Date: Friday, October 12th
Time: 10:45am – 11:30am
Speakers: James Ezick, Reservoir Labs

Gaining Insight from Bro Logs through Pattern Discovery with Tensor Decompositions

The vast quantities of network metadata being captured by Bro-enabled appliances is opening exciting new possibilities for machine-learning-driven big data analytics and workflows. One emerging approach is based on the application of tensor decompositions – a new paradigm where pattern discovery is the starting point rather than the end goal of the analysis. We present a scalable workflow based on the newest ideas in the field that enables the unsupervised discovery of deep, cross-dimensional patterns within linked metadata. This approach creates a virtual roadmap for comprehending large-volume log data without the need for heroic up-front training or feature engineering. Tensor decompositions have been demonstrated to provide critical insight into what is actually taking place on a network. In this presentation, we will show how we have used tensor methods to complement signature-based deployments and discover patterns of activity from Bro logs indicative of:

- Malware beaconing
- Distributed port scans evolving to machine takeover
- Scans for printers or IoT devices
- DNS-based data exfiltration/insider threat
- Abuse of control and/or backchannel message streams
- Exploitation of application-specific port vulnerabilities
- Broken and/or misconfigured network services
- Network policy violations

Attendees will receive an introduction to tensor decompositions as a practical tool for network activity analysis. This will include insight into tensor methods as a rapidly evolving technology - one that provides an approach to unsupervised machine learning that discovers coherent patterns and derives real value from the sort of rich network metadata collected by Bro. Attendees will also gain an understanding of recent progress toward addressing barriers to automating workflows and deploying the technology in enterprise environments. The

presentation will be structured in three parts: (1) Introduction to Tensor Methods, (2) Applications, and (3) Practical Considerations.
In Part 1: Introduction to Tensor Methods, we will provide a gentle introduction to tensor methods, covering the basic concepts and terminology. The introduction will frame tensor methods as a form of unsupervised machine learning, describe their unique value proposition in terms of finding multidimensional patterns, and draw contrasts to other unsupervised methods used today such as k-means clustering and DBSCAN.

In Part 2: Applications, we will walk through examples of the applications of tensors to linked metadata, with a focus on Bro network traffic logs. From these examples, we will illustrate discovery of both benign and anomalous network behaviors. In addition, we will demonstrate how patterns revealed through tensor analysis can be the starting point for deeper explorations using tools such as Splunk. The focus of this section will be on illustrating what is achievable with tensor methods.

In Part 3: Practical Considerations, we will discuss experiences and set expectations for deploying tensor methods in an enterprise environment. This will include practical guidance on tensor construction, computational requirements, and result interpretation. We will share lessons learned from our own environment and give suggestions on how to automate tensor methods. The presentation will conclude with pointers to resources for those interested in further engagement.