ENHANCE KEYSTROKE DYNAMIC FOR ONLINE BASED SYSTEM

TEH TECK GUAN

MASTER OF COMPUTER SCIENCE (INTERNETWORKING TECHNOLOGY)

2017

Faculty of Information and Communication Technology

ENHANCE KEYSTROKE DYNAMIC FOR ONLINE BASED SYSTEM

Teh Teck Guan

Master of Computer Science (Internetworking Technology)

2017

ENHANCE KEYSTROKE DYNAMIC FOR ONLINE BASED SYSTEM

TEH TECK GUAN

A dissertation submitted
in fulfillment of the requirements for the degree of Master of Computer Science (Internetworking Technology)

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2017

DECLARATION

I declare that this dissertation entitle "Enhance Keystroke Dynamic for Online Based System" is the result of my own research except as cited in the references. The dissertation has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : ……

Name : Teh Teck Guan

Date : ……

APPROVAL

I hereby declare that I have read this dissertation and in my opinion this dissertation is sufficient in terms of scope and quality for the award of Master of Computer Science.

Signature : …….…….…….

Supervisor Name : Dr. Siti Rahayu Selamat

Date : ...……..……..

DEDICATION

First of all, I would like to take this chance to thank my parent (Teh Chin Yong and Gan Geok Suat) and siblings for their continually support, understanding, patience and encouragement throughout my Master of Science journey.

I also would like to take this chance to thank UTeM lecturer, especially Project Supervisor, Dr. Siti Rahayu, for her guidance and assistant on the project especially on the project problem definition and problem solving and also general view on this project. Secondly, I also take the chance to thank Prof. Dr. Burairah bin Hussin on project documentation.

I also would like to take the opportunity that to thank my employer, Infineon which fully sponsor in monetary support. Hence, I can fully focus on my Master study and completed my Msc thesis on schedule.

I also dedicate this MSc thesis to my colleagues and friends who have supported me along journey of the process which give me the idea and guidance.

ABSTRACT

Current password authentication system was proven not secure enough to protect the information from intruders. A number of researches have been done for the past 25 years, with various methods been used. However, various researches have its own method of collection data and analysis the data. The False Reject Rate and False Accept Rate can be high up to 30 % but overall averages have around 5%. One of the methods suggests is enhancing the current system using keystroke dynamics. Keystroke dynamics is a type of biometric authentication that not requires any special hardware. Beside it is easy to use, as the same routine as normal password authentication. Furthermore, keystroke dynamic has a great potential to replace the existing password system as it doesn't require any major changes at the current system, with only add in some more fields for the system and have supported software to use capture and match the user's typing patterns. Therefore, this research proposed an authentication system using keystroke dynamics in order to prevent the system from intruders. A system is developed that consist of two parts which are enrollment and verification. In enrollment part, the user will register into the system by typing necessary information including username, password and the standard phrase. In this part, the user typing time is recorded and a reference template is created. Then, the verification part will take placed in which the match of the similarity between the login and the reference template will be verified based on the threshold set in order to grant the user accessing the system. The proposed system is utilized the Microsoft.NET framework, keyboard and Microsoft IIS. A prototype is developed that consists of 3 main modules, namely Enrollment, Client/Server Connection and, Verification and Retraining. In Enrollment module, the user's typing data will be collected during the user registration process. The Client/Server Connection module provides a communication medium between the client program and also the system server. Then the user will be identified in the Verification and Retraining Module by comparing the login data with reference template. Any successful login data into reference data is recorded to increase the efficiency of the system. Based on the testing, the system proved that keystroke dynamic authentication system were able to implement in client/server environment. In future, the system can be improved by enhancing the security, performance, and user interface.

ABSTRAK

Sistem pengesahan kata laluan semasa telah terbukti tidak cukup selamat untuk melindungi maklumat daripada penceroboh. Banyak kajian telah dilakukan sejak 25 tahun yang lalu, dengan pelbagai kaedah telah digunakan. Walau bagaimanapun, kebanyakkan penyelidikan mempunyai kaedah tersendiri dalam pengumpulan dan analisis data. Kadar False Reject dan False Accept boleh mencapai kadar yang tinggi sehingga 30% tetapi purata keseluruhan kira-kira 5%. Salah satu kaedah yang dicadangkan untuk mempertingkatkan sistem tersebut adalah dengan menggunakan Keystroke dynamics. Keystroke dynamics adalah sejenis pengesahan biometrik yang tidak memerlukan apa-apa perkakasan khas. Selain itu, ia adalah mudah untuk digunakan, rutin yang sama seperti pengesahan kata laluan biasa. Tambahan pula, Keystroke dynamics mempunyai potensi yang besar untuk menggantikan sistem kata laluan yang sedia ada kerana ia tidak memerlukan apa-apa perubahan besar pada sistem semasa, dengan hanya menambah beberapa lagi area pada sistem dan dan mendapat sokongan daripada perisian yang digunakan untuk menyokong perisian untuk digunakan dalam menangkap dan memadankan corak menaip pengguna. Oleh itu, kajian ini mencadangkan satu sistem pengesahan menggunakan Keystroke dynamics untuk mengelakkan sistem daripada penceroboh. Satu sistem dibangunkan yang terdiri daripada dua bahagian iaitu Enrolmen dan Pengesahan. Dalam bahagian Enrolmen, pengguna akan mendaftar ke dalam sistem dengan menaip maklumat yang diperlukan termasuk nama pengguna, kata laluan dan frasa piawai. Dalam bahagian ini, masa menaip pengguna direkodkan dan template rujukan dicipta. Kemudian, padanan terhadap persamaan antara login dan template rujukan dilakukan dan akan disahkan berdasarkan had yang ditetapkan untuk memberikan pengguna mencapai sistem. Sistem yang dicadangkan menggunakan rangka kerja Microsoft.NET, papan kekunci dan Microsoft IIS. Prototaip dibangunkan yang terdiri daripada 3 modul utama iaitu Pendaftaran, Sambungan Client / Server dan, Pengesahan dan Latihan Semula. Dalam modul Pendaftaran, data bagaimana pengguna menaip akan diambil semasa proses pendaftaran pengguna. Manakala dalam modul Sambungan Client / Server modul, ianya menyediakan medium komunikasi antara program pelanggan dan juga pelayan sistem. Pengguna akan dikenal pasti dalam Modul Pengesahan dan Latihan Semula dengan membandingkan data login dengan template rujukan. Sebarang data yang berjaya akan direkodkan ke dalam data rujukan untuk meningkatkan kecekapan sistem. Berdasarkan ujian, sistem itu membuktikan Sistem Pengesahan Keystroke dynamics dapat dilaksanakan dalam persekitaran pelanggan / pelayan. Pada masa akan datang, sistem boleh diperbaiki dengan meningkatkan keselamatan, prestasi, dan antara muka pengguna.

ACKNOWLEDGEMENT

First and foremost, I would like to take this chance to thanks all individual that have involved in this project for their assistance, especially Project Supervisor, Dr. Siti Rahayu, for your guidance, supervision and assistant on the project especially on the project problem definition and problem solving and also general view on this project.

Secondly, I would like to extend my thanks to the all lectures and colleagues for their time, guidance, and support during my studies and help toward my postgraduate studies. I would also like to thanks to my employer, Infineon for their funding support to my study.

Last but not least, my special thanks to all my colleagues and friends to be pilot user to test the prototype of Keystroke Dynamic system and always encourage each other's. Do not forget that to those who had indirectly contributed to this project, I gave my greatest thanks for your support.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

EER - Equal Error Rate

FRR - False Rejection Rate

FAR - False Acceptance Rate

CHAPTER 1

INTRODUCTION

1.1 Introduction

Nowadays more and more sensitive data have been stored and processed by computer systems. Thus there is a need to increase the security of the system to secure the important data. Normal authentication systems at present are not full proof. Common methods to break current authentication system including brute force attack, password dictionary and etc. Most of the current systems only have one-layer protection, in which is the password for those online systems. Thus if the password has been stolen it means the system is at risk to be breached.

Electronic devices which have internet access are an essential part of modern society. Especially, internet has drastically changing our lives and making our work and daily lives more convenient and ease the human work. However, it also brings us a big concern in information security. We are depending so much on computers and internet to store and process sensitive data. It has become extremely necessary to secure them from intruders. Our personal information is suffering from more risks than ever before if compare to last few year back.

To protect our data, we need to verify and identity of the user that he or she is legal to use authentication and identification techniques. For user authentication and identification in computer based applications, it is necessary to use simple, inexpensive and unobtrusive device. A user can be defined who is trying to access the internet and access

the information on a computer by using Keyboard or touch screen. At present, the password is widespread and widely used to prevent user account from being invaded.

But too many ways used to decipher password and once cracked. There may be significant economic losses of the users would be caused. For such crime on the internet could lead to a number of serious damages and it is yet more challenging to be prevented. Therefore, to handle such problem which we need to deal with, we urgently need one reliable way to protect our privacy.

Keystroke Dynamics is an important biometric solution for person authentication. Based upon keystroke dynamics, propose and design an embedded password protection device, develops an online system, collects two public databases for promoting the research on keystroke authentication, exploits and characterize keystroke dynamics, and provides benchmark results of three popular classification algorithms, one-class support for online system and Gaussian classifier.

In order to counter these types of problems, one of the method that been suggested is implementing biometric authentication. Biometric Authentication is a type of authentication method that uses the human's characteristics to identify the user. It is hard to forge human characteristics compare to forge password or ID card, therefore it will enhance the system security with to identify the user biometrically. This chapter describes about the research background, research problem, research question, research objective, research method, research scope and expected research outcome. This sub chapter is followed by detail explanations and will be summarized in the last section.

1.2 Research Background

There are two distinct meanings for biometric. Bio means living creature and Metric meaning the standard of measure an object quantitatively. The process of automatic

identification of a person based on one person physiological or behavioural characteristics is called biometric. Therefore, the meaning of biometrics can be definite as the science and technology of evaluating and statistically analysing biological data. The measurement of data derived from direct measurement of a part of the human body is called physiological characteristics. Fingerprints, retina, facial image, iris and hand geometry are leading physiological biometrics. Behavioural characteristics are based on an action taken by a person. Behavioural biometrics are based on measurements of data derived from an action, and indirectly measure characteristics of the human body. Signatures, voice recordings (which also has a physiological component), and keystroke rhythms are leading behavioural biometric technologies. The terms "Biometrics" and "Biometry" have been used since early in the 20th century to refer to the field of development of statistical and mathematical methods applicable to data analysis problems in the biological sciences.

Proper biometric use is very application dependent and case by case. Certain biometrics will be better than others based on the required levels of convenience and security. There are no single biometric which will meet all the requirements of every possible application.

Although the use of biometrics such as face, fingerprint and signature can improve the security very well and convenient. But such techniques usually require additional tools to protect a device, which induces an extra cost in comparison with the password technique. The use of keystroke dynamics which detects the typing pattern of an individual can be an alternative ways to enhance security without extra cost. As it can be obtained using the existing systems as basic device needed such as the standard keyboard.

Biometrics is able to operate online or offline depending on the needs of the application. For online system it requires to recognize the user and response immediately. It must use a fully automated system with a live scanner. On the other hand, offline system

doesn't require immediate recognition and it can use a semi-automated system with offline scanner

1.3 Research Problem

The research problem in this research is enhancing the keystroke dynamic authentication online based system that identifies the user using statistical approach. The system should be improving current password authentication system in terms of security, performance, effectiveness and also user friendliness. This project will aim to enhance robust performance in terms of Equal Error Rate (EER) in authentication. Another word to said that to reduce EER value.

As today, the most common way to enforce authentication is by password, personal identification number (PIN) or other predetermined passcode. Before a user want to perform any intended activity online, he/she is required to enter his/her username and credentials. Unfortunately, it also have many flow which make it vulnerable to hacking although a normal username/password access control effective to a certain extent.

As we know that a good password hard to hack must have certain rules. Example: include at least eight characters, some of which capital letters and special characters (e.g., @, ?, !). Regrettably, a hard-to-hack passwords are also hard-to-remember. Subsequently, many users choose passwords that relate to their private lives. Example: birthday date, vehicle number, IC number, pet's name, parent's or kids' names. As a result, this making them easy to hack. Furthermore, many users write their passwords on a note which may be intercepted by hackers. This so-called memory obstacle also leads most users to use the same username and password in several web sites. Thus, a hacker revealing a users' password from a non-secure website will gain access to many of the websites that the user has access. Hacker is hacking into some of user's bank website, may incur money lost to

the user. Due to these drawbacks, password-based user authentication methods provide only partial protection against hackers. Thus, they need to be complemented by additional authentication such as physiological and behavioral biometrics. Keystroke Dynamic authentication is one of the most common and low cost behavioral biometrics in the market now.

1.4 Research Question

Referring to the Research Problem in Section 1.3, three research questions are formed to represent the research problems which are:

RQ1. How to achieve a low Equal Error Rate (EER) in order to maintain robust performance?

RQ2. How to analyse the Keystroke Dynamic authentication is working for web based system?

RQ3. How to evaluate the Keystroke Dynamic authentication system?

1.5 Research Objective

Based on the research questions formulated in Section 1.4, the research objectives are developed as follows:

RO1. To enhance robustness of authentication by using keystroke dynamic methods between False Rejection Rate (FRR) and False Acceptance Rate (FAR).

RO2. To analyse keystroke dynamics authentication system that identifies the user using statistical approach via client-server connection.

RO3. To evaluate keystroke dynamic system from a prototype which build.

1.6 Research Method

Based on the research objective formulated in Section 1.5, the research method has developed as follows:

RM1. Preprocessing the literature survey is used to obtain information about current approaches in keystroke dynamic to predict the best FRR and FAR.

RM2. Using the statically algorithm with the real reference dataset on matching and retraining technique for client-server connection model.

RM3. Collect and compare the real user data from the prototype keystroke dynamic system.

1.7 Research Contribution

Based on the research method formulated in Section 1.6, the research contribution has developed as follows:

RC1. Improving current password authentication system in terms of security, performance, effectiveness and also user friendliness

RC2. Enhanced and identify the system statically/ learning algorithm using the real user reference dataset on matching and retrain technique client-server connection model

RC3. A simple of data collection from a prototype system for data analysis and future improvement.

1.8 Research Scope

In order to achieve the Research Objectives, this research will be focused on some issues as stated below:

RS1. To ensure current password authentication system able to be enhance the security, performance, effectiveness and also user friendliness by using keystroke dynamic.

RS2. To choose only 1 type keystroke dynamic method for statiscal, matching and retrain algorithm for client-server connection.

RS3. A collection of user score (from EER) from prototype system and keep for future improvement

1.9 Research Outcomes

Based on the research, by seeing the pattern of data collection from the prototype system, it can proof that reduce the EER value which aim to enhance robust performance in terms of Equal Error Rate (EER) in authentication. Hence, it will improve the password authentication system in term of security and performance wise.

However, there are a few challenges and open areas of research that should be addressed in order to make this an effective biometric. Keystroke dynamics has a strong behavioural basis which should be explored to understanding of the motor behavior during typing. Using these concepts, models could be built to better understand the processes involved in typing. An understanding of how different people or groups of people type may provide insight into patterns in biometric features such as age, gender and environment. This might help in the development of better classifiers which could improve the accuracies of existing systems.

1.10 Summary

In this chapter it is clearly stated the research problem and the intention to conduct the research study on various biometric authentication, keystroke dynamic methods,