Ain Shams University
Faculty of Engineering
Department of Computer and Systems Engineering

Defending Attacks on Cloud Computing

By
Eman Ahmed Abd El-Azim
Bachelor of Science Degree majoring in Computer Engineering – College of Engineering – Qatar University, 2008

A THESIS
SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

DEPARTMENT OF COMPUTER AND SYSTEMS ENGINEERING

Supervised By
Prof. Dr./ Ayman Mohammad Bahaa-Eldin
Dr./ Mohammed Ali Sobh
Dr./ Amin Abdul Wahab Sorrour

Cairo, Egypt
2017

Faculty of Engineering
Department of Computer and Systems Engineering

Name : Eman Ahmed Abd El-Azim
Thesis : Defending Attacks on Cloud Computing
Degree : Master of Science in Electrical Engineering – Computer and Systems Engineering

Examiners Committee

Name, Title, and Affiliate	**Signature**
Prof. Dr./ Samy Sayed Abdo Ghoniemy Computer Systems British University, Cairo, Egypt (Examiner)	……
Prof. Dr./ Hoda Korashy Mohamed Ismail Computer and Systems Engineering Ain Shams University, Cairo, Egypt (Examiner)	……
Prof. Dr./ Ayman Mohammad Bahaa-Eldin Computer and Systems Engineering Ain Shams University, Cairo, Egypt (Supervisor)	……
Dr./ Mohammed Ali Sobh Computer and Systems Engineering Ain Shams University, Cairo, Egypt (Supervisor)	……

Date: / /

Abstract

Eman Ahmed Abd El-Azim

Defending Attacks on Cloud Computing

Master of Science in Electric Engineering – Computer and Systems Engineering Dissertation

Ain Shams University, 2017

Malwares are increasing rapidly. The nature of distribution and effects of malwares attacking several applications requires a real-time response. Therefore, a high performance detection platform is required. In this thesis, Hadoop is utilized to perform static binary search and detection for malwares and viruses in portable executable files deployed mainly on the cloud. Hadoop was chosen as it is a software platform that allows designing applications capable of handling huge data amounts in a parallel manner in large clusters. The thesis presents an approach used to map the portable executable files to Hadoop compatible files. The Boyer–Moore-Horspool Search algorithm is modified to benefit from the distribution of Hadoop. The performance of the proposed model is evaluated using a standard virus database and the system is found to outperform similar platforms.

Keywords:
Cloud computing, Security issues, Malware, Static Binary Search, BMH, Hadoop.

Acknowledgements

"In the name of Allah, the Most Gracious and the Most Merciful."

Foremost, Alhamdulillah, all praises to Allah, the Almighty for the strengths and the blessings for the completion of the master's thesis. This thesis is accomplished in its current form due to the assistance and guidance of my supervisors: Dr. Ayman Mohammed Bahaa, Dr. Mohammed Ali Sobh and Dr. Amin Abdul Wahab. I would like to express my sincere gratitude to their efforts, patience, advices and continuous support which helped me a lot throughout this research. Sincere thanks to all my family, it would not have been done without your support, encouragement and prayers. Eventually, special thanks to my father, I know your spirit is always with us.

Statement

This dissertation is submitted to Ain Shams University for the degree of Master of Science in Electrical Engineering in Computer and Systems Engineering

Department of Computer and Systems Engineering

The work included in this thesis was out by the author at Computer and Systems Engineering, Ain Shams University.

No part of this thesis has been submitted for a degree or qualification at other university or institution.

Date : / / 2017

Signature :

Name : Eman Ahmed Abd El-Azim

Table of Contents

List of Publications

1. Cloud Computing and Malware Issues: A survey paper
2. A Cloud-based Malware Detection Framework: A paper

List of Tables

List of Figures and Illustrations

Chapter 1: Introduction

1.1 Overview and Thesis Main Concern

Malware stands for "malicious software." It is any software program that is created to perform harmful actions or unwanted by a computer's legitimate user. It has a variety of forms as: computer viruses, which are the most familiar type, worms, spyware, adware, trojans, ransomware and Botnets.

Having Malwares spreading widely on the internet, every host faces the risks of malware attacks. Since cloud environment inherited internet properties, cloud environment is vulnerable to Malwares. Most users transfer or share a vast amount of small files across the cloud as PE files, images. Many of these files could be infected by different Malware types. Hence, a demand to make researchers study and improve different techniques for scanning of files across clouds in a fast manner. Malware-detection techniques can be commonly classified into three categories: signature-based-detection, anomaly-based-detection and code-emulation [1] [2] [3]. A further explanation is presented in the next chapter for the cloud environment, difference between Malware types, detection techniques and some efforts carried using these techniques by other researchers.

The problem of detecting Malware in cloud environment is a performance and detection accuracy problem. This thesis handled this issue by utilizing Hadoop framework to speed up the detection rate. It presents the problems faced that affect detection performance and accuracy and how it was handled. In the next section, a brief description for the detection technique chosen, the type of files used in testing, the framework used and previous efforts done by other researchers using same framework.

1.2 Methodology

The detection technique handled in this thesis is signature-based detection for binary-executable files (e.g. in MSDOS: the .exe files and in Windows: the PE-files, …etc.). Considering the assumption that there are many of them and there is a need to process them in fast and parallel manner, hadoop platform is utilized to perform this detection. Apache Hadoop is "a free software framework for having a distributed storage and processing for huge datasets of computer-clusters" [4]. A further explanation for Hadoop architecture and versions is represented in the third chapter and its practical tests.

Researchers as [5] handled this problem using hadoop environment. But they proposed architecture of how hadoop framework uses its daemons to cooperate in scanning without further details. Other Researchers [6] emphasized how malwares have spread widely in the past few years. Where Endgame received 20M samples of malwares, McAfee receives 10M malware samples in 2012 and VirusTotal receives almost 600k unique files per day. These numbers proves how malware samples increased rapidly in the internet which needs making more studies on