Anomaly detection

Combined with unsupervised and supervised learning, an intranet dweb search engine can assist with anomaly detection (Denial of Service attacks, access from suspicious locations, probing of strange port addresses, unexpected downloads, viruses, spam) by ingesting router, switch, and other logs, and then applying predictive analysis techniques to incoming requests.

- Cluster data points within similar behavioural groups using inter-request time, number of unique endpoints and IP location (or content address (IPNS) or agent identifier).
- Train a classifier with labelled traffic data (request attributes) from historical malicious requests.
- Test on logs or data stores with unlabelled traffic data.
- Validate against data provided by third parties (coalition partners).

Current network monitoring systems provide data with a high degree of dimensionality, making large-scale application of machine learning approaches to improve the detection and classification of network attacks possible. Such widely adopted use usually consists of incorporating traditional machine learning models, for which a set of expertly handcrafted features is required to pre-process the data prior to training the models. This works for certain scenarios works, but …

- There is a systematic lack of a set of input features for a specific task, making generalisation and benchmarking of different approaches hard.
- Networking data is highly dynamic and static handcrafted features fail over time.
- The feature engineering process takes a lot of time and energy to come to a set of features that works best for a specific issue.

Deep learning models can complement conventional approaches, using different representations of the input data. Key is the ability of such models to learn feature representations from raw, non-processed input data.

Resources

- Deep in the dark: enhancing malware traffic detection with deep learning, tryo labs, 2019
- Deep in the Dark — Deep Learning-Based Malware Traffic Detection without Expert Knowledge, May 2019, Gonzalo Marín, Germán Capdehourat, Pedro Casas
- Deep in the Dark - Deep Learning-based Malware Traffic Detection without Expert Knowledge, Gonzalo Marín (IIE-FING, Universidad de la República), youtube
- Exploring Adversarial Examples in Malware Detection, Octavian Suciu, Scott E. Coulland, Jeffrey Johns

---

ir, se, search, project, tools