# Biocryptography

The Future of User Authentication?

Jayme Woogerd

December 12, 2014

## Abstract

In traditional cryptography, authentication relies on some sort of shared knowledge, usually a secret password or token. In today's wired world, these types of password authentication systems are pervasive – however, they face several fundamental problems, chief of which that they cannot distinguish between genuine users and attackers using stolen credentials. Biometric systems solve these problems by using physical characteristics to uniquely identify genuine users. However, these systems come with their unique set of vulnerabilities.

At the intersection of biometrics and traditional cryptography lies biocryptography, the study of cryptographic methods for securing biometric systems. This project explores the state-of-the-art strategies in biocryptography to provide systems that are both accurate and robust to attacks.

## 1 Introduction

In traditional cryptography, authentication relies on some sort of shared knowledge, usually a secret password or token. In today's wired world, these types of password authentication systems are pervasive – however, they face several fundamental problems, chief of which that they cannot distinguish between genuine users and attackers using stolen credentials. Additionally, users are forced to manage multiple accounts with multiple passwords.

Biometric systems solve these problems by using characteristics like fingerprints, irises, even ear shape to uniquely identify genuine users. However, these systems come with their unique set of vulnerabilities: what happens if a user's biometric data is compromised? A password can be easily reset...but a fingerprint certainly cannot. At the intersection of biometrics and traditional cryptography lies biocryptography, the study of specialized cryptographic methods for securing biometrics systems. This paper explores the state-of-the-art strategies in biocryptography to provide systems that accurately authenticate genuine users and are robust to traditional biometric attacks.

## 2 To the Community

According to the results of an online registration and password study conducted by Janrain in 2012, 58 percent of adults have five or more unique passwords associated with their online logins and 30 percent of people have more than ten unique passwords they need to remember. Additionally, almost 2 in 5 (37%) have to ask for assistance on their user name or password for at least one website per month [Janrain, 2012].

The current models for user authentication have created a ridiculous situation: on the one hand, experts implore us that a password should be long, random (and therefore not memorable) and unique so as to be resistant to cracking, but on the other hand, we have to create, keep track of, and periodically change credentials for *every single* service we sign up for. Anecdotally, for most people convenience trumps security and they end

up using (and reusing) weak passwords over and over again.

The situation is exacerbated by the fact that the entities we trust to secure our passwords are not secure themselves: if anything it seems the frequency with which we are hearing about major security breaches (e.g. LinkedIn, Target, Sony) is increasing. From the LinkedIn breach on June 5, 2012 alone, hackers gained access to 6.5 million user passwords – and had cracked and posted in cleartext more than 60% of the unique passwords by the next day [Vijayan, 2012].

Thus, the original motivation for this project was my own frustration with the current state of user authentication systems. It's deeply unsatisfying that so much of the onus is on the user to manage and secure her own credentials. Therefore, this project explores an alternative to a password-based authentication system: biometric systems and the biocrytographic methods used to secure them. I wanted to know if biometric systems as seen in popular culture (e.g. *Minority Report*) are even feasible and if so, in what ways are they better than current traditional authentication systems.

# 3 Cryptography

In cryptography, the basic idea is to allow two entities to send and receive messages to each other securely and confidentially in the presence of a third-party, or *adversary*. This is done through the process of data *encryption*, that is, transforming the message into an unreadable form to anyone who does not know how to decrypt it. Traditionally, the cryptographic process involves using an *encryption algorithm* and a *cryptographic key* to turn an

original message in *plaintext* into a scrambled, unreadable one, i.e. in *cypertext*.

### 3.1 Encryption Algorithms

In modern cryptography, there are two main flavors of encryption algorithms: symmetric- or private-key encryption and asymmetric or public-key encryption. Symmetric-key encryption algorithms use the same cryptographic key and algorithm to both encrypt plaintext into cyphertext and decrypt the cyphertext back into plaintext. Among the most widely used symmetric-key algorithm is the Data Encryption Standard (DES). Asymmetric-key systems involve pairs of keys, which are generated together: a public key for encryption and a private key for decryption. As the name suggests, the public key is shared publicly and can be used by someone else to encrypt data to send to you. Data encrypted with this public key can only be decrypted using the private key, which you keep secret. A common symmetric-key algorithm is RSA, named for its creators Rivest, Shamir, and Adleman [Xi and Hu, 2010].

## 4 Authentication Systems

In general, there are four main goals of modern cryptography: 1) confidentiality, 2) data integrity, 3) non-repudiation, and 4) authentication [Xi and Hu, 2010]. The last goal, authentication, concerns itself with verifying claims of identity. In the context of sending a message, the sender and receiver should be able to confirm the other's identity and the origin

of the message. Authentication is distinguished from *authorization*, which is the process of giving a party access to a system or data based on the confirmation of their identity.

In a knowledge-based authentication systems, a user's identity is verified via a piece of knowledge: a password, pass phrase, or a personal identification number (PIN). However, there are several shortcomings to this method: 1) As illustrated above, knowledge such as passwords and PINs can be easily forgotten, 2) Passwords and PINs can be guessed using social engineering 3) Even encrypted passwords are easily cracked with brute-forced and/or wordlist attacks 4) Passwords and PINs are easy to distribute and share in plaintext and 5) A password-based system cannot distinguish a genuine user from an attacker using stolen or forged credentials.

## 4.1 Biometric Authentication Systems

Rather than relying on a shared secret or key, in biometric systems, authentication relies on the physiological or behavior features of a person. Genuine users are recognized using characteristics like fingerprints, irises, or even ear shape [Xi and Hu, 2010].

In general, biometric systems comprise of five separate elements. A *biometric sensor* reads in the biometric information and usually does some quality checking on the sample, e.g. a fingerprint reader. The *feature extractor* consumes this raw biometric information and pulls out a relevant feature set (or template) that represents the data. In a fingerprint system, these features would include minutiae details that characterize the fingerprint. The *matcher* or *matching model* matches this sample template against

a pre-stored template and comes up with a score, i.e. the extent to which the sample matches the pre-stored template. Finally, it is common for systems to use a *database* to store known templates.

Biometric systems can be more reliable than traditional password-based systems because biometric features cannot be lost or forgotten and they are difficult to copy or forge and to share or distribute [Fengling Han, 2007]. However, they are not without their own set of issues including those of accuracy (i.e. false positive and negative matches), security (e.g. unlike a password it is impossible to replace a person's stolen biometric information), and privacy. Moreover, biometric systems are not impervious to attack, a biometric system comes with its own set of vulnerabilities.

## 4.2 Biometric System Attacks

Biometric systems come with their own set of unique attack vectors; according to Ratha et al, they can be categorized into eight types [Ratha et al., 2001]:

1. Fake biometric: attacker uses a reproduction of the biometric, e.g. a fake fingerprint
2. Replay attack: attacker replays an old recorded signal, e.g. presents old copy of fingerprint image to the system
3. Override feature extract: Attacker plants a Trojan horse in the feature extractor so that it produces a feature set that the attack chose
4. Override matcher: Attacker tampers with matcher to produce artificially high or low match rates
5. Tamper with the feature representation: Attacker replaces extracted feature set with a different, synthesized one

6. Tamper with stored templates: Attacker tampers with database holding the templates, e.g modifies a template to result in fraudulent authorization for an individual
7. Attack communication channel: Attack tampers with the templates when the are enroute from storage database to the matcher
8. Decision override: Recognition system works as expected but attacker changes final authentication decision at the last step

# 5 Biocryptographic Methods and Applications

According to Xi and Hu, among the types of attacks against biometric systems, those targeting templates can be hard to detect and can cause the most damage. [Xi and Hu, 2010] Therefore, for a system to be secure, biometric templates should always be encrypted, both when stored and during the matching process. However, because of the characteristics of biometric data, traditional methods that use non-smooth functions, like DES and RSA, cannot be used for encryption [Fengling Han, 2007]. For example, given a feature set derived from a fingerprint, even tiny variances in the plaintext feature set will produce wildly different encrypted feature sets, making it impossible to do feature matching using encrypted templates.

## 5.1 Template Encryption

In general, to encrypt a template, $T$ we use a secret key, $K_E$ and an encryption algorithm, $E$, such that the encrypted version, $C$ is given by:

$$\mathrm{C} = \mathrm{E}\ (\mathrm{T},\ K_E)$$

Then, to decrypt, we apply a decryption algorithm, $D$ to $C$ and a decryption key $K_D$ to get the template back:

T = D (C, $K_D$).

One biocryptographic technique is called key binding, in which the secret key and the biometric data (i.e. the template) are combined to produce an artifact in which both the template and the key are hidden. This artifact can then be shared publicly since it is computationally infeasible to decrypt the artifact directly [Xi and Hu, 2010].

### 5.1.1 Fingerprint Fuzzy Vault

A key-binding scheme proposed by Juels and Sudan is called the fuzzy vault algorithm [Juels and Sudan, 2006]. This algorithm is especially applicable to biometric data because it is *error-tolerant* and invariant to order. That is, data can be encoded using a set of values (e.g. a biometric feature set), and then unlocked with a *different* set of values as long as there is some threshold of overlap between the sets. Order invariance means that it does not matter which of the sets is used for locking and which is used for unlocking.

The basic algorithm works as follows: given a secret key and a template (i.e. feature set), first encode the key as the coefficients of a polynomial function, *p(x)*. Apply *p(x)* to every value in the feature set to produce a set of points that genuinely describes the polynomial. Then, to obscure the template data, generate a "chaff" set, points that are within the domain and range but do not lie on the polynomial. Finally, combine the two sets and scramble the order – this set of points is the "fuzzy vault". To decrypt the fuzzy vault, use the feature set provided by a user. If enough of the points match the set used to encode the data within a given error, the polynomial can be reconstructed and thus the secret key revealed [Xi and Hu, 2010,

Juels and Sudan, 2006].

As a proof of concept, I have implemented a simple "biometric" authentication system using the fuzzy vault algorithm to encrypt identity/fingerprint pairs. The code is available publicly at `https://github.com/jwoogerd/fuzzy_vault`.

## 6 Conclusion

Password-based authentication systems, though pervasive, are problematic in that users must manage many sets of credentials, creating an incentive to use and reuse weak, easy to crack passwords. Furthermore, a password-based system cannot distinguish between a genuine user and an attacker using stolen credentials. Biometric systems solve these problems by using physiological characteristics to uniquely identify genuine users, but come with their own set of challenges and vulnerabilities. Since many traditional cryptographic methods are unsuited to biometric data, the field of biocryptography explores the specialized cryptographic methods for securing biometrics systems.

## References

[buf, ] Fuzzy vault. `https://wiki.cse.buffalo.edu/cse545/content/fuzzy-vault`.

[Das, 2013] Das, R. (2013). *An Introduction to Biocryptography.* `http://www.nationalhomelandsecurityknowledgebase.com/cln/news/2013/11221.aspx`.

[Das, 2014] Das, R. (2014). *Biometric Technology: Authentication, Biocryptography, and Cloud-Based Architecture*. CRC Press.

[Fengling Han, 2007] Fengling Han, Jiankun Hu, X. Y. Y. W. (2007). Fingerprint images encryption via multi-scroll chaotic attractors. *Applied Mathematics and Computation*.

[Janrain, 2012] Janrain (2012). Online Americans Fatigued by Password Overload Janrain Study Finds. `http://janrain.com/about/newsroom/press-releases/online-americans-fatigued-by-password-overload-janrain-study-finds/`.

[Juels and Sudan, 2006] Juels, A. and Sudan, M. (2006). A fuzzy vault scheme. *Des. Codes Cryptography*, 38(2):237–257.

[Ratha et al., 2001] Ratha, N. K., Connell, J. H., and Bolle, R. M. (2001). An analysis of minutiae matching strength. In *Proc. 3rd AVBPA*, pages 223–228.

[Vijayan, 2012] Vijayan, J. (2012). Hackers crack more than 60% of breached LinkedIn passwords. `http://www.computerworld.com/article/2504078/cybercrime-hacking/hackers-crack-more-than-60--of-breached-linkedin-passwords.html`.

[Xi and Hu, 2010] Xi, K. and Hu, J. (2010). Bio-cryptography. In *Handbook of Information and Communication Security*, pages 129–157.