University of Windsor

Research Project

One-Time Biometric Authentication Using Keystroke Biometrics

Written and Conducted by:	**Stephen Arsenault**
Supervisor by:	**Dr. Sherif Saad, PhD**
Date Submitted:	**April 14, 2018**
Program of Study (Major):	**Bachelor of Computer Science (Honours)**
Term/Year:	**fall 2017 – winter 2018**

Table of Contents

List of Figures

1. Introduction

1.1. Access Control

Access control can be generally defined as “[p]rotection of system resources against unauthorized access.”[1] Authorization is defined as “approval that is granted to a system entity to access a system resource.”[1] An individual who regularly interacts with information technology is often required to authorize themselves tens of times a day. A few examples of information technology resources requiring authorization include personal computers, smartphones, and websites. These examples involve a human individual being authorized to access a system resource. The individual can more generally defined as a subject which is defined as “an active entity that accesses a passive object to receive information from, or data about an object.”[2] In this definition, the personal computer, smartphone or website are considered the passive object. A subject is not necessarily a human. One example where a non-human subject requires authorization is when a smartphone application must be authorized to access certain resources such as a camera hardware, or your email.

In order to access a system resource (i.e. to be authorized), the subject must be identified and authenticated. Identification could involve typing in a username while authentication involves typing in a password. These two steps (identification and authentication) must always be performed together in an access control system [2]. In some instances, such as smartphones where it is common to only have one subject as the user, only the authentication step is performed by the user because the identity is assumed. The identity of each subject authorized to use a system must be unique. In many cases, identity is not required to be private. Any information used for the authentication step (a password for example) must not be publicly accessible however [2].

1.1.1. Authentication Factors

There are many different techniques for authentication. Some common examples of authentication techniques include passwords, keys, and fingerprints. Each of the aforementioned examples are members of the three types of *authentication factors*.

Type 1 authentication factor involves the subject providing something they know. Different systems implement this factor differently. Depending on the complexity of the system, the type 1 factor allows for differing levels of strength. Personal computers allow for a password of variable length which includes uppercase and lower case letters, numbers and symbols. In a Microsoft Windows environment, policies can require users to maintain a certain complexity ensuring that the possible number of passwords is 218,340,105,584,896 [3]. Automatic Teller Machines (ATMs) typically restrict the password to a sequence of four numbers which greatly reduces the variability in passwords (only 10,000 possible passwords) from one subject to another. In some cases, such as online banking websites for Canadian banks, password requirements significantly limit the variability of password sequences by not allowing special characters, limiting password length to short passwords, or not checking the case of alphabetic characters [5]. Evidently, some implementations of type 1 authentication are stronger than others. In all cases, if knowledge of a subject's type 1 authentication information (their password for example) becomes public, an imposter may be able to access the system.

Type 2 authentication involves something the subject possesses. Common examples are smartcards, and hardware tokens. Bank Machines in Canada use smartcard banking cards both as a form of identification and authorization. In this case, authentication is a combination of type 2 (the smartcard) and type 1 (the four digit pin) authentication factors. When a type 2 authentication factor is used alone, it can present significant weakness if it is not properly protected against theft by an imposter.

Type 3 authentication involves something that the subject is or some way the subject behaves. Biometric authentication is a type 3 authentication factor. A common example of this is fingerprint or facial recognition used for personal computers and smartphones. Type 3 biometric authentication can be further broken down into physiological and behavioural biometrics. This factor covers a broad range of differing methods of implementation which will be elaborated on later. It will be seen that some forms of type 3 authentication can be circumvented.

1.1.2. Multi-Factor Authentication

Using more than one authentication is commonly known as multifactor authentication. It is generally accepted that multifactor authentication improves the security of a system. Following are some examples of multifactor authentication involving type 1 and type 2 factors.

Bank cards combine type 1 and type 2 by require the user to swipe their bank card or insert it into a smart card reader (type 2) in addition to entering a pin (type 1).

Many websites (Google, Facebook, LinkedIn, etc.) allow their users to require input of both a static user-defined password (type 1) in addition to a onetime password. A onetime password can be generated by a hardware token, a software time-based onetime password generator [4], or by texting a onetime password to a user's cellphone at login time. This provides type 2 authentication as the user requires physical access to an asset that will provide a valid onetime password at login time.

Type 3 authentication is generally considered to be the strongest form of authentication, stronger than either type 2 or type 1 [2]. It is possible to combine the authentication factors in any configuration. One arbitrary example of a highly secure authentication scheme could involve a combination of type 1 and type 2 authentication for an online web service which sends a onetime password to an application on a smartphone which requires type 3 authentication to access the onetime password.

2. Biometric Technologies

Biometric technologies allow a computer system to take samples of human biometrics, perform some computation, and determine whether or not to authenticate the subject providing the sample. There are numerous biometric measures that human subjects can provide. Biometric traits are broken down into two categories: physiological biometrics, and behavioural biometrics. The measurement of biometric traits requires some hardware that captures the biometric, and an algorithm for processing the samples, storing data, and creating a model. Many factors can affect how reliable a biometric system is. Depending on the algorithm, the same biometric sample from the same hardware may have different levels of accuracy for different biometric systems. Hardware, and quality of sample provided also effect reliability.

2.1. Overview

A biometric system can be used to authenticate a subject. It will accept a biometric measure, a sample, through a piece of hardware. A set of features will be extracted from the sample. This set of features will processed by some algorithm and it will determine whether or not to authenticate a user.

2.1.1. Physiological Biometrics

Physiological biometrics include: fingerprint, iris, and face shape, among others. These biometrics are measures of a physical trait of the subject that should be unique to each individual. This category of biometrics can be referred to as *something that the subject is*. They are widely used in the operating systems of smartphones, laptop and desktop computers, and can be used in physical access control systems.

Registration for these biometric systems usually requires a small number of samples from a subject after which point samples of the same biometric can be used to verify future authentication attempts. In the example of a fingerprint biometric system, a particular finger or set of fingers may be used to provide samples during registration, while in face shape systems would collect a number of samples of the subject.

With some systems for physiological biometrics, reproductions of a subject's samples can be used to gain access to the system. One example is the use of gummy bear candies to copy a fingerprint and relay that image to a fingerprint scanner. In some cases, once a biometric such as a fingerprint is stolen, it becomes a less desirable biometric unless the system can somehow accurately reject all reproductions of the biometric measure.

2.1.2. Behavioural Biometrics

Behavioural biometrics include: keystroke dynamics, voice, and gait, among others. These biometrics are measures of a behavioural trait of the subject that should be unique to each individual. This category of biometrics can be referred to as *something that the subject does*. These biometrics are less often used, although the use of voice in access control systems has increased in recent years.

Registration can sometimes require many more samples than a physiological biometric system. One type of behavioural biometric system will get a number of samples of the same behaviour. For example, a subject would repeat a phrase, or type the same text, a number of times. The samples would be used to create a model of the behaviour and future samples of the same phrase or text would be compared with the model to generate a score. If the score is above a threshold, access is granted. The other type of system intends to be able to properly classify samples of unfamiliar phrases or texts based on previous observations. The latter type of system typically requires a large number of samples upon registration.

As with physiological biometrics, some behavioural biometric systems are susceptible to reproductions. A recording of a voice, or a replay of a subjects keystroke pattern could be used to gain access to a system. Behavioural biometric systems are also vulnerable to replay attacks.

2.1.3. Good Biometric Traits

Given the challenge of having reproductions of biometrics, or “replay” attack vectors on these biometric systems, it is desirable to be able to classify samples of new phrases or texts from a subject each time they authenticate. Physical biometrics typically cannot be used in this manner. This investigations intends to assert that the ability of behavioural biometrics to continuously use unfamiliar data to authenticate users poses an advantage over biometrics which are static in nature (like fingerprints for example).

3. Problem Statement

Access control systems are always vulnerable to attack. The different authentication factors present different challenges for malicious users however. Although each factor has its advantages and disadvantages, type 3 biometric authentication has some particular advantages over type 1 and 2.

3.1. Authentication Factors and Challenges

Type 1 authentication factors generally take the form of a password. The limitations and complexity requirements for a password will create an upper and lower bound on the number of possible passwords that are valid for a particular system, which puts a limit on the number of brute force attempts required for an imposter to gain access. If an imposter acquires the password through any means, the access control system cannot determine whether the password has been typed by an imposter or the true subject. Furthermore, it is often the case that users will use fairly uncomplicated passwords and the complexity will reside close to the lower bound for password complexity. Passwords are often found to be written down, reused, only slightly different from pervious passwords, and may contain information related to the identity of the user.

Attackers can acquire passwords from database breaches and quickly try these passwords and user name combinations to gain access to the accounts of users who reuse the same credentials for multiple access control systems. Type 1 authentication presents some significant downfalls.

Type 2 authentication factors, “something that the subject possesses”, typically takes the form of a smartcard, USB token, or token generator. While smartcards and USB tokens need to be physically connected to the access control system in order to authenticate a login, token generators rely on the generation of random passwords to authenticate a user. These are referred to as one-time passwords.

A one-time password can be generated in a variety of ways, however the principle is that each time the user logs in they are required to acquire their one-time password and enter it at login time. If the password is correct, than access is granted. In the case of a security token which is battery powered, typically a clock will generate a new random one-time password using a seeded random number generator at a regular interval. The access control system will have the same seed and random number algorithm allowing it to verify that the logon was performed by someone with physical access to the security token.

Another scheme for type 2 authentication involves sending the user a one-time password on an out-of-band channel, such as the cellphone SMS network. In this case, when the user attempts to login, a one-time password is sent to their mobile phone, meaning whoever is logging in must have access to the incoming SMS messages for that user.

Type 2 presents some new difficulties for attackers as they typically need to acquire a physical asset in order to login. In the case of smartcards and USB tokens, the attacker may also require access to specific hardware compatible with the type 2 factor. There are however vectors of attack for one-time passwords. Malware on smartphones presents an attack vector in which an imposter can acquire a one-time password without the true subject being aware.

Type 3 authentication factors, “something that the subject is or some way the subject behaves”, take the form of physiological or behavioural biometrics. Despite the factor being a physical or behavioural trait of the subject, it can still be stolen. Fingerprints can be reproduced in a variety of ways, and face shape systems can be spoofed using pictures of the subject.

One attempt to remediate the use of reproductions is a liveness check which involves ensuring that the fingerprint sensor can detect a pulse, or the face moves and blinks. If liveness checks are circumvented however, reproductions of the subject’s biometrics, specifically the reuse of samples which have previously granted access, can still be used by an imposter to gain access to the system. This demonstrates a weakness in physiological biometrics. Just as replay attacks apply to physiological biometrics, behavioural biometrics such as keystroke dynamics can be captured and replayed to gain access to a system.

3.2. A Robust Biometric Authentication System

A novel idea for eliminating the need for liveness checks and the risk of replay attacks is to combine behavioural biometrics with a one-time password. Specifically, we will focus on creating a one-time biometric authentication system using keystroke dynamics. Such a system would provide the subject with a one-time password in the form of a short text which the user types. As with a typical one-time password system, the user must enter the exact password provided to them. The keystroke pattern would be captured while the user is entering the one-time password and the keystroke pattern would have to match the subject’s model for access to be granted.

This authentication system would provide robust security by significantly increasing the complexity of performing replay-attacks. An attacker would need to be able to accurately reproduce the typing behaviour of a subject for a text which the subject may have never typed.

3.3. The Problem Under Investigation

For a subject to authenticate with the system they must enroll by providing at least one sample. Typically multiple samples are required during enrollment into a keystroke biometric system. Once a subject has enrolled they can authenticate. Authentication starts with the subject providing a sample through the keyboard. The keystroke pattern is recorded by the system and classified by the system. If the classification matches the identity of the subject requesting access than access is granted and the sample may be integrated into the model and model may change.

Furthermore, this type of keystroke dynamic system would need to classify typing behaviour with a relatively small number of samples to model the behaviour. It would take parts of previously sampled text and combine them with unfamiliar texts in order to perform both verification and sampling each time the subject logs in. Part of the one-time password would be previously typed text while the remainder would be unfamiliar text. When the subject is granted access, the keystroke pattern captured during login would be integrated into the model.

Model management needs special consideration in this system. The system must model typing behaviour accurately with a small set of samples from the subjects. To accomplish this, the model will need to occasionally integrate new samples as the subject's typing behaviour changes over time. It is expected that the subject will authenticate with this system while in a variety of different mental and physical states which will effect typing behaviour. The system should be able to account for the variety of different typing behaviours displayed by each user. This will be discussed further in the results section.

4. Literature Review

There is strong research into biometric authentication as a whole, with many promising applications and realized access control systems. Keystroke dynamics is somewhat under researched when compared with technologies such as face shape, iris, or fingerprint biometrics. Keystroke dynamics research also lacks a generally agreed upon standard for collection of samples. Different feature sets are used by different researchers. There is not a wealth of publicly accessible datasets that can be used to benchmark studies against one another.

4.1. Collecting Data

The method of data collection is of great importance and will affect how rigorous of an assessment a biometric system will receive. The number of samples collected is of concern. There must be a sufficiently large amount of data collected to represent actual expected usage patterns for the application. The typing behaviour of a subject will change depending on the mood and energy level of the subject, among other things [6]. Data should be collected over multiple sessions to provide some variation in the subject's mental and physical state.

The paper “Gaussian Mixture Modeling of Keystroke Patterns for Biometric Applications” by Danoush Hosseinzadeh and Sridhar Krishnan [6] dedicates a significant section to a “keystroke protocol” which attempts to lay out some best practices and standards for data collection, among other things. If recommends that samples be collected over a number of sessions so that changes in physical and mental state are incorporated into the data.

Data collected for “Comparing Anomaly-Detection Algorithms for Keystroke Dynamics [7] was collected from 51 users over 8 sessions with 50 samples per subject per session for a total of 400 samples per subject. Sessions were at least one day apart from each other. This adheres to the practices prescribed in [6]. Furthermore, this dataset is publicly available for use by other researchers. More on this paper below.

4.2. Comparison of Keystroke Systems

There are a number of papers which focus on the comparison of keystroke systems developed by different researchers [6, 7, 8, 9]. The dataset from [7] has a large number of users and should capture the natural variability in subject mental and physical state, making it an excellent dataset for comparison of biometric systems and research into the technology in general. [7] also uses the dataset to make comparisons between 14 systems from various places in the literature. [8] reviews the subject in general with a focus on the variety of specific research areas including special pressure sensitive keyboards and mobile phones.

4.3. Biometric One-Time Password

Although there is some research which focuses on the use of a one-time password for biometric authentication [10], there does not appear to be any research looking into the use of a one-time password for keystroke biometrics. This novel area of research presents some unique challenges however. The system must be able to model the subject’s typing behaviour for short texts with a very small number of samples. The system will also need a reliable way to generate unique one-time passwords which contain previously sampled text and new text for future one-time passwords.

5. Methodology

This research project focuses on the feasibility of continually updating a subject’s model. This is intended to represent typical usage of an access control system. Initially, the user will enroll in the system with a number of samples. Thereafter, each time the user successfully logs in, the sample is added to the model. This will allow the model to adapt to changes in the user’s typing behaviour over time.

5.1. Dataset

The CMU benchmark dataset [7] will be used for this research as it provides a large number of samples collected from many users in a controlled environment over multiple sessions. The samples provided in this dataset are all of the same ten character password '.tie5Roanl'. Since one-time passwords should be around the same length, this will serve as a good text for testing the continuous updating mechanism of the one-time password system.

5.2. Metrics and Analysis

The reliability of a model must be measured each time the model is updated to track metrics such as true acceptance rate (TAR), false acceptance rate (FAR), and their respective complements, false rejection rate (FRR), and true rejection rate (TRR). These metrics can be used to show how the overall effectiveness of the system changes over time as the number of login attempts increases.

5.3. Evaluation of Samples

Samples will be evaluated using the K nearest neighbour algorithm with Manhattan distance. Multiple values of k will be used in the analysis. Additionally, the initial number of samples required before the first authentication attempt will be varied.

6. Experimental Design

6.1. General Overview

Subjects must enroll in the system prior to any authentication by the system, obviously. A number of samples, s, is taken from each subject and used to train the model during enrollment. Thereafter, until the remaining samples are exhausted, intervals of ten samples are classified using the trained KNN. Properly classified samples will be added to the model, improperly classified samples are not used again, and each subject has a number of randomly selected samples used to calculate the false acceptance rate. The imposter samples are selected from the set of samples which have not been used for training or testing.

The process of taking ten samples per subject and calculating TAR and FAR values will be referred to as a *trial*. The term *trial-set(s,k)* will be used to refer to the set of all trials on a model initially trained with s samples with k neighbours for the KNN classifier. An *experiment* will be used to refer to all of the trial-sets in some particular experiment. There are three experiments which perform different actions at the trial level or use a subset of the dataset. Each experiment is repeated for $k \in \{1,3,5,7,9\}$ and some values of s.

6.2. CMU Benchmark Dataset

The CMU Benchmark Dataset contains 400 typing samples from 51 users. These were collected 50 samples at a time over eight sessions that were at least one day apart. When referring to samples, a subject's first sample (sample 1) was the first sample provided on the first day, while the 52nd sample (sample 52) was the second sample collected on the second day. It is noted that all data was collected using the same system. The skill of subjects varied from two finger typing to touch typing.

6.3. Experiment 1

The first ten samples from each subject are used to train the model. Thereafter, ten samples per user are classified. Correctly classified samples are flagged to include in the model. The FAR is calculated using 100 randomly selected samples that have not been included in the model, have not been used for testing, and were not collected from the user whose FAR is being calculated. The model is updated with the correctly classified samples and the next ten samples per user are classified. This procedure is repeated for $k \in \{1,3,5,7,9\}$, and the values averaged over those five trials.

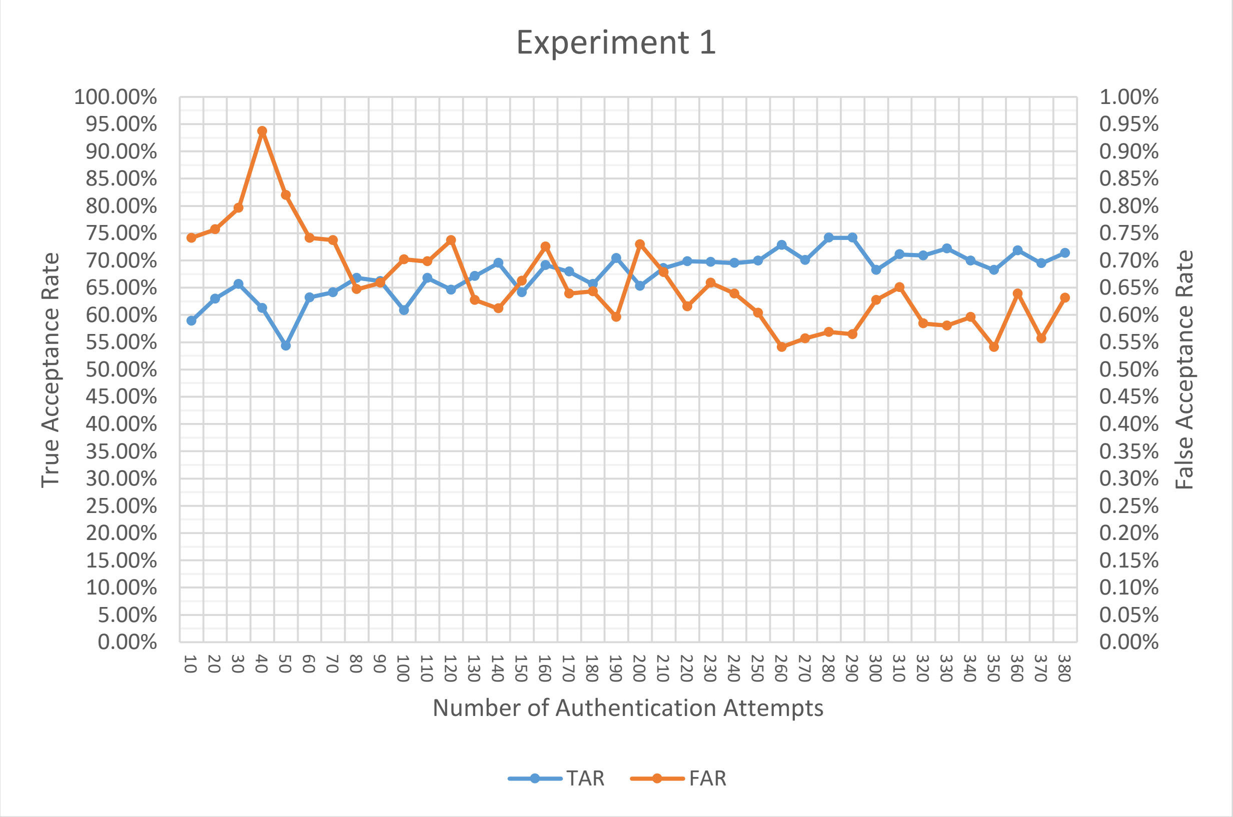

Figure 1 A slight increase in TAR as Authentication Attempts increase, and a notable decrease in TAR at the beginning of each collection session

Figure 1 shows that as the number of authentication attempts increases, the TAR improves from 60% but never exceeds 75%. The FAR slowly improves over time as well. There is a notable decrease in the TAR for 50, 100, and 150. The trend continues to a lesser effect for 200, 250, 300, and 350. At those points, the samples being used for authentication attempts are the first 10 samples from days 2 through 8. The data shows that a model trained on data sampled on some particular date may have difficulty recognizing typing samples from a later date. In particular, subjects may type erratically for the first few samples of a collection session. This is explored further in a later experiment.

6.4. Experiment 2

This experiment was identical to experiment 1 with the exception of using 100 randomly selected samples for calculating the FAR instead of using 100.

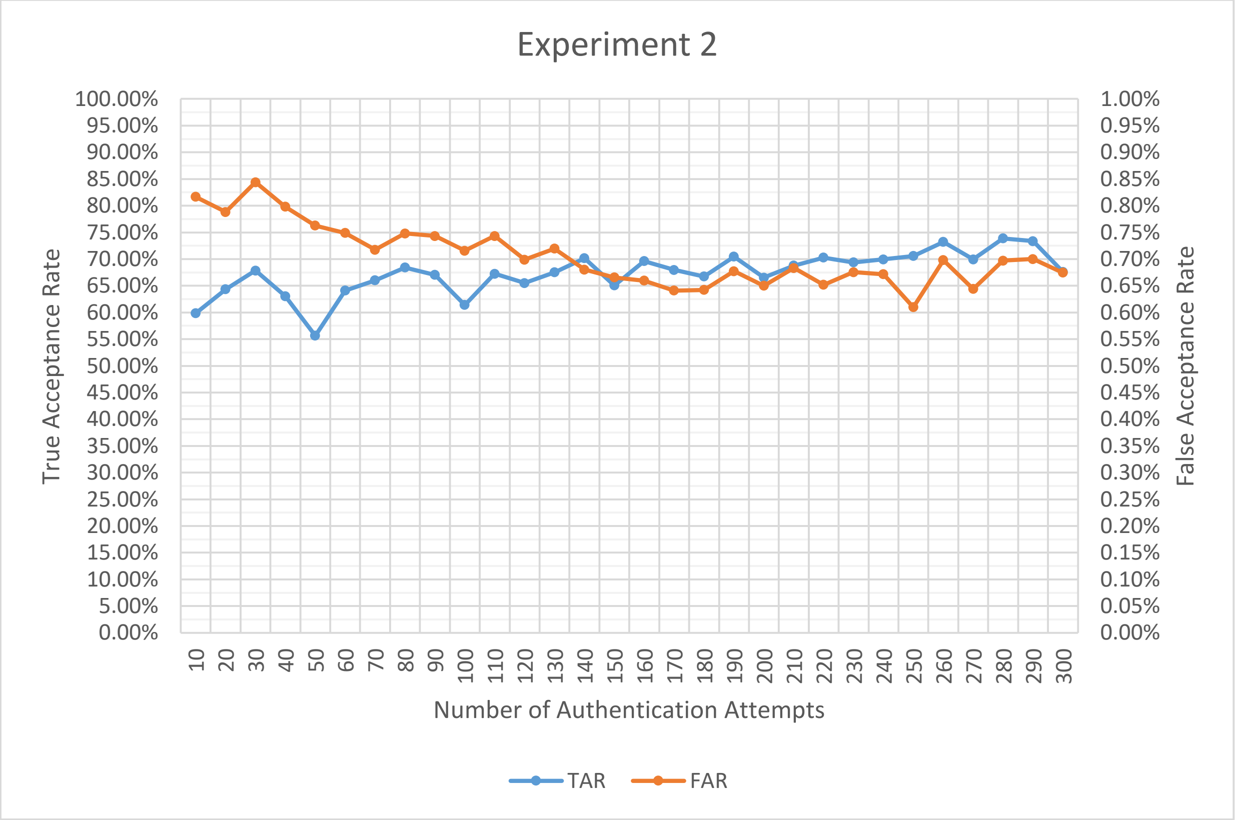

Figure 2

6.5. Experiment 3

This experiment is identical to experiment 2 with the exception of never updating the model after the first 10 samples are used to train initially.

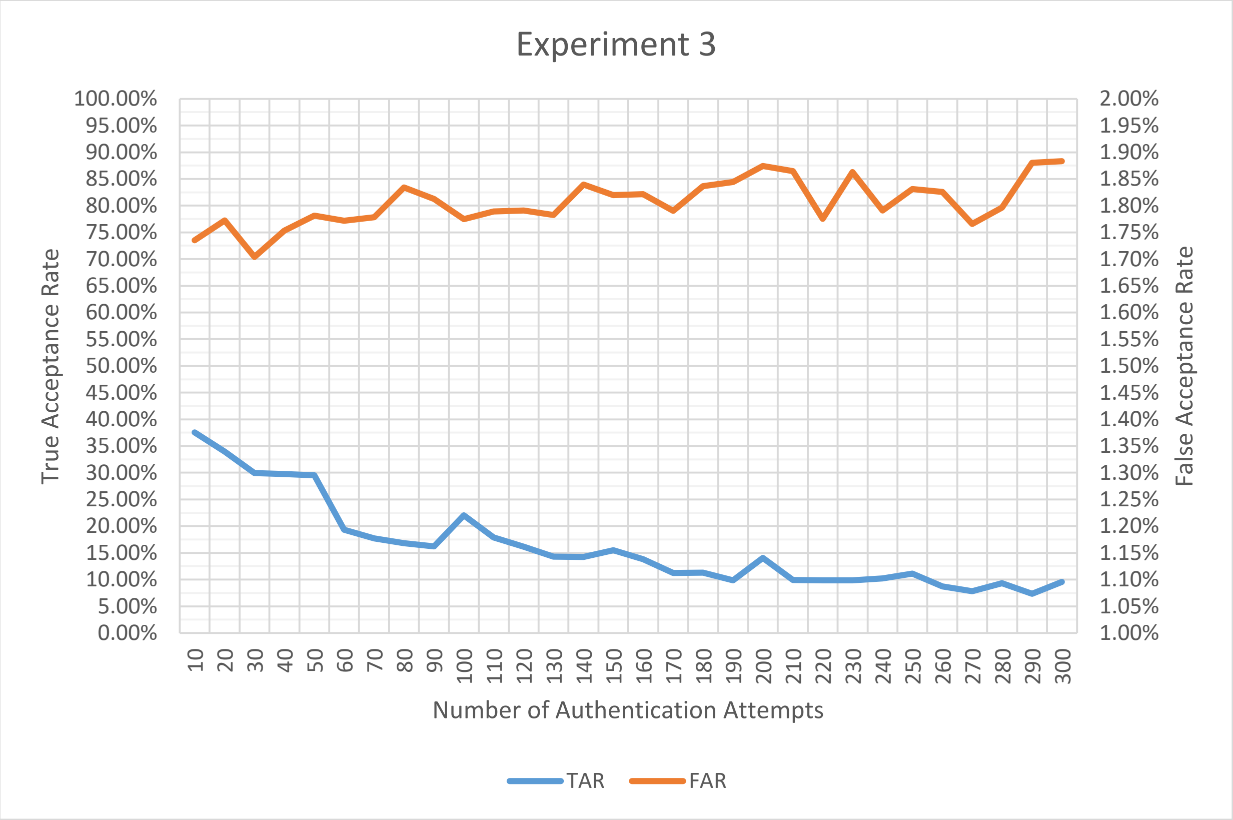

Figure 3 Increased TAR for the first 10 samples of each session when the model is only trained on the first 10 samples of the dataset

Figure 3 shows that not updating the model with new samples leads to a decrease in the TAR as the number of authentication attempts increases. It should be noted that contrary to experiments 1 and 2, at 50, 100, … , 250, 300, there is an improvement in TAR. This could be an indication that despite the first few samples in a session being erratic compared with the majority of later samples, the first few samples of each session are all similarly erratic leading to an increase in the TAR when the model is only trained with the first few samples of the first session and the authentication attempts are being made with the first few samples of a session.

7. Summary

Overall the experiments show that subjects type a particular way when they are typing the first few samples during a collection session. Experiment 3 shows that when a model is only trained using data from the beginning of a session and not updated, than it recognizes the first few samples of each session best. This is a desirable outcome as most real applications for a biometric authentication system will require substantially fewer than 50 samples per day. Currently, there is not much persuasive evidence showing that updating the model in a simplistic manner will lead to improved TAR and FAR rates.

These results demonstrates the need for a dataset which can more accurately represent the variations in typing behaviour expected in a real access control system. Such a dataset could be elicited from subjects each time they unlock their computer. Eliciting a sample from participants each time they unlock their computer may capture the variation in typing behaviour much more accurately than a session in which the subject will inevitably improve over the duration of the session.

References

[1] Internet Security Glossary, Version 2. (n.d.). Retrieved April 15, 2018, from https://tools.ietf.org/html/rfc4949#page-9

[2] Stewart, J. M., Chapple, M., & Gibson, D. (2015). CISSP (ISC)2 Certified Information Systems Security Professional Official Study Guide. Wiley.

[3] B. (n.d.). Password must meet complexity requirements (Windows 10). Retrieved April 15, 2018, from https://docs.microsoft.com/en-us/windows/device-security/security-policy-settings/password-must-meet-complexity-requirements

[4] TOTP: Time-based One-time Password Algorithm. (n.d.). Retrieved April 15, 2018, from https://tools.ietf.org/html/rfc6238

[5] Why Canada's banks have weaker passwords than Twitter or Google. (2017, March 26). Retrieved April 15, 2018, from https://www.theglobeandmail.com/technology/digital-culture/why-canadas-banks-have-weaker-passwords-than-twitter-or-google/article18325257/

[6] Hosseinzadeh, D., & Krishnan, S. (2008). Gaussian Mixture Modeling of Keystroke Patterns for Biometric Applications. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38(6), 816-826. doi:10.1109/tsmcc.2008.2001696

[7] Killourhy, K. S., & Maxion, R. A. (2009). Comparing anomaly-detection algorithms for keystroke dynamics. 2009 IEEE/IFIP International Conference on Dependable Systems & Networks. doi:10.1109/dsn.2009.5270346

[8] Bhatt, S., & Santhanam, T. (2013). Keystroke dynamics for biometric authentication — A survey. 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering. doi:10.1109/icprime.2013.6496441

[9] Zhong, Y., & Deng, Y. (2015). A Survey on Keystroke Dynamics Biometrics: Approaches, Advances, and Evaluations. Gate to Computer Science and Research Recent Advances in User Authentication Using Keystroke Dynamics Biometrics, 1-22. doi:10.15579/gcsr.vol2.ch1

[10] Jenkins, J., Shelton, J., & Roy, K. (2017). One-time password for biometric systems: Disposable feature templates. SoutheastCon 2017. doi:10.1109/secon.2017.7925304