Masterarbeit

Predicting SSH keys in Open SSH Memory dumps

A report by

Rascoussier, Florian Guillaume Pierre

Prüfer
Prof. Dr. Michael Granitzer
Christofer Fellicious
Prof. Dr. Pierre-Edouard Portier

August 28, 2023

Abstract

As the digital landscape evolves, cybersecurity has become an indispensable focus of IT systems. Its ever-escalating challenges have amplified the importance of digital forensics, particularly in the analysis of heap dumps from main memory. In this context, the Secure Shell protocol (SSH) designed for encrypted communications, serves as both a safeguard and a potential veil for malicious activities. This research project focuses on predicting SSH keys in OpenSSH memory dumps, aiming to enhance protective measures against illicit access and enable the development of advanced security frameworks or tools like honeypots.

This Masterarbeit is situated within the broader SmartVMI project, a collaborative research initiative with the objective to advance artificial intelligence-based mechanisms for attack detection and digital forensics. Specifically, this work seeks to build upon existing research on key prediction in OpenSSH heap dumps. Utilizing machine learning algorithms, the study aims to refine feature extraction techniques and explore innovative methods for effective key detection. The objective is to accurately predict the presence and location of SSH keys within memory dumps. This work builds upon, and aims to enhance, the foundations laid by the SmartKex paper [0], enriching both the methodology and the results of the original research while exploring the untapped potential of newly proposed approaches.

This report encapsulates the progress of a year-long Master's thesis research project executed between October 2022 and October 2023. Conducted within the framework of the PhDTrack program between the University of Passau and INSA Lyon, the research has been supervised by Christofer Fellicious and Prof. Dr. Michael Granitzer from the University of Passau, as well as Prof. Dr. Pierre-Edouard Portier from INSA Lyon. It offers an in-depth discussion on the current state-of-the-art in key prediction for OpenSSH memory dumps, research questions, experimental setups, program development as well as discussing potential future directions.

Acknowledgements

A special acknowledgment goes to Christofer Fellicious, my engaged supervisor at the University of Passau, for his guidance, support and feedback during the Masterarbeit.

I want to express my sincere gratitude to my colleague and friend, Clément Lahoche, whose human and technical skills have been a great source of inspiration and motivation throughout this project; especially considering that we have been working on closely related subjects. It has been a great pleasure to share our ideas and insights, and to collaborate on the development of several programs necessary for the experimentations.

Another acknowledgments go to my esteemed supervisors Prof. Dr. Granitzer and Prof. Dr. Portier for their support and feedback during the Masterarbeit.

I would also like to express my sincere gratitude to all the persons that have helped me, even punctually, during the Masterarbeit with their valuable help, insights, discussions and contributions as well as all the persons involved in the PhDTrack program that made this Masterarbeit possible, including but not limited to:

- Lionel Brunie, Director of CS Department at INSA Lyon, that makes this PhDTrack program possible from the French side.
- Harald Kosch, Head of the Chair of Distributed Information Systems at the University of Passau, that makes this PhDTrack program possible from the German side.
- Natalia Lucari, PhDTrack coordinator at INSA Lyon, for her support and help during the PhD-Track program.
- Ophelie Coueffe, PhDTrack coordinator at the University of Passau, for her support and help during the PhDTrack program.
- Elöd Egyed-Zsigmond, PhDTrack coordinator at the University of Passau, for the subject selection and administrative support.
- All the other PhDTrack students for the great atmosphere, mutual help and the interesting discussions during almost two years.

Finally, my last acknowledgments go to my family and friends for their support and encouragements.

Contents

1 Introduction

Motivate your research and outline the research gap in this chapter. Why is your thesis relevant and what do you address, what has not been addressed before.

General Requirements to the thesis:

- 60 pages of content in this format. Content does not include table of content, lists, appendices etc.
- Proper scientific referencing
- Introduction and Background should be less than 50% of the thesis
- Images should be readable and in the proper size.

2 Research Questions

Write down and explain your research questions (2-5) The initial objective of this thesis is to answer the following research questions:

- RQ1: What is the state of the art in the field of security key detection in heap dump memory?
- RQ2: What are the challenges in the field of security key detection in heap dump memory?

3 Structure of the Thesis

Explain the structure of the thesis.

4 Example citation & symbol reference

For symbols look at [**latex_symbols_2017**].

5 Example reference

Example reference: Look at chapter 1, for sections, look at section 4.

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

6 Example image

Example figure reference: Look at Figure 1 to see an image. It can be `jpg`, `png`, or best: `pdf` (if vector graphic).

7 Example table

Table 1 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

Acronyms

DEL Directed Edge-labelled Graphs. 3

ER Entity Resolution. 5

GCN Graph Convolutional Networks. 11

GNN Graph Neural Network. 11

KE Knowledge Engineering. 5, 9

KG Knownledge Graph. i, 1

ML Machine Learning. 7

NLP Natural Language Processing. i

QA Quality Assurance. 5

RDF Resource Description Framework. 3, 5, 6

SPARQL SPARQL Protocol and RDF Query Language. 5

SSH Secure Shell Protocol. i

References

[0] Christofer Fellicious et al. „SmartKex: Machine Learning Assisted SSH Keys Extraction From The Heap Dump“. In: arXiv:2209.05243 (Sept. 2022). arXiv:2209.05243 [cs]. DOI: 10.48550/arXiv.2209.05243. URL: http://arxiv.org/abs/2209.05243.

Additional bibliography

[0] Lisa Ehrlinger and Wolfram Wöß. „Towards a Definition of Knowledge Graphs“. In: (2016), pp. 1–4.

[0] Gianluca Fiorelli. *Best of 2013: No 13 – Search in the Knowledge Graph era*. Accessed: 2023-06-12. 2013. URL: https://www.stateofdigital.com/search-in-the-knowledge-graph-era/.

[0] Jackson Gilkey. *Graph Theory and Data Science*. Accessed: 2023-05-25. 2019. URL: https://towardsdatascience.com/graph-theory-and-data-science-ec95fe2f31d8.

[0] Google. „Introducing the Knowledge Graph: Things, not strings“. In: *Google Blog* (May 2012). Accessed: 2023-06-16. URL: https://blog.google/products/search/introducing-knowledge-graph-things-not/.

[0] Paul Groth et al. „Knowledge Graphs and their Role in the Knowledge Engineering of the 21st Century“. In: *Dagstuhl Reports* 12.9 (2022). Report from Dagstuhl Seminar 22372. Specific usage: pp. 60-72, Subsection "3.2 A Brief History of Knowledge Engineering: A Practitioner's Perspective", pp. 60–120. DOI: 10.4230/DagRep.12.9.60.

[0] Marvin Hofer et al. „Construction of Knowledge Graphs: State and Challenges“. In: *arXiv preprint arXiv:2302.11509* (2023). URL: https://doi.org/10.48550/arXiv.2302.11509.

[0] Aidan Hogan et al. „Knowledge Graphs“. In: *ACM Comput. Surv.* 54.4 (July 2021). ISSN: 0360-0300. DOI: 10.1145/3447772. URL: https://doi.org/10.1145/3447772.

[0] Frederick Edward Hulme. *Proverb Lore: Many Sayings, Wise Or Otherwise, on Many Subjects, Gleaned from Many Sources*. E. Stock, 1902, p. 188.

[0] M.S. Jawad et al. „Adoption of knowledge-graph best development practices for scalable and optimized manufacturing processes“. In: *MethodsX* 10 (2023), p. 102124. ISSN: 2215-0161. DOI: https://doi.org/10.1016/j.mex.2023.102124. URL: https://www.sciencedirect.com/science/article/pii/S2215016123001255.

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

[0] Michelle Venables. *An Introduction to Graph Theory*. Accessed: 2023-06-12. 2019. URL: https://towardsdatascience.com/an-introduction-to-graph-theory-24b41746fabe.

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 28, 2023

Rascoussier, Florian Guillaume Pierre