UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Faculty of Information and Communication Technology

INSIDER THREATS DETECTION MODEL FOR EMAIL CONTENT USING STATISTICAL ANALYSIS

Nur Ameera Natasha Binti Mohammad

Master of Science in Information and Communication Technology

2022

INSIDER THREATS DETECTION MODEL FOR EMAIL CONTENT USING STATISTICAL ANALYSIS

NUR AMEERA NATASHA BINTI MOHAMMAD

A thesis submitted
in fulfillment of the requirements for the
Master of Science in Information and Communication Technology

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2022

DECLARATION

I declare that this thesis entitled “Insider Threats Detection Model For Email Content Using Statistical Analysis” is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : …………………………………………………………...

Name : Nur Ameera Natasha Binti Mohammad

Date : 1/3/2022

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Master of Science in Information and Communication Technology.

Signature : ……………………………………………………...

Supervisor Name : Dr. S.M Warusia Mohamed Bin S.M.M Yassin

Date : 1/3/2022

DEDICATION

To my beloved mother and father, Mohammad Bin Ismail and Enisabiren Binti Basir and those who have always by my side during my upside down, without whom none of my success would be possible.

ABSTRACT

An insider threat has become one of the most challenging malicious activities in cybersecurity defence systems in a contrast to outsider threats recently. Usually, IP theft, fraud and sabotage against legal information are three well-known types of insider threat. Since an insider threat usually expands and spread internally, no one could predict what, when and how exactly malicious insider launched their attacks. This is with a view of fact that an email becomes one of the primary targets of an internal threat as this medium is widely used by everyone to communicate, share, and exchange confidential information. Therefore, it is extremely important to understand the nature of insider threat behavior beforehand and construct an accurate detection model. Furthermore, every single keyword used in an email can reflect the behavior of an individual and can be used to determine their intentions, such as having a motive to launch an insider threat or not. Henceforth, an innovative approach is proposed in modelling insider threat detection in this work. In addition, various approaches such as scoring, Friedman, linear regression (R^2) and correlation coefficient applied to analyse an insider threat relationship between historical insider threats behavior and relevant extracted keywords from email content. Firstly, the email content filtered into three different factors that influence the characteristics of an insider such as motive, opportunity and capability, before calculating the scores for the entire insider's keywords. Next, the Friedman statistical used to determine the minimum differences between each extracted insider threats keywords that represent different insider threat factors (motive, opportunity, capability). Besides, linear regression applied to estimate the relationship of an insider threat from training keywords and testing keywords with allocating an anomaly score. Finally, the correlation coefficient approach used to determine how strong a relationship is between extracted insider threats keywords and insider threat behavior in this research. The proposed modelling approach has been evaluated using the benchmark dataset known as CERT that comprises a malicious email file. Throughout the experiment, the proposed insider threats detection approach has achieved a higher attack detection rate as well as minimized undetectable insider threats behavior as compared to the previous researcher works.

MODEL PENETAPAN ANCAMAN ORANG DALAM UNTUK KANDUNGAN E-MEL MENGGUNAKAN ANALISIS STATISTIK

ABSTRAK

Ancaman dalaman telah menjadi salah satu aktiviti berbahaya yang paling mencabar dalam sistem pertahanan keselamatan siber berbanding dengan ancaman luaran baru-baru ini. Biasanya, pencurian, penipuan dan sabotaj IP terhadap maklumat sulit adalah tiga jenis ancaman dalaman yang terkenal. Oleh kerana ancaman dalaman biasanya berkembang dan tersebar secara dalaman, tidak ada yang dapat meramalkan apa, bila dan bagaimana pelaku yang jahat melancarkan serangan mereka. Kenyataan ini dinyatakan bersama fakta bahawa e-mel menjadi salah satu sasaran utama ancaman dalaman kerana media ini digunakan secara meluas oleh setiap individu untuk berkomunikasi, berkongsi, dan bertukar maklumat sulit. Oleh itu, adalah sangat penting untuk memahami sifat tingkah laku ancaman dalaman terlebih dahulu dan membina model pengesanan yang tepat. Selanjutnya, setiap kata kunci yang digunakan dalam e-mel dapat menggambarkan tingkah laku individu dan dapat digunakan untuk menentukan niat mereka, iaitu melalui motif untuk melancarkan ancaman dalaman atau tidak. Sejauh ini, pendekatan inovatif dicadangkan dalam memodelkan pengesanan ancaman dalaman dalam karya ini. Di samping itu, pelbagai pendekatan iaitu pemarkahan, Friedman, regresi linear (R^2) dan pekali korelasi digunakan untuk menganalisis hubungan antara tingkah laku ancaman dalaman dengan kata kunci relevan yang diekstrak dari kandungan e-mel. Pertama, kandungan e-mel disaring menjadi tiga faktor berbeza yang mempengaruhi ciri ancaman dalaman seperti motif, peluang dan kemampuan, sebelum mengira skor untuk kata kunci keseluruhan ancaman dalaman. Seterusnya, statistik Friedman digunakan untuk menentukan perbezaan minimum antara setiap kata kunci ancaman dalaman yang diambil yang mewakili faktor ancaman yang berbeza (motif, peluang, kemampuan). Selain itu, regresi linear digunakan untuk menganggar hubungan ancaman dalaman dari kata kunci latihan dan kata kunci pengujian dengan memperuntukkan skor anomali. Akhirnya, pendekatan pekali korelasi digunakan untuk menentukan seberapa kuat hubungan antara kata kunci ancaman dalaman yang diekstrak dan tingkah laku ancaman dalaman melalui penyelidikan ini. Pendekatan pemodelan yang dicadangkan telah dinilai menggunakan set data penanda aras yang dikenali sebagai CERT yang terdiri daripada fail e-mel yang berniat jahat. Sepanjang eksperimen, pendekatan pengesanan ancaman dalaman yang dicadangkan telah mencapai kadar pengesanan serangan yang lebih tinggi serta meminimumkan tingkah laku ancaman dalaman yang tidak dapat dikesan dibandingkan dengan karya penyelidik sebelumnya.

ACKNOWLEDGMENTS

First and foremost, I would like to take this opportunity to express my sincere acknowledgement to my supervisor, Dr. S.M Warusia Mohamed Bin S.M.M Yassin and my co-supervisor, Professor Ts. Dr. Rabiah Binti Ahmad for all the supervision, support and suggestion towards the completion of this thesis and contributed to its technical content. I am also very grateful to authors of various articles on the internet who make me aware of the research currently ongoing in this field. Special thanks to my beloved parents, sister, family members and all my friends who have always been there to give support and helped me to complete this research work. Last but not least, I would like to thank everyone who had been to the crucial parts of the realization of this research.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATION

CITD	-	Corporate Insider Threat Detection
CMO	-	Capability Means Opportunity
DLP	-	Data Loss Prevention
GRU	-	Gate Recurrent Unit
HIE	-	Heterogeneous Information Ensemble
IP	-	Intellectual Property
IT	-	Information Technology
ITS	-	Insider Threats Score
K-NN	-	K-Nearest Neighbours
NMF	-	Non-negative Matrix Factorization
PCA	-	Principal Component Analysis
PS	-	Problem Statement
RO	-	Research Objectives
SOFIT	-	Sociotechnical and Organizational Factor for Insider Threats
SQL	-	Structured Query Language
SVM	-	Singular Value Decomposition
SVM	-	Singular Value Decomposition
TF-IDF	-	Term Frequency-Inverse Document Frequency

CHAPTER 1

INTRODUCTION

1.1 Introduction

Intellectual property and private information have become more difficult to protect as malicious threats nowadays emanate more from internal sources Gavai, Sricharan, Gunning, Hanley and Rolleston (2015). Known as insider threats, they are launched by malicious insiders who have circumvented any cybersecurity defence systems of organizations. Such malicious insiders come from within organizations themselves and could be current employees, retired or resigned employees, business collaboration partners, and anyone having knowledge of computer systems including the vulnerabilities and weaknesses of the organizations. IP theft, fraud and sabotage are types of insider threats launched by malicious insiders that create trust and financial issues as well as loss of reputation of the targeted entity (Jang-Jaccard, 2014);(UK, 2015);(Heneke et al., 2016); (Ekran, 2016);(Wood et al., 2016);(Scott and Spaniel, 2017);(Huff et al., 2019);(Biringer et al., 2019);(Lykou, 2019);(Ambrogi, 2019).

Therefore, it is important to understand and analyse insider threats from internal sources such as email platforms to mitigate the damage. Common platforms, such as emails are familiar and used by millions including malicious insiders to easily perform daily tasks. As such, communicating and sharing or exchanging legal and private information through emails make it easy for malicious insiders to access important information. In addition, emails also contain various keywords used for communicating. Thus, their content also becomes a medium for people, including potential malicious parties, to express and convey their feelings. This research proposes modelling an insider-threat detection approach for understanding insider-threat behaviour as reflected in relevant extracted keywords and in detecting the relationships between each of the elements involved using Friedman's statistical, linear regression, and correlation coefficients. In addition, insider threat behaviour is divided into three factor indicators, namely motive, capability, and opportunity to enable a better understanding of which behaviour will indicate the particular insider threat factor.

1.2 Motivation

Addressing insider threats has become a major cybersecurity issue in mitigating malicious threats. As malicious insiders are individuals who benefit from access to legal and private information and assets as well as have knowledge about the vulnerabilities in an organization, they can easily launch attacks without being detected by cybersecurity defence. This situation is worsened if an organization relies only on traditional security cybersecurity systems that exclude any insider threat detection mechanisms. Consequently, more updated and efficient cybersecurity defence is required especially in detecting and mitigating insider threats. Additionally, organizations rely heavily on email as an important and ubiquitous platform to facilitate their daily tasks. Therefore, email content is potentially open to malicious insiders for gathering legal and private information for their own benefit. Also, keywords used in email content can directly portray individuals, including insider threat, behaviour. Hence, this research applies various approaches that focus on detecting insider threat relationships and behaviour and relevant keywords used through Friedman's statistical, linear regression, and correlation coefficients. This research proposes a model for identifying the high attack-detection rate to detect insider threats.

1.3 Problem Statement

The number of insider threat issues to cybersecurity is growing rapidly and of major concern to the data and privacy of organizations. Due to the lack of studies on insider threat detection, proposing a solution could be difficult as the influencing factor of insider threats is difficult to be determine. This is because understanding and acknowledging such malicious behaviour needs more studies to define such threats more accurately. As this research focuses on insider threat detection on email content, it is important to understand and determine such activities and the relevant keywords that are used. However, the difficulty to determine the factor that influences insider threats by analysing extracted insider keywords is one of the issues that need to be overcome.

Moreover, detecting insider threats behaviour more accurately as insider threats keywords effected by the outlier is difficult as this is a time consuming and complex process. This phase needs to be done with accurate references from various resources to determine what indicators of insider threat behaviour is specifically reflected in each malicious keyword.

Table 1.1: Summary of Problem Statement

PS	Problem Statement.
PS1	Difficult to determine the factor that influences insider threat by analyzing extracted insider keywords.
PS2	Difficult to detect insider threat behaviour more accurately as insider threat keywords effected by the outlier.

1.4 Research Question

i. What are the current issues regarding insider threats?

ii. What method should be used to detect the behaviour of insider threats from relevant extracted insider threat keywords?

iii. How to accurately detect insider threat relationships?

1.5 Research Objective

Insider threats can be determined with the proposed relevant detection method. The process of understanding such malicious behaviour and the characteristics, types and impacts of insider threats can contribute towards proposing effective models for insider threat detection. Besides that, current issues of insider threats also should be investigated as such threats are daily found to have various ways and platforms to create harm. Besides, by investigate critical success factor of insider threats detection model, previously proposed detection approaches could be used as a reference to provide better and more effective insider threat detection. Therefore, the first research objective is to investigate critical success factors of insider threat detection model.

Through a deeper investigation of insider threats issues, it will be possible to have a better understanding of insider threat behaviour and lead to efficient detection approaches. In this research, several previously proposed insider threat detection approaches and mechanisms have been used to accurately detect insider threats. However, such attempts face many drawbacks as malicious insiders use various means to launch insider threats. Therefore, this research proposes modelling an insider threat detection approach by the application of a scoring and statistical-based mechanism that can analyse and accurately detect insider threat behaviour. This second research objective proposes modelling an insider threat detection approach that is updated and more effective in mitigating current insider threat issues. Table 1.2 provides a summary of the research objectives.

Table 1.2: Summary of Research Objective

RO	Research Objective.
RO1	To investigate critical success factors of insider threats detection model.
RO2	To propose insider threats detection model using scoring and statistical analysis to improve detection accuracy that influenced by outlier.

1.6 Research Scope

This research applies various approaches to detect insider threats.

1.6.1 Specific Tools

i. Use of Microsoft Excel to gather related insider threat data.

ii. Use of SQLyog to execute SQL query.

iii. Weka 3.8.3.

1.6.2 Specific Data Used

i. CERT dataset.

1.7 Research Contribution

Table 1.3 provides a summary of the research problem statement, research questions, research objectives, and research contribution of the proposed model of the insider threat detection approach.

Table 1.3: Summary of Problem Statement, Research Questions, Research Objectives and Research Contribution

Problem Statement	Research Questions	Research Objectives	Research Contribution
Difficult to determine the factor that influences insider threat by analyzing extracted insider keywords. Difficult to detect insider threat behaviour more accurately as	What are the current issues regarding insider threats? What method should be used to detect the behaviour of insider threats from relevant extracted insider-threat keywords?	To investigate critical success factors of insider threats detection model. To propose insider threats detection model using scoring and statistical analysis to improve detection accuracy	The proposed modelling of the insider threat detection approach is able to analyze and detect insider threat behaviour as reflected from relevant extracted

insider threat keywords effected by the outlier.	How to accurately detect insider threat relationships?	that influenced by outlier.	insider threat keywords. The scoring method applied in the proposed insider threat detection model can produce anomaly scores for attack profiling which is important to mark the detected insider threat behaviour more accurately.

In this research, the proposed modelling for the insider threats detection approach is based on earlier studies and covers three major factors, namely motive, capability and opportunity. Each of the relevant insider threats keywords is extracted from email content and indicates different factors representing the existence of insider behavior. When the relationship between insider threat behaviour and each extracted keyword is determined, the threat is defined as detected. In addition, the attack profiling and scoring method are also used in the proposed modelling approach to marking insider behaviour more accurately. The applied Friedman statistical, linear regression and correlation coefficient in this proposed model is able to better detect the relationship between insider threat behavior and the extracted insider threats keywords compared to previous studies.

1.8 Thesis Organization

This dissertation is divided into seven chapters.

Chapter 1: Introduction

This chapter describes the research and the activities to be developed. It covers the research motivation, problem statement, research questions, research objectives, research scope, and research contribution.

Chapter 2: Literature Review

This chapter discusses related work in a number of earlier studies and includes journals and books on the subject matter from 2010 to 2021. It explores current drawbacks, identifies gaps, and proposes solutions. The factors indicating insider threat behaviour (motive, capability and opportunity) and the impacts of insider threats are also covered.

Chapter 3: Research Methodology

This chapter discusses the framework methodology and each stage of the processes involved in conducting the research. In addition, insider threat dataset collection is also explained briefly.

Chapter 4: The Proposed Insider Threats Detection Model Approach

This chapter introduces the implementation phase of the proposed model on insider threat detection including the SQL query and data pre-processing process following data collection. The insider threats dataset is divided into the training and testing stages to produce better accuracy in attack detection. The framework for the proposed model of keyword analysis and extraction of relevant insider threats keywords, the Friedman statistical, linear regression, and correlation coefficient are also discussed.

Chapter 5: Results

Chapter 5 covers the performance evaluation including evaluation of factor indicator (motive, capability and opportunity), attack detection rates of the proposed model on insider threat detection, results and discussion of applied approaches, overall findings of the relationship between insider threats, and comparison of performance evaluation.

Chapter 6: Conclusion and Future Work

Chapter 6 includes the conclusion and future work on the proposed threat detection model approach and the contributions of the work.

1.9 Summary

This chapter included an introduction of insider threats which to express the importance of understanding and analyzing insider threats from internal sources such as email. As one of the common communication platforms, emails are used by millions including malicious insiders to perform daily tasks smoothly. Moreover, communicating and exchanging legal and private information through emails is able to make malicious insiders access important information. As well as emails containing various keywords used for communicating, the content becomes a medium to express and convey their feelings. Therefore, each keyword used represent emotion of many people. Furthermore, most organizations relied on email as important platform to simplify their daily tasks. This situation is worsened if organization relies only on traditional security cybersecurity systems that exclude insider threat detection mechanisms because malicious insiders are individuals who have benefit to access any private information and assets. With knowledge about the vulnerabilities in an organization, they can easily launch attacks without being detected by cybersecurity defense. Prior objectives of this research detecting insider threats from email, malicious behaviour of insiders that reflect from relevant extracted insider threats keywords are difficult as this is a time consuming and complex process. Therefore, various studies need to be done with accurate references from a number of resources in order to determine what indicators of insider threat behaviour is specifically reflected in each malicious keyword.