Karlsruhe Institute of Technology

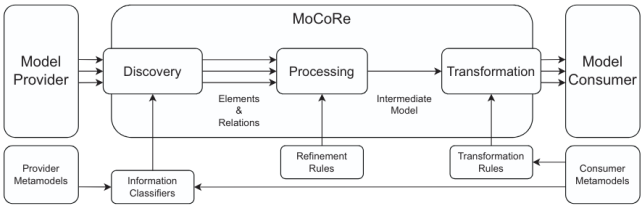
ENERGY LAB 2.0

Institute for Automation and Applied Informatics (IAI)

Bachelor's / Master's Thesis

Audit log data conversion and enhancement for DARPA TC Engagement CDM records

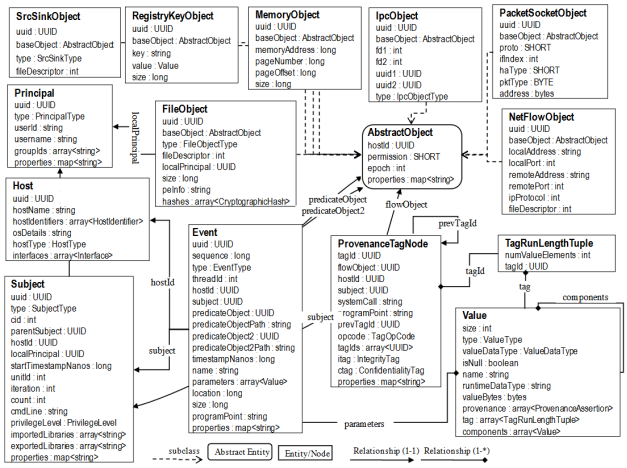
The de-facto standard datasets for evaluating APT attack detection approaches are the DARPA Transparent Computing (TC) Engagement 3 and 5 datasets. Many attack detection approach implementations are evaluated using these datasets and thus feature parsers for this specific data format. To integrate the dataset from our cyber range, we want to convert the raw audit data to the DARPA CDM format.



Gstür et al. *MoCoRe – A Generic Model-Driven Composition and Rule-Based Refinement Framework*, 2024

For this, different methods of data conversion and enrichment shall be implemented and evaluated. DARPA provides an API for their existing Kafka-based code to generate CDM records, which can be used to implement a naive data mapping utility.

Also, MoCoRe by Gstür et al. shall be used for data conversion and its data enrichment capabilities. With MoCoRe, additional data can be extracted and visualisations like network and provenance graphs be created.



Khoury et al. *An Event-based Data Model for Granular Information Flow Tracking*, 2020

Main Tasks:

- Familiarization with the TC Engagement datasets and data format and auditing systems for Linux and Windows
- Evaluating the use and benefit of ML for data extraction, conversion and validation
- Implementation of a data conversion utility
- Implementation of data enrichment and visualisation
- Evaluating and validation of the data

Education, experience and skills:

- Studies in Computer Science
- Python, C++ or Java programming skills
- Basic understanding of machine learning and data types
- Basic knowledge of system calls and network traffic

What we offer:

- Possibility of contributing to scientific publications
- Close supervision
- Thesis in English or German possible
- Opportunity to work with machine learning models and research data
- Student assistant position or familiarisation period possible

Richard Rudolph
Robust Secure Automation (RSA)
richard.rudolph@kit.edu

Moritz Gstür
Secure Energy Systems (SES)
moritz.gstuer@kit.edu

Institute for Automation und Applied Informatics (IAI)
Karlsruhe Institute of Technology, Campus North
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen