Wyatt Blair

EMT-678-WS: Big Data Technologies

10/25/2024

**Class 7 -Network Analytics for Streaming Traffic Analysis**

This paper examines modern techniques to identify malicious behavior on network traffic. The authors focus on two methods: Random Forest Classification (an offline solution) and Stream-GP Classification. The entire analysis centers around a particular benchmarking dataset called CTU-13. This dataset contains four classes: Background, Normal, Botnet, and Botnet C&C. The goal is to create a system which can correctly classify this dataset so it can be used for cybersecurity on an online service.

In the end, the authors summarize their findings by saying “Stream-GP based traffic analysis not only shed light into unknown traffic but also seems to generalize what it learns given the ground truth traffic better.” The authors also summarize their findings regarding network topology and traffic analysis; specifically, ground truth traffic and background traffic.