MULTI-SEED BASED AUTHENTICATION

BY

NADER NASSAR

A THESIS SUBMITTED TO

THE FACULTY OF SCHOOL OF COMPUTER SCIENCE AND INFORMATION SYSTEMS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PROFESSIONAL STUDIES IN COMPUTING

AT

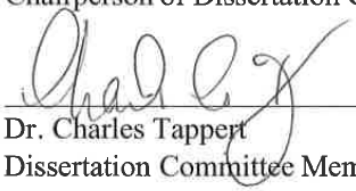
PACE UNIVERSITY

APRIL, 2016

We hereby certify that this dissertation, submitted by **Nader Nassar** satisfies the dissertation requirements for the degree of Doctor of Professional Studies in Computing and has been approved.

Dr. Li-Chiou Chen
Chairperson of Dissertation Committee

April 19, 2016
Date



Dr. Charles Tappert
Dissertation Committee Member

April 19, 2016
Date

Dr. Meikang Qiu
Dissertation Committee Member

April 19, 2016
Date

Seidenberg School of Computer Science and Information Systems

Pace University

Acknowledgements

I would like to thank my advisor, Dr. Li-Chiou Chen for all the effort she put and invested in my ideas and for her passion to research that drove me forward whenever I needed a push. I always walked into her office when I'm at cross roads and walked out with a shining light bulb on top of my head. Dr. Chen, THANK YOU FOR YOUR SUPPORT.

Also, I would like to thank IBM for sponsoring my doctorate, and for providing a platform that cultivates innovation and embraces the "what's next?" culture so that we always drive forward and never settle down with technology.

Finally, I would like to dedicate this humble effort in research to my parents, and to my family because I can't be who I am and where I am today without their support and prayers.

Abstract

Multi-Seed Base Authentication

by
Nader Nassar

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Professional Studies in Computing
April 2016

Multi-Seed Base Authentication, MSBA, an authentication protocol presented in this research, improves many of the shortcomings of the existing authentication protocols. In addition, it identifies new spaces of entropy sources that could enhance the generation of seeds and tokens in the pseudo random number generation domain.

This thesis includes a study and a protocol on using image file as an entropy source to generate the seeds used by the pseudo random number generators in order to produce tokens. A protocol to communicate dynamic multiple tokens between a client and a server in a challenge response fashion and also a study on the effect of using S/Key scheme on improving the security of the proposed protocol.

A set of experiments was conducted to analyze the proposed protocols. Using a set of 14,000 images of distinct individuals, we ran experiments using the proposed authentication protocol to study the accuracy and performance of the proposed authentication method. In addition, using the same set of images, we experimented the use of pixel-based entropy to generate the needed seeds.

The result of the experiments confirmed that MSBA is a dynamic and secure authentication protocol that alleviates many threats and usability concerns that are associated with most of the common authentication methods. The simulation conducted also asserted that the use of dynamic seeds generated from user data such as image files provides a larger entropy space because the dynamic generation of the seeds and the token proved to be more secure than both text and graphical based password solutions.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Motivation

The ubiquity of computing and the invasiveness of its reach resulted in a swarm of personal digital assets stored remotely. This raised the need to protect these resources and assets in order to ensure only the proper users have access. Ordean [157] defined the authentication process as the process responsible of linking humans to their digital resources. This process is set to start by the end users in order to gain the proper access to their resources. Research and studies proven that human behavior is predictable. Thus, it is proven that humans are somehow predictable and considered as a poor source of secret generation [71, 101, 102, 134]. As a result, traditional text-based passwords have a well-known weakness; because end users exhibit the tendency to resort to weak secrets and predictable passwords. In reaction to the bleak status quo of text-based passwords, multifactor authentication emerged in the enterprise to stress on "the secret you know" and the "secret you have" authentication approaches. Unfortunately, step up authentication came in different flavors, none of which made it easy on the end user. It was unanimously agreed upon that usability and security are in an inverse proportion relationship. The first motivation of this research was to define a new protocol that simplifies or even eliminates the use of traditional text base password and its associated drawbacks as we know it.

In addition, for any given web application, authentication mechanisms have three elements [authentication URL, username, and password]. Usually both URL and username are not considered part of the secret. URL in particular is public knowledge and any one can access it via the browser or API. Username, however, while it should be part of the secret, never treated as

such. With bare minimum social engineering efforts, to no effort at all, adversary could obtain the target's username. That leaves the adversary with one challenge to crack then gain access to the resource. Very little research was conducted in this space to protect all three pillars of authentication (URL, Username, and Password) [159, 160].

The second motivation for our research is to reinforce both username and URL as a part of the secret exchange with the backend and ensure that it is not static or guessable.

The ubiquity of mobile computing played a major role in the emerge and the spread of graphical passwords and all of its derivatives such as recognition based, (PassFace)[156], Pure Recall based, Draw A Secret (DAS), and cued recall based (PassMap) [33, 36, 38]. However, graphical passwords came about with a lot of its own set of problems and exploits added on top of the known exploits of the text-based passwords. Some of which are common between both techniques. Others are unique for graphical passwords only. The major drawback with graphical password is, beside the over the shoulder vulnerability, is the lack of entropy source. Whether recall based or cued based, entropy space is not as large enough for the user. This was a third motivation for our research in order to identify a graphical source of entropy that is large enough and yet doesn't share the same exploits identified in the graphical passwords.

Finally, with the huge momentum the Internet of Things, IoT, is gaining to be the next generation of computing, there is no identified authentication solution that utilizes the power of IoT as an additional layer of security. Yet, this research was motivated to a loosely coupled authentication model using an IoT to provide additional security to the end user if one segment of this authentication system is lost or compromised, the user would be still protected and not compromised.

1.2 Problem Statement

As listed in the motivation section, above, we can state that there is no common authentication solution that is robust enough to be a single point of entry to many different web applications. Moreover, username, URL and a form of a secret that is shared between server and the end user are always static. To initiate authentication, this secret becomes cumbersome when the user has to memorize many of them in order to gain access to different web sites. The purpose of this research is to evaluate the current problems and limitations of the authentication process (graphical, text, and multi factor) and to propose a new authentication method that does not rely on username and password. It supports dynamic user credentials and enable authentication to multiple different sites via one point of entry yet it is not a typical password vault or password manager.

1.3 Research Objectives

In this research we identified four main objectives described as follows:

Objective 1: Define a protocol that uses no static URL, username and password for authentication in a client-service environment.

Objective 2: Increase the security of the authentication process by providing a secure and reliable system that is easier for users and support multiple authentication sites from one entry point.

Objective 3: Propose a new image-based entropy source where the user can generate unique input seeds that could be used by different pseudo random number generators at the user's end point.

Objective 4: Prove that dynamic multiple tokens authentication approach is more secure than existing text-based and graphical-based authentication systems.

1.4 Methodology

A background review of authentication methods that are used in a client-server environment was conducted. First part of this research, we identified a new authentication method that utilizes multiple tokens generated from different seeds to complete a full authentication of the user and we call it "*Multi-Seed Based Authentication"*, MSBA. We researched the security of MSBA by creating a pilot that embodied the client server interaction via MSBA and simulated real life scenarios of user authentication on a large pool of users. The purpose of the study wass to examine the capabilities of MSBA and to affirm if it is able to identify users via multiple (dynamic) tokens only.

The second part in this thesis is to examine MSBA's flexibility to adapt and to integrate with further security protocols by modifying MSBA to implement S/Key scheme. In this section we intended to understand both aspects (flexibility and the security) of MSBA-S /key. To do so, we developed a client server implementation of the protocol and ran a simulation to study the security of the protocol from the client side. This simulation ran over the same large pool of users whom identified in the first part in order to examine if MSBA-S/key was able to identify the users accurately, and protect against different attacks.

The third part of our research is a study on the effectiveness of using image files as a new source of entropy for randomness. This is an integral component of our protocol to generate multiple yet unique seeds used the token generation. The data collected are also studied for uniqueness and performance.

The fourth part is about an implementation of seed management protocol to illustrate how to manage multiple seeds from one application and how the user can start from one entry point to

authenticate to multiple applications. We created an implementation of the seed management protocol and examined its applicability to integrate with the seed generation approach.

1.5 Thesis Overview and Organization

The thesis is organized into four sections starting with literature review, threat review, research details, and implementation.

First section: includes chapter 2 and chapter 3 which focuses on literature review. Chapter 2 includes a general review about authentication methodologies, cryptography, and authentication fundamentals. In addition to common cryptographic solution review used by various authentication mechanisms, the section concludes with a comparative analysis of our proposed solution and the closest approach to ours followed by a review of pseudo random number generators, PRNG, and the concept of pseudo random numbers. In addition, we illustrated the S/Key scheme and its use in cryptography and authentication.

Chapter 2 illustrates the challenges for the authentication process including the common threats and the vectors of attacks on the various elements involved in the authentication process. This is followed by illustrating usability reviews and challenges that face the users during authentication. Chapter 3, where we address the current issues in the known authentication solution. Chapter 3 reviews known threats in the current authentication solutions, then side by side highlight the differences between the proposed solution and the existing ones when it comes to exploits. Because usability is a major aspect with authentication, this chapter continues with illustrating issues revolves around usability review with most authentication systems. Chapter 3 is a literature review of mobile authentication techniques and its various approaches. Since our proposed authentication protocol depends on users' graphical input file to generate needed

secret. The second part of chapter 3 is dedicated to the review of the state of the art of graphical passwords. The chapter ends with a comparison between these methods and our proposed protocol.

Section two is dedicated to research conducted where the proposed authentication protocol is described in chapter 4 with details including the modified version of the protocol using S/key scheme in chapter 5.

Section three of this thesis is about the conducted experiments and data collection including discussion about the security and robustness of the protocol. This is explained in chapters 5 and chapter 6. The first part of the experiment is discussed in chapter 5 which describes tokens collision study using our protocol. While in chapter 6, we illustrated a series of experiments in regards to image file and seed generation approach using the unique colors entropy found in user's selected image. The implementation of the protocol is addressed in chapter 7 since we have different types of implementations, and we explained the topology and the elements (software, hardware) of the systems used during our experiments.

Section four is the conclusion of the thesis. Chapter 8 is an overall summary of the proposed protocol, where we detailed our contribution and iterated through the goals of the research. Also, in that chapter, we listed the areas which are not included in the research as well as the possible future extension of our research that can be adopted and forked off our thesis.

Chapter 2: Background on Authentication

2.1 Introduction

Authentication landscape has many solutions, but regardless of the authentication approach used, there are few fundamentals that are looked upon as pillars for any identity access validation and authentication process.

In this chapter, we will highlight some of the main concepts used in cryptography which are also used in our research. Starting with a broad brush, as we define cryptography and the concept of hashing, then detail the advantages and disadvantages of the two most common cryptography technics; symmetric and asymmetric key encryption. Still, in broad brush strokes, we define the two most general authentication models, secret based and token based beside the latest trend of token based which is known as Fast Identity Online / Universal Two Factor authenticators, FIDO/U2F. Finally, we conclude the chapter by illustrating a literature review to the password-less U2F Yubikey which provides similar functionality; however, we detail the difference between our research and FID/U2F solution.

In addition, we will detail how current authentication methods using random number generation, the use of pseudo random number generations, PRNG, in various authentication techniques. PRNG will be our Segway to describe how S/Key scheme is used to generate authentication protocol that stores no secret and no passwords locally which was a great idea, however it exhibited some vulnerabilities. Finally, we will

explain the various known techniques of key, seed, token, and password management, which we also introduce a new approach to enhance as a part of our overall research.

2.2 Cryptography Background

Vanstone and et al [118] defined cryptography as the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data origin authentication. Cryptography is not the only means of providing information security, but rather one set of techniques.

Modern cryptography concerns itself with the following four objectives:

1) Confidentiality (the information cannot be understood by anyone for whom it was unintended)

2) Integrity (the information cannot be altered in storage or transit between sender and intended receiver without the alteration being detected)

3) Non-repudiation (the creator/sender of the information cannot deny at a later stage his or her intentions in the creation or transmission of the information)

4) Authentication (the sender and receiver can confirm each other's identity and the origin/destination of the information)

2.2.1 Hash Functions

A hash function as defined in [118] is a function h which has, as a minimum, the following two properties:

A. Compression: h maps an input x of arbitrary finite bit-length, to an output h(x) of fixed bit length n.

B. Case of computation: given h and an input x, h(x) is easy to compute.

To facilitate further definitions, three potential properties are listed (in addition to ease of computation and compression for an unkeyed hash function h with inputs x, x' and outputs y, y'.

1. Pre-image resistance—for essentially all pre-specified outputs, it is computationally infeasible to find any input which hashes to that output, i.e., to find any pre-image x such that h(x') = y' when given any y for which a corresponding input is not known.

2. 2nd pre-image resistance—it is computationally infeasible to find any second input which has the same output as any specified input, i.e., given x, to find a 2nd pre-image X' = x such that h(x') = h(x).

3. Collision resistance—it is computationally infeasible to find any two distinct inputs x, x' which hash to the same output, i.e., such that h(x) = h(x').

2.2.2 Symmetric Key Cryptography

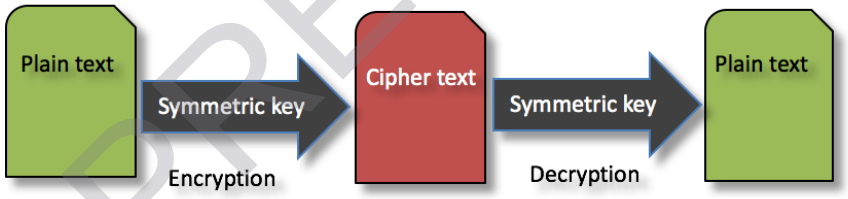


Figure 2.1 Symmetric key encryption

The most common type in cryptography and it is known as “shared secret key". This type is defined as a key that is used to encrypt/decrypt a message using symmetric cryptographic algorithms. Also it is used to provide integrity to the authentication process using message authentication codes which is known as Hash based Message Authentication Code (HMAC). A symmetric key is also called a secret key since the

same key is needed for the message sender to encrypt the message and for the message receiver to be able to decrypt it, figure 2.1.

2.2.3 Advantages and Disadvantages of Symmetric Key Cryptography

In symmetric key cryptography sender and receiver only have to specify the shared key in the beginning, and then they can begin to encrypt and decrypt messages between them using that key. It has several advantages and disadvantages in comparison to other techniques as:

- Advantages:

a) Simplicity: This type of encryption is easy to carry out. Users have to specify one secret and share the key; then they are able to use it to encrypt/decrypt messages.

b) Better fit for localized security. If the user intends to apply encryption for messages or own files which are stored locally, there is no need to create different keys. Single-key encryption is best fit for this type of operation.

c) Contained and controlled effect when the key is compromised: Given a different secret key is used for communication with every different party. In case of a key is compromised, only the messages between a particular pair of sender and receiver are affected. Communications with other parties are still secure.

d) Fast and suitable for longer data: Symmetric key encryption is much faster than asymmetric key encryption. Due to less computation complexity, it is suitable for longer messages.

e) Uses less computer resources: symmetric key encryption does not utilize a lot of computational resources when compared to public/private key encryption.

-Disadvantages:

a) Keys management: A good practice is to create a new shared key for communication with every different party. This creates a problem with the mechanism of how to manage and ensure the security of all keys.

b) Secret key exchange: Sharing the secret key in the set up or at the start of a process is a problem in symmetric key encryption. It has to be exchanged in a way to ensures it remains secret.

c) Questionable authenticity of message: Since both sender and receiver use the same key, messages cannot be verified to have come from a particular user. This is a common shortcoming where the end user who generates the message can't be authenticated.

2.2.4 Asymmetric Key Cryptography

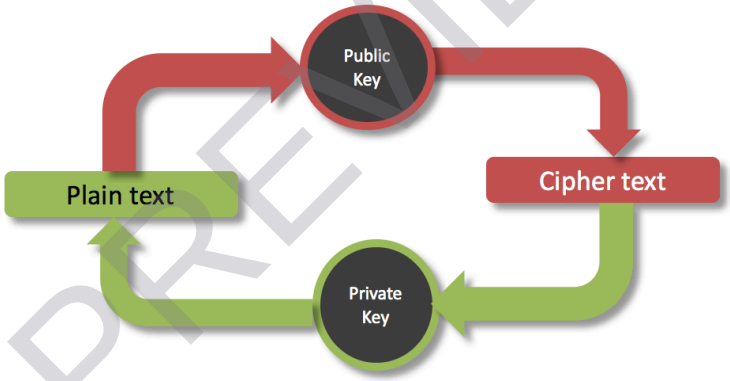


Figure 2.2 Public private key encryption

Asymmetric key cryptography is well known as Public/Private Key Cryptography, PKC. This type of encryption is defined as a pair of mathematically related keys used in asymmetric cryptography for authentication, digital signature, or key establishment. As the name indicates, the private key is used by the owner of the key pair, is kept secret, and should be protected at all times, while the public key can be published and used by the relying party to complete the protocol or invert the operations performed with the private key[117].