AN ENSEMBLE-BASED ANOMALY-BEHAVIOURAL CRYPTO-RANSOMWARE PRE-ENCRYPTION DETECTION MODEL

BANDER ALI SALEH AL-RIMY

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Computer Science)

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

MAY 2019

DEDICATION

This thesis is dedicated to my father and mother whose Du'a has been my guidance along the way. It is also dedicated to my sincere wife Asma'a; whose love, encouragement, sacrifice and patience have fuelled my resoluteness to finish this journey. To my sweetheart kids, Mohammed and Ayman, whose dreams and enthusiasm have inspired me.

ACKNOWLEDGEMENT

In the name of Allah, Most Gracious, Most Merciful and peace and blessings be upon his Messenger, Mohammed (Peace Be Upon Him). I thank Allah S.W.T for granting me perseverance and strength I needed to complete this thesis.

In preparing this thesis, I was in contact with many people, researchers, academicians and practitioners. They have contributed towards my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main supervisor, Professor Dr. Mohd Aizaini Maarof, for encouragement, guidance, advices and comments. His patience and considerate nature made him accessible whenever I needed his opinion and assistance. I am also very grateful to my co-supervisor Dr. Syed Zainudeen Mohd Shaid for his guidance, advices, motivation and friendship. I am indebted to both of them for teaching me how to identify interesting problems and how the research can be conducted correctly.

I would like to take advantage of the unique opportunity to express my greatest thanks from my heart to my parents, wife, kids, brothers and sisters for giving the unlimited morale supports and patience to complete my study. I would never forget their sacrifice that they have made for me. Their prayers and support have empowered my determination to make it to the end of this journey. Words are not enough to say thank you to my friend Dr. Fuad A. Ghaleb Alshameri for his advices, comments and constructive criticism. My sincere appreciation also extended to all my colleagues and friends who have in one way or the other added value to my life.

ABSTRACT

Crypto-ransomware is a malware that leverages cryptography to encrypt files for extortion purposes. Even after neutralizing such attacks, the targeted files remain encrypted. This irreversible effect on the target is what distinguishes crypto-ransomware attacks from traditional malware. Thus, it is imperative to detect such attacks during pre-encryption phase. However, existing crypto-ransomware early detection solutions are not effective due to inaccurate definition of the pre-encryption phase boundaries, insufficient data at that phase and the misuse-based approach that the solutions employ, which is not suitable to detect new (zero-day) attacks. Consequently, those solutions suffer from low detection accuracy and high false alarms. Therefore, this research addressed these issues and developed an Ensemble-Based Anomaly-Behavioural Pre-encryption Detection Model (EABDM) to overcome data insufficiency and improve detection accuracy of known and novel crypto-ransomware attacks. In this research, three phases were used in the development of EABDM. In the first phase, a Dynamic Pre-encryption Boundary Definition and Features Extraction (DPBD-FE) scheme was developed by incorporating Rocchio feedback and vector space model to build a pre-encryption boundary vector. Then, an improved term frequency-inverse document frequency technique was utilized to extract the features from runtime data generated during the pre-encryption phase of crypto-ransomware attacks' lifecycle. In the second phase, a Maximum of Minimum-Based Enhanced Mutual Information Feature Selection (MM-EMIFS) technique was used to select the informative features set, and prevent overfitting caused by high dimensional data. The MM-EMIFS utilized the developed Redundancy Coefficient Gradual Upweighting (RCGU) technique to overcome data insufficiency during pre-encryption phase and improve feature's significance estimation. In the final phase, an improved technique called incremental bagging (iBagging) built incremental data subsets for anomaly and behavioural-based detection ensembles. The enhanced semi-random subspace selection (ESRS) technique was then utilized to build noise-free and diverse subspaces for each of these incremental data subsets. Based on the subspaces, the base classifiers were trained for each ensemble. Both ensembles employed the majority voting to combine the decisions of the base classifiers. After that, the decision of the anomaly ensemble was combined into behavioural ensemble, which gave the final decision. The experimental evaluation showed that, DPBD-FE scheme reduced the ratio of crypto-ransomware samples whose pre-encryption boundaries were missed from 18% to 8% as compared to existing works. Additionally, the features selected by MM-EMIFS technique improved the detection accuracy from 89% to 96% as compared to existing techniques. Likewise, on average, the EABDM model increased detection accuracy from 85% to 97.88% and reduced the false positive alarms from 12% to 1% in comparison to existing early detection models. These results demonstrated the ability of the EABDM to improve the detection accuracy of crypto-ransomware attacks early and before the encryption takes place to protect files from being held to ransom.

ABSTRAK

perisian tebusan-kripto adalah malware yang memanfaatkan kriptografi untuk menyulitkan fail bagi tujuan pemerasan. Walaupun setelah serangan dineutralkan, fail yang disasarkan kekal tersulit. Kesannya yang tidak dapat dikembalikan kepada sasaran adalah apa yang membezakan serangan perisian tebusan-kripto dari serangan malware tradisional. Oleh itu, adalah penting untuk mengesan serangan tersebut semasa fasa pra-penyulitan. Walau bagaimanapun, penyelesaian pengesanan awal serangan perisian tebusan-kripto yang sedia ada tidak berkesan kerana penggunaan definisi sempadan fasa pra-penyulitan yang tidak tepat, data yang tidak mencukupi pada fasa tersebut dan pendekatan berasaskan penyalahgunaan yang menggunakan penyelesaian yang tidak sesuai untuk mengesan serangan yang baru. Oleh itu, penyelesaian tersebut menyumbang kepada kadar pengesanan yang rendah dan penggera palsu yang tinggi. Oleh itu, penyelidikan ini membincangkan isu-isu ini dengan membangunkan Model Pengesanan Pra-penyulitan Perilaku Anomali (EABDM) bagi mengatasi kekurangan data dan meningkatkan ketepatan pengesanan serangan perisian tebusan-kripto yang sedia diketahui dan yang baru. Dalam kajian ini, tiga fasa digunakan dalam pembangunan EABDM. Pada fasa pertama, skema Definisi Sempadan Pra-penyulitan Dinamik dan Pengekstrakan Ciri (DPBD-FE) telah dibangunkan dengan memasukkan model ruang maklum balas dan vektor Rocchio untuk membina vektor sempadan pra-penyulitan. Kemudian, teknik baru kekerapan terma-frequensi dokumen songsang yang lebih baik telah digunakan untuk mengekstrak ciri-ciri dari data perilaku sampel yang dijana semasa kitar hayat fasa pra-penyulitan perisian tebusan-kripto. Pada fasa kedua, teknik Maksimum Minimum Pemilihan Ciri Maklumat Bersama Tertingkat (MM-EMIFS) digunakan untuk memilih ciri-ciri maklumat yang ditetapkan, dan mencegah limpahan yang disebabkan oleh dimensi data yang tinggi. MM-EMIFS menggunakan Teknik Peningkatan Beransur-ansur Pekali Lebihan (RCGU) yang dibangunkan untuk mengatasi masalah kekurangan data semasa fasa pra-penyulitan dan meningkatkan anggaran ciri-ciri yang penting. Pada fasa akhir, teknik yang dipertingkatkan yang disebut penambahan penyarungan (iBagging) telah dicadangkan untuk membina subset data tambahan bagi kesatuan pengesanan berasaskan perilaku dan anomali. Teknik Dipertingkat Pemilihan Subruang Separa-Rawak (ESRS) kemudian digunakan untuk membina subruang pelbagai yang bebas hingar bagi setiap subset data tambahan itu. Berdasarkan subruang tersebut, pengklasifikasi asas dilatih bagi setiap kesatuan. Kedua-dua kumpulan itu menggunakan pengundian majoriti untuk menggabungkan keputusan pengkelas asas. Selepas itu, keputusan anomali digabungkan menjadi kesatuan perilaku yang akan memberikan keputusan muktamad. Penilaian eksperimen menunjukkan bahawa, skema DPBD-FE mengurangkan nisbah sampel perisian tebusan-kripto yang sempadan pra-penyulitannya tidak terjawab dari 18% ke 8% berbanding dengan kerja yang ada. Selain itu, ciri-ciri yang dipilih oleh teknik MM-EMIFS meningkatkan ketepatan pengesanan dari 89% ke 96% berbanding dengan teknik sedia ada. Begitu juga secara purata, model EABDM meningkatkan ketepatan pengesanan dari 85% hingga 97.88% dan mengurangkan penggera positif palsu dari 12% ke 1% berbanding model pengesanan awal yang sedia ada. Hasil tersebut menunjukkan keupayaan EABDM untuk meningkatkan ketepatan pengesanan serangan perisian tebusan-kripto awal sebelum penyulitan berlaku, melindungi fail daripada disulitkan untuk tebusan.

TABLE OF CONTENTS

TITLE		PAGE
DECLERATION		**ii**
DEDICATION		**iii**
ACKNOWLEDGMENT		**iv**
ABSTRACT		**v**
ABSTRAK		**vi**
TABLE OF CONTENTS		**vii**
LIST OF TABLES		**xiii**
LIST OF FIGURES		**xv**
LIST OF ABBREVIATIONS		**xviii**
LIST OF APPENDICES		**xix**
CHAPTER 1	**INTRODUCTION**	**1**
1.1	Overview	1
1.2	Problem Background	3
1.3	Problem Statement	8
1.4	Research Aim	10
1.5	Research Objectives	10
1.6	Research Scope	11
1.7	Significance of Research	12
1.8	Research Methodology	12
1.9	Research Contribution	13
1.10	Thesis Organization	14
CHAPTER 2	**LITERATURE REVIEW**	**17**
2.1	Introduction	17
2.2	Ransomware Overview	17
2.3	Ransomware Attacks	19

2.3.1 Ransomware Evolution 20
2.3.2 Ransomware Threat Success Factors 22
2.3.2.1 Ransomware Infection Vectors 22
2.3.2.2 Ransomware Enabling Technologies 23
2.3.3 Ransomware Attacks Phases 25
2.3.4 Ransomware Types 28
2.3.4.1 Scareware 29
2.3.4.2 Detrimental Ransomware 30
2.4 Research in Crypto-ransomware 31
2.4.1 Crypto-ransomware Analysis 32
2.4.1.1 Static Analysis 33
2.4.1.2 Dynamic Analysis 34
2.4.2 Crypto-ransomware Detection 38
2.4.2.1 Strctural-Based Detection 38
2.4.2.2 Behavioral-Based Detection 39
2.4.3 Zero-Day Attacks Detection 41
2.5 Crypto-ransomware Detection Approaches 46
2.5.1 Data Centric-Based Approach 46
2.5.2 Process Centric-Based Approach 48
2.5.2.1 Event-Based Detection 48
2.5.2.2 Machine Learning-Based Detection 49
2.6 Related Techniques Used to Build Crypto-ransomware Early Detection Models 53
2.6.1 Features Extraction Techniques 55
2.6.2 Features Selection Techniques 57
2.6.3 Detection Techniques 60
2.7 Limitations of Existing Research in Crypto-ransomware Early Detection 62
2.7.1 Limitations Related to Pre-encryption Features Extraction 62
2.7.2 Limitations Related to Features Selection 63
2.7.3 Limitations Related to Detection Model Design 64

	2.8	Discussion and Research Directions	66
	2.9	Summary	69
CHAPTER 3	**RESEARCH METHODOLOGY**		**71**
	3.1	Introduction	71
	3.2	Problem Situation and Solution Concept	71
	3.3	Dataset	73
	3.4	The Architecture of the Proposed Model	74
		3.4.1 Phase 1: A Dynamic Pre-encryption Boundary Definition and Features Extraction Scheme	76
		3.4.2 Phase 2: Redundancy Coeffcient Gradual Upwieghting-Baded Mutual Information Feature Selection Technique	77
		3.4.3 Phase 3: An Ensemble-Based Anomaly-Behavioral Crypto-ransomware Pre-encryption Detection Model	78
	3.5	Experimental Enivronment Setup	80
	3.6	Evaluation Metrics	80
	3.7	Results Analysis and Validation	82
	3.8	Summary	82
CHAPTER 4	**DYNAMIC PRE-ENCRYPTION BOUNDARY DEFINITION AND FEATURES EXTRACTION SCHEME**		**85**
	4.1	Introduction	85
	4.2	Overview on the Dynamic Pre-encryption Boundary Definition and Features Extraction Scheme	85
		4.2.1 Raw Data Collection	87
		4.2.2 Dynamic Pre-encryption Boundary Definition Technique for Pre-encryption Data Extraction	88
		4.2.2.1 Building the Initial Subsets	88
		4.2.2.2 Pseudo Rocchio Relevance Feedback	89
		4.2.3 Annotated Term Frequency-Inverse Document Frequency Technique for Pre-encryption Features Extraction	92
	4.3	The Experimental Results	97
		4.3.1 The Experimental Enivironment Setup	97

	4.3.2	The Dataset	97
	4.3.3	Experimental Results of Dynamic Pre-encryption Boundary Definition Technique	98
	4.3.4	Experimental Results of Annotated Term Frequency-Inverse Document Frequency	103
4.4		Discussion	107
	4.4.1	The Dynamic Pre-encryption Boundary Definition Technique	108
	4.4.2	Annotated Term Frequency-Inverse Document Frequency Technique	107
4.5		Summary	110

CHAPTER 5 REDUNDANCY COEFFICIENT GRADUAL UPWEIGHTING-BASED MUTUAL INFORMATION FEATURES SELECTION TECHNIQUE **113**

5.1		Introduction	113
5.2		Overview on Redundancy Coefficient Gradual Upweighting-Based Mutual Information Features Selection	114
	5.2.1	Enhanced Mutual Information Features Selection using Redundancy Coefficient Gradual Upweighting Technique	116
	5.2.2	Maximum of Minimum-Based Enhanced Mutual Information Features Selection Technique	120
5.3		The Experimental Results	124
	5.3.1	The Dataset and Experimental Enviornment	124
	5.3.2	Experimental Results of Enhanced Mutual Information Features Selection Technique	124
	5.3.3	Experimental Results of Maximum of Minimum-Based Mutual Information Features Selection Technique	127
	5.3.4	Comparison With the Related Techniques	129
5.4		Discussion	136
	5.4.1	The Enhanced Mutual Information Features Selection Technique	137

5.4.2 Maximum of Mininmum-Based Enhanced Mutual Information Features Selection Technique 138

5.5 Summary 140

CHAPTER 6 AN ENSEMBLE-BASED ANOMALY-BEHAVIORAL CRYPTO-RANSOMWARE PRE-ENCRYPTION DETECTION MODEL 143

6.1 Introduction 143

6.2 Description of the Anomaly-Behavioral Crypto-ransomware Early Detection Model 143

6.2.1 Incremental Bagging Technique 146

6.2.2 Behavioral Crypto-ransomware Early Detection Module 150

6.2.3 Anomaly-Behavioral Detection Model 152

6.2.4 Enhanced Anomaly-Behavioural Detection Model 153

6.2.4.1 Enhanced Semi-Random Subsapce Selection Technique 153

6.2.4.2 Crypto-ransomware Detection 155

6.3 The Experimental Results 158

6.3.1 Experimental Enivironmental Setup 158

6.3.2 Detection Model Training 158

6.3.3 Experimental Results of Behavioural Crypto-ransomware Early Detection Model 159

6.3.4 Experimental Results of the Anomaly-Behavioural Detection Model 161

6.3.5 Experimental Results of the Enhanced Anomaly-Behavioural Detection Model 163

6.4 Discussion 168

6.4.1 Behavioural Crypto-ransomware Early Detection Module 168

6.4.2 Anomaly-Behavioural Detection Model 170

6.4.3 Enhanced Anomaly-Behavioural Detection Model 171

6.5 Summary 172

CHAPTER 7 CONCLUSION **175**

7.1 Introduction 175

7.2 Problems and Objectives Revisited 175

7.3 Research Contributions 177

7.3.1 Dynamic Pre-encryption Boundary Definition 177

7.3.2 Annotated Term Frequency-Inverse Document Frequency 178

7.3.3 Redundancy Coefficient Gradual Upweighting Technique 178

7.3.4 Incremental Bagging Technique 179

7.3.5 Enhanced Semi-Random Subspace Selection Technique 179

7.4 Recommendations for Future Work 180

7.5 Closing Notes 181

REFERENCES **183**

LIST OF TABLES

TABLE NO.	TITLE	PAGE
Table 2.1	Type of data used in the analysis	36
Table 2.2	Crypto-ransomware families used by this study	43
Table 2.3	Limitations of existing crypto-ransomware early detection solutions	68
Table 3.1	Summary of problem situation and solution concept	72
Table 3.2	Crypto-ransomware families used by this study	74
Table 3.3	Research plan	79
Table 4.1	The pre-encryption boundary vector	98
Table 4.2	Experimental evaluation results of the different classifiers trained using the pre-encryption dataset	100
Table 4.3	Experimental results of the aTF-IDF on pre-encryption dataset (extracted using the proposed DPBD technique in the previous section) with different classifiers	104
Table 4.4	Significance test (t-test) results between the proposed aTF-IDF technique and TF-IDF technique with different classification algorithms	107
Table 5.1	Experimental results of the EMIFS on pre-encryption dataset (DS-pre) with different sizes of features sets used to train several classifiers	125
Table 5.2	Experimental results of the EMIFS on DS1 with different sizes of features sets used to train several classifiers	126
Table 5.3	Experimental results of the EMIFS on DS2 with different sizes of features sets used to train several classifiers	126
Table 5.4	Experimental results of the EMIFS on DS3 with different sizes of features sets used to train several classifiers	126
Table 5.5	Experimental results of the MM-EMIF on pre-encryption dataset (DS-pre) with different sizes of features sets used to train several classifiers	128
Table 5.6	Experimental results of the MM-EMIF on DS1 with different sizes of features sets used to train several classifiers	128

Table 5.7	Experimental results of the MM-EMIF on DS2 with different sizes of features sets used to train several classifiers	128
Table 5.8	Experimental results of the MM-EMIF on DS3 with different sizes of features sets used to train several classifiers	129
Table 5.9	Significance test (t-test) results between the proposed EMIFS technique and the related techniques with different classification algorithms	135
Table 5.10	Significance test (t-test) results between the proposed MM-EMIFS technique and the related techniques with different classification algorithms	135
Table 6.1	The accuracy (ACC), F1, precision, detection rate (DR) and false positive rate (FPR) results of BCED using different percentages of API calls	160
Table 6.2	The significance test (t-test) of RS against SS in terms of accuracy (ACC), F1, precision, detection rate (DR) and false positive rate (FPR)	161
Table 6.3	The accuracy (ACC), F1, precision, detection rate (DR) and false positive rate (FPR) results of ABDM model using different percentages of API calls	162
Table 6.4	The significance test (t-test) of ABDM against BCED in terms of accuracy (ACC), F1, precision, detection rate (DR) and false positive rate (FPR)	163
Table 6.5	The accuracy (ACC), F1, precision, detection rate (DR) and false positive rate (FPR) results of EABDM model using different proportions of API calls	164
Table 6.6	The significance test (t-test) of EABDM against ABDM in terms of accuracy (ACC), F1, precision, detection rate (DR) and false positive rate (FPR)	165

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
Figure 1.1	Scenario describing the problem	5
Figure 1.2	Research phases	13
Figure 2.1	The number of users attacked by ransomware	20
Figure 2.2	Crypto-Ransomware vs. Locker-Ransomware attacks in the period between Jul-2014 and Jun-2015	21
Figure 2.3	Novel ransomware attacks between 2010 and 2017	21
Figure 2.4	Ransomware articles between 1994 and 2018	23
Figure 2.5	Ransomware attacks phases	26
Figure 2.6	Ransomware types	29
Figure 2.7	Existing research in crypto-ransomware	32
Figure 2.8	Generic design of crypto-ransomware early detection models	54
Figure 2.9	Cause-Effect diagram illustrating the limitations in existing crypto-ransomware early detection solutions	62
Figure 3.1	Overall model architecture	75
Figure 4.1	The general architecture of the proposed Dynamic Pre-encryption Boundary Definition and Features Extraction (DPBD-FE) scheme	87
Figure 4.2	The Dynamic Pre-encryption Boundary Definition (DPBD) technique	91
Figure 4.3	The pseudo code of the proposed Pre-encryption Boundary Definition and Features Extraction scheme	92
Figure 4.4	The Pseudo code of the proposed aTF-IDF technique	95
Figure 4.5	Flow chart of the proposed aTF-IDF technique	96
Figure 4.6	Comparison between the proposed DPBD with related techniques	99
Figure 4.7	The comparison of the classification accuracy between the dataset extracted using proposed DPBD (DSPre) and datasets of the related works DS1, DS2 and DS3	101

Figure 4.8	The comparison of the classification precision between the dataset extracted using proposed DPBD (DSPre) and datasets of the related works DS1, DS2 and DS3	101
Figure 4.9	The comparison of the classification recall between the dataset extracted using proposed DPBD (DSPre) and datasets of the related works DS1, DS2 and DS3	102
Figure 4.10	The comparison of the roc_auc between the dataset extracted using proposed DPBD and datasets of the related works DS, DS2 and DS3	102
Figure 4.11	The comparison of the F1 measure between the dataset extracted using proposed DPBD and datasets of the related works DS, DS2 and DS3	103
Figure 4.12	The comparison of the detection accuracy between the proposed aTF- IDF and TF-IDF	105
Figure 4.13	The comparison of F1 measure between the proposed aTF-IDF and TF- IDF	105
Figure 4.14	The comparison of the detection precision between the proposed aTF- IDF and TF-IDF	106
Figure 4.15	The comparison of the detection recall between the proposed aTF-IDF and TF-IDF	106
Figure 4.16	The comparison of the roc_auc between the proposed aTF-IDF and TF- IDF	107
Figure 5.1	The general architecture of the proposed Redundancy Coefficient Gradual Upweighting-Based Mutual Information Features Selection Technique	115
Figure 5.2	Pseudo code of the proposed EMIFS technique	119
Figure 5.3	EMIFS Technique	120
Figure 5.4	The proposed MM-EMIFS technique	122
Figure 5.5	Pseudo code of MM-EMIFS technique	123
Figure 5.6	The comparison between proposed EMIFS, MM-EMIFS and the related techniques using the pre-encryption dataset (DR-pre)	130
Figure 5.7	Detection accuracy comparison of the proposed MM-EMIFS with MIFS and mRMR implemented using the maximum of minimum approximation approach	133
Figure 5.8	The comparison of the detection accuracy between the proposed EMIFS and MM-EMIFS with the related techniques using different crypto-ransomware datasets (DS1, DS2 and DS3)	136

Figure 6.1	The general architecture of the proposed Ensemble-Based Anomaly-Behavioural Crypto-ransomware Pre-encryption Detection model	146
Figure 6.2	The flow chart of iBagging technique	148
Figure 6.3	The pseudo code of iBagging technique	149
Figure 6.4	The components of the behavioral module	151
Figure 6.5	The pseudo code of ESRS technique	154
Figure 6.6	Enhanced Semi-Random Subspace Selection (ESRS) technique	154
Figure 6.7	The process of crypto-ransomware detection	155
Figure 6.8	Pseudo code of the process of crypto-ransomware detection process	156
Figure 6.9	The anomaly-behavioral crypto-ransomware early detection model	157
Figure 6.10	Average comparison of accuracy, F1, precision, DR and FPR between subspace selection (SS) technique employed by BCED and random selection (RS) technique	161
Figure 6.11	Comparison of the averaged accuracy, F1, precision, DR and FPR between ABDM and BCED	163
Figure 6.12	A comparison of averaged accuracy, F1, precision, DR and FPR between SS, SRS and ESRS	165
Figure 6.13	A comparison of averaged accuracy, F1, precision, DR and FPR between ABDM and EABDM	165
Figure 6.14	A comparison of accuracy (a), F1 (b), precision (c), DR (d) and FPR (e) between EABDM, BCED and related techniques	166

LIST OF ABBREVIATIONS

ABDM	-	Anomaly Behavioural Detection Model
API	-	Application Program Interface
aTF-IDF	-	Annotated Term Frequency-Inverse Document Frequency
AV	-	Anti-Virus
BCEDM	-	Behavioural Crypto-ransomware Early Detection Model
DPBD	-	Dynamic Pre-encryption Boundary Definition
DPBD-FE	-	Dynamic Pre-encryption Boundary Definition and Feature Extraction Scheme
C&C		Command and Control
DT	-	Decision Tree
EMIFS	-	Enhanced Mutual Information Feature Selection
ESRS	-	Enhanced Semi-Random Subspace
FBI	-	Federal Bauru of Investigation
FPR	-	False Positive Rate
iBagging	-	Incremental Bagging
IR		Improvement Ratio
LR	-	Logistic Regression
MIFS	-	Mutual Information Feature Selection
ML	-	Machine Learning
MLP	-	Multi-layer Perceptron
MM-EMIFS	-	Maximum of Minimum Enhanced Mutual Information Feature Selection
RaaS		Ransomware-as-a-Service
RF	-	Random Forests
RSS	-	Random Subspace Selection
SRS	-	Semi-Random Subspace
SS	-	Subspace Selection
SVM	-	Support Vector Machine
TF	-	Term Frequency
TF-IDF	-	Term Frequency-Inverse Document Frequency

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	List of Part of the Hashes of Crypto-ransomware Samples Used by This Research	198
B	Sample OF Crypto-ransomware Trace File	202
C	Detailed comparison results of the proposed Dynamic Pre-encryption Boundary Definition and Features Selection scheme	203
D	The Comparison of the Detection Accuracy Between the Proposed EMIFS and MM-EMIFS with the Related Techniques using Different Crypto-ransomware Datasets (DS1, DS2 and DS3)	205
E	Part of Subspace Built using Subspace Selection	210
F	List of Publications	211

CHAPTER 1

INTRODUCTION

1.1 Overview

The rapid proliferation of internet technologies and online services is accompanied with several cybersecurity concerns that impedes the momentum of such technologies and obstructs the full integration of those services into people's daily life and business. Malicious software; also called malware; is one of those concerns that compromise the confidentiality, integrity and availability of the data in the computer systems (Nong *et al.*, 2004). Since its first occurrence on early 1970s, several types of malware have been witnessed in the wild such as Viruses, Worms, Trojans, Spyware and Ransomware. Ransomware is a malware category that locks user data and files and demands ransom to release them (Azmoodeh *et al.*, 2017; Yalew *et al.*, 2017; Yaqoob *et al.*, 2017; Chen *et al.*, 2018).

Ransomware history dates back to the late 1980s when the first sample called AIDS was released. Since then, ransomware has become a major threat that intimidates the accessibility to user and business data (Gomez-Hernandez *et al.*, 2018). By creating ransomware, the attackers have introduced the extortion concept into cyberspace (Caporusso *et al.*, 2019). Due to the monetary motivation, adversaries have been tempted to develop many variants of ransomware which explains the dominance of ransomware in the threat landscape recently (Homayoun *et al.*, 2017; Cusack *et al.*, 2018; Hampton *et al.*, 2018; Kao and Hsiao, 2018).

Not only are individuals targeted by ransomware attacks, but also business and governmental institutions (Cohen and Nissim, 2018). In 2014, the attackers earned around $3 million through ransomware attacks (Homayoun *et al.*, 2017). According to the reports, $352 million were paid by victims around the world in 2015 to the attackers in order to unlock their data (Cohen and Nissim, 2018). In 2016, up to $220K was spent in Indiana county only to recover from ransomware attacks (Cohen and Nissim, 2018). Inability to access data is not the only ramification that ransomware victims incur, the damage could also include downtime costs, loss of money and reputation (Azmoodeh *et al.*, 2017).

There are two types of ransomware, namely locker-ransomware and crypto-ransomware (Cohen and Nissim, 2018; Gomez-Hernandez *et al.*, 2018). While the former locks the user's device and/or resources, the latter employs the cryptography mechanism of the underlying operating system to encrypt user-related data and files (Chen *et al.*, 2018; Gonzalez and Hayajneh, 2018). Contrary to locker-ransomware attacks whose effect can easily be mitigated, the effect of crypto-ransomware attacks persist even after detection and removal and; in many cases; the victim has no choice but to pay the ransom in order to get the decryption key (Gomez-Hernandez *et al.*, 2018). With the help of Ransomware-as-a-Service (RaaS), cryptography, and the difficult-to-trace cyber-currency technologies like Bitcoin, it becomes easy and feasible for even novice attackers to develop and distribute their own crypto-ransomware (Gomez-Hernandez *et al.*, 2018; Moussaileb *et al.*, 2018). Consequently, the rate of crypto-ransomware attacks has increased dramatically in recent years (Kharraz *et al.*, 2015; Everett, 2016; Kharraz *et al.*, 2016).

Two main characteristics distinguish crypto-ransomware from other types of malware, namely the benign-alike behaviour and the irreversible nature of the attack (Scaife *et al.*, 2016; Sgandurra *et al.*, 2016; Kharraz *et al.*, 2018; Lokuketagoda *et al.*, 2018). By targeting user-related files using the system legitimate cryptography applications and APIs, the behaviour of crypto-ransomware resembles the behaviour of benign programs. Similarly, the employment of cryptography leaves the targeted files inaccessible even after detecting and removing the causing crypto-ransomware. Once crypto-ransomware encrypts the targeted resource, it is difficult to regain the access without holding the decryption key (Homayoun *et al.*, 2017; Cabaj *et al.*, 2018; Chen *et al.*, 2018). Such irreversibility entails the early detection to effectively confront crypto-ransomware attacks (Homayoun *et al.*, 2017; Yaqoob *et al.*, 2017; Gomez-Hernandez *et al.*, 2018; Rhode *et al.*, 2018).

The goal of this study is to propose an enhanced early detection solution able to detect crypto-ransomware attacks at the early phases of their execution lifecycle. To be effective, it is imperative that such detection takes place early before the encryption is carried out (Gomez-Hernandez *et al.*, 2018). Such period can be referred to as pre-encryption phase which begins from the moment when crypto-ransomware starts installing itself in the victim's machine and lasts until the first call of any of cryptography-related APIs.

1.2 Problem Background

Several studies have been conducted to detect crypto-ransomware attacks. These studies could be categorized into data-centric and process-centric approaches. Data-centric approach monitors user data and files subjected to attack and raises the alarm when it discovers a suspicious change in those files. Several techniques such as decoy technique, entropy and similarity measures were employed to monitor the file structure before and after it gets accessed (Kharraz *et al.*, 2016; Mbol *et al.*, 2016; Shahriari, 2016; Song *et al.*, 2016; Gomez-Hernandez *et al.*, 2018). However, this approach does not distinguish between the changes that have been carried out by benign programs from those caused by crypto-ransomware, which lead to high rate of false alarms (Scaife *et al.*, 2016; Morato *et al.*, 2018; Moussaileb *et al.*, 2018). More importantly, this approach does not fully protect user data from being held to ransom as it sacrifices part of the data; which could be more valuable to victim than the remaining data; before detection (Scaife *et al.*, 2016; Sotelo Monge *et al.*, 2018). Thus, data-centric approach is not effective for crypto-ransomware early detection.

Process-centric approach monitors the behaviour of the running process so as to discover the suspicious patterns. Several studies like Shahriari (2016); Chen and Bridges (2017); Chen *et al.* (2017); Cohen and Nissim (2018) have employed such approach and acquired different types of behavioural data by which, machine learning classifiers like Random Forests and Naïve Bayes have been trained. However, most of those studies follow malware detection approach that depends on the entire runtime data; which include pre-encryption and post-encryption data; to detect the attacks

(Mehnaz *et al.*, 2018; Rhode *et al.*, 2018). Such approach assumes the availability of the entire data at detection time (Rhode *et al.*, 2018). Thus, they are not suitable for crypto-ransomware early detection where the data of the instance in question are not fully available.

Monitoring the computational resources used and/or dealt with by ransomware processes is another type of process-centric approach. That is, one or more resources in the user machine like CPU, network, I/O buffer and memory are observed, and the alarm raises when some events related to ransomware and/or cryptography were encountered. Maltester is one of those solutions proposed by Cabaj *et al.* (2015) to detect the infection chain of Cryptowall ransomware family via introspecting the network traffic. Likewise, Cabaj *et al.* (2018); Cusack *et al.* (2018) have proposed detection solutions based on monitoring the network traffic between the infected devices and ransomware's command and control (C&C) server. In their study, Kharraz *et al.* (2016) proposed UNVEIL that observes I/O access patterns and file system activities. Similarly, Song *et al.* (2016) put forward a model that monitors CPU, I/O and device's memory in order to detect the suspicious activities caused by ransomware. However, the reliance on ad-hoc events leads to high rate of false alarms as those events are not mutually exclusive to crypto-ransomware and some benign programs raise such events as well (Morato *et al.*, 2018). Additionally, those events could happen after the encryption takes place, which renders this approach ineffective for the early detection (Kharraz *et al.*, 2016). To be effective, it is essential that the detection takes place during early phases before the attack starts the main sabotage, which is the encryption in crypto-ransomware's case.

To early detect crypto-ransomware attacks effectively, the detection solutions need to be able to accurately identify known and novel attacks on time, i.e. before the encryption takes place (Sgandurra *et al.*, 2016; Homayoun *et al.*, 2017; Gomez-Hernandez *et al.*, 2018; Homayoun *et al.*, 2019). This could be achieved by focusing on the pre-encryption phase, i.e. the phase in the crypto-ransomware lifecycle that precede the encryption's starting point. However, detecting crypto-ransomware at early phases of its attack is challenging (Alam *et al.*, 2018). Several factors contribute to such challenge including the static definition of the pre-encryption phase boundary, the insufficient information about the attack at this early phase, the high dimensional data and the inability to detect novel (zero-day) attacks (Das *et al.*, 2016a; Morato *et al.*, 2018; Nissim *et al.*, 2018; Rhode *et al.*, 2018). Figure 1.1 summarizes the challenges of the existing works along with the current status, gaps and desired solutions.

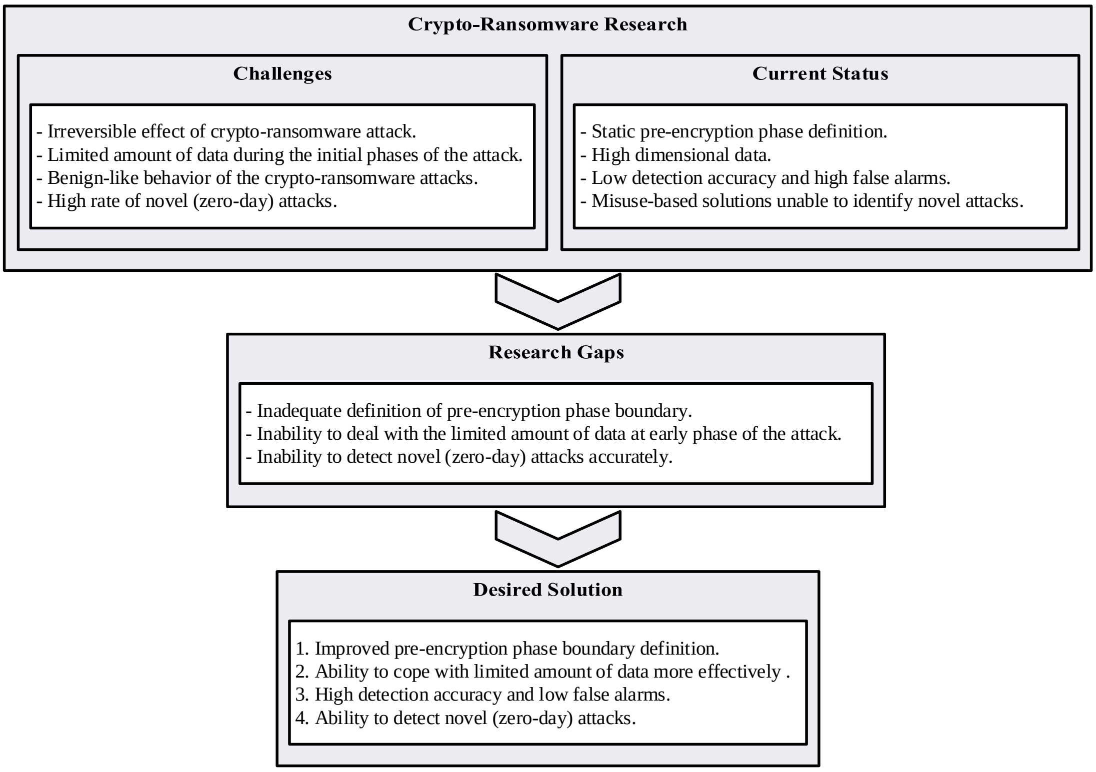

Figure 1.1: Scenario describing the problem

For detection model to carry out the early detection, it needs to be trained on the data that represent the early phases of the attacks' lifecycle. The idea of building detection models using the early data extracted during the onset of crypto-ransomware attacks was introduced by Sgandurra *et al.* (2016). To define the amount of data required, authors proposed fixed time-based thresholding by which, the data captured during the first 30 seconds of ransomware instance runtime were collected and used to build an early detection model. Likewise, Homayoun *et al.* (2017) and Rhode *et al.* (2018) used the same approach but with decreasing the threshold into 10 seconds and 1 second respectively. However, the fixed time-based thresholding implies that all instances start the encryption before the specified time. This does not hold for many crypto-ransomware attacks as the time for the main sabotage to start varies among different instances due to the obfuscation techniques employed by those instances, which create different attack behaviours (Das *et al.*, 2016a; Kharraz *et al.*, 2016; Nissim *et al.*, 2018). Therefore, the fixed thresholding could miss the encryption starting point and; consequently; the captured data would not accurately represent the pre-encryption phase of crypto-ransomware attacks, which adversely affects the ability of detection solutions to identify the attacks before the encryption takes place. As such, more accurate pre-encryption boundary definition approach that can cope with the dynamic nature of crypto-ransomware behaviour is needed.

The small amount of data captured during the initial phases of the attack is one of the issues that early detection solutions face, which causes poor detection accuracy (Rhode *et al.*, 2018). This issue exacerbates with high dimensional feature space caused by features extraction methods like n-gram adopted by most of detection solutions (Peng *et al.*, 2016; Sgandurra *et al.*, 2016; Ye *et al.*, 2017; Stiborek *et al.*, 2018a). Such high dimensional data renders the model prone to overfitting, which degrades the detection accuracy (Reineking, 2016; Fallahpour *et al.*, 2017; Li *et al.*, 2017). Several features selection approaches could be used to address this issue including similarity-based, statistical-based, sparse-learning-based and information theory-based techniques (Fallahpour *et al.*, 2017; Li *et al.*, 2017). Characterized by having no assumption about the distribution of the underlying data, information theory-based features selection techniques have been utilized by several malware and ransomware detection solutions as well as many other selection tasks (Liu *et al.*, 2009; Sgandurra *et al.*, 2016; Wang *et al.*, 2017b; Ye *et al.*, 2017). These techniques try to enhance a trade-off between the relevancy and redundancy terms by adjusting the values of redundancy coefficients (Brown *et al.*, 2012; Li *et al.*, 2017). Those coefficients are adjusted either statically or dynamically (Battiti, 1994; Yang and Moody, 1999; Hanchuan *et al.*, 2005; Brown *et al.*, 2012; Che *et al.*, 2017). Nevertheless, selecting a static value for those parameters is difficult and need to be set experimentally (Brown *et al.*, 2012; Che *et al.*, 2017). On the other hand, the dynamic adjustment of these coefficients changes the belief in the redundancy term at each iteration inversely proportional to the current size of the selected features set (Brown *et al.*, 2012). While this approach is suitable for data with full observations of the attacks' patterns, it hinders the ability of goal function to estimate features significance accurately when dealing with only small portion of data that contain limited amount of observed attacks' patterns (Bennasar *et al.*, 2015; Che *et al.*, 2017). Consequently, the selected set could include redundant and irrelevant features given the limited amount of attack patterns as it is the case in the early detection where the entire characteristics of the attack have not been observed yet (Das *et al.*, 2016a; Che *et al.*, 2017). Therefore, an improved technique that can overcome the limitation in pre-encryption data and estimate features significance more accurately is needed.

The lack of enough data at the initial phases of the attack also adversely affects the accuracy of the detection solutions (Das *et al.*, 2016a; Nissim *et al.*, 2018; Rhode *et al.*, 2018). That is, incomplete observations lead to sparse data with which, weak classifiers are created (Wei *et al.*, 2017; Ryu *et al.*, 2018). In addition, existing solutions were built based on the premise that the data required for the detection is complete and ready to use at detection time, which does not hold for the early detection tasks while the attack is underway and the data are not fully available (Das *et al.*, 2016a). Furthermore, the design of those solutions does not reflect the progression of crypto-ransomware behaviour during the time of the attack, which renders those solutions unable to early detect the attacks accurately (Alrawashdeh and Purdy, 2016; Das *et al.*, 2016a). Some studies have employed ensemble learning to overcome such weakness and boost the detection accuracy (Homayoun *et al.*, 2017; Rhode *et al.*, 2018). It turned out that the accuracy of ensemble-based models rely on the accuracy of the individual components of the ensemble (base classifiers) and the diversity among those components (Mao *et al.*, 2017). However, the random sampling employed by the ensemble techniques to consolidate the diversity might produce weak base classifiers based on suboptimal subspaces with many noisy and irrelevant features which, consequently, degrades the overall accuracy of the ensemble (Yang *et al.*, 2010; Aburomman and Reaz, 2017; Koziarski *et al.*, 2017). As such, an enhanced ensemble-based model that builds the data subsets in a way that reflects the attack progression as well as improves the diversity-relevancy trade-off among its different components is needed.

The monetary motivation increased the rate of novel (zero-day) crypto-ransomware attacks, which explains the dominance of ransomware in the threat landscape recently (Ahmadian and Shahriari, 2016; Kaspersky, 2016; Symantec, 2016a; Homayoun *et al.*, 2017; Cusack *et al.*, 2018; Gomez-Hernandez *et al.*, 2018; Hampton *et al.*, 2018; Kao and Hsiao, 2018). The inability to identify novel (zero-day) attacks is one of the main limitations of most of existing crypto-ransomware detection solutions (Cohen and Nissim, 2018). Existing crypto-ransomware early detection solutions are misuse-based (Sgandurra *et al.*, 2016; Homayoun *et al.*, 2017). As such, these solutions are deemed ineffective as they are not able to detect the previously unseen attacks due to the reliance on pre-defined signatures extracted from known crypto-ransomware instances statically (structural-based) or dynamically (behavioural-based) (Mercaldo *et al.*, 2016; Morato *et al.*, 2018; Homayoun *et al.*, 2019). Although the behavioural detection approaches can detect the variants with common known signature, they are unable to detect those whose signatures are not previously seen (Liao *et al.*, 2013; Creech and Hu, 2014; Gandotra *et al.*, 2014; Ganame *et al.*, 2017; Turaev *et al.*, 2018). Thus, adopting the anomaly detection approach is needed to overcome such limitation.

1.3 Problem Statement

Detecting crypto-ransomware at early phases of its attack is challenging due to several issues that render existing solutions not effective. The first issue is that these solutions employ fixed time-based thresholding to define the pre-encryption phase boundary, which is not suitable given the dynamic nature of crypto-ransomware behaviour. As such, the static threshold could miss the encryption's starting point for many instances. Consequently, the captured data do not accurately represent the pre-encryption phase of crypto-ransomware attacks. The second issue is the limited amount of data observed at the early phases of the attack. The impact of this issue is twofold. On the one hand, it obstructs the accurate estimation of features significance during the feature selection process, which leads to the inclusion of many redundant and irrelevant features in the selected set. On the other hand, it provides the detection model with incomplete patterns, which hinders the ability of existing solutions to detect the attacks accurately. The third issue is that the existing solutions are unable to identify novel (zero-day) attacks accurately due to the misuse nature that those solutions have been built based on.

The focus of this research is to clearly define the boundary of pre-encryption phase of crypto-ransomware lifecycle from which, the related features are extracted and selected based on the small portion of the data available at this phase. Such data and features are then used to build a model able to early detect the known and novel attacks more effectively with high accuracy and low false alarms. The research hypothesis is stated as follows.

The effectiveness of crypto-ransomware early detection can be improved by dynamically defining the pre-encryption phase of the attack from which, the data and features are extracted and selected; and used to derive incremental subsets by which, the design of detection model is improved to compensate the lack of enough data at early phases of the attack's lifecycle, which in turns increases the detection accuracy and decreases the false alarms.

To prove the research hypothesis, the following are the research questions that will be addressed:

(i) How to extract the features related to the dynamic pre-encryption phase of crypto-ransomware attacks?

(ii) What is the suitable technique to estimate the features significance given the limitation in the data observed during the pre-encryption phase of crypto-ransomware attacks such that, only relevant and non-redundant features can be selected, and data dimensionality can be reduced?

(iii) How to build an early detection model that overcomes the data limitation at the early phases of crypt-ransomware attacks and accurately detect novel and known attacks early?

1.4 Research Aim

The aim of this research is to propose and develop an Ensemble-Based Anomaly-Behavioural Crypto-ransomware Pre-encryption Detection model, which dynamically defines the boundary of pre-encryption phase of the attacks from which, the data and features are extracted and selected, and used to train an enhanced anomaly-behavioural ensemble-based model able to early detect novel and known attacks more accurately.

1.5 Research Objectives

The objectives of the research are:

(i) To propose an enhanced feature extraction scheme, by integrating a dynamic thresholding-based boundary definition technique with an annotation-based features extraction technique, in order to improve the boundary definition of pre-encryption phase of crypto-ransomware attacks and extract its related features which increases the detection accuracy.

(ii) To propose an enhanced feature selection technique, by integrating an improved redundancy term calculation technique into the goal function, in order to enhance features significance estimation and filter out the redundant/irrelevant features, which reduces the data dimensionality and increases the detection accuracy.

(iii) To develop an anomaly-behavioural early detection model, by incorporating the techniques proposed in (i) and (ii) into an enhanced anomaly-misuse-based ensemble, in order to compensate the limitation of the pre-encryption data, which increases the detection accuracy of novel and known attacks and reduces the false alarms.

1.6 Research Scope

This research study is limited to the following:

(i) Crypt-ransomware samples used for the research were acquired from http://www.viresshare.com, which is a public malware repository used by many researchers (Sgandurra *et al.*, 2016; Chen *et al.*, 2017; Hasan and Rahman, 2017; Lu *et al.*, 2017; Rhode *et al.*, 2018; Turaev *et al.*, 2018). The samples are current (at the time of this research) and could be found in the wild.

(ii) Ground truth data were obtained through VirusTotal (http://www.virustotal.com/), which provides scan results for all malware categories including crypto-ransomware. Similar to related works, more than 55 different Anti-Virus (AV) engines were involved in such scan (Wang *et al.*, 2018; Zimba *et al.*, 2018b; Zhang *et al.*, 2019).

(iii) In this research, Windows X86's benign/malicious programs were utilized to conduct the experiments (Shashidhar, 2017; Hampton *et al.*, 2018; Rhode *et al.*, 2018).

(iv) This research used crypto-ransomware samples that leverage API calls to carry out their attacks and leaves an artefact in the trace file, as this is the common approach for detecting ransomware as well as malware attacks (Ki *et al.*, 2015; Sgandurra *et al.*, 2016; Hampton *et al.*, 2018; Jung and Won, 2018; Moussaileb *et al.*, 2018; Zimba *et al.*, 2018a).

(v) The dynamic analysis was carried out in a Cuckoo Sandbox analysis platform, as it is one of the most popular analysis platforms used in malicious code analysis studies including those related to crypto-ransomware (Sgandurra *et al.*, 2016; Maniath *et al.*, 2017; Genç *et al.*, 2018; Rhode *et al.*, 2018; Popli and Girdhar, 2019).

(vi) This study includes neither remedial nor response actions to the detection of crypto-ransomware.

1.7 Significance of the Research

The research is important and significant as it addresses several real-world problems in the field of malicious programs research. Among those addressed are:

(i) The irreversible effect of encryption employed by crypto-ransomware families renders it imperative to detect such attacks before they carry out the encryption.

(ii) The number of crypto-ransomware released is ever increasing that entails having an accurate zero-day attack detection mechanism.

(iii) The study advances the body of knowledge in cybersecurity by introducing techniques that can cope with the lack of enough data at early phases of the attacks. This might as well be useful for future research into confronting not only ransomware but also other similar attacks at the early phases, which consolidates the protection of both personal and business data.

(iv) The proposed model contributes to advance the knowledge by introducing the dynamic pre-encryption boundary definition, redundancy gradual coefficient upweighting and incremental bagging concepts into ensemble learning techniques which further enhances the classification performance especially in the absence of enough information as it is the case in the ransomware early detection tasks.

1.8 Research Methodology

The research methodology is described in detail in Chapter 3. The proposed approach for this study includes three phases as shown in Figure 1.3. In the first phase, the dynamic pre-encryption boundary definition and features extraction scheme is designed and implemented. The second phase proposes and implements the redundancy coefficient gradual up-weighting technique for features selection process. In phase 3, the anomaly-behavioural crypto-ransomware early detection is proposed and implemented based on the data and features prepared in the phases 1 and 2.

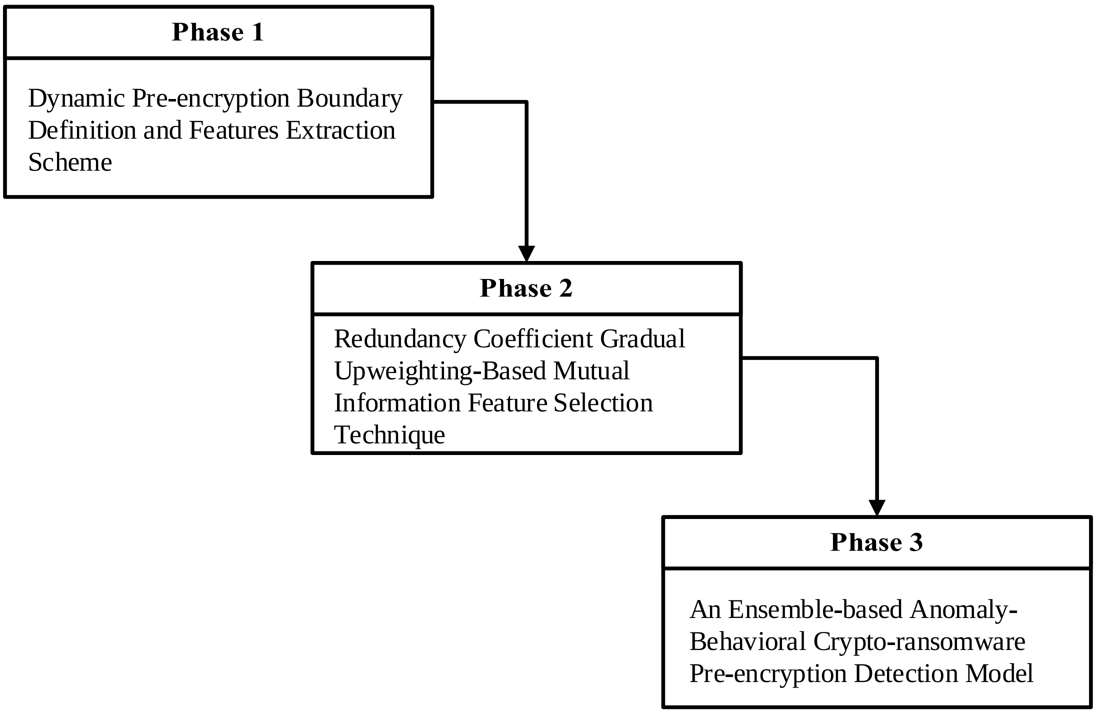

Figure 1.2: Research phases

1.9 Research Contribution

The main contribution of this research is an anomaly-behavioural crypto-ransomware early detection model able to identify the imminent encryption attacks. This contribution has been achieved by several enhancements carried out on the different components of the model as follows.

(i) An effective features extraction scheme able to extract the features relevant to pre-encryption phase of crypto-ransomware lifecycle, which includes:

a. A dynamic pre-encryption boundary definition technique to track the encryption starting points for all crypto-ransomware instances.

b. An annotated term frequency-inverse document frequency technique able to distinguish the general-purpose APIs given the absence of full runtime data.

(ii) A redundancy coefficient gradual up-weighting (RCGU) technique to improve relevancy-redundancy trade-off calculation in the feature selection process.

(iii) An enhanced anomaly-behavioural ensemble-based detection model which includes:

a. An incremental bagging (iBagging) ensemble technique for training data subsets preparation, which compensates the lack of enough data at the early phases of crypto-ransomware attacks.

b. An enhanced semi-random subspace selection (ESRS) technique to improve the diversity among the ensemble's features subspaces while maintaining high relevant features within each subspace.

c. A stack-based hybridization method between anomaly and behavioural detection approaches which improves the zero-day detection accuracy of the entire model without compromising the low false rate of the behavioural module.

1.10 Thesis Organization

In this chapter, the general idea of this research, problem background as well as the problem formalization has been presented along with research questions and objectives. The rest of this thesis is organized as follows.

Chapter 2 provides the theoretical foundation of the research into crypto-ransomware early detection. It provides a comprehensive and thorough investigation to the state-of-the-art solutions in the context of crypto-ransomware early detection. It also summarizes the current research issues and directions. The research methodology adopted by this study is described in Chapter 3. In addition, this chapter elaborates on the research frameworks along with dataset description and evaluation metrics. In Chapter 4, the design and implementation of the dynamic pre-encryption boundary definition and features extraction scheme is discussed. Chapter 5 presents the design and implementation of the redundancy coefficient gradual up-weighting technique. In Chapter 6, the design and implementation of the anomaly-behavioural crypto-ransomware early detection model is elaborated. This thesis is concluded with Chapter 7 which elucidates research objectives revisiting, research findings, contributions and implications and provides suggestions for future work.

REFERENCES

Abaid, Z., Sarkar, D., Kaafar, M. A., and Jha, S. (2016, 7-10 Nov. 2016). *The Early Bird Gets the Botnet: A Markov Chain Based Early Warning System for Botnet Attacks.* Paper presented at the 2016 IEEE 41st Conference on Local Computer Networks (LCN), 61-68.

Aburomman, A. A., and Reaz, M. B. I. (2017). A survey of intrusion detection systems based on ensemble and hybrid classifiers. *Computers & Security, 65*, 135-152.

Aghdam, M. H., and Kabiri, P. (2016). Feature Selection for Intrusion Detection System Using Ant Colony Optimization. *IJ Network Security, 18*(3), 420-432.

Ahmadian, M. M., and Shahriari, H. R. (2016, 7-8 Sept. 2016). *2entFOX: A framework for high survivable ransomwares detection.* Paper presented at the 2016 13th International Iranian Society of Cryptology Conference on Information Security and Cryptology (ISCISC), 79-84.

Ahmadian, M. M., Shahriari, H. R., and Ghaffarian, S. M. (2015). *Connection-monitor & connection-breaker: A novel approach for prevention and detection of high survivable ransomwares.* Paper presented at the 12th International ISC Conference on Information Security and Cryptology, ISCISC 2015, 79-84.

Alam, M., Bhattacharya, S., Mukhopadhyay, D., and Chattopadhyay, A. (2018). RAPPER: Ransomware Prevention via Performance Counters. *arXiv preprint arXiv:1802.03909*.

Alazab, M., Layton, R., Venkataraman, S., and Watters, P. (2010). *Malware detection based on structural and behavioural features of api calls.* Paper presented at the 1st International Cyber Resilience Conference, Edith Cowan University, Perth Western Australia, 1-10.

Alhawi, O. M. K., Baldwin, J., and Dehghantanha, A. (2018). Leveraging machine learning techniques for windows ransomware network traffic detection, *Advances in Information Security* (Vol. 70, pp. 93-106): Springer New York LLC.

Ali, A., Waleed., and Shamsuddin, S. M. (2012). *Intelligent web proxy caching based on supervised machine learning.*

Alrawashdeh, K., and Purdy, C. (2016, 18-20 Dec. 2016). *Toward an Online Anomaly Intrusion Detection System Based on Deep Learning.* Paper presented at the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 195-200.

Andronio, N., Zanero, S., and Maggi, F. (2015). HELDROID: Dissecting and detecting mobile ransomware. In H. Bos, G. Blanc and F. Monrose (Eds.), *18th International Symposium on Research in Attacks, Intrusions, and Defenses, RAID 2015* (Vol. 9404, pp. 382-404): Springer Verlag.

Aragorn, T., YunChun, C., YiHsiang, K., and Tsungnan, L. (2016). Deep Learning for Ransomware Detection. *IEICE Technical Report; IEICE Tech. Rep., 116*(282), 87-92.

Azmoodeh, A., Dehghantanha, A., Conti, M., and Choo, K.-K. R. (2017). Detecting crypto-ransomware in IoT networks based on energy consumption footprint. *Journal of Ambient Intelligence and Humanized Computing, 9*(4), 1141-1152.

Bai, J. R., and Wang, J. F. (2016). Improving malware detection using multi-view ensemble learning. *Security and Communication Networks, 9*(17), 4227-4241.

Banescu, S., Wuchner, T., Salem, A., Guggenmos, M., Ochoa, M., and Pretschner, A. (2015, 20-22 Oct. 2015). *A framework for empirical evaluation of malware detection resilience against behavior obfuscation.* Paper presented at the 2015 10th International Conference on Malicious and Unwanted Software (MALWARE), 40-47.

Battiti, R. (1994). Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks, 5*(4), 537-550.

Belaoued, M., and Mazouzi, S. (2015). A Real-Time PE-Malware Detection System Based on CHI-Square Test and PE-File Features. In A. Amine, L. Bellatreche, Z. Elberrichi, E. J. Neuhold and R. Wrembel (Eds.), *Computer Science and Its Applications: 5th IFIP TC 5 International Conference, CIIA 2015, Saida, Algeria, May 20-21, 2015, Proceedings* (pp. 416-425). Cham: Springer International Publishing.

Bennasar, M., Hicks, Y., and Setchi, R. (2015). Feature selection using Joint Mutual Information Maximisation. *Expert Systems with Applications, 42*(22), 8520-8532.

Bhardwaj, A., Avasthi, V., Sastry, H., and Subrahmanyam, G. (2016). Ransomware Digital Extortion: A Rising New Age Threat. *Indian Journal of Science and Technology, 9*, 14.

Bhardwaj, A., Subrahmanyam, G., Avasthi, V., and Sastry, H. (2015). Ransomware: A Rising Threat of new age Digital Extortion. *arXiv preprint arXiv:1512.01980*.

Biryukov, A., and Pustogarov, I. (2015, 17-21 May 2015). *Bitcoin over Tor isn't a Good Idea.* Paper presented at the 2015 IEEE Symposium on Security and Privacy, 122-134.

Bridges, L. (2008). The changing face of malware. *Network Security, 2008*(1), 17-20.

Brown, G., Pocock, A., Zhao, M. J., and Luján, M. (2012). Conditional likelihood maximisation: A unifying framework for information theoretic feature selection. *Journal of Machine Learning Research, 13*, 27-66.

Cabaj, K., Gawkowski, P., Grochowski, K., and Kosik, A. (2016a, 11-14 Sept. 2016). *Developing malware evaluation infrastructure.* Paper presented at the 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), 981-989.

Cabaj, K., Gawkowski, P., Grochowski, K., and Osojca, D. (2015). Network activity analysis of CryptoWall ransomware. *Przeglad Elektrotechniczny, 91*(11), 201-204.

Cabaj, K., Gregorczyk, M., and Mazurczyk, W. (2016b). Software-Defined Networking-based Crypto Ransomware Detection Using HTTP Traffic Characteristics. *arXiv preprint arXiv:1611.08294*.

Cabaj, K., Gregorczyk, M., and Mazurczyk, W. (2018). Software-defined networking-based crypto ransomware detection using HTTP traffic characteristics. *Computers & Electrical Engineering, 66*, 353-368.

Cabaj, K., and Mazurczyk, W. (2016). Using Software-Defined Networking for Ransomware Mitigation: The Case of CryptoWall. *Ieee Network, 30*(6), 14-20.

Cai, D., Zhang, C., and He, X. (2010). *Unsupervised feature selection for multi-cluster data*. Paper presented at the Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining.

Canali, D., Lanzi, A., Balzarotti, D., Kruegel, C., Christodorescu, M., and Kirda, E. (2012). *A quantitative study of accuracy in system call-based malware* *detection.* Paper presented at the 21st International Symposium on Software Testing and Analysis, ISSTA 2012, Minneapolis, MN, 122-132.

Canfora, G., Iannaccone, A. N., and Visaggio, C. A. (2014). Static analysis for the detection of metamorphic computer viruses using repeated-instructions counting heuristics. *Journal in Computer Virology, 10*(1), 11-27.

Caporusso, N., Chea, S., and Abukhaled, R. (2019). *A Game-Theoretical Model of Ransomware.* Paper presented at the International Conference on Applied Human Factors and Ergonomics, Cham, 69-78.

Che, J., Yang, Y., Li, L., Bai, X., Zhang, S., and Deng, C. (2017). Maximum relevance minimum common redundancy feature selection for nonlinear data. *Information Sciences, 409*(Supplement C), 68-86.

Chen, J., Wang, C. H., Zhao, Z. M., Chen, K., Du, R. Y., and Ahn, G. J. (2018). Uncovering the Face of Android Ransomware: Characterization and Real-Time Detection. *Ieee Transactions on Information Forensics and Security, 13*(5), 1286-1300.

Chen, K., Zhang, Z., Long, J., and Zhang, H. (2016). Turning from TF-IDF to TF-IGM for term weighting in text classification. *Expert Systems with Applications, 66,* 245-260.

Chen, Q., and Bridges, R. A. (2017). Automated Behavioral Analysis of Malware A Case Study of WannaCry Ransomware. *arXiv preprint arXiv:1709.08753*.

Chen, Z.-G., Kang, H.-S., Yin, S.-N., and Kim, S.-R. (2017). *Automatic Ransomware Detection and Analysis Based on Dynamic API Calls Flow Graph*. Paper presented at the Proceedings of the International Conference on Research in Adaptive and Convergent Systems.

Chittooparambil, H. J., Shanmugam, B., Azam, S., Kannoorpatti, K., Jonkman, M., and Samy, G. N. (2019). *A Review of Ransomware Families and Detection Methods*, Cham, 588-597.

Choi, K., Scott, T., and LeClair, D. (2016). Ransomware Against Police: Diagnosis of Risk Factors via Application of Cyber-Routine Activities Theory. *Int J Forensic Sci Pathol, 4*(7), 253-258.

Choudhary, S. P., and Vidyarthi, M. D. (2015). *A Simple Method for Detection of Metamorphic Malware using Dynamic Analysis and Text Mining.* Paper presented at the Procedia Computer Science, 265-270.

Christensen, J. B., and Beuschau, N. (2017). Ransomware detection and mitigation tool.

Cohen, A., and Nissim, N. (2018). Trusted detection of ransomware in a private cloud using machine learning methods leveraging meta-features from volatile memory. *Expert Systems with Applications, 102*, 158-178.

Conti, M., Gangwal, A., and Ruj, S. (2018). On the economic significance of ransomware campaigns: A Bitcoin transactions perspective. *Computers & Security, 79*, 162-189.

Continella, A., GUAGNELLI, A., ZINGARO, G., DE PASQUALE, G., BARENGHI, A., ZANERO, S., et al. (2016). ShieldFS: The Last Word In Ransomware Resilient Filesystems.

Creech, G., and Hu, J. (2014). A Semantic Approach to Host-Based Intrusion Detection Systems Using Contiguousand Discontiguous System Call Patterns. *IEEE Transactions on Computers, 63*(4), 807-819.

Cusack, G., Michel, O., and Keller, E. (2018). *Machine Learning-Based Detection of Ransomware Using SDN.* Paper presented at the Proceedings of the 2018 ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization.

Daku, H., Zavarsky, P., and Malik, Y. (2018, 1-3 Aug. 2018). *Behavioral-Based Classification and Identification of Ransomware Variants Using Machine Learning.* Paper presented at the 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 1560-1564.

Damodaran, A., Troia, F. D., Visaggio, C. A., Austin, T. H., and Stamp, M. (2015). A comparison of static, dynamic, and hybrid analysis for malware detection. *Journal of Computer Virology and Hacking Techniques*, 1-12.

Das, S., Liu, Y., Zhang, W., and Chandramohan, M. (2016a). Semantics-Based Online Malware Detection: Towards Efficient Real-Time Protection Against Malware. *Ieee Transactions on Information Forensics and Security, 11*(2), 289-302.

Das, S., Xiao, H., Liu, Y., and Zhang, W. (2016b). *Online malware defense using attack behavior model.* Paper presented at the Circuits and Systems (ISCAS), 2016 IEEE International Symposium on, 1322-1325.

del Rey, A. M. (2015). Mathematical modeling of the propagation of malware: A review. *Security and Communication Networks, 8*(15), 2561-2579.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*: John Wiley & Sons.

Egele, M., Scholte, T., Kirda, E., and Kruegel, C. (2012). A survey on automated dynamic malware-analysis techniques and tools. *ACM Computing Surveys (CSUR), 44*(2), 6.

Everett, C. (2016). Ransomware: To pay or not to pay? *Computer Fraud and Security, 2016*(4), 8-12.

Fallahpour, S., Lakvan, E. N., and Zadeh, M. H. (2017). Using an ensemble classifier based on sequential floating forward selection for financial distress prediction problem. *Journal of Retailing and Consumer Services, 34*, 159-167.

Fan, Y., Ye, Y., and Chen, L. (2016). Malicious sequential pattern mining for automatic malware detection. *Expert Systems with Applications, 52*, 16-25.

FBI. (2015). Criminals Continue To Defraud And Extort Funds From Victims Using Cryptowall Ransomware Schemes. *Public Service Anouncements* Retrieved 09-11-2016, 2016, from https://www.ic3.gov/media/2015/150623.aspx

Feng, Y., Liu, C., and Liu, B. (2017). *Poster: A New Approach to Detecting Ransomware with Deception*. Paper presented at the 38th IEEE Symposium on Security and Privacy.

Fujino, A., Murakami, J., and Mori, T. (2015). *Discovering similar malware samples using API call topics.* Paper presented at the 2015 12th Annual IEEE Consumer Communications and Networking Conference, CCNC 2015, 140-147.

Galal, H. S., Mahdy, Y. B., and Atiea, M. A. (2015). Behavior-based features model for malware detection. *Journal of Computer Virology and Hacking Techniques, 12*(2), 59-67.

Galal, H. S., Mahdy, Y. B., and Atiea, M. A. (2016). Behavior-based features model for malware detection. *Journal of Computer Virology and Hacking Techniques, 12*(2), 59-67.

Ganame, K., Allaire, M. A., Zagdene, G., and Boudar, O. (2017). Network Behavioral Analysis for Zero-Day Malware Detection – A Case Study, *1st International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, ISDDC 2017* (Vol. 10618 LNCS, pp. 169-181): Springer Verlag.

Gandotra, E., Bansal, D., and Sofat, S. (2014). Malware analysis and classification: A survey. *Journal of Information Security, 2014*.

Ganesh, N., Troia, F. D., Corrado, V. A., Austin, T. H., and Stamp, M. (2016). *Static Analysis of Malicious Java Applets*. Paper presented at the Proceedings of the 2016 ACM on International Workshop on Security And Privacy Analytics.

Gazet, A. (2010). Comparative analysis of various ransomware virii. *Journal in Computer Virology, 6*(1), 77-90.

Genç, Z. A., Lenzini, G., and Ryan, P. (2018). Security Analysis of Key Acquiring Strategies Used by Cryptographic Ransomware. *Advances in Cybersecurity 2018*.

George, N., and Vinod, P. (2015). *Opcode position a ware metamorphic malware detection: Signature vs histogram approach.* Paper presented at the 2nd International Conference on Computing for Sustainable Global Development, INDIACom 2015, 1011-1017.

Gomez-Hernandez, J. A., Alvarez-Gonzalez, L., and Garcia-Teodoro, P. (2018). R-Locker: Thwarting ransomware action through a honeyfile-based approach. *Computers & Security, 73*, 389-398.

Gonzalez, D., and Hayajneh, T. (2018). *Detection and prevention of crypto-ransomware*. Paper presented at the 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2017, 472-478.

Gostev, A., Unuchek, R., Garnaeva, M., Makrushin, D., and Ivanov, A. (2016). IT THREAT EVOLUTION IN Q1 2016: Kaspersky Lab.

Grill, M., Pevný, T., and Rehak, M. (2017). Reducing false positives of network anomaly detection by local adaptive multivariate smoothing. *Journal of Computer and System Sciences, 83*(1), 43-57.

Hampton, N., Baig, Z., and Zeadally, S. (2018). Ransomware behavioural analysis on windows platforms. *Journal of Information Security and Applications, 40*, 44-51.

Hanchuan, P., Fuhui, L., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(8), 1226-1238.

Hansen, S. S., Larsen, T. M. T., Stevanovic, M., and Pedersen, J. M. (2016, 15-18 Feb. 2016). *An approach for detection and family classification of malware based on behavioral analysis.* Paper presented at the 2016 International Conference on Computing, Networking and Communications (ICNC), 1-5.

Hasan, M. M., and Rahman, M. M. (2017, 22-24 Dec. 2017). *RansHunt: A support vector machines based ransomware analysis framework with integrated feature set.* Paper presented at the 2017 20th International Conference of Computer and Information Technology (ICCIT), 1-7.

He, X., Cai, D., and Niyogi, P. (2006). *Laplacian score for feature selection.* Paper presented at the Advances in neural information processing systems, 507-514.

Homayoun, S., Dehghantanha, A., Ahmadzadeh, M., Hashemi, S., and Khayami, R. (2017). Know Abnormal, Find Evil: Frequent Pattern Mining for Ransomware Threat Hunting and Intelligence. *IEEE Transactions on Emerging Topics in Computing*.

Homayoun, S., Dehghantanha, A., Ahmadzadeh, M., Hashemi, S., Khayami, R., Choo, K. K. R., et al. (2019). DRTHIS: Deep ransomware threat hunting and intelligence system at the fog layer. *Future Generation Computer Systems-the International Journal of Escience, 90*, 94-104.

Huan, L., and Setiono, R. (1995, 5-8 Nov. 1995). *Chi2: feature selection and discretization of numeric attributes.* Paper presented at the Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence, 388-391.

Ioanid, A., Scarlat, C., and Militaru, G. (2017). *The Effect of Cybercrime on Romanian SMEs in the Context of Wannacry Ransomware Attacks.* Paper presented at the 12th European Conference on Innovation and Entrepreneurship ECIE 2017, 307.

Jabbar, M. A., Aluvalu, R., and Reddy, S. S. S. (2017). *Cluster Based Ensemble Classification for Intrusion Detection System*. Paper presented at the Proceedings of the 9th International Conference on Machine Learning and Computing.

Jain, K. (2015). Script Kiddies can Now Create their Own Ransomware using This Kit. Retrieved 09-11-2016, 2016, from http://thehackernews.com/2015/08/ransomware-creator-toolkit.html

Joldzic, O., Djuric, Z., and Vuletic, P. (2016). A transparent and scalable anomaly-based DoS detection method. *Computer Networks, 104*, 27-42.

Jung, S., and Won, Y. (2018). Ransomware detection method based on context-aware entropy analysis. *Soft Computing, 22*(20), 6731-6740.

Jurek, A., Hong, J., Chi, Y., and Liu, W. (2017). A novel ensemble learning approach to unsupervised record linkage. *Information Systems, 71*(Supplement C), 40-54.

Kao, D. Y., and Hsiao, S. C. (2018, 11-14 Feb. 2018). *The dynamic analysis of WannaCry ransomware.* Paper presented at the 2018 20th International Conference on Advanced Communication Technology (ICACT), 159-166.

Kaspersky. (2016). Ksn Report: Ransomware In 2014-2016: Kaspersky Lab.

Kaur, R., and Singh, M. (2014). A Survey on Zero-Day Polymorphic Worm Detection Techniques. *IEEE Communications Surveys & Tutorials, 16*(3), 1520-1549.

Kaur, R., and Singh, S. (2016). A survey of data mining and social network analysis based anomaly detection techniques. *Egyptian Informatics Journal, 17*(2), 199-216.

Kharraz, A., Arshad, S., Mulliner, C., Robertson, W., and Kirda, E. (2016). UNVEIL: A Large-Scale, Automated Approach to Detecting Ransomware. *Proceedings of the 25th Usenix Security Symposium*, 757-772.

Kharraz, A., and Kirda, E. (2017). Redemption: Real-time Protection Against Ransomware at End-Hosts.

Kharraz, A., Robertson, W., Balzarotti, D., Bilge, L., and Kirda, E. (2015). Cutting the gordian knot: A look under the hood of ransomware attacks. In F. Maggi, M. Almgren and V. Gulisano (Eds.), *12th International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment, DIMVA 2015* (Vol. 9148, pp. 3-24): Springer Verlag.

Kharraz, A., Robertson, W., and Kirda, E. (2018). Protecting against Ransomware: A New Line of Research or Restating Classic Ideas? *Ieee Security & Privacy, 16*(3), 103-107.

Ki, Y., Kim, E., and Kim, H. K. (2015). A Novel Approach to Detect Malware Based on API Call Sequence Analysis. *International Journal of Distributed Sensor Networks, 11*(6), 659101.

Kim, D., Soh, W., and Kim, S. (2015). Design of Quantification Model for Prevent of Cryptolocker. *Indian Journal of Science and Technology, 8*(19).

Kim, G., Lee, S., and Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications, 41*(4, Part 2), 1690-1700.

Kim, H., Yoo, D., Kang, J., and Yeom, Y. (2017, 13-14 Nov. 2017). *Dynamic ransomware protection using deterministic random bit generator*. Paper presented at the 2017 IEEE Conference on Application, Information and Network Security (AINS), 64-68.

Kolodenker, E., Koch, W., Stringhini, G., and Egele, M. (2017). *PayBreak: Defense Against Cryptographic Ransomware*. Paper presented at the In Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security, 599-611.

Koziarski, M., Krawczyk, B., and Woźniak, M. (2017). The deterministic subspace method for constructing classifier ensembles. *Pattern Analysis and Applications*, *20*(4), 981-990.

Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., and Woźniak, M. (2017). Ensemble learning for data stream analysis: A survey. *Information Fusion, 37*(Supplement C), 132-156.

Kumar, C. U. O., Kishore, S., and Geetha, A. (2015). *Debugging using MD5 process firewall*. Paper presented at the 2014 International Conference on Contemporary Computing and Informatics, IC3I 2014, 1279-1284.

Kumar, S. M., and Kumar, M. R. (2013). Cryptoviral Extortion: A virus based approach. *International Journal of Computer Trends and Technology (IJCTT), 4*(5), 1149-1153.

Labs, M. (2018). McAfee Labs Threats Report. Retrieved 04/04/2019, from https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-dec-2017.pdf

Le Guernic, C., and Legay, A. (2017). *Ransomware and the Legacy Crypto API*. Paper presented at the Risks and Security of Internet and Systems: 11th International Conference, CRiSIS 2016, Roscoff, France, September 5-7, 2016, Revised Selected Papers, 11.

Lee, J. K., Moon, S. Y., and Park, J. H. (2016). CloudRPS: a cloud analysis based enhanced ransomware prevention system. *The Journal of Supercomputing*, 1-20.

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., et al. (2017). Feature Selection: A Data Perspective. *ACM Comput. Surv., 50*(6), 1-45.

Liao, H.-J., Richard Lin, C.-H., Lin, Y.-C., and Tung, K.-Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications, 36*(1), 16-24.

Lin, C.-T., Wang, N.-J., Xiao, H., and Eckert, C. (2015). Feature selection and extraction for malware classification. *Journal of Information Science and Engineering, 31*(3), 965-992.

Liu, H., Sun, J., Liu, L., and Zhang, H. (2009). Feature selection with dynamic mutual information. *Pattern Recognition, 42*(7), 1330-1339.

Liu, J., Zhao, S., and Wang, G. (2018). SSEL-ADE: A semi-supervised ensemble learning framework for extracting adverse drug events from social media. *Artif Intell Med, 84*, 34-49.

Lokuketagoda, B., Weerakoon, M., Madushan, U., Senaratne, A., and Abeywardena, K. (2018). R-Killer: An Email Based Ransomware Protection Tool. *World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering, 5*(2).

Lu, T., Zhang, L., Wang, S., and Gong, Q. (2017, 15-17 Dec. 2017). *Ransomware detection based on V-detector negative selection algorithm.* Paper presented at the 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), 531-536.

Luo, X., and Liao, Q. (2007). Awareness Education as the key to Ransomware Prevention. *Information Systems Security, 16*(4), 195-202.

Luo, X., and Liao, Q. (2008). Ransomware: A new cyber hijacking threat to enterprises. In *Handbook of Research on Information Security and Assurance* (pp. 1-6): IGI Global.

Maiorca, D., Mercaldo, F., Giacinto, G., Visaggio, C. A., and Martinelli, F. (2017). *R-PackDroid: API package-based characterization and detection of mobile ransomware*. Paper presented at the Proceedings of the Symposium on Applied Computing.

Maniath, S., Ashok, A., Poornachandran, P., Sujadevi, V. G., Sankar, A. U. P., and Jan, S. (2017, 26-27 Oct. 2017). *Deep learning LSTM based ransomware detection.* Paper presented at the 2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE), 442-446.

Mao, S., Lin, W., Chen, J., and Xiong, L. (2017). Optimising ensemble combination based on maximisation of diversity. *Electronics Letters, 53*(15), 1042-1044.

Mbol, F., Robert, J.-M., and Sadighian, A. (2016). An Efficient Approach to Detect TorrentLocker Ransomware in Computer Systems. In S. Foresti and G. Persiano (Eds.), *Cryptology and Network Security: 15th International Conference, CANS 2016, Milan, Italy, November 14-16, 2016, Proceedings* (pp. 532-541). Cham: Springer International Publishing.

McAfee. (2016a). 2016 Threats Predictions, *McAfee Labs*.

McAfee, L. (2016b). McAfee Labs Threats Report.

McAfee, L. (2016c). Understanding Ransomware and Strategies to Defeat It. In I. Security (Ed.).

McGinn, D., Birch, D., Akroyd, D., Molina-Solana, M., Guo, Y., and Knottenbelt, W. J. (2016). Visualizing Dynamic Bitcoin Transaction Patterns. *Big Data, 4*(2), 109-119.

Mehetrey, P., Shahriari, B., and Moh, M. (2016). *Collaborative Ensemble-Learning Based Intrusion Detection Systems for Clouds.* Paper presented at the Collaboration Technologies and Systems (CTS), 2016 International Conference on, 404-411.

Mehnaz, S., Mudgerikar, A., and Bertino, E. (2018). RWGuard: A real-time detection system against cryptographic ransomware, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 11050 LNCS, pp. 114-136).

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., et al. (2016). A fistful of Bitcoins: characterizing payments among men with no names. *Commun. ACM, 59*(4), 86-93.

Mercaldo, F., Nardone, V., Santone, A., and Visaggio, C. A. (2016). Ransomware Steals Your Phone. Formal Methods Rescue It. In E. Albert and I. Lanese (Eds.), *Formal Techniques for Distributed Objects, Components, and Systems: 36th IFIP WG 6.1 International Conference, FORTE 2016, Held as Part of the 11th International Federated Conference on Distributed Computing Techniques, DisCoTec 2016, Heraklion, Crete, Greece, June 6-9, 2016, Proceedings* (pp. 212-221). Cham: Springer International Publishing.

Miao, Q., Liu, J., Cao, Y., and Song, J. (2015). Malware detection using bilayer behavior abstraction and improved one-class support vector machines. *International Journal of Information Security, 15*(4), 361-379.

Milošević, N. (2013). History of malware. *arXiv preprint arXiv:1302.5392*.

Mohurle, S., and Patil, M. (2017). A brief study of Wannacry Threat: Ransomware Attack 2017. *2017, 8*(5), 3.

Moore, C. (2016, 2-4 Aug. 2016). *Detecting Ransomware with Honeypot Techniques.* Paper presented at the 2016 Cybersecurity and Cyberforensics Conference (CCC), 77-81.

Morato, D., Berrueta, E., Magaña, E., and Izal, M. (2018). Ransomware early detection by the analysis of file sharing traffic. *Journal of Network and Computer Applications, 124,* 14-32.

Moussaileb, R., Bouget, B., Palisse, A., Le Bouder, H., Cuppens, N., and Lanet, J. L. (2018). *Ransomware's early mitigation mechanisms*. Paper presented at the 13th International Conference on Availability, Reliability and Security, ARES 2018. Retrieved 27 August 2018 through 30 August 2018, from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85055288709&doi=10.1145%2f3230833.3234691&partnerID=40&md5=fd3fa38ed1fb15bb45641bb3db029589

Mustaca, S. (2014). Are your IT professionals prepared for the challenges to come? *Computer Fraud and Security*, *2014*(3), 18-20.

Nadir, I., and Bakhshi, T. (2018, 3-4 March 2018). *Contemporary cybercrime: A taxonomy of ransomware threats & mitigation techniques.* Paper presented at the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), 1-7.

Nauman, M., Azam, N., and Yao, J. T. (2016). A three-way decision making approach to malware analysis using probabilistic rough sets. *Information Sciences, 374*, 193-209.

Naval, S., Laxmi, V., Rajarajan, M., Gaur, M. S., and Conti, M. (2015). Employing Program Semantics for Malware Detection. *IEEE Transactions on Information Forensics and Security, 10*(12), 2591-2604.

Nie, F., Huang, H., Cai, X., and Ding, C. H. (2010). *Efficient and robust feature selection via joint ℓ2, 1-norms minimization.* Paper presented at the Advances in neural information processing systems, 1813-1821.

Nie, F., Xiang, S., Jia, Y., Zhang, C., and Yan, S. (2008). *Trace ratio criterion for feature selection.* Paper presented at the AAAI, 671-676.

Nieuwenhuizen, D. (2017). A behavioural-based approach to ransomware detection.

Nissim, N., Lapidot, Y., Cohen, A., and Elovici, Y. (2018). Trusted system-calls analysis methodology aimed at detection of compromised virtual machines using sequential mining. *Knowledge-Based Systems, 153*, 147-175.

Nong, Y., Yebin, Z., and Borror, C. M. (2004). Robustness of the Markov-chain model for cyber-attack detection. *IEEE Transactions on Reliability, 53*(1), 116-123.

O'Gorman, G., and McDonald, G. (2012). *Ransomware: a growing menace*: Symantec Corporation.

Onan, A., and Korukoğlu, S. (2017). A feature selection model based on genetic rank aggregation for text sentiment classification. *Journal of Information Science, 43*(1), 25-38.

Paganini, P. (2015). Tox, how to create your ransomware in 3 steps. Retrieved 09-11-2016, 2016, from http://securityaffairs.co/wordpress/37180/cyber-crime/tox-ransomware-builder.html

Paik, J.-Y., Shin, K., and Cho, E.-S. (2016). *Poster: Self-Defensible Storage Devices based on Flash memory against Ransomware.* Paper presented at the Proceedings of IEEE Symposium on Security and Privacy.

Pandey, S. K., and Mehtre, B. M. (2015). *Performance of malware detection tools: A comparison.* Paper presented at the 2014 IEEE International Conference on Advanced Communication, Control and Computing Technologies, ICACCCT 2014, 1811-1817.

Parikh, D., and Polikar, R. (2007). An Ensemble-Based Incremental Learning Approach to Data Fusion. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 37*(2), 437-450.

Pathak, P., and Nanded, Y. M. (2016). A Dangerous Trend of Cybercrime: Ransomware Growing Challenge. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 5*(2), 371-373.

Peddabachigari, S., Abraham, A., Grosan, C., and Thomas, J. (2007). Modeling intrusion detection system using hybrid intelligent systems. *Journal of Network and Computer Applications, 30*(1), 114-132.

Peng, H., Wei, J., and Guo, W. (2016). *Micro-architectural Features for Malware Detection.* Paper presented at the Conference, 48-60.

Pluskal, O. (2015). *Behavioural malware detection using efficient SVM implementation.* Paper presented at the Research in Adaptive and Convergent Systems, RACS 2015, 296-301.

Poonia, A. S., and Singh, S. (2015). *Malware detection by token counting.* Paper presented at the 2014 International Conference on Contemporary Computing and Informatics, IC3I 2014, 1285-1288.

Popli, N. K., and Girdhar, A. (2019). *Behavioural Analysis of Recent Ransomwares and Prediction of Future Attacks by Polymorphic and Metamorphic Ransomware*, Singapore, 65-80.

Prakash, K. P., Nafis, T., and Sankar Biswas, D. S. (2017). Preventive Measures and Incident Response for Locky Ransomware. *2017, 8*(5), 4.

Prelipcean, D. B., Popescu, A. S., and Gavrilut, D. T. (2016). *Improving Malware Detection Response Time with Behavior-Based Statistical Analysis Techniques.* Paper presented at the Proceedings - 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2015, 232-239.

Quinkert, F., Holz, T., Hossain, K., Ferrara, E., and Lerman, K. (2018). RAPTOR: Ransomware Attack PredicTOR. *arXiv preprint arXiv:1803.01598*.

Ray, O., Hicks, S., and Moyle, S. (2017). Using ILP to Analyse Ransomware Attacks.

Reineking, T. (2016). Active classification using belief functions and information gain maximization. *International Journal of Approximate Reasoning, 72*(Supplement C), 43-54.

Rhee, J., Riley, R., Lin, Z., Jiang, X., and Xu, D. (2014). Data-Centric OS Kernel Malware Characterization. *IEEE Transactions on Information Forensics and Security, 9*(1), 72-87.

Rhode, M., Burnap, P., and Jones, K. (2018). Early-stage malware prediction using recurrent neural networks. *Computers & Security, 77,* 578-594.

Richet, J.-L. (2015). Extortion on the Internet: the Rise of Crypto-Ransomware.

Rossow, C., Dietrich, C. J., Grier, C., Kreibich, C., Paxson, V., Pohlmann, N., et al. (2012, 20-23 May 2012). *Prudent Practices for Designing Malware Experiments: Status Quo and Outlook.* Paper presented at the 2012 IEEE Symposium on Security and Privacy, 65-79.

Ryu, D., Lee, K., and Baik, J. (2018). Location-based Web Service QoS Prediction via Preference Propagation to address Cold Start Problem. *IEEE Transactions on Services Computing*, 1-1.

Salton, G., and Buckley, C. (1997). Improving retrieval performance by relevance feedback. *Readings in information retrieval, 24*(5), 355-363.

Scaife, N., Carter, H., Traynor, P., and Butler, K. R. (2016). *CryptoLock (and Drop It): Stopping Ransomware Attacks on User Data*. Paper presented at the Distributed Computing Systems (ICDCS), 2016 IEEE 36th International Conference on.

Sgandurra, D., Muñoz-González, L., Mohsen, R., and Lupu, E. C. (2016). Automated Dynamic Analysis of Ransomware: Benefits, Limitations and use for Detection. *arXiv preprint arXiv:1609.03020*.

Shahriari, M. M. A. H. R. (2016). *2entFOX: A Framework for High Survivable Ransomwares Detection*. Paper presented at the Information Security and Cryptology (ISCISC), 2015 12th International Iranian Society of Cryptology Conference on Information Security and Cryptology.

Shashidhar, N. K. (2017). Ransomware Analysis and Defense-WannaCry and the Win32 environment. *International Journal of Information Security Science, 6*(4), 57-69.

Shijo, P. V., and Salim, A. (2015). *Integrated static and dynamic analysis for malware detection.* Paper presented at the Procedia Computer Science, 804-811.

Shim, K. A. (2016). A Survey of Public-Key Cryptographic Primitives in Wireless Sensor Networks. *Ieee Communications Surveys and Tutorials, 18*(1), 577-601.

Shukla, M., Mondal, S., and Lodha, S. (2016). *POSTER: Locally Virtualized Environment for Mitigating Ransomware Threat*. Paper presented at the Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security.

Singh, T., Di Troia, F., Corrado, V. A., Austin, T. H., and Stamp, M. (2015). Support vector machines and malware detection. *Journal of Computer Virology and Hacking Techniques, 12*(4), 203-212.

Sittig, D. F., and Singh, H. (2016). A Socio-technical Approach to Preventing, Mitigating, and Recovering from Ransomware Attacks. *Applied Clinical Informatics, 7*(2), 624-632.

Soltani, S., Seno, S. A. H., Nezhadkamali, M., and Budiarto, R. (2014). A survey on real world botnets and detection mechanisms. *International Journal of Information and Network Security, 3*(2), 116.

Song, S., Kim, B., and Lee, S. (2016). The Effective Ransomware Prevention Technique Using Process Monitoring on Android Platform. *Mobile Information Systems, 2016*.

Sotelo Monge, M. A., Vidal, J. M., and García Villalba, L. J. (2018). *A novel self-organizing network solution towards crypto-ransomware mitigation.* Paper presented at the 13th International Conference on Availability, Reliability and Security, ARES 2018, 48.

Spagnuolo, M., Maggi, F., and Zanero, S. (2014). Bitiodine: Extracting intelligence from the bitcoin network. In R. Safavi-Naini and N. Christin (Eds.), *18th International Conference on Financial Cryptography and Data Security, FC 2014* (Vol. 8437, pp. 457-468): Springer Verlag.

Stiborek, J., Pevny, T., and Rehak, M. (2018a). Multiple instance learning for malware classification. *Expert Systems with Applications, 93*, 346-357.

Stiborek, J., Pevny, T., and Rehak, M. (2018b). Probabilistic analysis of dynamic malware traces. *Computers & Security, 74*, 221-239.

Subedi, K. P., Budhathoki, D. R., and Dasgupta, D. (2018). *Forensic Analysis of Ransomware Families using Static and Dynamic Analysis*. Paper presented at the 018 IEEE Security and Privacy Workshops (SPW).

Suditu, N., Fran, #231, and Fleuret, o. (2012). *Iterative relevance feedback with adaptive exploration/exploitation trade-off*. Paper presented at the Proceedings of the 21st ACM international conference on Information and knowledge management.

Symantec. (2015). The evolution of ransomware, *SECURITY RESPONSE*: Symantec Corporation.

Symantec. (2016a). Internet Security Threat Report. *Symantec*.

Symantec. (2016b). Ransomware and Businesses 2016. In J.-P. P. Dick O'Brien, Scott Wallace (Ed.), *An ISTR Special Report*: Symantec Corporation.

Tailor, J. P., and Patel, A. D. (2017). A Comprehensive Survey: Ransomware Attacks Prevention, Monitoring and Damage Control. *International Journal of Research and Scientific Innovation (IJRSI), IV*(VIS), 116-121.

Tandon, A., and Nayyar, A. (2019). *A Comprehensive Survey on Ransomware Attack: A Growing Havoc Cyberthreat*, Singapore, 403-420.

Tang, A., Sethumadhavan, S., and Stolfo, S. J. (2014). Unsupervised anomaly-based malware detection using hardware features, *17th International Symposium on Research in Attacks, Intrusions, and Defenses, RAID 2014* (Vol. 8688 LNCS, pp. 109-129). Gothenburg: Springer Verlag.

Taylor, M. (2017). *Ransomware Detection Using Machine Learning and Physical Sensor Data.* Unpublished M.S., Southern Methodist University, Ann Arbor.

Taylor, M. A., Smith, K. N., and Thornton, M. A. (2017). *Sensor-based Ransomware Detection*. Paper presented at the Future Technologies Conference (FTC) 2017.

Tripathy, A., Agrawal, A., and Rath, S. K. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications, 57*, 117-126.

Turaev, H., Zavarsky, P., and Swar, B. (2018). *Prevention of ransomware execution in enterprise environment on windows os: Assessment of application whitelisting solutions.* Paper presented at the 1st International Conference on Data Intelligence and Security, ICDIS 2018, 110-118.

Ucci, D., Aniello, L., and Baldoni, R. (2017). Survey on the Usage of Machine Learning Techniques for Malware Analysis. *arXiv preprint arXiv:1710.08189*.

Uppal, D., Sinha, R., Mehra, V., Jain, V., and Ieee. (2014). Malware Detection and Classification Based on Extraction of API Sequences. *2014 International Conference on Advances in Computing, Communications and Informatics (Icacci)*, 2337-2342.

Usha, M., and Kavitha, P. (2016). Anomaly based intrusion detection for 802.11 networks with optimal features using SVM classifier. *Wireless Networks*, 1-16.

Usman, L., Prayudi, Y., and Riadi, I. (2017). Ransomware analysis based on the surface, runtime and static code method. *Journal of Theoretical and Applied Information Technology, 95*(11), 2426-2433.

Van Nhuong, N., Yen Nhi, V. T., Cam, N. T., Phu, M. X., and Dang Tan, C. (2015). *SSSM-semantic set and string matching based malware detection.* Paper presented at the 7th IEEE Symposium on Computational Intelligence for Security and Defense Applications, CISDA 2014, 1-6.

Villalba, L. J. G., Orozco, A. L. S., Vivar, A. L., Vega, E. A. A., and Kim, T.-H. (2018). Ransomware Automatic Data Acquisition Tool. *IEEE Access*, 1-1.

Vinayakumar, R., Soman, K. P., Velan, K. K. S., and Ganorkar, S. (2017). *Evaluating shallow and deep networks for ransomware detection and classification.* Paper presented at the 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017, 259-265.

Wagner, M., Fischer, F., Luh, R., Haberson, A., Rind, A., Keim, D., et al. (2015). *A Survey of Visualization Systems for Malware Analysis.* Paper presented at the Eurographics Conference on Visualization (EuroVis) State of The Art Reports, 105-125.

Wan, Y.-L., Chang, J.-C., Chen, R.-J., and Wang, S.-J. (2018). *Feature-Selection-Based Ransomware Detection with Machine Learning of Data Analysis.* Paper presented at the 3rd International Conference on Computer and Communication Systems, 85-88.

Wang, J., Wei, J. M., Yang, Z., and Wang, S. Q. (2017a). Feature Selection by Maximizing Independent Classification Information. *IEEE Transactions on Knowledge and Data Engineering, 29*(4), 828-841.

Wang, P., and Wang, Y.-S. (2015). Malware behavioural detection and vaccine development by using a support vector model classifier. *Journal of Computer and System Sciences, 81*(6), 1012-1026.

Wang, X., Yang, Y., Zeng, Y., Tang, C., Shi, J., and Xu, K. (2015). *A Novel Hybrid Mobile Malware Detection System Integrating Anomaly Detection With Misuse Detection*. Paper presented at the Proceedings of the 6th International Workshop on Mobile Cloud Computing and Services.

Wang, Y., Wang, J., Liao, H., and Chen, H. (2017b). An efficient semi-supervised representatives feature selection algorithm based on information theory. *Pattern Recognition, 61*(Supplement C), 511-523.

Wang, Z., Wu, X., Liu, C., Liu, Q., and Zhang, J. (2018, 18-21 June 2018). *RansomTracer: Exploiting Cyber Deception for Ransomware Tracing.* Paper presented at the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), 227-234.

Watson, M. R., Shirazi, N. U. H., Marnerides, A. K., Mauthe, A., and Hutchison, D. (2016). Malware Detection in Cloud Computing Infrastructures. *IEEE Transactions on Dependable and Secure Computing, 13*(2), 192-205.

Wei, J., He, J., Chen, K., Zhou, Y., and Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications, 69,* 29-39.

Woźniak, M., Graña, M., and Corchado, E. (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion, 16*(Supplement C), 3-17.

Xu, Y., Wu, C., Zheng, K., Wang, X., Niu, X., and Lu, T. (2017). Computing Adaptive Feature Weights with PSO to Improve Android Malware Detection. *Security and Communication Networks, 2017*, 14.

Xue, L., and Sun, G. (2015). Design and implementation of a malware detection system based on network behavior. *Security and Communication Networks, 8*(3), 459-470.

Xue, Y., Wang, J., Liu, Y., Xiao, H., Sun, J., and Chandramohan, M. (2015). *Detection and classification of malicious JavaScript via attack behavior modelling*. Paper presented at the Proceedings of the 2015 International Symposium on Software Testing and Analysis. from http://delivery.acm.org/10.1145/2780000/2771814/p48-xue.pdf?ip=161.139.222.59&id=2771814&acc=ACTIVE%20SERVICE&key=69AF3716A20387ED%2EC758BA176ED44BB8%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=597235143&CFTOKEN=93856099&__acm__=1459740818_f8bf4649cd4bea30427c69ff8442b8fc

http://delivery.acm.org/10.1145/2780000/2771814/p48-xue.pdf?ip=161.139.222.59&id=2771814&acc=ACTIVE%20SERVICE&key=69AF3716A20387ED%2EC758BA176ED44BB8%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&CFID=601420615&CFTOKEN=85591093&__acm__=1460869848_8e33d583515df407e89dfb3c1dd2dab7

Yalew, S. D., Maguire, G. Q., Haridi, S., and Correia, M. (2017). *Hail to the Thief: Protecting data from mobile ransomware with ransomsafedroid.* Paper presented at the 2017 IEEE 16th International Symposium on Network Computing and Applications, NCA 2017, 1-8.

Yan, P., and Yan, Z. (2017). A survey on dynamic mobile malware detection. *Software Quality Journal*, 1-29.

Yang, H., and Moody, J. (1999). *Feature selection based on joint mutual information.* Paper presented at the Proceedings of international ICSC symposium on advances in intelligent data analysis, 22-25.

Yang, M., Bao, J., and Ji, G. L. (2010). *Semi-random subspace sampling for classification.* Paper presented at the 2010 6th International Conference on Natural Computation, ICNC'10, Yantai, Shandong, 3420-3424.

Yang, T., Yang, Y., Qian, K., Lo, D. C.-T., Qian, Y., and Tao, L. (2015a). Automated Detection and Analysis for Android Ransomware. 1338-1343.

Yang, T., Yang, Y., Qian, K., Lo, D. C. T., Qian, Y., and Tao, L. (2015b). *Automated detection and analysis for android ransomware.* Paper presented at the 17th IEEE International Conference on High Performance Computing and Communications, IEEE 7th International Symposium on Cyberspace Safety and Security and IEEE 12th International Conference on Embedded Software and Systems, HPCC-ICESS-CSS 2015, 1338-1343.

Yang, X., Lo, D., Xia, X., and Sun, J. (2017). TLEL: A two-layer ensemble learning approach for just-in-time defect prediction. *Information and Software Technology*, *87*(Supplement C), 206-220.

Yang, Y., Shen, H. T., Ma, Z., Huang, Z., and Zhou, X. (2011). *l2, 1-norm regularized discriminative feature selection for unsupervised learning.* Paper presented at the IJCAI proceedings-international joint conference on artificial intelligence, 1589.

Yaqoob, I., Ahmed, E., Rehman, M. H. U., Ahmed, A. I. A., Al-Garadi, M. A., Imran, M., et al. (2017). The rise of ransomware and emerging security challenges in the Internet of Things. *Computer Networks*, *129*, 444-458.

Ye, Y., Li, T., Adjeroh, D., and Iyengar, S. S. (2017). A Survey on Malware Detection Using Data Mining Techniques. *ACM Comput. Surv.*, *50*(3), 1-40.

Young, A., and Yung, M. (1996). *Cryptovirology: Extortion-based security threats and countermeasures.* Paper presented at the Security and Privacy, 1996. Proceedings., 1996 IEEE Symposium on, 129-140.

Young, A. L. (2005). Building a Cryptovirus Using Microsoft's Cryptographic API. In J. Zhou, J. Lopez, R. H. Deng and F. Bao (Eds.), *Information Security: 8th International Conference, ISC 2005, Singapore, September 20-23, 2005. Proceedings* (pp. 389-401). Berlin, Heidelberg: Springer Berlin Heidelberg.

Young, A. L. (2006). Cryptoviral extortion using Microsoft's Crypto API. *International Journal of Information Security*, *5*(2), 67-76.

Yu, B., Fang, Y., Yang, Q., Tang, Y., and Liu, L. (2018). A survey of malware behavior description and analysis. *Frontiers of Information Technology & Electronic Engineering, 19*(5), 583-603.

Zahra, A., and Shah, M. A. (2017). *IoT based ransomware growth rate evaluation and detection using command and control blacklisting.* Paper presented at the ICAC 2017 - 2017 23rd IEEE International Conference on Automation and Computing: Addressing Global Challenges through Automation and Computing, 15-22.

Zakeri, M., Faraji Daneshgar, F., and Abbaspour, M. (2015). A static heuristic approach to detecting malware targets. *Security and Communication Networks, 8*(17), 3015-3027.

Zhang-Kennedy, L., Assal, H., Rocheleau, J., Mohamed, R., Baig, K., and Chiasson, S. (2018). *The aftermath of a crypto-ransomware attack at a large academic institution.* Paper presented at the 27th {USENIX} Security Symposium ({USENIX} Security 18), 1061-1078.

Zhang, H. Q., Xiao, X., Mercaldo, F., Ni, S. G., Martinelli, F., and Sangaiah, A. K. (2019). Classification of ransomware families with machine learning based on N-gram of opcodes. *Future Generation Computer Systems-the International Journal of Escience, 90*, 211-221.

Zhang, M., Xu, B. Y., and Wang, D. X. (2016). An Anomaly Detection Model for Network Intrusions Using One-Class SVM and Scaling Strategy. In S. Guo, X. Liao, F. Liu and Y. Zhu (Eds.), *Collaborative Computing: Networking, Applications, and Worksharing, Collaboratecom 2015* (Vol. 163, pp. 267-278). New York: Springer.

Zhang, P., and Tan, Y. (2015). *Hybrid concentration based feature extraction approach for malware detection.* Paper presented at the 2015 28th IEEE Canadian Conference on Electrical and Computer Engineering, CCECE 2015, 140-145.

Zimba, A. (2017). Malware-Free Intrusion: A Novel Approach to Ransomware Infection Vectors. *International Journal of Computer Science and Information Security, 15*(2), 317.

Zimba, A., Simukonda, L., and Chishimba, M. (2017a). Demystifying Ransomware Attacks: Reverse Engineering and Dynamic Malware Analysis of WannaCry for Network and Information Security. *Zambia ICT Journal, 1*(1), 35-40.

Zimba, A., Wang, Z., and Chen, H. (2017b, 22-24 July 2017). *Reasoning crypto ransomware infection vectors with Bayesian networks.* Paper presented at the 2017 IEEE International Conference on Intelligence and Security Informatics (ISI), 149-151.

Zimba, A., Wang, Z., and Simukonda, L. (2018a). Towards Data Resilience: The Analytical Case of Crypto Ransomware Data Recovery Techniques. *International Journal of Information Technology and Computer Science, 10*(1), 40-51.

Zimba, A., Wang, Z. S., and Chen, H. S. (2018b). Multi-stage crypto ransomware attacks: A new emerging cyber threat to critical infrastructure and industrial control systems. *Ict Express, 4*(1), 14-18.