

**MALWARE THREAT SCORE FRAMEWORK USING
STATIC AND DYNAMIC FEATURES**

By
Ali Suwanda
21851013

MASTER'S DEGREE
in
MASTER OF INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY



SWISS GERMAN UNIVERSITY
The Prominence Tower
Jalan Jalur Sutera Barat No. 15, Alam Sutera
Tangerang, Banten 15143 - Indonesia

June 2020

**MALWARE THREAT SCORE FRAMEWORK USING
STATIC AND DYNAMIC FEATURES**

By

Ali Suwanda

21851013

MASTER'S DEGREE

in

MASTER OF INFORMATION TECHNOLOGY

FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY



SWISS GERMAN UNIVERSITY
The Prominence Tower
Jalan Jalur Sutera Barat No. 15, Alam Sutera
Tangerang, Banten 15143 - Indonesia

June 2020

Revision after the Thesis Defense on July 17th 2020

STATEMENT BY THE AUTHOR

I hereby declare that this submission is my own work and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgement is made in the thesis.

Ali Suwanda

Student

Date

Approved by:

Dr. Charles Lim, BSc., MSc.

Thesis Advisor

Date

SWISS GERMAN UNIVERSITY

Dr. Ir. Lukas, MAI.

Thesis Co-Advisor

Date

Dr. Maulahikmah Galinium S.Kom., M.Sc.

Dean of Faculty of Engineering and
Information Technology

Date

ABSTRACT

MALWARE THREAT SCORE FRAMEWORK USING STATIC AND DYNAMIC FEATURES

By

Ali Suwanda

Dr. Charles Lim, BSc., MSc. , Advisor

Dr. Ir. Lukas, MAI. , Co-Advisor

SWISS GERMAN UNIVERSITY

In recent study malware threat becomes more targeted over three years. With those indications, malware threat became a massive risk to the organization therefore, it is important for the organization to improve their situational awareness in cyber security context. It is important for the management to take a decision and create awareness to the organizations hence, threat score can provide high-level information of malware threat that can support strategic and decision making process in the organization. This study focus on to analyze data to create a threat score framework and measure their threat using static and dynamic analysis.

Keywords: Static Analysis, Dynamic Analysis, Threat Score.



SWISS GERMAN UNIVERSITY

DEDICATION

I would like to dedicate this research project to my family and my beloved country
Indonesia.



ACKNOWLEDGEMENT

I would like to express my deepest gratitude to Dr. Charles Lim, BSc., MSc. and Dr. Ir. Lukas, MAI. as my thesis advisor and co-advisor for their relentless support and guidance during my research project. I wish to thank for Mr. Reischaga Mohajit as my discussion partner and support throughout this research. It is because of their contributions this thesis report and the whole research project can arrive at this point.

I would like to thank my family, especially to my brother Mr. Yohanes Christiano for their moral supports throughout my pursue of master degree in SGU. Finally, I would like to thank all of my direct supervisor, my team at work and my friends especially Batch 23 MIT SGU for their companionship, and to the countless number of people who have helped me throughout this research project, either directly or indirectly.

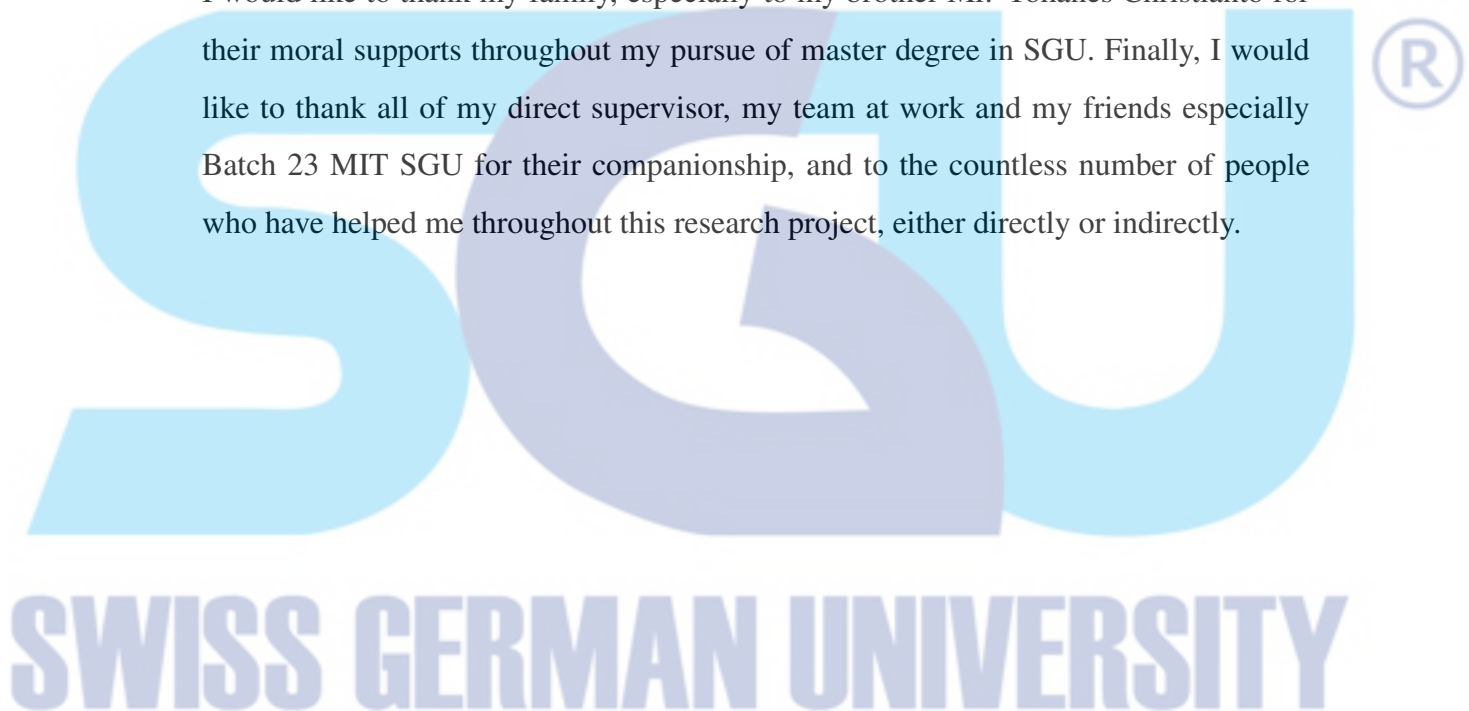


TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR	1
ABSTRACT	2
DEDICATION	4
ACKNOWLEDGEMENT	5
TABLE OF CONTENTS	9
LIST OF FIGURES	10
LIST OF TABLES	12
1. INTRODUCTION	13
1.1 Background	13
1.2 Problem Statement	14
1.3 Research Objective	14
1.4 Research Question	15
1.5 Scope and Limitation	15
1.6 Significance of Study	15
1.7 Hypothesis	15
1.8 Thesis Structure	15
2. LITERATURE REVIEW	17
2.1 Malware Definition and Brief History	17
2.2 Malware Analysis	19
2.2.1 Static Analysis	19

2.2.1.1	Determining Type of File	19
2.2.1.2	Malware Fingerprint	20
2.2.1.3	String Extraction	20
2.2.1.4	PE Header Inspection	20
2.2.1.5	File Obfuscation Determination	22
2.2.1.6	Validation Using Anti-Virus Scanning	23
2.2.2	Dynamic Analysis	23
2.2.2.1	Manual and Automated Dynamic Analysis	24
2.2.2.2	Debugging	24
2.2.2.3	Virtualized-Based Sandboxes	25
2.2.2.4	Emulation-Based Sandboxes	26
2.2.2.5	Bare-Metal Sandboxes	26
2.3	Static and Dynamic Analysis Limitation	27
2.3.1	Malware Obfuscation	27
2.3.1.1	Entry-Point Obscuring	27
2.3.1.2	Encryption	29
2.3.1.3	Polymorphism	30
2.3.1.4	Metamorphism	31
2.3.1.5	Packers	31
2.3.2	Malware Evasion Techniques	32
2.3.2.1	Anti-Sandboxing	32
2.3.2.2	Anti-Debugging	32
2.3.2.3	Evasion Techniques Classification	33
2.3.2.4	Anti-Sandboxing Detection Dependent	34
2.3.2.5	Anti-Sandboxing Detection Independent	34
2.3.2.6	Anti-Debugging Detection Dependent	35
2.3.2.7	Anti-Debugging Detection Independent	36
2.4	Threat Score Framework	37
2.5	Graph Theory	39
2.6	K-Means Theory	39
2.7	Related Works	40

3. RESEARCH METHODS 43

3.1	Research Framework	43
3.1.1	Static Analysis	44
3.1.2	Dynamic Analysis	47
3.1.3	Final Threat Score	49
4.	RESULTS AND DISCUSSIONS	50
4.1	Experiment Setup	50
4.1.1	Hardware Setup	50
4.1.2	Software Setup	51
4.2	Dataset	52
4.3	Preliminary experiment	52
4.3.1	Graph Analysis	53
4.3.2	Weighted Score for Static and Dynamic Analysis Score Experiment	54
4.4	Results and Analysis	55
4.4.1	Static Threat Scoring Result	57
4.4.2	Dynamic Threat Scoring Result and Analysis	58
4.4.2.1	API Call Extraction	58
4.4.2.2	Threat Score Calculation	58
4.4.3	Final Threat Scoring Result	59
4.4.4	Observation	59
4.5	Summary	60
5.	CONCLUSIONS AND RECOMMENDATIONS	61
5.1	Conclusions	61
5.2	Recommendations	61
5.3	Future Works	61
	GLOSSARY	62
	REFERENCES	67
	Appendix A. Malware Families	68
	Appendix B. Benign Dataset	69

Appendix C. Malware Samples 70

**Appendix D. Snippet of Extracted JSON File From Cuckoo - WannaCrypt
Sample 86**

Appendix E. Source Code 99

Appendix F. Submitted Paper Based On This Thesis 118



LIST OF FIGURES

1.1	Total Malwarebytes detection consumer and business in 2018	13
1.2	Total 10 countries with their biggest threats	14
2.1	Typical infected host execution flow (Elisan, 2015)	28
2.2	Typical infected host execution flow (Elisan, 2015)	29
2.3	Malware encryption process during infection (Elisan, 2015)	29
2.4	Malware encryption process during execution (Elisan, 2015)	30
2.5	Encrypted malware will be same in memory although it looks different due to polymorphism (Elisan, 2015)	31
2.6	Metamorphic malware infections differ in memory and on disk (Elisan, 2015)	31
2.7	From left to right: A simple graph, a multi graph, and a directed graph .	39
3.1	Research framework	43
3.2	Static analysis process and threat score calculation	44
3.3	Dynamic analysis process and threat score calculation	47
4.1	A graph of WannaCrypt sample based on API category	54
4.2	Static analysis score evaluation result	55
4.3	Dynamic analysis score evaluation result	56
4.4	Final threat score evaluation result	56

LIST OF TABLES

2.1	PE Sections	21
2.2	Different debugging approaches based on the needed functionality	25
2.3	Anti-Debugging Techniques	33
2.4	Thesis framework and approaches comparison	42
4.1	Hardware that are used in the research	50
4.2	Software that are used in the research	51
4.3	Malware Dataset 2018, 2019, 2020 based on Malware Trend Reports	52
4.4	Weighted Score Experiment Result	55
4.5	Number of Clustered Malware/Benignware Samples in Percentage	56
4.6	Example of Static Analysis Score Calculation	57
A.1	Malware Families	68
B.1	Benign Dataset	69
C.1	Malware Dataset	70
C.2	Malware Dataset - Continued	71
C.3	Malware Dataset - Continued	72
C.4	Malware Dataset - Continued	73
C.5	Malware Dataset - Continued	74
C.6	Malware Dataset - Continued	75
C.7	Malware Dataset - Continued	76
C.8	Malware Dataset - Continued	77
C.9	Malware Dataset - Continued	78
C.10	Malware Dataset - Continued	79
C.11	Malware Dataset - Continued	80
C.12	Malware Dataset - Continued	81

C.13 Malware Dataset - Continued 82
C.14 Malware Dataset - Continued 83
C.15 Malware Dataset - Continued 84
C.16 Malware Dataset - Continued 85

