UNIVERSITI PUTRA MALAYSIA

METHOD OF EVENT RECONSTRUCTION IN DIGITAL INVESTIGATION AND ITS VISUALIZATION

MOHD TAUFIK ABDULLAH

METHOD OF EVENT RECONSTRUCTION IN DIGITAL INVESTIGATION AND ITS VISUALIZATION

By

MOHD TAUFIK ABDULLAH

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requiremenst for the Degree of Doctor of Philosophy

January 2011

DEDICATION

I would like to dedicate my work to my beloved wife; Wan Sakiah Wan Oman,
my sons; Muhammad Syamsi, Abdul Muhaimin, and Muhammad Afifuddin,
my daughter; Nur Wahidah and Ajlaa Bazilah
and my family.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

METHOD OF EVENT RECONSTRUCTION IN DIGITAL INVESTIGATION AND ITS VISUALIZATION

By

MOHD TAUFIK ABDULLAH

January 2011

Chairman : Associate Professor Ramlan Mahmod, PhD

Faculty : Computer Science and Information Technology

A reconstruction of sequences of events that leads to a suspicious incident is an important phase in digital forensics investigation. Event reconstruction answers the question concerning the existence of digital object within computer at any particular time either triggered by an event or an effect of an event. Various event reconstruction techniques are used for representing the sequence of event that caused presence of the digital objects.

The reconstruction of events in digital investigations is fairly complicated. Unaided reasoning is usually insufficient to comprehensively analyze the sequence of events to identify suspect, apprehend the guilty and defend the innocent. Most present techniques lacks of thoroughness, relevancy, and user friendliness. A development of a sound technique which could reduce the possibility of reasoning errors and hence increases the effectiveness of the analysis is crucial.

This research defines a new method of event reconstruction which associates the capability to handle infinite set of incident scenarios, determine the relevancy of witness statements, and visualize all possibilities of incident scenarios. This study proposed a new method for representing the functionality of system under investigation as well as evidential statements. Some previous works only represent the functionality of the system under investigation as Finite State Machine (FSM). In the proposed method, the functionality of the system under investigation is represented as FSM whereby witness statement is represented as regular expression. An algorithm is developed to derive a Deterministic Finite Automaton (DFA) that accepts computations of FSM that represent the functionality of system under investigation. Similarly, the regular expression is transformed into another DFA using standard algorithms. Finally, the two DFAs are intersected to produce another DFA known as Diagram of Digital Event Reconstruction and Analysis (DDERA).

Having both the functionality of system under investigation and evidential statement represented as DFAs, the event reconstruction is reduced to the problem of automata intersection. The proposed method of event reconstruction in this research has an ability to represent infinite sets of incident scenarios. Therefore, it is capable of handling problematic even transition graphs with loops. Moreover, it allows relevancy checking among given statements themselves as well as against the representation of the functionality of system under investigation. Visualization of all possible scenarios of incident in graphical manner facilitates efficient insight gaining into digital evidence**.** Above all, the whole research formalizes and automates digital forensic analysis into a new horizon.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

KAEDAH PEMBINAAN SEMULA URUTAN PERISTIWA DALAM PENYIASATAN DIGITAL DAN PENGGAMBARANNYA

Oleh

MOHD TAUFIK ABDULLAH

Januari 2011

Pengerusi : Profesor Madya Ramlan Mahmod, PhD

Fakulti : Sains Komputer dan Teknologi Maklumat

Pembinaan semula urutan peristiwa yang memberi petunjuk ke arah sesuatu kejadian yang mencurigakan adalah satu fasa yang mustahak di dalam penyiasatan forensiks digital. Pembinaan semula akan menjawab persoalan berkenaan dengan kewujudan objek digital di dalam komputer pada suatu masa tertentu sama ada dicetuskan oleh suatu peristiwa atau kesan daripada suatu peristiwa. Pelbagai teknik pembinaan semula peristiwa yang digunakan untuk mewakilkan urutan peristiwa yang menyebabkan satu objek digital wujud.

Membina semula peristiwa dalam penyiasatan digital agak rumit. Penaakulan tanpa bantuan biasanya tidak mencukupi untuk mengupas secara menyeluruh urutan peristiwa tersebut untuk mengenal pasti orang yang disyaki, memahami orang yang bersalah dan membela orang yang tidak bersalah. Kebanyakan teknik yang ada kurang kesempurnaan, kerelevanan dan ramah pengguna. Pembinaan satu teknik

yang kukuh yang dapat mengurangkan kebarangkalian kesilapan penaakulan dan seterusnya meningkatkan keberkesanan analisis adalah sangat penting.

Penyelidikan ini mentakrifkan satu kaedah baharu pembinaan semula peristiwa yang menggabungkan keupayaan untuk mengendalikan set senario kejadian tak terhingga, menentukan kerelevanan kenyataan saksi dan dapat menggambarkan segala kemungkinan senario kejadian.

Kajian ini mencadangkan satu kaedah baharu untuk menggambarkan fungsian sistem yang sedang disiasat dan kenyataan keterangan. Beberapa karya lepas hanya menggambarkan fungsian sistem yang sedang disiasat dengan mesin automata terhingga. Dalam kaedah yang dicadangkan fungsian sistem yang sedang disiasat digambarkan dengan mesin berkeadaan terhingga dan kenyataan saksi digambarkan dengan ungkapan nalar. Satu algoritma dibangunkan untuk menerbitkan satu automata berketentuan terhingga yang menerima pengiraan mesin berkeadaan terhingga yang menggambarkan fungsian sistem yang sedang disiasat. Begitu juga, ungkapan nalar diubah bentuk ke dalam automata berketentuan terhingga menggunakan algoritma lazim. Akhir sekali, dua automata berketentuan terhingga itu disilang untuk menghasilkan satu automata berketentuan terhingga yang lain yang kenali sebagai ‘Diagram of Digital Event Reconstruction and Analysis’.

Memiliki kedua-dua fungsian sistem yang sedang disiasat dan kenyataan keterangan yang digambarkan sebagai automata berketentuan terhingga, pembinaan semula peristiwa diturunkan ke masalah persilangan automata. Kaedah pembinaan semula peristiwa yang dicadangkan di dalam penyelidikan ini berkemampuan untuk

menggambarkan set senario kejadian tak terhingga. Oleh sebab itu, ia berkeupayaan mengendalikan masalah walaupun graf peralihan mempunyai gelung. Tambahan pula, ia membolehkan penyemakan kerelevanan sesama kenyataan dan juga terhadap perwakilan fungsian sistem yang sedang disiasat. Penggambaran segala senario kejadian secara bergrafik memudahkan dalam memahami bukti digital dengan sempurna. Kesemua di atas, keseluruhan penyelidikan merumus dan mengautomatikkan analisis forensiks digital ke satu ufuk baru.

ACKNOWLEDGEMENTS

I could not have completed this research work without endless guidance, help, blessings and motivation from Allah the Almighty. I also extend my sincere gratitude to a number of people, who deserve special thanks. Foremost of all, I would like to express my deep and sincere gratitude to my supervisory committee, Associate Professor Dr Ramlan Mahmod, Professor Dr Abdul Azim Abd. Ghani, and Professor Dr Abdullah Mohd Zain for their guidance, support, constructive advice, insight, and helpful suggestions throughout the year. A special thank goes also to Dr Pavel Gladyshev at the Department of Computer Science, University College Dublin for his valuable suggestions and insight regarding my project and for his comments in relation to several drafts. I also extend my special thank to Mr. Mohamad Afendee Mohamed for his help and comments.

Finally, I would like to express my deepest gratitude for the constant support, understanding, sacrifice, patience and love that I received from my beloved wife, sons and daughters, without which this thesis would not have been possible.

I certify that an Examination Committee has met on 25 January 2011 to conduct the final examination of Mohd Taufik b Abdullah on his degree thesis entitled "Method of Event Reconstruction in Digital Investigation and Its Visualization" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Examination Committee were as follows:

Md. Nasir Sulaiman, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Ali Mamat, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Hamidah Ibrahim, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Frederick Charles Piper
Emeritus Professor
Royal Holloway (University of London)
Information Security Group, Egham, Surrey
TW20 0EX, UK
(External Examiner)

NORITAH OMAR, PhD
Associate Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Ramlan Mahmod, PhD
Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Abdul Azim Abd. Ghani, PhD
Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Abdullah Mohd Zin, PhD
Professor
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
(Member)

_____

HASANAH MOHD GHAZALI, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

I declare that the thesis is my own work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently submitted for any other degree at Universiti Putra Malaysia or at any other institution.

MOHD TAUFIK ABDULLAH

Date: 25 January 2011

TABLE OF CONTENTS

		Page
DEDICATION		ii
ABSTRACT		iii
ABSTRAK		iv
ACKNOWLEDGEMENTS		v
APPROVAL		vi
DECLARATION		vii
LIST OF TABLES		xvi
LIST OF FIGURES		xviii
LIST OF ABBREVIATIONS		xxii
CHAPTER		
1	**INTRODUCTION**	
1.1	Background	1.1
1.2	Problem Statement	1.5
1.3	Objectives of Research	1.7
1.4	Scope of Research	1.8
1.5	Contributions of Research	1.9
1.6	Organization of the Thesis	1.10
2	**EVENT RECONSTRUCTION IN COMPUTER FORENSICS**	
2.1	Introduction	2.1
2.2	Digital Evidence	2.2
2.2.1	Classes of Digital Evidence	2.2
2.2.2	Digital Information Context	2.3
2.2.3	Digital Information Obscurity	2.4
2.2.4	Automation of Digital Information Interpretation	2.4
2.2.5	Risk of Contaminated Information	2.4
2.3	Definition of Digital Investigation	2.5
2.4	Digital Investigation Process Models	2.5
2.5	Examination and Analysis Techniques	2.7
2.5.1	Search Methods	2.7
2.5.2	Reconstruction of Events	2.9
2.5.3	Time Analysis	2.16
2.6	The Need of Event Reconstruction Theory in Computer Forensics	2.18
2.7	State of the Art Event Reconstruction	2.18
2.7.1	Attack Trees	2.19
2.7.2	Visual Investigative Analysis	2.21
2.7.3	Multilinear Event Sequencing	2.24
2.7.4	Why-Because Analysis	2.27
2.7.5	Root Cause Analysis	2.29
2.7.6	Finite State Machine Approach to Digital Event Reconstruction	2.30
2.7.7	A Hypothesis-based Approach to Digital Forensic Investigation	2.33

2.8 Summary 2.36

3 **FINITE STATE MACHINE**
3.1 Introduction 3.1
3.2 Concepts of Finite State Machine Theory 3.1
3.2.1 Alphabets 3.2
3.2.2 Strings 3.2
3.2.3 Language 3.3
3.3 Finite State Machine Model of Computation 3.4
3.3.1 Basic Finite State Machine (FSM) Model and its Variations 3.5
3.3.2 System Models Creation 3.8
3.3.3 Intersection of Finite State Machine 3.11
3.3.4 Finite Computations Analysis 3.11
3.4 Regular Expression 3.12
3.4.1 Equivalence of Regular Expression and Finite Automata 3.12
3.4.2 Converting Regular Expression to FA 3.13
3.4.3 Converting a Regular Expression to Non-determine FA 3.14
3.4.4 Eliminating Epsilon Transitions 3.15
3.5 Summary 3.17

4 **RESEARCH METHODOLGY**
4.1 Introduction 4.1
4.2 Problem Identification 4.1
4.3 Data Requirement 4.2
4.4 Determination of Performance Measurement 4.4
4.4.1 Effective of Event Reconstruction 4.5
4.4.2 Efficiency of Event Reconstruction 4.6
4.4.3 Legal Admissibility of Event Reconstruction 4.6
4.5 Design of a Method of Event Reconstruction 4.7
4.6 Implementation and Generation of Results 4.8
4.7 Analysis and Documentation 4.9
4.8 Summary 4.9

5 **DESIGN OF FSM MODEL OF DIGITAL EVENT RECONSTRUCTION**
5.1 Introduction 5.1
5.2 Formalization of Digital Event Reconstruction Problem 5.2
5.2.1 Representing the Knowledge of the System Functionality as an FSM 5.3
5.2.2 DFA Model of Computation of an FSM System Model 5.4
5.2.3 Representing the Evidence 5.5
5.3 Digital Event Reconstruction Framework 5.15
5.4 Construction of an FSM System Model 5.16
5.5 Construction of DFA Accepting Computations of an FSM System Model 5.23
5.6 Construction of a DFA of Evidence 5.27
5.7 Computing the intersection of DFAs 5.32
5.8 Summary 5.33

6 **DESIGN OF VISUALIZATION OF DIGITAL EVENT RECONSTRUCTION**
6.1 Introduction 6.1
6.2 Generating DFA Graph *Dot* Script 6.2
6.2.1 Generating *Dot* Script for Labeling the Nodes 6.3
6.2.2 Generating *Dot* Script for Arcs 6.7
6.2.3 Generating *Dot* Script for DFA Graph 6.10
6.3 Visualizing the DFA Graph 6.13
6.4 Summary 6.15

7 **RESULTS AND DISCUSSION**
7.1 Introduction 7.1
7.2 Experiment Design 7.2
7.2.1 Data Preparation 7.3
7.2.2 Performance Measurement for Analysis 7.3
7.2.3 Result Generated 7.8
7.3 Result and Analysis from Case Study 1: Two-bit Counter System Analysis 7.8
7.3.1 Formalization of System Behavior 7.9
7.3.2 Proving the Correctness of DFA C1 Representation for Two-bit Counter System Behavior 7.11
7.3.3 Formalization of Evidence 7.12
7.3.4 Proving the Correctness of Representation of Examiner Claim 7.13
7.3.5 Intersection of DFAs 7.14
7.3.6 Proving the Correctness of DFAs Intersection 7.15
7.4 Result and Analysis Case Study 2: Networked Printer Analysis 7.21
7.4.1 The Dispute 7.21
7.4.2 The Investigation 7.22
7.4.3 The Analysis 7.22
7.4.4 Informal Analysis with a State Machine 7.23
7.4.5 Formalization of System Behavior 7.25
7.4.6 Formalization of Evidence 7.26
7.4.7 Intersection of DFAs 7.31
7.4.8 Query Information from Graph 7.33
7.4.9 Proving the Correctness of DFA System Behavior, DFAs Representation of Witness Statement, and DFAs of Intersection 7.33
7.5 Result and Analysis Case Study 3: Blackmail Analysis 7.40
7.5.1 The Forensic Examination 7.41
7.5.2 Formalization of Last Cluster of the File 7.23
7.5.3 Formalization of Evidence 7.48
7.5.4 Intersection of DFA of Last Cluster and DFA of Evidence 7.50
7.5.5 Proof the Correctness of DFA 7.55
7.6 Comparison with Existing Method 7.63
7.7 Discussion 7.67
7.8 Summary 7.69

8 **CONCLUSIONS AND FUTURE WORK**
8.1 Introduction 8.1
8.2 Conclusion 8.1
8.3 Future Work 8.4

8.3.1 Investigating New Ways for Constructing Evidence Model 8.5
8.3.2 Investigating New Approach for Constructing System Model 8.5
8.3.3 Trusted Computing 8.5

REFERENCES R.1
APPENDICES A.1
A Basic Mathematic Objects and Notations of a Finite State Machine Theory A.2
B Tables of Arbitrary Computations A.6
BIODATA OF STUDENT B.1
LIST OF PUBLICATIONS C.1