Cybersecurity + NLP

Petros Lambropoulos
Cybersecurity Analytics Software Engineer
NLP Enthusiast

1. Intrusion Detection

- Advanced Persistent Threat (APT)

  is a stealthy computer network threat actor, typically a nation state or state-sponsored group, which gains unauthorized access to a computer network and remains undetected for an extended period.

- Remote Access Tool (RAT)

  is a malware program that includes a back door for administrative control over the target computer.

- 365 days is the mean response time of a Security Operation Center (SOC) on APT’s

  a facility that houses an information security team responsible for monitoring and analyzing an organization’s security posture on an ongoing basis.

2. Modern NLP

- ImageNet moment has arrived for LM’s

  A statistical language model (LM) is a probability distribution over sequences of words. Given such a sequence, say of length m, it assigns a probability to the whole sequence.

- Unsupervised pre-training LM’s

  Large open unlabeled text corpora are used to train massive LM’s and condition their weights to realistic language distributions.

- Fine Tuning pre-trained LM’s

  Effective transfer learning methods that can be applied to any task in NLP, and introduce techniques that are key for fine-tuning a state-of-the-art language model.

3. Initial efforts show promise

- Recurrent Neural Network Language Models for Open Vocabulary Event-Level Cyber Anomaly Detection (Tuor et al 2018)

  Detecting anomalous behavior in computer and network logs, one that largely eliminates domain-dependent feature engineering employed by existing methods, by treating system logs as threads of interleaved "sentences" (event log lines) to train online unsupervised neural network language models.

- ATTACK2VEC: Leveraging Temporal Word Embeddings to Understand the Evolution of Cyberattacks (Shen et al 2019)

  Uses temporal word embeddings to model how attack steps are exploited in the wild, and track how they evolve.

4. Challenges for open work

- Large diverse unlabeled datasets of network traffic
- Small labeled dataset’s of malicious activity
- Well defined downstream tasks to compare on
- Leaderboards, SOTA etc for performing models
- Releases of large pre-train models like in NLP
- Preservation of privacy on open data/models