

ANALYZING INSIDER THREATS BASED ON DNS NETWORK TRAFFIC
(CASE STUDY IN ORGANIZATION XYZ)

By
Kris Ivan Santosa
12112012

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



SWISS GERMAN UNIVERSITY
EduTown BSDCity
Tangerang 15339
Indonesia

August 2016

ANALYZING INSIDER THREATS BASED ON DNS NETWORK TRAFFIC
(CASE STUDY IN ORGANIZATION XYZ)

By
Kris Ivan Santosa
12112012

BACHELOR'S DEGREE
in

INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY



SWISS GERMAN UNIVERSITY
EduTown BSDCity
Tangerang 15339
Indonesia

August 2016

Revision after the Thesis Defense on 21 July 2016

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.

Kris Ivan Santosa

Student

Date

Approved by:

Charles Lim, M.Sc., ECSA, ECSP, ECIH, CEH, CEI

Thesis Advisor

Date

Alva Erwin, ST, M.Sc., MTI

Thesis Co-Advisor

Date

Dr. Ir. Gembong Baskoro, M.Sc

Dean

Date

Kris Ivan Santosa

ABSTRACT

ANALYZING INSIDER THREATS BASED ON DNS NETWORK TRAFFIC (CASE STUDY IN ORGANIZATION XYZ)

By

Kris Ivan Santosa

Charles Lim, M.Sc., ECSA, ECSP, ECIH, CEH, CEI, Advisor

Alva Erwin, ST, M.Sc., MTI, Co-Advisor

SWISS GERMAN UNIVERSITY



The Internet is a media for people to communicate with each other. The Internet also full of threats for it's users. This makes security one of the problem of the Internet. This is also one of the problem for organizations that use the Internet.

Organization have two types, external threats and insider threats. External threats are threats that came from outside of the system and insider threats are threats that came from inside of the system. Most of organizations prioritize external threats over insider threats. Although insider threats are the dominant in security breaches and the number of insider breaches are increasing.

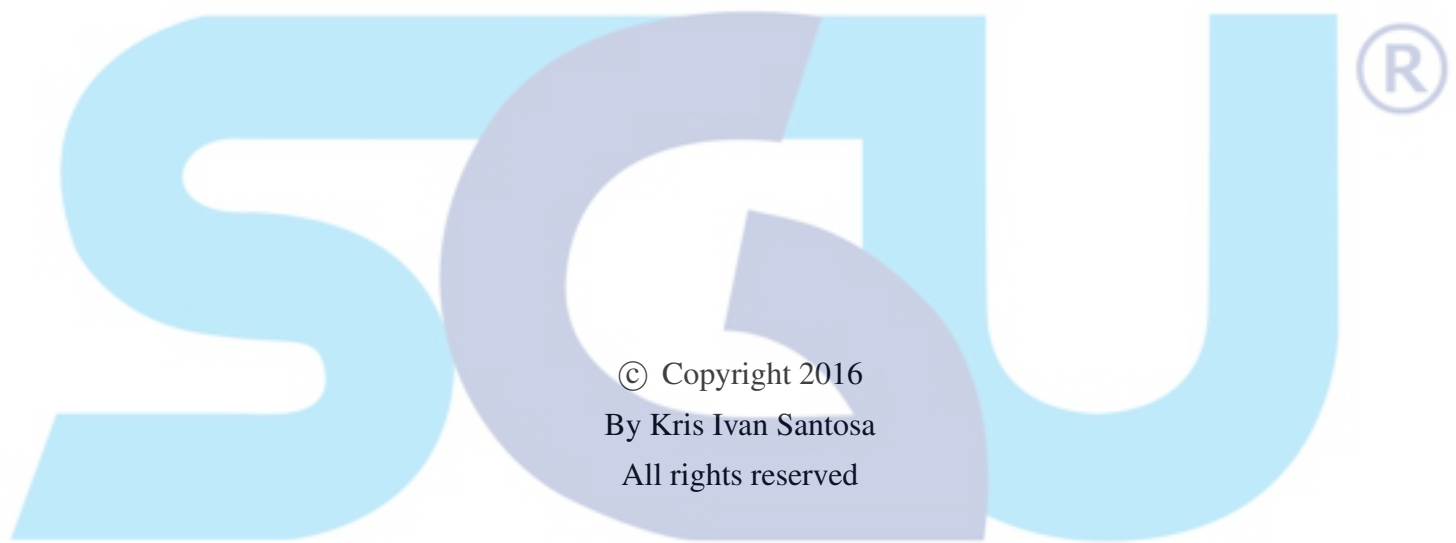
DNS is one of the main function of Internet. One of DNS function is to resolve domain name to IP address. Most of the user use DNS to be able to connect to the Internet including malicious hackers. DNS can also be used to detect insider threat using the features of insider threats which can detect unknown insider threats. This research aims to detect insider threats using DNS based detection.

The features of the insider threats will be extracted from the raw DNS queries. These features will be preprocessed to remove the unused data and will be clustered. From the clusters, it can show the features of insider threats.

This research is able to suspect the clusters. The result is that there may exist insider threats in organizations and the most frequent suspects of insider threats are botnet which categorized as misuse in insider threat classification. There also some clusters that benign but abnormal traffic that still have few features of insider threats. The recommendation for the insider threats mitigation in organizations are using features of botnets to filter the DNS packets and block the domain once it reach certain threshold.

Keywords: clustering, insider threat, DNS, machine learning, data mining





SWISS GERMAN UNIVERSITY

DEDICATION

I dedicate this work to my parents that always supporting me and give their best to take care of me.

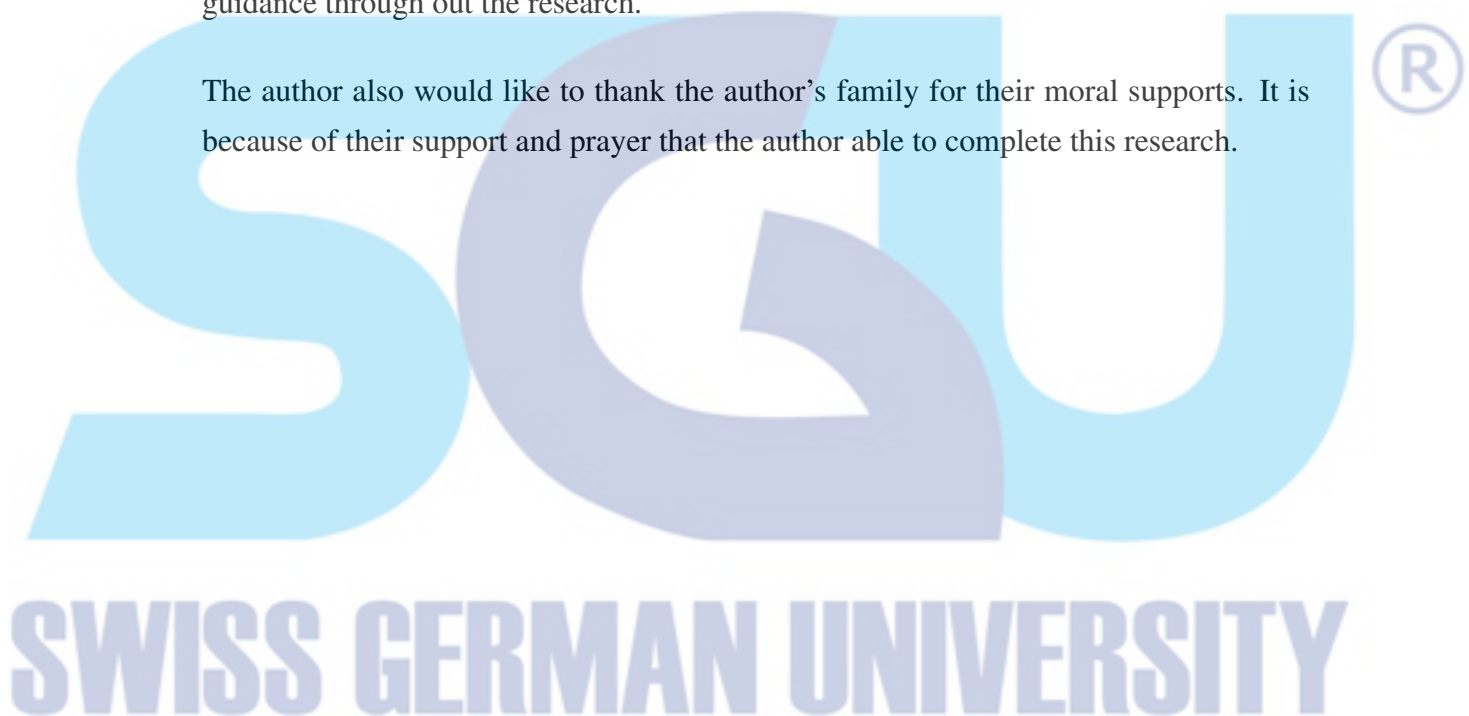


ACKNOWLEDGEMENTS

The author would like to express deepest gratitude to Mr. Charles Lim, MSc., ECSA, ECSP, ECIH, CEH, CEI and Mr. Alva Erwin, ST, M.Sc., MTI for their guidance during the completion of this research. With their support, this research can reach this point.

The author is also grateful for guidance and support from his friends, seniors during this research development. Thanks to the numerous support of his friends and seniors that the author could complete this research. Especially Mario Marcello M.Kom for his guidance through out the research.

The author also would like to thank the author's family for their moral supports. It is because of their support and prayer that the author able to complete this research.



Contents

STATEMENT BY THE AUTHOR	2
ABSTRACT	3
DEDICATION	6
ACKNOWLEDGEMENTS	7
TABLE OF CONTENTS	10
LIST OF FIGURES	11
LIST OF TABLES	12
1 INTRODUCTION	13
1.1 Research Background	13
1.2 Research Problem	15
1.3 Research Objective	15
1.4 Research Questions	15
1.5 Hypotheses	15
1.6 Significance of Studies	16
1.7 Research Scope	16
1.8 Document Structure	16
2 LITERATURE REVIEW	17
2.1 Internet	17
2.2 Internet Security Threat	17
2.3 Internet Security Threats to Organization	18
2.4 Insider Threats	18
2.5 Insider Threat Classification	19
2.6 Domain Name System	20
2.6.1 DNS Capture	21
2.7 Malware	22
2.7.1 Malware Traffic	22
2.7.2 Botnet	22
2.7.3 Features of Botnet	22

2.7.3.1	Daily Similarity	23
2.7.3.2	Number of distinct IP address	24
2.7.3.3	Number of distinct countries	24
2.7.3.4	Average TTL value	24
2.7.3.5	Numerical char length domain name	24
2.7.3.6	Silent IP	25
2.7.3.7	Very low frequency query	25
2.7.3.8	Average Packet length	25
2.8	Non-Malicious Traffic	25
2.8.1	Content Delivery Network	25
2.8.2	Feature of Content Delivery Network	26
2.8.2.1	Low TTL	26
2.8.3	Normal DNS Traffic	27
2.8.4	Features of Normal DNS Traffic	27
2.8.4.1	TTL	27
2.8.4.2	IP address	27
2.8.4.3	Country	27
2.9	Machine Learning	27
2.9.1	Genetic Algorithm	28
2.9.2	Expectation-Maximization Algorithm	29
2.10	Related Work	30
3	RESEARCH METHODOLOGY	34
3.1	Research Methodology	34
3.2	Research Framework	34
3.3	Data Collection	35
3.4	Pre-processing	35
3.5	Clustering	36
3.6	Analysis	36
4	EXPERIMENTAL SETUP	38
4.1	Experiment Environment	38
4.2	Data Collection	39
4.3	Pre-processing	40
4.3.1	First Filter	40
4.3.2	Second Filter	40
4.3.3	Feature Selection	41
4.4	Clustering	41

4.5	Results	42
4.5.1	Week 1	43
4.5.2	Week 2	45
4.5.3	Week 3	45
4.5.4	Week 4	46
4.6	Analysis	47
4.7	Counter Measure	48
5	CONCLUSION AND FUTURE WORK	49
5.1	Conclusions	49
5.2	Recommendation	49
5.3	Future Work	49
	GLOSSARY	51
	REFERENCES	55
	CURRICULUM VITAE	56



SWISS GERMAN UNIVERSITY