Malware Investigation

Law Enforcement Agencies must cope with evidence from multiple different sources, a shortage of skilled cybercrime agents and inefficient data exchange between stakeholders.

Among the digital offences most frequently investigated, there are phishing attacks aimed at spreading financial services-targeted malware. Such malicious software is normally sold to, and used by, multiple thieves, which vastly outnumber the amount of developers.

Arresting malware programmers, albeit more challenging, could have a cascade effect on the Online Banking Malware (OBM) ecosystem. Therefore, novel investigation techniques balancing cybercrime domain knowledge and computer processing power should be researched.

Applying Ontologies

Semantic technologies could enhance relationship discovery and hypothesis testing within the OBM domain.

For instance, supporting analysts in reasoning among a large volume of scattered, supposedly unrelated evidence.

By defining concepts and their relationships, properties and value types, an ontology allows:

- Analysts and computer software to share a common domain understanding, enhancing interaction and information exchange;
- The reuse of domain knowledge: e.g. file hash concepts are also useful for investigating other cybercrimes;
- To easily devise domain assumptions: e.g. “malware family X has at least 3 of their embedded files identical”.

Main Objectives

- Link different criminal organizations
- Identify malware developers
- Facilitate future data integration
- Optimize cyber evidence retrieval

Current and Future Work

A working prototype with a reduced set of the most relevant concepts relating to the malware (contact medium, URL, keyword and hash) is being implemented.

Next, concepts from entities (contact, nickname, user login and messages exchanged) will be incorporated into the implementation.

Future work will:

- Load real data into the knowledge base to validate and improve the inference rules’ efficiency and efficacy;
- Refine the approach to malware string matching by analysing previous research and sandbox platforms;
- Research Natural Language Processing techniques to automate entity extraction and provide content-based categorisation;
- Create an online resource to foster OBM ontology discussion and improvement among Law Enforcement Agencies.

References

- Anti-Phishing Working Group, ‘Phishing activity trends report - 1st quarter 2014’;