Accepted Manuscript

DRTHIS: Deep ransomware threat hunting and intelligence system at the fog layer

Sajad Homayoun, Ali Dehghantanha, Marzieh Ahmadzadeh, Sattar Hashemi, Raouf Khayami, Kim-Kwang Raymond Choo, David Ellis Newton

PII: S0167-739X(17)32846-7
DOI: https://doi.org/10.1016/j.future.2018.07.045
Reference: FUTURE 4363

To appear in: *Future Generation Computer Systems*

Received date : 12 December 2017
Revised date : 9 May 2018
Accepted date : 22 July 2018

Please cite this article as: S. Homayoun, A. Dehghantanha, M. Ahmadzadeh, S. Hashemi, R. Khayami, K.R. Choo, D.E. Newton, DRTHIS: Deep ransomware threat hunting and intelligence system at the fog layer, *Future Generation Computer Systems* (2018), https://doi.org/10.1016/j.future.2018.07.045

DRTHIS: Deep Ransomware Threat Hunting and Intelligence System at the Fog Layer

Sajad Homayoun[1], Ali Dehghantanha[2], Marzieh Ahmadzadeh[1], Sattar Hashemi[3], Raouf Khayami[1], Kim-Kwang Raymond Choo[4], David Ellis Newton[5]

Abstract

Ransomware, a malware designed to encrypt data for ransom payments, is a potential threat to fog layer nodes as such nodes typically contain considerably amount of sensitive data. The capability to efficiently hunt abnormalities relating to ransomware activities is crucial in the timely detection of ransomware. In this paper, we present our *Deep Ransomware Threat Hunting and Intelligence System (DRTHIS)* to distinguish ransomware from goodware and identify their families. Specifically, *DRTHIS* utilizes *Long Short-Term Memory (LSTM)* and *Convolutional Neural Network (CNN)*, two deep learning techniques, for classification using the softmax algorithm. We then use 220 *Locky*, 220 *Cerber* and 220 *TeslaCrypt* ransomware samples, and 219 goodware samples, to train *DRTHIS*. In our evaluations, DRTHIS achieves an F-measure of 99.6% with a true positive rate of 97.2% in the classification of ransomware instances. Additionally, we demonstrate that *DRTHIS* is capable of detecting previously unseen ransomware samples from new ransomware families in a timely and accurate manner using ransomware from the *CryptoWall*, *TorrentLocker* and *Sage* families. The findings show that 99% of *CryptoWall* samples, 75% of *TorrentLocker* samples and 92% of *Sage* samples are correctly classified.

Keywords: Crypto-ransomware, ransomware detection, ransomware family detection, deep learning, Long Short-Term Memory, Convolutional Neural Network.

1. Introduction

Ransomware is a recent threat that has affected a number of industries and countries [1], and is reportedly the fastest growing malware type [2, 3]. Today's ransomware is a sophisticated threat affecting users all around the world. The first wave of 'misleading' applications appears in 2005. Specifically, performance enhancement tools or fake spyware removal tools (e.g. RegistryCare, PerformanceOptimizer and SpySherriff) designed to mainly target Windows computers and their users, claimed that there is a critical performance/security issue in the victim's computer and recommended the user to buy an additional program to eliminate the problem. Since then, a more disruptive form of extortion emerged which disables access and control of the computer by locking up the computer from being use. There has been a recent shift to the use of ransomware, where data in the infected computers are being encrypted for ransom.

In the literature, there are two main types of ransomware, namely: *Locker* and *Crypto* ransomware. Lockers deny users' access without generally making any changes to the data stored on the system, while crypto-ransomware encrypts all or selected data based on predefined file formats (e.g. *.pdf and *.doc) using a (strong) cryptography algorithm such as AES or RSA [4]. After the victim's data have been encrypted, the victim is presented with the ransom payment instructions in order to obtain a decryption key and recover their data.

Unsurprisingly, ransomware has attracted the attention of security researchers and practitioners. For example, *ransomwaretracker.abuse.ch*[6] tracks major ransomware families, such as *Locky, Cerber, TeslaCrypt, CryptoWall, TorrentLocker* and *Sage*. *Locky* ransomware is usually distributed via phishing e-mails that contain Microsoft Word Office documents with embedded malicious macros, which will subsequently result in the download of the ransomware [5]. *Cerber* ransomware is often distributed via exploit kits [6], and has the capability to en-

Email addresses: S.Homayoun@sutech.ac.ir (Sajad Homayoun), A.Dehghantanha@sheffield.ac.uk (Ali Dehghantanha), Ahmadzadeh@sutech.ac.ir (Marzieh Ahmadzadeh), S_Hashemi@shirazu.ac.ir (Sattar Hashemi), Khayami@sutech.ac.ir (Raouf Khayami), Raymond.Choo@fulbrightmail.org (Kim-Kwang Raymond Choo), D.E.Newton@salford.ac.uk (David Ellis Newton)

[1]Department of Computer Engineering and Information Technology, Shiraz University of Technology, Shiraz, Iran.

[2]Department of Computer Science, University of Sheffield, Sheffield, U.K.

[3]Department of Computer Engineering, Shiraz University, Shiraz, Iran.

[4]Department of Information Systems and Cyber Security and Department of Electrical and Computer Engineering, The University of Texas at San Antonio, San Antonio, TX 78249, USA.

[5]Department of Computer Science, School of Computing, Science and Engineering, University of Salford, Salford, U.K.

[6]https://ransomwaretracker.abuse.ch/tracker/