



UNIVERSITI PUTRA MALAYSIA

***CHARACTER PROPERTY METHOD WITH BIOMETRIC MULTIFACTOR
AUTHENTICATION FOR ARABIC TEXT STEGANOGRAPHY***

NUUR ALIFAH BINTI ROSLAN

FSKTM 2020 29



**CHARACTER PROPERTY METHOD WITH BIOMETRIC MULTIFACTOR
AUTHENTICATION FOR ARABIC TEXT STEGANOGRAPHY**

By

NUUR ALIFAH BINTI ROSLAN

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

June 2018

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



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

**CHARACTER PROPERTY METHOD WITH BIOMETRIC MULTIFACTOR
AUTHENTICATION FOR ARABIC TEXT STEGANOGRAPHY**

By

NUUR ALIFAH BINTI ROSLAN

June 2018

Chairman : Nur Izura Udzir, PhD
Faculty : Computer Science and Information Technology

Text steganography is an ancient means of secret communication that uses the text hiding process to conceal a message and, when combined with cryptography, enhances its level of security. However, it is limited in its ability to optimize embedded data capacity with a high perceptual transparency level that will also not raise suspicion when written. Besides that, other concerns are active attacks by intruders which are a crucial security issue in the transmission of the shared secret key that enables the receiver to extract the secret information. Also, such attacks can be infected through a fake identity that allows the receiver to modify the secret information thus degrading its integrity. To overcome these drawbacks, we propose the Character Property method, which uses the basic properties of the Arabic Text such as dots, calligraphy typographical proportions, and sharp-edges to hide the secret message using a table index mapping technique to optimize data capacity with high perceptual transparency to avert suspicion. We apply biometric multi factor authentication to enhance the security of the transmitted shared secret key used to extract the stego-text. The designed biometric multi factor authentication has a liveness detection feature to spot a receiver's fake identity. The biometric multi factor authentication is implemented through a custom Arduino smart watch with a fingerprint and heartbeat sensor as a proof of concept device which increases capacity in hiding the secret message by up to 23.5% compared to the previous methods. Since the designed method does not affect the stego-text appearance, its 1.0 Jaro Similarity score as compared to the other methods proves the high transparency of the stego-text. The biometric device evaluation results in a false rejection rate of only 4% while the false acceptance rate is 0%. The results are significant for the liveness detection with 0% results for both false acceptance of fake inputs (FerrFake) and

false rejection of live subject (FerrLive) compared with a fingerprint-only biometric authentication approach which has a high percentage of up to 13% of false acceptance of fake inputs (FerrFake). To conclude, the Character Property method with biometric multi factor authentication provides an optimum embedded data capacity and a high level of perceptual transparency in hiding secret information together with a high level of user authorization that offers the liveness detection of users. This method with biometric multifactor authentication offers a new perspective on Arabic text steganography to cover both passive and active attack issues.



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

**KAEDAH CIRI AKSARA DENGAN PENGESAHAN BIOMETRIK MULTI
FAKTOR UNTUK STEGANOGRAFI TEKS ARAB**

Oleh

NUUR ALIFAH BINTI ROSLAN

Jun 2018

Pengerusi : Nur Izura Udzir, PhD
Fakulti : Sains Komputer dan Teknologi Maklumat

Tulisan steganografi adalah kaedah purba untuk berkomunikasi secara rahsia, yang menyembunyikan mesej rahsia, jika dikombinasikan dengan kriptografi, akan meningkatkan tahap keselamatannya. Walau bagaimanapun, kemampuannya terhad untuk mengoptima kapasiti data dengan tahap ketelusan persepsi yang tinggi yang juga tidak akan menimbulkan syak bila ditulis. Selain itu, kebimbangan lain adalah serangan luar oleh penceroboh yang merupakan isu keselamatan yang serius semasa pemindahan kunci rahsia yang dikongsi, yang membolehkan penerima mengekstrak maklumat rahsia tersebut. Serangan sebegini juga boleh dilakukan melalui identiti palsu yang membolehkan penerima mengubah maklumat rahsia tersebut yang merosakkan integritinya. Bagi mengatasi masalah ini, kami mencadangkan kaedah Ciri Aksara yang menggunakan ciri asas teks Arab seperti titik, perkadaran tipografi kaligrafi, dan hujung-tajam untuk menyembunyikan mesej rahsia menggunakan teknik pemetaan indeks jadual untuk mengoptimumkan kapasiti data dengan ketelusan persepsi yang tinggi untuk mengelakkan syak. Kami menggunakan pengesanan multi faktor biometrik untuk meningkatkan keselamatan pemindahan kunci rahsia yang dikongsi yang digunakan untuk mengekstrak teks-stego. Pengesanan multi faktor biometrik ini mempunyai ciri pengesanan hidup bagi mengesan identiti palsu penerima. Dalam eksperimen kami, pengesanan multi faktor biometrik ini diimplementasi pada jam tangan pintar Arduino suai langgan yang mengandungi pengesan cap jari dan degupan jantung sebagai peranti prototaip yang meningkatkan kapasiti menyembunyikan mesej rahsia sehingga 23.5% berbanding kaedah terdahulu. Memandangkan kaedah yang direkabentuk ini tidak memberi kesan kepada rupa teks-stego, skor 1.0 Kesamaan Jaronya berbanding kaedah lain membuktikan ketinggian ketelusan teks-stego tersebut. Penilaian peranti biometrik ke atas prototaip kami tersebut

menunjukkan kadar penolakan palsu adalah hanya 4% manakala kadar penerimaan palsu adalah 0%. Keputusan signifikan bagi pengesanan hidup dengan keputusan pengecaman 0% untuk kedua-dua penerimaam palsu bagi input palsu (Ferrfake) dan penolakan palsu bagi subjek hidup (FerrLive) berbanding pendekatan biasa pengesanan biometrik menggunakan pengecaman cap jari sahaja yang mempunyai peratusan tinggi sehingga 13% penolakan palsu bagi input palsu (FerrFake). Sebagai kesimpulan, kaedah Ciri Aksara dengan pengesanan multi faktor biometrik ini mampu memberikan kapasiti data terbenam yang optima dan tahap ketelusan persepsi yang tinggi dalam menyembunyikan maklumat rahsia dengan tahap sah kuasa pengguna yang tinggi yang menawarkan pengesanan pengguna hidup. Kaedah dengan pengesanan multi faktor biometrik ini memberikan perspektif baharu dalam steganografi teks Arab untuk merangkumi isu-isu serangan aktif dan pasif.

ACKNOWLEDGEMENTS

Praise to Almighty Allah for His shower of blessings, I am able to complete this study. I wish to express my sincere and deep gratitude to my former Supervisor Professor Dr. Ramlan bin Mahmod, my Supervisors Assoc. Prof Dr. Nur Izura Udzir and my co-supervisors, Assoc. Prof Dr. Zuriati Ahmad Zulkarnain and Dr. Mohd Izuan Hafez Ninggal for their guidance, patience, and support in helping me overcome the various obstacles I faced during my research.

My special gratitude goes to my dearest friends who were always there to motivate each other, extend a helping hand whenever needed, and to share their thoughts despite their own busy schedules and research.

My sweetest appreciation goes to my dearest family, especially my beloved mother and life-coach, for her prayers and my father for his affectionate support and encouragement. Finally, I wish to thank my siblings and other family members for their understanding and good wishes and constantly helping me to be strong, especially in difficult times.

I certify that a Thesis Examination Committee has met on 29 June 2018 to conduct the final examination of Nuur Alifah binti Roslan on her thesis entitled “Character Property Method with Biometric Multifactor Authentication for Arabic Text Steganography” 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 Thesis Examination Committee were as follows:

Mohd Taufik b Abdullah, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Nor Fazlida binti Mohd Sani, PhD

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

Sharifah Mumtazah Syed Ahmad Abdul Rahman, PhD

Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Miss Laiha Mat Kiah, PhD

Professor
Faculty of Computer Science and Information Technology
Universiti Malaya Malaysia
(External Examiner)

ZURIATI AHMAD ZUKARNAIN, PhD

Professor Ts. and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date : 02 October 2020

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:

Nur Izura binti Udzir, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Zuriati Ahmad Zukarnain, PhD

Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member of Supervisory)

Mohd Izuan Hafez Bin Ninggal, PhD

Senior Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member of Supervisory)

ZALILAH MOHD SHARIFF, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 08 October 2020

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: Nuur Alifah Binti Roslan, GS35552

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of Chairman
of Supervisory
Committee: _____

Signature: _____
Name of Member of
Supervisory
Committee: _____

Signature: _____
Name of Member of
Supervisory
Committee: _____

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF TABLES	xiii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xvi
CHAPTER	
1 INTRODUCTION	
1.1 Introduction	1
1.2 Problem Statement	2
1.3 Research Objectives	3
1.4 Research Contributions	4
1.5 Research Scope	5
1.6 Thesis Organization	5
2 LITERATURE REVIEW	
2.1 Introduction	7
2.2 Steganography in Information Hiding	7
2.2.1 Steganography Protocols	10
2.2.2 Capacity, Robustness and Perceptual Transparency in Steganography	11
2.3 Text Steganography	12
2.4 Text Steganography in various languages and scripts of writing	17
2.4.1 Arabic Text Steganography	
2.4.2 Previous work in Arabic Text Steganography	17
2.5 Performance Measurement in Text Steganography	20
2.5.1 Capacity Measurement	26
2.5.2 Perceptual Transparency Measurement	26
2.5.3 Robustness Measurement	27
2.6 Hybrid Steganography	27
2.6.1 Steganography Biometric Hybrid	30
2.7 Biometric Authentication	32
2.7.1 Biometric Authentication	34
2.7.2 Fake Fingerprint issued & livness elements for fingerprint	38
2.8 Biometric Multifactor Authentication with liveness detection	39
2.9 Research Gap & Character Property Method with a Biometric Multifactor Authentication	41

3	MATERIALS AND METHODS / METHODOLOGY	
3.1	Introduction	44
3.2	Research Methodology Process	44
3.2.1	Phase 1: Problem Identification & Requirement Analysis	46
3.2.2	Phase 2: Design & Implementation	46
3.2.3	Phase 3: Experiments & Evaluation	47
3.3	Summary	49
4	CHARACTER PROPERTY METHOD WITH A BIOMETRIC MULTIFACTOR AUTHENTICATION SYSTEM DESIGN & IMPLEMENTATION	
4.1	Introduction	50
4.2	Character Property with Biometric Multifactor Authentication System Design	50
4.2.1	Biometric Multifactor Authentication System Design	52
4.2.2	Character Property Method System Design	55
4.3	The Implementation of Character Property with Biometric Multifactor Authentication System Design	66
4.3.1	The Authentication Process	68
4.3.2	The Hiding Process	72
4.3.3	The Retrieving Process	75
4.5	Summary	76
5	EXPERIMENTAL DESIGN	
5.1	Introduction	77
5.2	Text Steganography Experimental Design	77
5.2.1	Capacity Experimental Design	78
5.2.2	Perceptual Transparency Experimental Design	82
5.2.3	Robustness Experimental Design	83
5.3	Biometric Multifactor Authentication Experimental Design	83
5.3.1	Data Logger Comparison Experiment	83
5.3.2	Biometric Multifactor Authentication Evaluation	84
5.4	Summary	90
6	RESULTS & DISCUSSIONS	
6.1	Introduction	91
6.2	Arabic Text Steganography Evaluation Results & Analysis	91
6.2.1	Capacity Evaluation Results & Analysis	91
6.2.2	Perceptual Transparency Evaluation Results & Analysis	96
6.2.3	Robustness Evaluation Results & Analysis	97

6.3	Biometric Multifactor Authentication Evaluation Results & Analysis	97
6.3.1	The Hardware's Data Logger Results	97
6.3.2	Biometric Multifactor Authentication Evaluation	98
6.4	Summary	100
7	CONCLUSION & FUTURE WORK	
7.1	Introduction	101
7.2	Conclusion	101
7.3	Future Works	102
	REFERENCES	103
	APPENDICES	110
	BIODATA OF STUDENT PUBLICATION	112
		113

LIST OF TABLES

Table		Page
2.1	Reflection of English Alphabet Text Steganography	14
2.2	Word mapping method	15
2.3	Advantages and disadvantages of Text Steganography methods	16
2.4	The Arabic Alphabets	18
2.5	The Arabic Alphabets with four types of shapes in writing.	19
2.6	Hiding secret bit using Kashida	20
2.7	Arabic Letters with the respect of the number of the points	21
2.8	Relation of Shifting And Distance Of The Letter Format	21
2.9	DNA Code mapping with Arabic Diacritics	24
2.10	Primitive Structural method	25
2.11	Advantages and disadvantages of Arabic Text Steganography methods	28
2.12	DNA Encryption Key	30
2.13	Hybrid Steganography	32
2.14	Type of authentication and its independent credentials	32
4.1	Character Property	56
4.2	Random Unique Number	63
4.3	Generated random unique number table	75
5.1	The Cover Text with different length data set	79
5.2	The secret message with different length data set	80
5.3	Experiment 2 Cover Text Data Set	81
5.4	Perceptual transparency data set comparison	82
5.5	Robustness evaluation	83
5.6	Multi-Factor Cross Check Authentication	86
5.7	FAR & FRR results calculation reference	87
6.1	Group of the basic shapes for Primitive Structural method	93
6.2	Jaro Score of the string pars of the cover file (C) and stego file (S)	96
6.3	96	96
6.4	Robustness evaluation result	97
6.5	FRR & FAR results	99
6.6	Threshold	99
6.7	FFR & FAR Results with threshold	99

LIST OF FIGURES

Figure		Page
2.1	A classification of Information Hiding techniques	8
2.2	Modern Steganography	9
2.3	The Prisoner's Problem	10
2.4	General Steganography Protocol	11
2.5	Magic Triangle	12
2.6	Categories of Text Steganography	13
2.7	Arabic Diacritics	18
2.8	Vertical Shifting Point	21
2.9	The right image is the reverse Fathah	22
2.10	Biometric authentication system architecture	33
2.11	Fingerprint characteristic	35
2.12	Some common minutiae types	35
2.13	Fingerprint recognition system block diagram	36
2.14	Finger print spoofing method	38
2.15	Fingerprint and RFID multifactor authentication system	41
3.1	Research Methodology Process	45
3.2	Design & Implementation	46
3.3	Evaluation and Analysis	48
4.1	Character Property with Biometric Multifactor Authentication framework data flow diagram	51
4.2	Biometric Multifactor Authentication Character Property Algorithm Flow Chart	52
4.3	Enrolment Module Algorithm	53
4.4	Verification Module Algorithm	53
4.5	Biometric Authentication Algorithm	55
4.6	Arabic script characteristic for Character Property method	57
4.7	Architecture of Character Property method	57
4.8	Hiding Module Flow Chart	58
4.9	Validation Process Pseudocode.	59
4.10	Bit Hiding Process Flow Chart	60
4.11	Character Property Algorithm	60
4.12	Character Property Algorithm process with examples	61
4.13	Maximum Positioning Algorithm	61
4.14	Maximum Positioning Algorithm with examples	62
4.15	Secret Key Embedding Algorithm	64
4.16	The Retrieving Module Flow Chart	65
4.17	The Retrieving process algorithm	65
4.18	The Smartwatch Biometric Arduino system diagram	67
4.19	The Smartwatch Biometric Arduino device	67
4.20	Character Property Method software system interface	68
4.21	Bluetooth connector and the Smartwatch Biometric Device	68
4.22	User Interface for Bluetooth Connection	69
4.23	User Enrollment Interface	69
4.24	Step by Step of the authentication process	70
4.25	Successfully Registered User	70
4.26	The User Authentication	71

4.27	Step by step of the User Authentication	72
4.28	Character Property Identification	73
4.29	Sorted Cover Text	74
4.30	Retrieving Secret Message	76
5.1	The fixed Secret Message data set	79
5.2	The fixed cover text	80
5.3	Oximeter Fingertips Pulse	84
5.4	Finger Plasticine mold and silicon model	88
5.5	Fake Fingerprint Making Process	89
6.1	Fixed Secret message with different Cover Text for Character Property method	93
6.2	Fixed Secret message with different Cover Text for Primitive Structural method	96
6.3	Different Secret message with Fixed Cover Text for Character Property method	96
6.4	Different Secret message with Fixed Cover Text for Primitive Structural method	97
6.5	Character Property method Capacity evaluation comparison with the Isolated Character method and the Run Length Code method.	99
6.6	Heart Beat reading comparison between Arduino heartbeat sensor and the Oximeter Pulse sensor	99
6.7	Error message for the modified inputs	99

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic
DCT	Discrete Cosine Transform
DFT	Discrete Frequency Transform
DWT	Discrete Wavelet Transform
OCR	Optical Character Recognition
LSB	Least Significant Bits
CSS	Cascading Style Sheets
HTML	Hypertext Markup Language
EOL	End of Line
LRMCA	Linear Reversible Memory Cellular Automata
AES	Advanced Encryption Standard
MSCUKAT	Maximizing Steganography Capacity Using "Kashida" In Arabic Text
MD5	Message Digest 5
PIN	Personal Identification Number
PKI	Public Key Infrastructure
ATM	Automated Teller Machine
RFID	Radio Frequency Identification

CHAPTER 1

INTRODUCTION

1.1 Introduction

Steganography is a sub-discipline in information hiding relating to the implementation of secret communications. It is an ancient form of secret communication that was developed in Greece around 440 B.C. It is defined as “covered writing” and derived from the ancient Greek root word *steganos* (στεγανός) meaning “covered or protected”, and *graphei* (γραφή) meaning “writing”. Throughout history, most steganography applications were meant for military purposes that employed this covert form to communicate and to relay important secret information.

The first example of this ancient technique was used by a Greek general who wrote secret information on the shaved head of a trusted messenger and sent him to the receiver once his hair had fully grown back to completely cover the message. The technique evolved over time and during World War 2, the German army used invisible ink to write secret information in their covert communications (Fridrich, 2009).

The basic steganography elements comprise of a secret message and a cover medium in which various techniques are used to embed the former into the latter. The steganography media will be the result of the combination between the secret message and the cover medium. Then, the secret communication is completed by sending the steganography media to the other parties with a key or technique to reveal the secret message. Only the recipient knows the keys or the technique used to conceal the secret message. On the other hand, the stego object has to have a high degree of perceptual transparency to not raise any public suspicion during transmission.

The main goal of steganography is to have secret public information, the capacity or place to hide the secret information, and a robust steganography media against attacks by eavesdroppers. Besides the effectiveness of the steganography technique, another aspect that needs to be considered is *The Prisoners' Problem* which is a classical explanation concerning attacks through the transmission of a secret message. Therefore, the main issue is to design an efficient steganography algorithm to achieve a secure communication that has these main goals as a benchmark (Petitcolas, Anderson, & Kuhn, 1999).

Moreover, there is the need to consider possible active attacks in order to have high security in steganography communication. This can be achieved via a hybrid steganography having other layers of information and sub disciplines. Modern steganography has made a huge impact on secret digital communication since most users nowadays employ digital communications. In modern steganography, the cover medium is related to the digital media such as images, text, audio, video, and even animation. This offers greater benefits in steganography application which is of much importance nowadays due to the widespread need to protect private and intellectual properties especially for high profile persons rather than only focusing on military purposes.

The implementation of modern research steganography had broadened its benefits through applications in secret communication such as in gadgets that can be worn. These gadgets are mainly through affordable fitness bands and general-purpose smartwatches like the Apple Watch and Android Wear. These allow for the use of user-installable apps such as smartphones, which have a broad range of applications such as voice calls, contactless payments, unlocking the doors of cars or houses,

measuring vital signs, and for healthcare purposes (Blasco, Chen, Tapiador, Peris-Lopez, & Peris, 2016). According to some estimates, wearable sales will rise to 100 million units by 2020 (Paul Lee, Duncan Stewart, 2014).

1.2 Problem Statement

Text steganography is an ancient technique that has evolved over the years especially with the digital media. This category of steganography is the most difficult compared to other mediums especially image steganography due to the difficulty in finding redundant information in a file (Ali, 2010). Therefore, most text steganography approaches focus on optimizing embedded secret information capacity with the high perceptual transparency level of the stego-text which will not raise suspicions when published publicly. This is the main issue in creating a Text Steganography algorithm which is having a high capacity of the embedded secret information.

Text as an ancient information and communication medium has a large scope and variety of writing forms. Various techniques in text steganography have been introduced using different approaches based on the structure of the written script from different languages such as Thai (Samphaiboon & Dailey, 2008), Chinese (Xinmei Sun, Peng Meng, Yun Ye, & Liusheng Hang, 2010), and Hindi (Ali, 2010). These related works use various scripts of writing to optimize the capacity

of hiding the secret information. However, besides optimizing the capacity of the embedded secret information, the next concern is a high perceptual transparency of the stego-text to avoid suspicious of the steganography communication because, once its suspicious the steganography communication will collapse and this passive attacks is one of the main issue concerned in steganography (Johnson, Duric, & Jajodia, 2001). It is a challenge in having an optimum capacity with a high perceptual transparency text steganography algorithm as we highlighted these as our first main issue.

Therefore, since 2006, Arabic text steganography have offered the benefits of the scripts itself such as dots (M. H. Shirali-Shahreza & Shirali-Shahreza, 2006), La Words (M. Shirali-Shahreza, 2008), kashidah extension (Al-Nazer & Gutub, 2009), diacritics (A. A. Gutub, Ghouti, Elarian, Awaideh, & Alvi, 2010), and sharp edges (N.A Roslan, Mahmud, & Udzir, 2011). The Arabic script has a huge potential to be discovered since it has many potential characteristics for embed more secret information with a high perceptual transparency.

The Run Length encoding method by (Kadhem, 2016) both present a high capacity but have issues fragile by Optical Character Recognition system (OCR) where the removal of the empty spaces in the sentences results in the loss of the hidden secret information. Previous method by Khadim (2014) presented a mapping algorithm between DNA and diacritics based on the B+ Tree method and the Isolated present a high perceptual transparency as well as Word Map method by Bhattacharyya (2010). However, both have a low capacity in hiding the embedded secret information.

Abbasi (2015) presented an isolated character approach. The designed algorithm uses the Unicode approach replacement as a method to hide and provide high perceptual transparency in publishing the stego-text. However, the embedded secret information's data capacity is limited to the isolated character which is a waste over the other potential cover text which may collapse via the statistical text steganography steganalysis.

Secondly, the algorithm has shared secret keys for embedding and extracting the secret message. This provides a huge opportunity for intruders to manipulate the communication by creating a fake identity. Through such active attacks, the steganography communication can fail when the stego text and keys are modified. Through the critical review, besides the mentioned critical issues in text steganography, we identified the active attacks such as the fake identity of both users of the steganography communication. Therefore, through our hybrid steganography review a hybrid steganography with Biometric Multifactor Authentication using fingerprint with a liveness detection. Since the fake fingerprint issues for fingerprint authentication will be a main concern for the authentication layer (Sequeira & Cardoso, 2015).

1.3 Research Objectives

The main goal of this research is to propose a new framework of text steganography with biometric multifactor authentication, which optimizes the capacity of the embedded secret communication and a high level of perceptual transparency for stego-text. Therefore, the objectives of this research are as follows:

1. To propose an Arabic text steganography algorithm which optimizes the capacity of the embedded secret information.
2. To propose an Arabic text steganography method with a high level of perceptual transparency for stego-text besides having high capacity of the embedded secret communication.
3. To propose a biometric multifactor authentication text steganography to prove the liveness of the users in order to address the fake identity issue.

To achieve the above objectives, our design applied the following features as the main modules:

- a) **To propose a Characteristic Property Module**
The design method addresses the optimizing capacity of the embedded secret information capacity issue which arouses suspicions in text steganography by using Arabic character properties such as sharp edges, dots, and typo proportions.
- b) **To propose a Biometric Multifactor Authentication Module**
The new approach presented is the biometric element factor which is the heartbeat pulse combined with fingerprint biometric authentication to prove the liveness of the authorized user.

1.4 Research Contribution

This research will contribute to the following areas of study:

- a) Character Property Method with Biometric Multifactor Authentication framework for high level security of the Arabic text steganography. The framework consists of Character Property Method which is to optimize the capacity of the embedded secret information with a high level of perceptual transparency mostly to counter passive attacks and a biometric multifactor authentication module for a high level of security communication as a countermeasure to active attacks.
- b) The Character Property Method Algorithms an Arabic text steganography algorithm which optimizes the embedded capacity of

secret information with a high level of perceptual transparency of the stego-text. The Character Property method uses the index mapping technique to embed the secret information with the cover text. We designed it with a random position to hide secret bits using the basic characteristic properties of Arabic text such as dots, typo proportion, and sharp-edges which offer a huge capacity to hide secret bits per Arabic character. The method has two secret keys to conceal the secret message. Since our method is designed with symmetric key steganography, we have to create a holistic layer of security to overcome the issue of fake identity which we considered as an active attack.

- c) A new approach of Biometric Multi-Factor Authentication using a combination of fingerprint and heartbeat to prove the liveness of the authorize personnel and as a countermeasure to the fake identity issue. The proposed biometric multifactor module using the fingerprint will not only authenticate the sender and the receiver, but also indicate the liveness of the users by employing a new element of the authentication factor, that is, the user's heartbeat. This is due to concerns over spoofing attacks with fake samples of the fingerprint (Sequeira & Cardoso, 2015).

1.5 Research Scope

This research focuses on the Arabic Text steganography algorithm which optimized the capacity of the embedded secret information and increase the perceptual transparency of the cover text only and the robustness against the stego-text modification only.

This research not include the new design of authentication framework therefore the performance measurement such as efficiency and accuracy will not include. Since our research are hardware based of biometric multifactor authentication, the evaluation for the authentication part only cover the user acceptance evaluation only.

1.6 Thesis Organization

This thesis comprises seven chapters. **Chapter 1** provides the introduction of the thesis including the research problems, research objectives, and contributions of the research.

Chapter 2 has background information on work related to hiding and steganography, and then narrows down to Arabic text steganography. The literature review also includes the basics of Arabic character properties, the

hybrid type of text steganography, and a brief on the fundamentals of biometric multifactor authentication.

Chapter 3 shows the methodology and process used in the research. It explains the steps taken from the beginning till the final part of the research.

Chapter 4 introduces the proposed design of the Character Property method. The structure and design of the method consists of two main parts, i.e., the text steganography method and the biometric multifactor authentication. This chapter includes the implementation of the methods via a designed system.

Chapter 5 describes the experimental design and evaluation in testing the proposed design. It also explains how the experimental design is used to select the testing of input data and used data set. This chapter will include the performance measurement for text steganography and the evaluation of the designed biometric multifactor authentication. The results and the analysis of the conducted experiment will be discussed in Chapter 6.

Chapter 6 presents the results of the analysis section will cover the analysis for text steganography and the biometric multifactor authentication as based the conducted experiment in Chapter 5.

Finally, **Chapter 7** presents the conclusions of the research work carried out in this thesis. In addition, some future directions for exploration are proposed

REFERENCES

- Abbasi, Tabassum, A., Naqvi, S. N., Khan, A., & Ahmad, B. (2016). Urdu Text Steganography: Utilizing Isolated Letters Urdu text steganography: Utilizing isolated letters. *International Journal on Information Technologies and Security*, 9(1), 85–100.
- Abhishek, K., Roshan, S., Kumar, P., & Ranjan, R. (2013). A comprehensive study on multifactor authentication schemes. In *Advances in Computing and Information Technology* (pp. 561–568). Berlin, Heidelberg: Springer.
- Adale, D. A. (2011). *Hybrid Information security models: crypto-steg and steg-crypto systems*. Howard University.
- Agarwal, M. (2013). Text Steganographic Approaches: A Comparison. *International Journal of Network Security & Its Applications*, 5(1), 91–106.
- Al-Najjar, Y., & Sheta, A. (2008). Minutiae extraction for fingerprint recognition. In *Systems, Signals and Devices, 2008. IEEE SSD 2008. 5th International Multi-Conference* (pp. 1–5). IEEE.
- AL-Ani, Z. K., & A.A.Zaidan, B. B. Z. H. O. A. (2010). Overview: Main Fundamentals for Steganography. *Journal of Computing*, 2(3), 158–165.
- Al-Nazer, A., & Gutub, A. (2009). Exploit kashida adding to Arabic e-Text for high capacity steganography. In IEEE (Ed.), *hird International Conference on Network and System Security* (pp. 447–451).
- Ali, A. E. (2010). A New Approach to Hindi Text Steganography Using Matraye, Core Classification And HHK Scheme. In *Seventh International Conference on Information Technology* (pp. 1223–1224). IEEE.
- Azmi, A., & Alsaiari, A. (2010). Arabic Typography: A Survey. *International Journal of Electrical & Computer Sciences*, 9, 16–22.
- B.Sivaranjani, & N.Radha. (2017). Securing Patient's Confidential Information using ECG Steganography. In *2nd International Conference on Communication and Electronics Systems* (pp. 540–544). IEEE.
- Bajwa, I. S., & Riasat, R. (2011). A new perfect hashing based approach for secure stegnograph. In *Digital Information Management (ICDIM), 2011 Sixth International Conference* (pp. 174–178). IEEE.
- Banerjee, Indradip, Bhattacharyya, S., Mukherjee, S., & Sanyal, G. (2015). Biometric steganography using face geometry. In *TENCON 2014-2014 IEEE Region 10 Conference* (Vol. 2015–Janua, pp. 1–6). IEEE.
- Basilio-Ramirez, J., Perez-Meana, H., & Ponomaryov, V. (2016). Multifactor authentication system based on biometrics and radio frequency identification. In *Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW), 2016 9th International Kharkiv Symposium*

(pp. 1–4). IEEE.

- Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2010). A novel approach of secure text based steganography model using word mapping method (WMM). *International Journal of Computer and Information Engineering*, 4(2), 96–103.
- Blasco, J., Chen, T. M., Tapiador, J., Peris-Lopez, P., & Peris, P. (2016). A Survey of Wearable Biometric Recognition Systems. *ACM Computing Surveys (CSUR)*, 49(3), 1–43.
- Cheddad, A., Condell, J., Curran, K., & Kevitt, P. M. (2010). A skin tone detection algorithm for an adaptive approach to steganography. *Signal Processing*, 89(12), 2465–2478.
- Choi, H., Kang, R., Choi, K., & J. Kim. (2007). Aliveness Detection of Fingerprints using Multiple Static Features. In *World Academy of Science, Engineering and Technology* (Vol. 22).
- Dubey, R., Saxena, A., & Gond, S. (2015). An Innovative Data Security Techniques Using Cryptography and Steganographic Techniques. *International Journal of Computer Science and Information Technologies*, 6(3), 2175–2183.
- El-abad, M., Charrier, C., El-abad, M., Charrier, C., Systems, B., & El-abad, M. (2014). Evaluation of Biometric Systems. *IEEE International Conference on Hand-Based Biometrics (ICHB)*, 149–169.
- Espinoza, M., & Champod, C. (2011). Risk evaluation for spoofing against a sensor supplied with liveness detection. *Forensic Science International*, 204(1–3), 162–168.
- Espinoza, M., Champod, C., & Margot, P. (2011). Vulnerabilities of fingerprint reader to fake fingerprints attacks. *Forensic Science International*, 204(1–3), 41–49.
- Fridrich, J. (2009). *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press.
- Gil, Y., Moon, D., Pan, S., & Chung, Y. (2002). Fingerprint Verification System Involving Smart Card. In *Information Security and Cryptology - ICISC*.
- Girgis, M. R., Mahmoud, T. M., & Abd-El-Hafeez, T. (2007). An Approach to Image Extraction and Accurate Skin Detection from Web Pages. *World Academy of Science, Engineering and Technology*.
- Go, W., Lee, K., & Kwak, J. (2014). Construction of a secure two-factor user authentication system using fingerprint information and password. *Journal of Intelligent Manufacturing*, 25(2), 217–230.
- Gustavus J. Simmons. (1985). The prisoners problem and the subliminal channel. In *Advances in Cryptology* (pp. 51–67). Springer.

- Gutub, A. A., Ghouti, L. M., Elarian, Y. S., Awaideh, S. M., & Alvi, A. K. (2010). Utilizing Diacritic Marks for Arabic Text Steganography. *Kuwait Journal of Science & Engineering (KJSE)*, 37(1), 89–109.
- Gutub, A., Al-alwani, W., & Mahfoodh, A. Bin. (2010). Improved Method of Arabic Text Steganography Using the Extension “Kashida” Character. *Bahria University Journal of Information & Communication Technology*, 3(1), 68–72.
- Gutub, A., Al-Haidari, F., Al-Kahsah, K. M., & Hamodi, J. (2010). e-Text Watermarking: Utilizing “Kashida” Extensions in Arabic Language Electronic Writing. *Journal of Emerging Technologies in Web Intelligence*, 2(1), 48–55.
- Harshitha, K. M., & Vijaya, P. A. (2012). Secure Data Hiding Algorithm Using Encrypted Secret Message. *International Journal of Scientific and Research Publications*, 2(6), 1–4.
- Hart, J. (2015). Normal resting pulse rate ranges. *Journal of Nursing Education and Practice*, 5(8), p95.
- Hori, M., & Okamoto, H. (2012). Heart rate as a target of treatment of chronic heart failure. *Journal of Cardiology*, 60(2), 86–90.
- Indrayani, R., Nugroho, H. A., Hidayat, R., & Pratama, I. (2017). Increasing the security of MP3 steganography using AES Encryption and MD5 hash function. In *Proceedings - 2016 2nd International Conference on Science and Technology-Computer, ICST 2016*.
- Jain, K., A., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20.
- Jia, H., & Cao, K. (2012). The research on the preprocessing algorithm for fingerprint image. In *IEEE Symposium on Electrical & Electronics Engineering* (pp. 163–166).
- Johnson, N. F., Duric, Z., & Jajodia, S. (2001). *Information Hiding: Steganography and Watermarking-Attacks and Countermeasures*. Boston: Kluwer Academic Publisher.
- Kadhem, S. M. (2016). Text Steganography Method Based On Modified Run Length Encoding. *Iraqi Journal of Science*, 57(3), 2338–2347.
- Karnan, M., Akila, M., & Krishnaraj, N. (2011). Biometric personal authentication using keystroke dynamics: A review. *Applied Soft Computing Journal*, 11(2), 1565–1573.
- Kaur, M., Singh, M., Girdhar, A., & Sandhu, P. S. (2008). Fingerprint Verification System using Minutiae Extraction Technique. *World Academy of Science, Engineering and Technology*, 46.

- Kermanidis, K. (2011). Capacity-rich Knowledge-poor Linguistic steganography. *Journal of Information Hiding and Multimedia Signal Processing*, 2(3), 247–258.
- Khadim, E. A., Abdulwahab, H. B., & Kadhem, S. M. (2014). Proposed Approach for Steganography in Arabic Text Based on B+ Tree , DNA Coding and Arabic Diacritics, 2(12), 954–965.
- Krieg, M., & Rogmann, N. (2015). Liveness Detection in Biometrics. In *Biometrics Special Interest Group (BIOSIG), 2015 International Conference*. IEEE.
- Le, H., & Bui, T. D. (2009). Online fingerprint identification with a fast and distortion tolerant hashing. *Journal of Information Assurance and Security*, 4, 117–123.
- Lunji Qiu. (2014). Sensor Tehnology. In *2014 IEEE 9th Conference on Industrial Electronics and Applications (ICIEA)* (pp. 1433–1436).
- Luo, X., Tian, J., & Wu, Y. (2000). minutia matching algorithm in fingerprint verification. In *International Conference on Pattern Recognition* (pp. 833–836).
- Majumder, A., & Changder, S. (2013). A Novel Approach for Text Steganography: Generating Text Summary Using Reflection Symmetry. *Procedia Technology*, 10, 112–120.
- Maltoni D, Jain AK, Maio D, & Prabhakar S. (2009). *Handbook of Fingerprint Recognition*. Springer.
- Mohamed, T. S. (2014). Information and Knowledge Management Security of Multifactor Authentication Model to Improve Authentication Systems, 4(6). Retrieved from www.iiste.org
- Nikam, S., & S. Agarwal. (2008). Ridge let-based fake fingerprint detection. *Neurocomputing*, 72, 2491–2506.
- Ntalianis, K., & Tsapatsoulis, N. (2015). Remote Authentication via Biometrics: A Robust Video-Object Steganographic Mechanism Over Wireless Networks. *Ieee Transactions On Emerging Topics In Computing*, 4(1), 156–174.
- Parvez, M. T., & Mahmoud, S. a. (2013). Offline arabic handwritten text recognition. *ACM Computing Surveys*, 45(2), 1–35.
- Paul Lee, Duncan Stewart, J. B. (2014). *Deloitte TMT Prediction 2014 Technical Report*. London, United Kingdom.
- Petitcolas, F. a. P. F., Anderson, R. J., & Kuhn, M. G. (1999). Information hiding-a survey. In *Proceedings of the IEEE* (Vol. 87, pp. 1062–1078).
- Rajiv Mukherjee. (2007). *Indexing Techniques for Fingerprint and Iris Databases*. West Virginia University.

- Rasmussen, K. B. (2014). Authentication Using Pulse-Response Biometrics. In *NDSS* (pp. 23–26).
- Robert Batie Jr., Yair Levy, Steven Furnell, & Peixiang Liu. (2015). Improving User Authentication with Fingerprint Biometrics and Biometric Personal Identification Number (BIO-PINTM) as a Multi-Factor Authentication Mechanism. In *Pre-ICIS Workshop on Information Security and Privacy (SIGSEC)*.
- Roslan, N. A., Mahmud, R., Udzir, N. I., & Zurkarnain, Z. A. (2014). Primitive structural method for high capacity text steganography. *Journal of Theoretical and Applied Information Technology*, 67(2).
- Roslan, N. A., Mahmud, R., & Udzir, N. U. R. I. (2011). Sharp-Edges Method in Arabic Text Steganography. *Journal of Theoretical and Applied Information Technology*, 33(1).
- S. D. Pandya, P. V. Virparia. (2009). Testing various similarity metrics and their permutations with clustering approach in context free data cleaning. *Int. Journal of Computer Science and Security*, 3, 344–350.
- Samir, P., & Bandyopadhyay, K. (2010). A Method for Public Key Method of Steganography. *International Journal of Computer Applications*, 6(3), 4–7.
- Samphaiboon, N., & Dailey, M. N. (2008). Steganography in Thai text. In *5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008* (Vol. 1, pp. 133–136).
- Saravanan, D., Doni, A., & Ajith, A. (2013). Image Information Hiding: An Survey. *The SIJ Transactions on Computer Science Engineering & Its Applications (CSEA)*, 1(1), 1–7.
- Sequeira, A. F., & Cardoso, J. S. (2015). Fingerprint liveness detection in the presence of capable intruders. *Sensors (Switzerland)*, 15(6), 14615–14638.
- Shirali-Shahreza, M. (2008). A New Persian / Arabic Text Steganography Using “ La ” Word. *Advances in Computer and Information Sciences and Engineering*, 2008.
- Shirali-Shahreza, M. H., & Shirali-Shahreza, M. (2006). A New Approach to Persian/Arabic Text Steganography. In *5th IEEE/ACIS International Conference on Computer and Information Science and 1st IEEE/ACIS International Workshop on Component-Based Software Engineering, Software Architecture and Reuse (ICIS-COM SAR'06)* (pp. 310–315). IEEE.
- Shirali-Shahreza, M. H., & Shirali-Shahreza, M. (2010). Arabic/Persian text steganography utilizing similar letters with different codes. *The Arabian Journal for Science and Engineering*, 35(1), 213–222.

- Sousedik, C., & Busch, C. (2014). Presentation attack detection methods for fingerprint recognition systems: a survey. *IET Biometrics*, 3(4), 219–233.
- Suhad Malalla, A., & Shareef, M. R. (2016). Improving Hiding Security of Arabic Text Steganography by Hybrid AES Cryptography and Text Steganography. *Journal of Engineering Research and Application*, 6(65), 2248–962260. Retrieved from www.ijera.com
- Sujata, S., College, W., & Shefali, S. S. (2013). Cellular Automata For Crypt-Steganography. *International Journal of Advanced Technology & Engineering Research*, 3(1), 73–78.
- Tan, B., & Schuckers, S. (2008). New approach for liveness detection in fingerprint scanners based on valley noise analysis. *Journal of Electronic Imaging*, 17(1).
- Trewin, S., Swart, C., Koved, L., Martino, J., Singh, K., & Ben-david, S. (2012). Biometric Authentication on a Mobile Device : A Study of User Effort , Error and Task Disruption. In *Proceedings of the 28th Annual Computer Security Applications Conference* (pp. 159–168). ACM.
- Tripathi, K. P. (2011). A Comparative Study of Biometric Technologies with Reference to Human Interface, 14(5), 10–15.
- Vidhya, P. (2012). Text Steganography Using Public Key Cryptosystem in CSS. *International Journal of Computer Applications in ...*, 22(3), 234–237.
- W. E. Winkler. (1999). *The state of record linkage and current research problems*. Internal Revenue Service Publication.
- Wadehn, F., Carnal, D., & Loeliger, H. A. (2015). Estimation of heart rate and heart rate variability from pulse oximeter recordings using localized model fitting. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (Vol. 2015, pp. 3815–3818). IEEE.
- Wai, E., & Khine, M. (2011). Syntactic Bank-based Linguistic Steganography Approach. *International Conference on Information Communication and Management IPCSIT*, 15, 108–113.
- Wang, F., & Gao, G. (2011). Embedded Fingerprint Identification System Based on DSP Chip. *Advances in Computer Science, Intelligent System and Environment*, 595–599.
- Wang, Z., Zhao, X., Wang, H., & Cui, G. (2013). Information hiding based on DNA steganography. In *Software Engineering and Service Science (ICSESS), 2013 4th IEEE International Conference on* (pp. 946–949). IEEE.
- Wayman, J., Jain, A., Maltoni, D., & Maio, D. (n.d.). An Introduction to Biometric Authentication Systems.

Xinmei Sun, Peng Meng, Yun Ye, & Liusheng Hang. (2010). Steganography in Chinese Text. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)* (Vol. 8, pp. 651–654). IEEE.

Zainal, N. I., Sidek, K. A., Gunawan, T. S., Manser, H., & Kartiwi, M. (2014). Design and development of portable classroom attendance system based on Arduino and fingerprint Biometric. In *Information and Communication Technology for The Muslim World (ICT4M), 2014 The 5th International Conference* (pp. 1–4). IEEE.

Zhao, Q., Zhang, L., Zhang, D., & Nan Luo. (2008). Adaptive Pore Model for Fingerprint Pore Extraction. In *International Conference of Pattern Recognition 2008* (pp. 1–4). IEEE.



BIODATA OF STUDENT

Nuur Alifah Roslan was born in Malaysia on 24th December 1983. She received a Bachelor Degree in Computer Science (2006), and M.Sc. In Security In Computing (2011) from the Universiti Putra Malaysia (UPM). His professional working experience includes 2 years of service as programmer and a research assistant. She is currently a PhD candidate at the Faculty of Computer Science and Information Technology, UPM and receiving a scholarship from Tenaga Akademik Muda (TAM) UPM. Her current research interests include, steganography, information hiding and multimedia security.

PUBLICATION

Journal

Roslan, N. A., Mahmud, R., Udzir, N. I., & Zurkarnain, Z. A. (2014). Primitive Structural Method For High Capacity Text Steganography. *Journal Of Theoretical & Applied Information Technology*, 67(2).





UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : First Semester 2020/2021

TITLE OF THESIS / PROJECT REPORT :

CHARACTER PROPERTY METHOD WITH BIOMETRIC MULTIFACTOR AUTHENTICATION
FOR ARABIC TEXT STEGANOGRAPHY

NAME OF STUDENT: NUUR ALIFAH BINTI ROSLAN

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (v)

CONFIDENTIAL

(Contain confidential information under Official Secret Act 1972).

RESTRICTED

(Contains restricted information as specified by the organization/institution where research was done).

OPEN ACCESS

I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

PATENT

Embargo from _____ until _____
(date) (date)

Approved by:

(Signature of Student)
New IC No/ Passport No.:

Date :

(Signature of Chairman of Supervisory Committee)
Name:

Date :

[Note : If the thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]