Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Prof. Dr. Christoph Meinel
Internet-Technologies & -Systems

Master Project Winter Term 2019/2020

Towards Successful Use Cases of Big Data Security Analytics

Background

Nowadays, it becomes usual that more and more organizations collect and store log data generated from the IT infrastructure across the entire corporate network. The challenges have been turned into how the large amount of data can be efficiently analyzed in terms of deriving meaningful values and insights. Within this master project, students are expected to work with the data collected from two real world cases and contribute with advanced analytical approaches helping take full advantage of the data and enhancing the security of enterprise network.

Case 1: The Cloud Lifecycle Management (CLM) of a modern software company consists of the whole procedure of cloud-based software development, testing, deployment, and maintenance. The available data includes the information about Lifecycle Management (LM) processes, tools' performance, infrastructure resources usage, deployed content, operated applications, trouble tickets, incidents, OS, etc. It is always expected that better understanding on these data can help reduce MTTD (mean time to detect) and MTTR (mean time to resolve), automate operations, as well as prevent potential vulnerabilities, threats and attacks.

Case 2: DNS logs generated by corporates' DNS servers are responsible for domain name resolution within the internal network. Due to the fact that DNS traffic is typically allowed by firewalls in most of organizations, it is extensively abused by cybercriminals (e.g., drive-by download, bots communication with command-and-control servers, etc.), hence leading to the popularity of DNS log analysis in the security domain. It is expected to deep and efficient investigations on DNS data (queries and responses) can be carried out to find innovative ways to detect known as well as unknown attacks against DNS servers in a faster manner.

.

Deliverables

- Advanced big data analytical approaches and threat detection/hunting approaches implemented, tested and documented
- Intermediate and final presentations demos
- Technical Report and ideally Scientific paper

Requirements

- Knowledge of
  - Network/software security as well as IT operations and management.
  - (Big) Data science and engineering
- Experiences with data science technologies (e.g., MapReduce, Apache Spark, TensorFlow, etc.)
- Programming Languages: Scala, Python, R

Contacts

- Prof. Dr. Christoph Meinel, Dr. Feng Cheng, Pejman Najafi
  - Email: security-analytics@hpi.de Room: H-1.13 / Tel: +331-5509-519