BENOIT HAMELIN, Tutte Institute for Mathematics and Computing

Telemetry representation and interactive labeling to facilitate cyber defense

Cyber defense processes involve the processing and analysis of large volumes of telemetry data. Many phenomena of interest are described over tangled subsequences of telemetry streams, and cyber analysts experience difficulty making out the data patterns that drive their investigations. Behaviours over telemetry streams form a heavy-tailed distribution, precluding the proxy approach of anomaly detection; and there exists no labeling information to discern patterns of interest. This work presents tools for the interactive annotation of such collective telemetry phenomena by characterizing them from mutual similarity and time persistence. The annotation process and outcome is demonstrated through the example of host-based sensor telemetry.