By continuing to browse this site you agree to us using cookies as described in About Cookies

Wiley Online Library

Log in / Register

Go to old article view

Get access

Statistical Analysis and Data Mining: The ASA Data Science Journal Explore this journal >

Volume 10, Issue 3
June 2017
Pages 182–193

View issue TOC
Special Issue:
CoDA 2016 Special Issue: Selected Papers from the Conference on Data Analysis 2016 – Part I
Original Article

APT malware static trace analysis through bigrams and graph edit distance

Alexander D. Bolton, Christine M. Anderson-Cook

First published:
17 May 2017 Full publication history

DOI:
10.1002/sam.11346 View/save citation

Cited by (CrossRef):
0 articles Check for updates Citation tools

Am score 3

Abstract

Research and business organizations are vulnerable to attack by malware, particularly advanced persistent threat malware tailored for a specific target. Malware identification is made more difficult because samples can be subtly altered to avoid detection by methods that check for an identical match to known code. Different versions of an original piece of malware form a malware family. When new malicious software is identified, reverse engineers seek to identify its origin and purpose. Knowing whether new malware is from a known family or a previously unobserved family aids the efficiency of reverse engineers. This article presents a three-stage method to classify new malware into a family by comparing its similarity to existing static traces, and assigning it to the most similar family. First, a fast filtering method creates a shortlist of samples with some similarity to the new malware, using a simple bigram comparison of the instructions. The second stage takes the call graph view of the shortlisted static traces and uses simulated annealing to estimate the graph edit distance, a measure of dissimilarity between graphs. Finally, a random forest classifier combines the previous two results to predict the family to which a new sample belongs. The paper also considers how to detect when malware is from a new family.

Get access to the full text of this article

» Article Information

Related content

Articles related to the one you are viewing

The articles below have been selected for you based on the article you are currently viewing.

Improving malware detection using multi-view ensemble learning

Jinrong Bai, Junfeng Wang

24 August 2016

Bypassing system calls–based intrusion detection systems

Ishai Rosenberg, Ehud Gudes

16 November 2016

A static Android malicious code detection method based on multi-source fusion

Yao Du, Xiaoqing Wang, Junfeng Wang

30 March 2015

Classifying malwares for identification of author groups

Jiwon Hong, Sanghyun Park, Sang-Wook Kim, Dongphil Kim, Wonho Kim

31 July 2017

Structural classification and similarity measurement of malware

Hongbo Shi, Tomoki Hamagami, Katsunari Yoshioka, Haoyuan Xu, Kazuhiro Tobe, Shigeki Goto

27 September 2014

WILEY

Browse Publications
Browse by Subject
Resources
Help & Support
Cookies & Privacy
Terms of Service
About Us
Wiley Job Network
Advertisers & Agents

Powered by Wiley Online Library Copyright © 1999 - 2018 John Wiley & Sons, Inc. All Rights Reserved