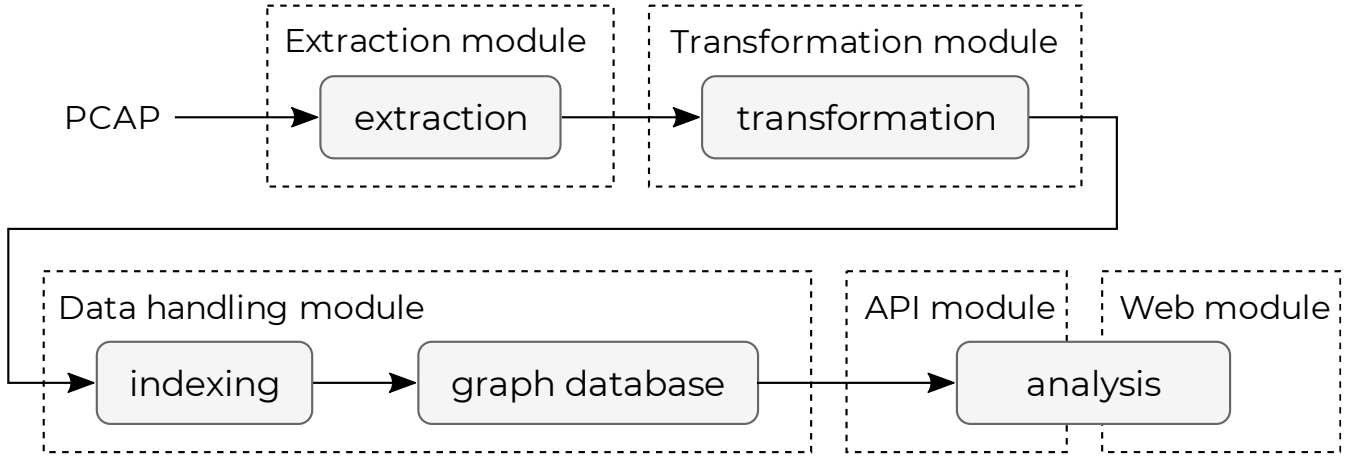
GRANEF: Utilization of a Graph Database for Network Forensics

SECRYPT 2021 – Poster Presentations

Milan Cermak and Denisa Sramkova

Masaryk University, Brno, Czech Republic

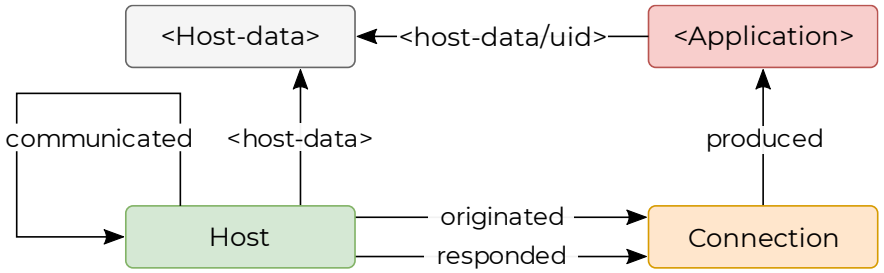
Toolkit Design

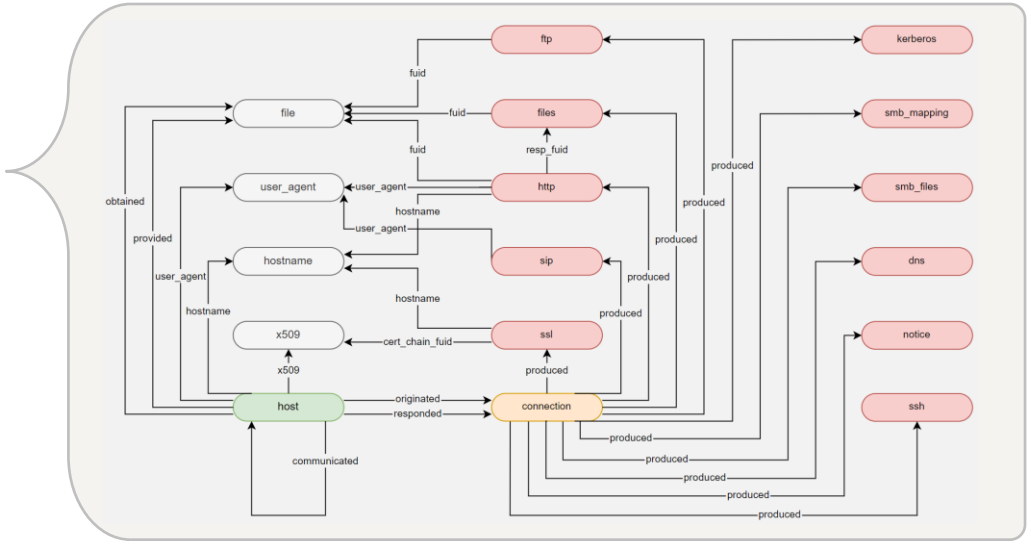


- The toolkit's core consists of a graph database Dgraph (https://dgraph.io/) that stores transformed information from network traffic captures extracted by Zeek (https://zeek.org/) network security monitor.
- Custom Python scripts control all modules to ease toolkit setup and usage.
- Modules are implemented as Docker containers.
- Web interface visualizes data as an interactive relationship diagram.

Database Schema

- The schema follows the format of Zeek logs, preserves their relation, and eases extension.
- All edges are directional but allow reverse processing.
- Host – a device with IP address observed in the network traffic capture.
- Host-data – data related to the host extracted from network traffic (hostname, certificate, …).
- Connection – information about individual network connections (statistics, flags, …).
- Application – application data extracted from the connection (DNS, HTTP, TLS, …).

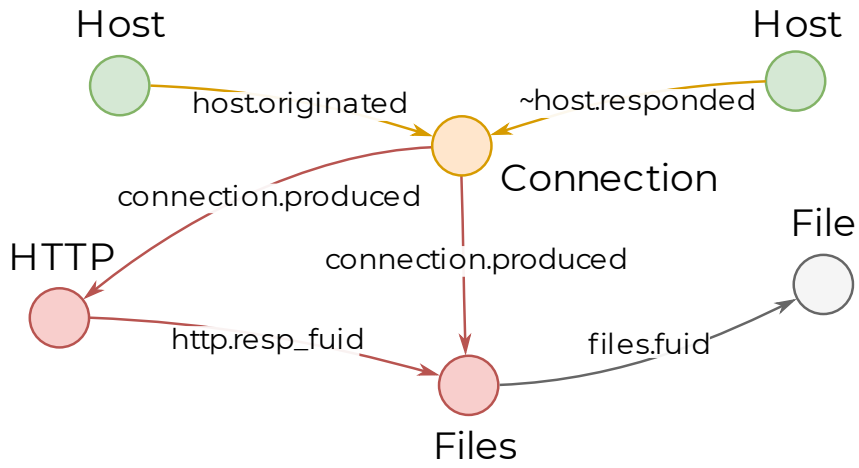




Data Analysis

- Example of a DQL query (Dgraph Query Language) with a selection of TCP connections with a file transfer from a local network:

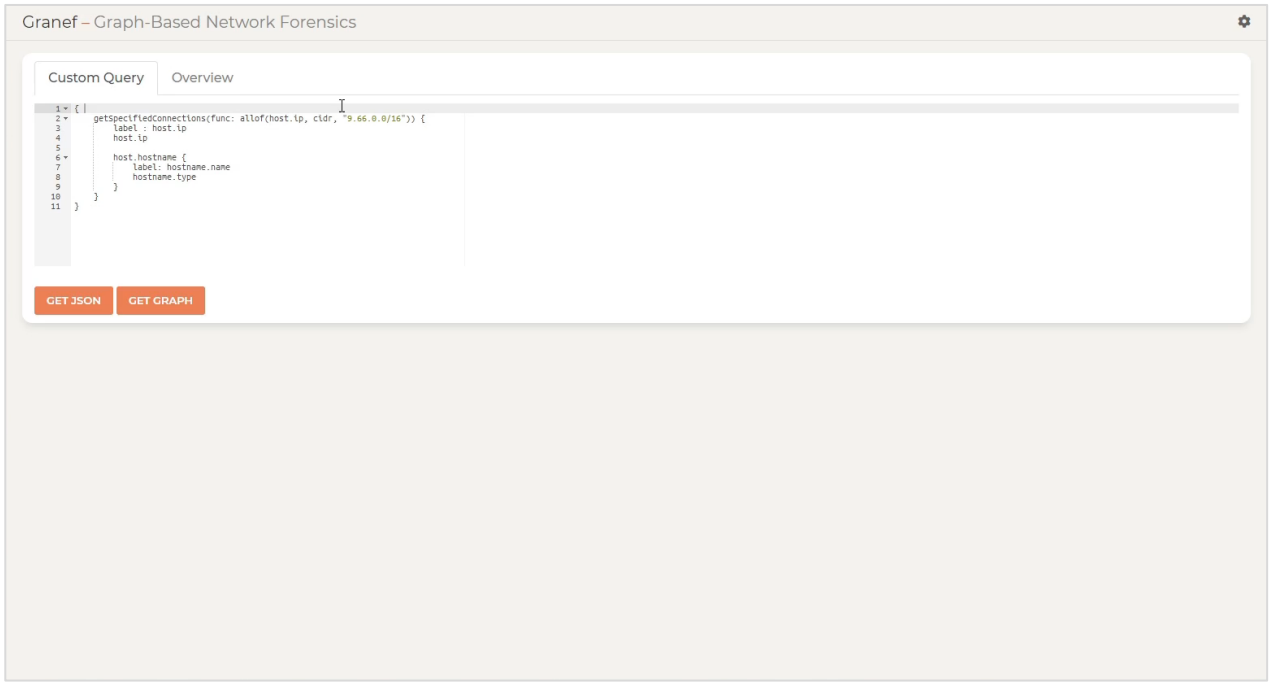
```
{getConn(func: allof(host.ip, cidr, "192.168.0.0/16"))
  name : host.ip
  host.originated @filter(eq(connection.proto, "tcp"))
    expand(Connection)
    connection.produced {
      expand(_all_)
      files.fuid { expand(File) }
    }
    ~host.responded { responded_ip : host.ip }
  }
}}
```

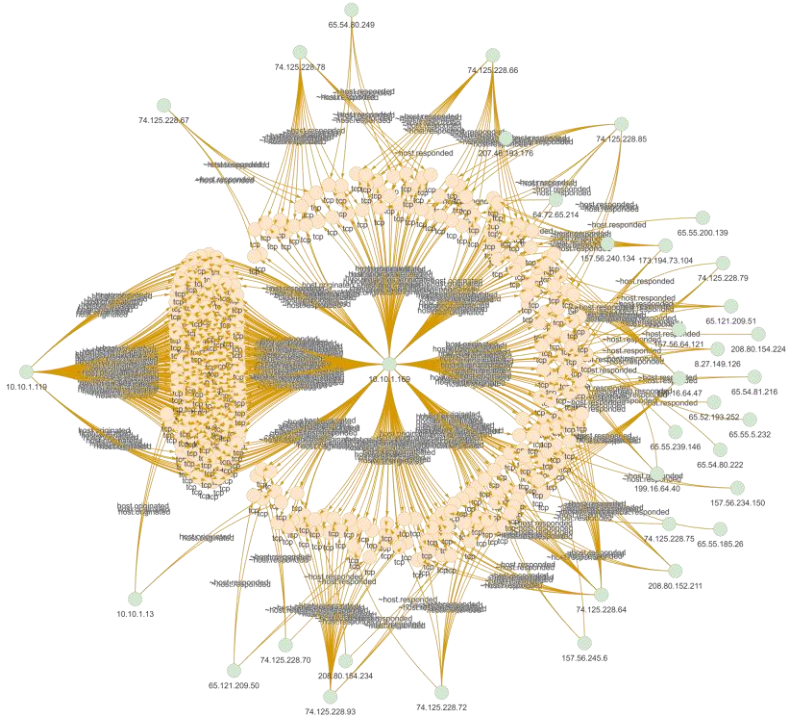


- The toolkit contains an abstract layer API with common analysis functions to ease data investigation (e.g., node neighbors' discovery, data filtering, connections overview).
- Results are provided as JSON or visualized in an interactive relationship diagram.
- Visualization uses a force-directed graph layout and allows nodes aggregation to show large relationship diagrams while preserving a simple overview of the data.

Network Forensics

- The interactive relationship visualization allows the analyst to get details about any selected node, go into the graph's depth, and gain new observations.
- The API includes additional functionality to get initial insights about network connections and perform anomalies detection.





Conclusion

- Graph-based network forensics is a new approach to network traffic data analysis.
- It follows the typical way of human thinking and perception of the characteristics of the surrounding world.
- We introduced the GRANEF toolkit utilizing the Dgraph database that stores transformed Information from network traffic captures extracted by the Zeek network security monitor.
- The toolkit contains an abstract layer API with common analysis functions to ease investigation and support exploratory analysis using an interactive relationship diagram.
- Our approach is not only the new method of network data storage and analysis, but it is also a shift of mindset that allows us to perceive network traffic in a new way.

Thank you for your attention!

If you have any questions, please contact me at cermak@ics.muni.cz

Sharing and Automation for Privacy Preserving Attack Neutralization

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833418.