

1163-15-1341

**Michael A. Ostroski\*** (michaelostroski1@gmail.com). *Dynamic Graph Edge Clustering: The Art of Conversation... Mining*. Preliminary report.

One approach to network security is attempting to find anomalies in network traffic. Netflow data provides records for how a network communicates and a natural way to model this data is with a dynamic multi-graph. We often know which nodes in our graph are important and need to answer the question of "when" we should be paying attention to those nodes. To accomplish this we look for a way to group the record data itself, effectively clustering the edges of the graph in order to identify anomalies. An interesting cluster of edges might be, for example, a pattern of behavior that is emblematic of a breach in security.

The Line Graph of a graph  $G$  is another graph that represents the adjacencies between edges of  $G$ . Forming the line graph for record data and clustering would result in a grouping of records. Unfortunately, the line graph can be several orders of magnitude larger than the  $G$ , which is unacceptable for many data sets. We present a scalable approach to edge clustering that uses an approximate line graph, filtered by time, as a way to capture potential causal relationships between records. Analysis of the time-filtered line graph captures higher-level phenomenology that are not explicitly coded in the data, suggesting that it is an effective model for these problems. (Received September 15, 2020)