Ph.D. in Information Technology:

Carminati, Ferroni, and Polino Final Dissertations

DEIB Seminar Room

February 17th, 2017

2.00 pm

First Ph.D. presentation and discussion:

Michele CARMINATI – XXIX Cycle

“Internet Banking Fraud Analysis and Detection”

Advisor: Prof. **Stefano Zanero**

Abstract:

This thesis work details our research on Internet banking fraud and financial malware analysis. Our work tries to correlate frauds to malware campaigns to provide a sound detection of the two key aspects of modern online frauds: anomalous transactions and malware activity.

Our contributions regarding Internet banking fraud analysis is inspired and rooted around the idea of constructing user’s profiles from historical data to detect suspicious deviations from them. We propose a modular fraud analysis and detection platform, able of big data behavioral analysis and where analyst feedback is put together with state-of-the-art machine learning techniques to build an end-to-end active fraud analysis system. The system has two key features: the first is an unsupervised decision-support and fraud-analysis system capable of automatically ranking transactions based on the risk of being fraudulent by means of a threefold analysis. Besides demonstrating its effectiveness, we evaluate the influence, on the detection quality, of the granularity at which spending habits are modeled and its security against mimicry attacks. The second is a supervised learning module for fraud detection and parameter auto-tuning. The results of the evaluation on a real dataset show that our solution is capable of efficiently defending financial institution against sophisticated attacks. To mitigate the lack of data in this research area, we describe a tool for generating synthetic online banking transactions whose properties are borrowed from the real dataset in our possession, through a statistical modeling of each feature distribution.

Our contributions regarding financial malware analysis focus on studying the client-side behavior of financial Trojans that perform WebInjects. We propose a platform that can analyze advanced banking Trojans and to generate behavioral signatures of the WebInject component thanks to a web-page differential analysis. This research has the goal of characterizing the malicious activity of banking Trojan and then use the acquired knowledge to find a methodology to detect and counter them.

Second Ph.D. presentation and discussion:
Matteo FERRONI – XXIX Cycle
“Enabling power- awareness for multi-tenant systems”
Advisor: Prof. **Marco Domenico Santambrogio**

Abstract:

Power consumption has become a major concern for almost every digital system: from the smallest embedded devices to the biggest data centers, energy and power budgets are always constraining the performance of the system. Moreover, the actual power consumption of these systems is strongly affected by their current "working regime" (e.g., from idle to heavy-load conditions, with all the shades in between), which depends on the guest applications they host, as well as on the external interactions these are subject to. It is then difficult to make accurate predictions on the power consumed by the whole system over time, when it is subject to constantly changing operating conditions: a self-aware and goal-oriented approach to resource allocation may then improve the instantaneous performance of the system, but still the definition of energy saving policies remains not trivial as far as the system is not really able to learn from experience in real world scenarios.

In this context, this thesis proposes a holistic power modeling framework that a wide range of energy and power constrained systems can use to profile their energy and power consumption. Starting from the preliminary experience developed on power consumption models for mobile devices during my M.Sc. thesis, I designed a general methodology that can be tailored on the actual system's features, extracting a specific power model able to describe and predict the future behavior of the observed entity. This methodology is meant to be provided in an "as-a-service" fashion: at first, the target system is instrumented to collect power metrics and workload statistics in its real usage context; then, the collected measurements are sent to a remote server, where data is processed using well known techniques (e.g., Principal Components Analysis, Markov Decision Chains, ARX models, etc.); finally, an accurate power model is built as a function of the metrics monitored on the instrumented system. The generalized approach has been validated in the context of power consumption models for multi-tenant virtualized infrastructures, outperforming results from the state of the art. Finally, the experience developed on power consumption models for server infrastructures led me to the design of a power-aware and QoS-aware orchestrator for multi-tenant systems. On the one hand, I propose a performance-aware power capping orchestrator in a virtualized environment, that aims at maximizing performance under a power cap. On the other hand, I bring the same concepts into a different approach to multi-tenancy, i.e., containerization, thus moving the first steps towards power-awareness for Docker containers orchestration, laying the basis for further research work.

Third Ph.D. presentation and discussion:
Mario POLINO– XXIX Cycle
"Automated Malware Behavioral Analysis"
Advisor: Prof. **Stefano Zanero**

Abstract:

Automated Malware Behavioral Analysis Malicious programs are a constant modern threat to everyone. To be able to defend ourselves from this menace, we need updated tools capable of quickly and efficiently analyze those programs. Our research focused on the development of such tools. First of all, we approached the task of automated behavioral malware analysis. By developing an unsupervised system to identify common behavioral pattern in malware binaries. The behaviors that our system extracts carry also static information in form of control-flow graph based fingerprints. Then, Jackdaw associates semantic information to the behaviors, to create a descriptive summary that helps the analysts, especially the inexperienced ones. All produced information can be easily browsed through a visualization tools that we implemented.

The approach presented in this thesis exploits machine learning techniques to identify similar malware samples and graph mining to extrapolate the aforementioned behaviors. We tested our system on a dataset of 2136 distinct binaries, including both malicious and benign libraries and executables. We compared the behaviors extracted automatically against a ground truth of 44 behaviors created manually by expert analysts founding 77.3% of them. To be able to perform the aforementioned analysis this tool need to unobfuscated access to the malware binary. However, most malware nowadays implement some sort of packing techniques. This led us to develop a generic unpacker that is able to unpack binary for 63% of randomly collected samples. We also explored the possibility to develop a dynamic protection framework that can be used to defend PIN, one of the most used and supported DBI, against anti-instrumentation attacks. Starting from the techniques discovered in literature, we classified them and implemented a set of countermeasure

as generic as possible to defeat them. The framework was tested with three main test cases: eXait, a tool which aims to detect DBI exploiting different techniques, Obsidium, a very complete packer known to employ anti- instrumentation attacks, and PEspin, another packer which employs self- modifying code that could crash the DBI framework. In every case, our tool was able to avoid PIN from being detected, permitting the analysis of the original protected program.

PhD Committee:

Prof. Stefano Zanero, DEIB – Politecnico di Milano
Prof. Fatma Seda Memik, Northwestern University
Prof. Wejing Rao, University of Illinois at Chicago