Chair of Network Architectures and Services
Department of Informatics
Technical University of Munich

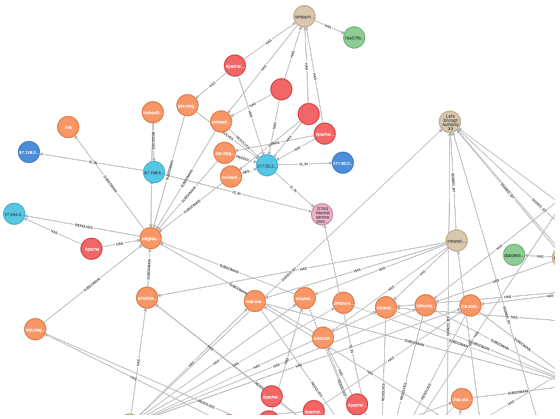
Thesis M.Sc.

IDP, Guided Research

Revealing Organizational Structures in an Internet-wide TLS Graph

Motivation

With Internet-wide TLS scans a huge amount ($\approx 200\text{GB/scan}$) of data can be collected. This data has a lot of potential to be evaluated with graph algorithms. E.g TLS certificates contain a list of domain names, these domain names resolve to IP addresses, we collect certificates from scanned IP addresses and certificates can sign other certificates. This and external data can be used to construct a TLS graph similar to the Webgraph used by the Pagerank algorithm from Google. If a server administrator controls several domains and IP addresses, it is likely we can directly see this in the graph.



However, in the Internet strange things happen and these relations are not as clearly defined as one would hope. For example what happens if an IP address seems to be shared between multiple organisations? We can observe this from cloud providers or CDNs.

To automatically reveal organisational sturctures in the Graph, an algorithm grouping multiple nodes together needs to be carefully designed. If done in a meaningful way, this could reveal valuable information like every domain belonging to Google or clusters of malicious servers.

Your Task

- Extend an Apache Spark parser to transform TLS scans into a graph
- Load the Graph into Neo4J and manually inspect subgraphs from different organisations or companies and collect a set of patterns people use in the Internet
- Design an Apache Spark algorithm to group nodes in the graph, that likely belong together according to a defined criteria and the collected patterns
- Evaluate the algorithm by looking at hand picked examples

Technologies

- Apache Spark with Graphframes
- Python 3, Scala

Contact

Markus Sosnowski sosnowski@net.in.tum.de
Patrick Sattler sattler@net.in.tum.de
Juliane Aulbach aulbach@net.in.tum.de

http://go.tum.de/509123