ACTIONABLE THREAT INTELLIGENCE GENERATION BASED ON DARKNET TRAFFIC ANALYSIS

By

Ryandy

11602002

BACHELOR'S DEGREE

in

INFORMATION TECHNOLOGY

ENGINEERING AND INFORMATION TECHNOLOGY

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

July 2020

Revision after the Thesis Defense on 14th July 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 this thesis.

SGU ® SWISS GERMAN UNIVERSITY

Ryandy

_______________________________ _______________

Student Date

Approved by:

Dr. Charles Lim, B.Sc, M.Sc.

_______________________________ _______________

Thesis Advisor Date

Kalpin Erlangga Silaen, M.Kom, CISSP, CEH, ECIH, ECSA, CHFI, ISO 27001 LI.

_______________________________ _______________

Thesis Co-Advisor Date

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

_______________________________ _______________

Dean Date

ABSTRACT

ACTIONABLE THREAT INTELLIGENCE GENERATION BASED ON DARKNET TRAFFIC ANALYSIS

By
Ryandy
Dr. Charles Lim, B.Sc, M.Sc., Advisor
Kalpin Erlangga Silaen, M.Kom, CISSP, CEH, ECIH, ECSA, CHFI, ISO 27001 LI.,
Co-Advisor

SWISS GERMAN UNIVERSITY

SWISS GERMAN UNIVERSITY

The rise of Cyber security threat is evolving rapidly because the advancement of adversary and the adversary payload. In this era a solo fighter Cyber security defender is not an option and more and more Cyber Security Defender join hand to eliminate the emerging threat and eliminate the usage of the same payload to compromise system this joint work is realized by implementing Cyber Threat intelligent and in this thesis the main purpose is to automatically generate Actionable Cyber Threat Intelligence that is able to capture emerging threat by deploying passive monitoring, trap, deception tool into the darknet environment where supposed there shouldn't have a connection incoming and outgoing to the system, because of that the traffic that is coming to the darknet environment should be treated as an attack. This research categorize the result of the experiment based on the honeypot-based security threats, to generate into the Cyber Threat Information.

Keywords: Actionable Threat Intelligence, Darknet Monitoring, Threat categorization, Analysis of Threat, Traffic analysis, Malware analysis

© Copyright 2020
By Ryandy
All rights reserved

DEDICATION

I dedicate this to my Family.

SWISS GERMAN UNIVERSITY

ACKNOWLEDGEMENT

First of all I would like to thank God because with his word I can finish my thesis without problems. I would also like to express my biggest gratitude to my advisor Dr. Charles Lim and my Co Advisor Mr. Kalpin Erlangga Silaen for the continous mental support and technical support that is provided and also for the time that is sacrificed to support me in doing my thesis. without their knowledge and guidance this thesis may will not be completed

I would like also to thank Mr. Mario Marcello and Mr. Yohanes Syailendra for the technical support and for the guidance

SWISS GERMAN UNIVERSITY

TABLE OF CONTENTS

STATEMENT BY THE AUTHOR 2

ABSTRACT 3

DEDICATION 5

ACKNOWLEDGEMENT 6

CONTENTS 11

LIST OF FIGURES 12

LIST OF TABLES 13

1 INTRODUCTION 14
1.1 Research Background . . . 14
1.2 Research Problem . . . 16
1.3 Research Objective . . . 16
1.4 Significance of Study . . . 16
1.5 Research Question . . . 17
1.6 Scope and Limitation . . . 17
1.7 Hypothesis . . . 17
1.8 Document Structure . . . 17

2 LITERATURE REVIEW 18
2.1 Internet and Cyber Security threat . . . 18
2.2 Darknet . . . 18
2.3 Honeypot . . . 19
2.3.1 Transport Service . . . 22
2.3.1.1 XML & XMPP . . . 22
2.3.1.2 HPFeeds and Mnemosyne . . . 23
2.3.1.3 HPFeeds Vs XMPP . . . 24
2.3.2 Data Processing Service . . . 24
2.3.2.1 Python . . . 24

SWISS GERMAN UNIVERSITY

2.3.2.2 Flow-based Programming 25
2.4 Data Correlation Analysis 25
2.4.1 Graph Database 26
2.5 Threat analysis 26
2.5.1 Traffic 27
2.5.1.1 Threat Categorization 27
2.5.1.2 Whois 27
2.5.1.3 Virus Total 27
2.5.2 Payload 27
2.5.2.1 Static Analysis 28
2.5.2.1.1 PE Studio 28
2.5.2.1.2 Hybrid Analysis 28
2.5.2.2 Dynamic Analysis 28
2.5.2.2.1 Cuckoo 28
2.6 Cyber Attack Framework 29
2.6.1 Cyber Kill Chain 29
2.6.2 MITRE ATT&CK 29
2.7 Threat Intelligence 30
2.7.1 Threat Intelligence Standard 31
2.7.1.1 IODEF 31
2.7.1.2 IDMEF 31
2.7.1.3 CybOX 31
2.7.1.4 VERIS 31
2.7.1.5 OpenIOC 32
2.7.1.6 OTX 32
2.7.1.7 CAPEC 32
2.7.1.8 TAXII 32
2.7.1.9 MAEC 32
2.7.1.10 CIF 33
2.7.1.11 TLP 33
2.7.1.12 STIX 2.0 33
2.7.1.13 Comparison of 3 Top common use CTI Standard . . . 33
2.7.2 Actionable Intelligence 33
2.8 Threat Information Sharing Platform 34
2.8.1 MISP 34

SWISS GERMAN UNIVERSITY

2.9 Related Works 34
2.9.1 Honeypot 34
2.9.1.1 Honeypot for Generation of Cyber Threat Intelligence 34
2.9.1.2 Towards a collaborative architecture of Honeypots . . 35
2.9.2 Threat Analysis 36
2.9.2.1 Dynamic Malware analysis Using Cuckoo Sandbox . 36
2.9.2.2 Data Reduction for the Scalable Automated Analysis of Distributed Darknet Traffic 36
2.9.3 Threat intelligence 37
2.9.3.1 A survey on technical threat intelligence in the age of sophisticated cyber attacks 37
2.9.3.2 Automated Cyber Threat intellgence 37
2.9.3.3 Acing the IOC Game: Toward Automatic Discovery and Analysis of Open-Source Cyber Threat Intelligence 37

3 RESEARCH METHODOLOGY 39
3.1 Research Methodology 39
3.2 System Overview 40
3.3 Research Framework 40
3.3.1 Data Collection 41
3.3.2 Data Processing 41
3.3.2.1 Data Pre Processing 42
3.3.2.2 Data Extraction 42
3.3.3 Data Analysis& Enrichment 42
3.3.3.1 Threat Analysis 42
3.3.3.1.1 Traffic Analysis 43
3.3.3.1.2 Malware 44
3.4 Evaluation and Validation of Result 45

4 EXPERIMENTAL RESULTS 46
4.1 Experiment Setup 46
4.2 Data Collection 46
4.3 Data Pre-Process 48
4.4 Data Analysis and Extraction 51
4.4.1 Data Enrichment 55

SWISS GERMAN UNIVERSITY

4.4.2 Correlation Analysis . . . 57
4.4.2.1 Successful login with empty command (SCS001) query 57
4.4.2.2 Unsuccessful login (SCS002) query . . . 59
4.4.2.3 Persistence cpu (SCS005) query . . . 59
4.4.3 Threat Analysis . . . 59
4.4.3.1 Malware . . . 59
4.4.3.1.1 Static Analysis . . . 60
4.4.3.1.2 Dynamic Analysis . . . 60
4.5 Threat Information generation . . . 60
4.6 Evaluation and Validation of Result . . . 62
4.7 Analysis . . . 63
4.8 Summary . . . 65

5 CONCLUSION **68**
5.1 Recommendations . . . 68
5.2 Future Works . . . 69

GLOSSARY **70**

REFERENCES **75**

APPENDICES **63**

A Sample Data for Data Processing **76**
A.1 Mssqld . . . 76
A.2 Memcache . . . 76
A.3 Mysqld . . . 77
A.4 Mysqld dionaea capture . . . 77
A.5 httpd . . . 77
A.6 Httpd dionaea capture . . . 78
A.7 Pptpd . . . 78
A.8 pcap . . . 79
A.9 Mongod . . . 79
A.10 Ftpd . . . 80
A.11 Epmapper . . . 80
A.12 Mqttd . . . 80

SGU ® SWISS GERMAN UNIVERSITY

B Sample Data for data categorization 82

B.1 Service iptables stop and print the result of service iptables into the etc/rc/local . . . 82

B.2 #!/bin/sh; modify path and show the gcc.pid . . . 83

B.3 Usrbin/bfhhuf and do the secure copy of the usrlocal/bin . . . 83

C Sample Data for Data enhancement 84

C.1 SCS006 Data enhancement . . . 85

C.2 SCS008 Data enhancement . . . 86

D Sample Data for Data correlation query 87

D.1 Check all cowrie query . . . 87

D.2 Vommand is not zero . . . 87

D.3 Uname-a; nproc only . . . 87

D.4 Uname-a only . . . 87

D.5 Uname -s -m only . . . 87

D.6 Service iptables stop and print the result of service iptables into the etc/rc/local . . . 88

D.7 System scheduler . . . 88

CURRICULUM VITAE 89

SWISS GERMAN UNIVERSITY

List of Figures

1.1 Malware Infection Purplesec (2019). 14
1.2 Internet Growth internetworldstats.com (2020) 15

2.1 Classification of Honeypot 21
2.2 Honeypot Deployment Location 22
2.3 Cyber Kill Chain 29
2.4 Mitre Att&ck MITRE(2018) 30
2.5 Multi-Honeypot framework architecture introduced by Kumar et al. (2019) 35
2.6 Honeypot to improve effectiveness of SIEM framework architecture introduced by Kumar et al. (2020) 36

3.1 Research Methodology 39
3.2 System Overview 40
3.3 Research Framework 41
3.4 Data Analysis and Enhancement Flow 43
3.5 Data Evaluation Diagram 45

4.1 Example of static analysis submitted to the VirusTotal 61
4.2 Result of the Cuckoo sandbox for 79665f2a8fea28a4934427ffd59e1de3 file 62
4.3 Result of the Cuckoo sandbox for 79665f2a8fea28a4934427ffd59e1de3 dangerous 62
4.4 MISP Overview 63
4.5 MISP Event Information view 65
4.6 MISP Event detail 66
4.7 Threat Categorization tree for Dionaea and Cowrie 67
4.8 Generic attack sequence 67

SWISS GERMAN UNIVERSITY

List of Tables

2.1 Service of Dionaea and Cowrie 23
2.2 Comparison of HPFeeds Vs XMPP 24
2.3 Comparison of Python Vs Node-RED 26
2.4 Advantages and Disadvantages of 3 Top Standard 34
2.5 Related Works 38

4.1 Hardware Setup 46
4.2 Software Requirement 47
4.3 Data Collection 47
4.4 Total Attack of Cowrie and Dionaea Honeypot per day in June 2020 . . 48
4.5 Channel that is used by HPFeeds to differentiate Dionaea and Cowrie information 49
4.6 Important key 52
4.7 SCS005 Analysis example 53
4.8 SCS006 Analysis example 54
4.9 SCS008 Analysis example 55
4.10 Threat Categorization mapped into the MITRE ATT&CK Matrix 58
4.11 Threat Categorization Total 64

SWISS GERMAN UNIVERSITY