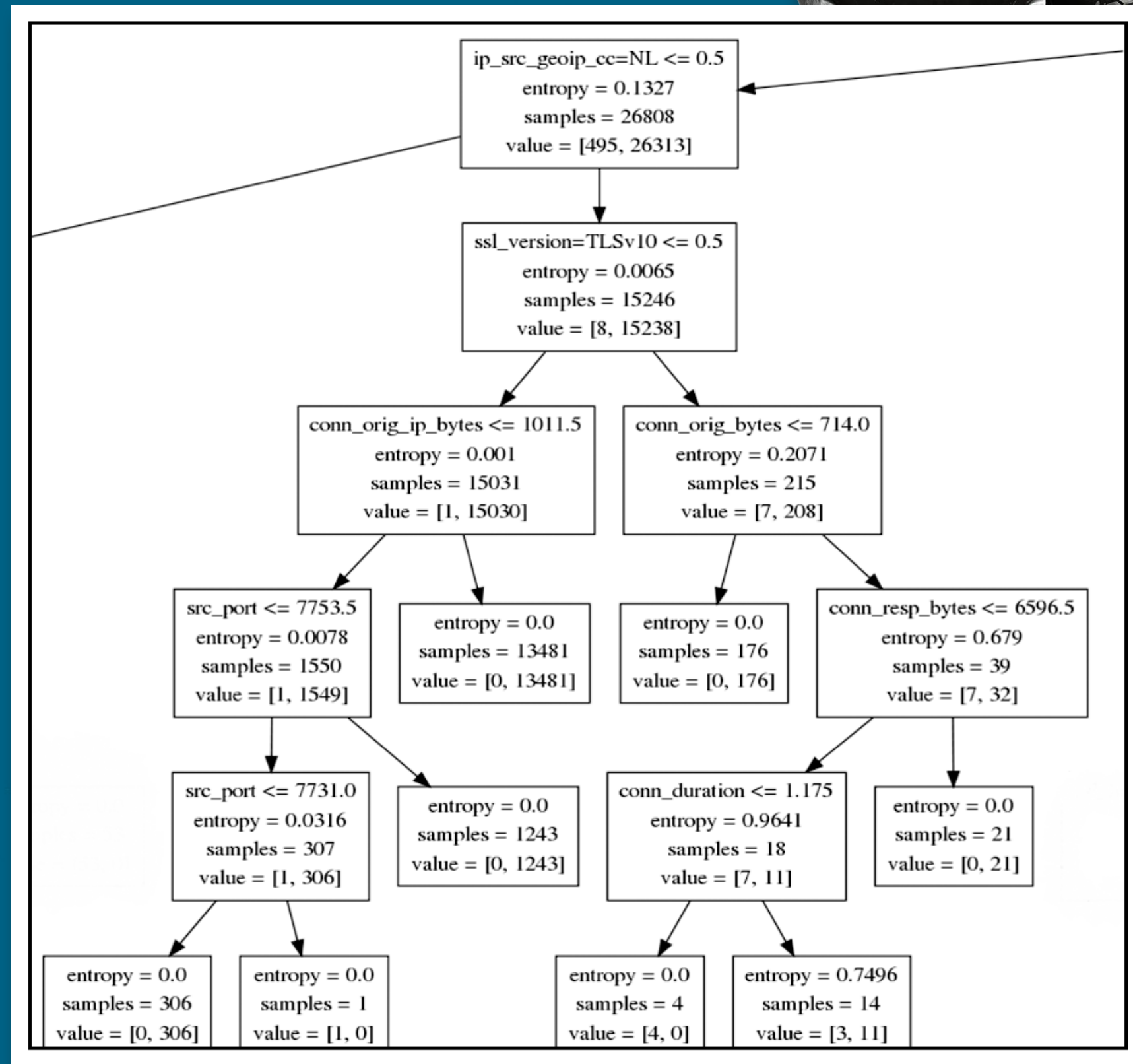
The goal of Warmachine was to take TLS network traffic and run it through a machine learning algorithm to identify key indicators of malicious traffic.

Objective:

- Write a Python script to put TLS traffic into a vector for the machine learning algorithm.
- Write a Python script to run the vector through sklearn's machine learning algorithm.

Approach:

- Started by finding out how to parse the data from the TLS traffic.
- Wrote a script to take the data from a text file and parse it into a Python library based on 46 fields within the data. The data was either malicious or non-malicious based off of the samples we had pulled from TLS traffic that was tagged by Tamizar. We then took these libraries and turned the data into a vector.
- Wrote a script to run the vector through sklearn's forest classifier and decision tree machine learning algorithm.
- Interpreted the output graph to write an analytic to find patterns indicative of bad TLS traffic and fed this information to Tamizar for classification and visualization.



Results:

- Found key indicators of bad traffic for Tamizar to analyze and visualize for hunters.
- Found previously unknown and unusual network traffic for my team to go over and investigate.
- Found 18 rules in total that could categorize the traffic as either bad or good, 14 being bad indicators and 4 being good.

Impact and Benefits:

- inform cybersecurity analysts on what to look for when identifying malicious TLS traffic.
- Provides a scoring function to help prioritize and tag relevant TLS traffic.