doi:10.20396/revpibic.v0i0.id

XXVI Congresso de Iniciação Científica Unicamp
17 a 19 de outubro Campinas | Brasil

Análise de *malware* em larga escala – Large Scale Malware Analysis

Giovanni Bertão*, Prof. Dr. Paulo Lício de Geus.

Abstract

Malware overview reports are valuable information to understand threats behavior and develop proper countermeasures. Currently, most of these studies are focused on either fine-grained, individual sample analysis or coarse-grained landscapes. On the one hand, only the first allows professionals to handle specific security breaches. On the other hand, only the second allows understanding threat scenario as a whole. We claim that a complete security treatment is only possible when combining both approaches. Therefore, this work presents an analysis of a large malware dataset, showing the distinctions between coarse-grained and fine-grained analysis results. It presents both a general threat scenario based on coarse-grained results as well as it details fine-grained results to identify particular malicious constructions to anticipate incident response of future threats.

Key words:
Malware Analysis, Program Tracing, Computer Security.

Introduction

Malware is a constant threat to modern computer systems. To counter such kind of threat, analysis procedures are employed, thus allowing vaccine development, remediation and enabling forensic procedures. Most of current malware analysis research is presented in two forms: i) a coarse-grained overview, highlighting only major aspects, discarding samples details; ii) a fine-grained, specific view, focusing on implementation details, but not stating the risk of such sample in the overall scenario. Such approaches are complementary and security analysis must consider both to provide a complete threat understanding. To support this claim, this work presents a comparison of both approaches to highlight their differences and how they can be integrated. For our evaluation, we considered a dataset of 135 thousand unique, real binary samples. They were all submitted to static and dynamic analysis[1] procedures and their results were analyzed in both coarse and fine-grained ways.

Results and Discussion

The coarse-grained analysis allowed us to draw a panorama of the entire dataset and to compare it with other scenarios, such as the Brazilian one[2]. In such, samples were presented as a mix of binaries and DLLs, relying on system native functions, with background activity and few system interaction. In comparison, our dataset was mainly composed by executables—with fell libraries—, also relying in system native libraries—with few external ones—, but using graphic user interfaces and presenting many system interactions.

Despite allowing drawing panoramas, the coarse-grained analysis is not able to explain samples' project decisions. For instance, its results identified each sample contacts, on average, 2 distinct IP addresses. Fine-grained analysis, in turn, is able to identify that a given sample contacted 16386 IPs whereas other ones accessed only a single one. It also explains that such distinction is due to samples distinct goals: The first sample is a ransomware which propagates by scanning the network, thus contacting many IPs. The second, in turn, despite also being a ransomware, presents a distinct spreading mechanism, thus contacting only its own Command and Control.

These results demonstrate that coarse-grained analysis allows us to draw a threat landscape but only a fine-grained analysis allows us to identify project decisions and rare constructions.

Conclusions

In this work, we presented analysis results of a large scale malware dataset. We compared the results on two approaches—coarse and fine grained—, highlighting their differences. The results showed that whereas coarse-grained analysis allowed drawing threat panoramas, only fine-grained analysis enabled understanding samples' internals. Therefore, only a combined approach allows complete security analysis treatment.

Acknowledgements

This work is supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Process 138239/2017-7.

[1] Botacin, M.;Afonso, V.;Grégio, A. e de Geus, P. Monitoração de comportamento de malware em sistemas operacionais windows nt 6.x de 64 bits, in SBSeg2014 - Artigos Completos, Belo Horizonte (MG), **2014**, http://www.lbd.dcc.ufmg.br/colecoes/sbseg/2014/0015.pdf.

[2] Botacin, M.;Grégio, A. e de Geus, P. Uma visão geral do malware ativo no espaço nacional da internet entre 2012 e 2015. **2015,**

https://siaiap34.univali.br/sbseg2015/anais/WFC/artigoWFC02.pdf.