Masterarbeits

Structure embeddings for OpenSSH heap dump analysis

A report by

Lahoche, Clément Claude Martial

Prüfer

Prof. Dr. Michael Granitzer

Christofer Fellicious

Prof. Dr. Pierre-Edouard Portier

August 21, 2023

Abstract

Acknowledgements

Contents

List of Figures

List of Tables

1 Introduction

Digital forensics is a linchpin in cybersecurity, enabling the extraction of vital evidence from devices like PCs. This evidence is key for detecting malware and tracing intruder activities. Analyzing a device's main memory is a go-to technique in this field. The fusion of machine learning promises to amplify and streamline these analyses.

With the rising need for encrypted communication, Secure Shell (SSH) protocols are now commonplace. However, these security-focused channels can inadvertently shield malicious actions, posing challenges to standard investigative approaches. Cutting-edge research offers solutions. The work in „SmartKex: Machine Learning Assisted SSH Keys Extraction From The Heap Dump“ [0] highlights how machine learning can boost the extraction of session keys from OpenSSH memory images. In a complementary vein, „SSHkex: Leveraging virtual machine introspection for extracting SSH keys and decrypting SSH network traffic“ [0] showcases the power of Virtual Machine Introspection (VMI) for direct SSH key extraction.

Inspired by „SmartKex: Machine Learning Assisted SSH Keys Extraction From The Heap Dump“, this thesis zeroes in on a central challenge: data embedding. While previous studies set the stage for key extraction, the data embedding technique, especially windowing, can be optimized. The design of data embeddings is pivotal for machine learning efficacy, especially in nuanced tasks like memory analysis. This research introduces fresh embedding strategies, aiming to refine extraction and unearth deeper memory snapshot patterns. Merging graph embeddings with advanced machine learning, the goal is to craft a sophisticated toolkit for OpenSSH heap dump studies, bridging digital forensics and machine learning.

2 Research Questions

Write down and explain your research questions (2-5)

3 Structure of the Thesis

Explain the structure of the thesis.

4 Example citation & symbol reference

For symbols look at.

Figure 1: Meaningful caption for this image

First column	Number column
Accuracy	0.532
F1 score	0.87

Table 1: Meaningful caption for this table

5 Example reference

Example reference: Look at chapter **??**, for sections, look at section **??**.

6 Example image

Example figure reference: Look at Figure **??** to see an image. It can be `jpg`, `png`, or best: `pdf` (if vector graphic).

7 Example table

Table **??** shows a simple table[1]

[1] Check `https://en.wikibooks.org/wiki/LaTeX/Tables` on syntax

8 Background

Introduce the related state-of-the-art and background information in order to understand the method developed in the thesis.

9 Methods

Describe the method/software/tool/algorithm you have developed here

10 Results

Describe the experimental setup, the used datasets/parameters and the experimental results achieved

11 Discussion

Discuss the results. What is the outcome of your experimetns?

12 Conclusion

Summarize the thesis and provide a outlook on future work.

A Code

B Math

C Dataset

Acronymes

SSH Secure Shell. 1

VMI Virtual Machine Introspection. 1

References

[0] Christofer Fellicious et al. „SmartKex: Machine Learning Assisted SSH Keys Extraction From The Heap Dump“. In: arXiv:2209.05243 (Sept. 13, 2022). arXiv: 2209.05243[cs]. URL: http://arxiv.org/abs/2209.05243 (visited on 08/17/2023).

[0] Stewart Sentanoe and Hans P. Reiser. „SSHkex: Leveraging virtual machine introspection for extracting SSH keys and decrypting SSH network traffic“. In: *Forensic Science International: Digital Investigation* 40 (Apr. 2022), p. 301337. ISSN: 26662817. DOI: 10.1016/j.fsidi.2022.301337. URL: https://linkinghub.elsevier.com/retrieve/pii/S2666281722000063 (visited on 08/17/2023).

Additional bibliography

[0] Vivek Gite. *How To Reuse SSH Connection To Speed Up Remote Login Process Using Multiplexing*. nixCraft. Aug. 20, 2008. URL: https://www.cyberciti.biz/faq/linux-unix-reuse-openssh-connection/ (visited on 10/21/2022).

[0] Weijie Huang and Jun Wang. „Character-level Convolutional Network for Text Classification Applied to Chinese Corpus“. In: arXiv:1611.04358 (Nov. 15, 2016). arXiv: 1611.04358[cs]. URL: http://arxiv.org/abs/1611.04358 (visited on 08/17/2023).

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich diese Masterarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe und alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, als solche gekennzeichnet sind, sowie, dass ich die Masterarbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt habe.

Passau, August 21, 2023

Lahoche, Clément Claude Martial