CERIAS Tech Report 2015-2
ErsatzPasswords Ending Password Cracking
by Mohammed H. Almeshekah, Christopher N. Gutierrez, Mikhail J. Atallah and Eugene H. Spafford
Center for Education and Research
Information Assurance and Security
Purdue University, West Lafayette, IN 47907-2086

ErsatzPasswords – Ending Password Cracking

Mohammed H. Almeshekah, Christopher N. Gutierrez,
Mikhail J. Atallah and Eugene H. Spafford

February 13, 2015

Abstract

In this work we present a simple, yet effective and practical, scheme to improve the security of stored password hashes rendering their cracking detectable and insuperable at the same time. We utilize a machine-dependent function, such as a physically unclonable function (PUF) or a hardware security module (HSM) at the authentication server. The scheme can be easily integrated with legacy systems without the need of any additional servers, changing the structure of the hashed password file or any client modifications. When using the scheme the structure of the hashed passwords file, *etc/shadow* or *etc/master.passwd*, will appear no different than in the traditional scheme.[1] However, when an attacker exfiltrates the hashed passwords file and tries to crack it, the only passwords he will get are the ersatzpasswords — the "fake passwords". When an attempt to login using these ersatzpasswords is detected an alarm will be triggered in the system that someone attempted to crack the password file. Even with an adversary who knows the scheme, cracking cannot be launched without physical access to the authentication server. The scheme also includes a secure backup mechanism in the event of a failure of the hardware dependent function. We discuss our implementation and provide some discussion in comparison to the traditional authentication scheme.

1 Introduction

Passwords are the most dominant form of online authentication and likely to remain so for a while despite their weaknesses. It thus behooves us to seek to protect them as much as possible. Within authentication servers, passwords are usually stored in a salted hashed format to prevent easy pre-image recovery. Nevertheless, an adversary who steals the list of hashed passwords can use brute-force to find a password p with a hash value $H(p)$ that equals the value stored for a given user. Later, the adversary can use p to impersonate the user at the authentication server.

There are a number of threats that come with the use of passwords. These threats fall into three main categories; technical, procedural and human related – these will be discussed in more detail in the following section. There have been a number of high-profile thefts of user passwords files in recent years. For example, Evernote reported the leakage of the hashed passwords for more than 50 million users [12]. Other attacks against Yahoo!, RockYou, LinkedIn and eHarmony has been reported [11] [26]. Furthermore, password cracking is often a precursor to more significant attacks as illustrated in [16].

The contribution of our work can be summarized in two main points: (i) we eliminate the possibility of any offline password cracking without physical access to the target's machine, (ii) when using this scheme the passwords' hashes file will appear no different than a traditional file and if an attacker uses traditional cracking tools to recover users' passwords he will "discover" fake passwords that will trigger an alarm when used. We refer to these fake passwords as "ersatz-passwords". There are somewhat similar schemes that have been proposed in the literature such as "Honeywords" [13] and "Failwords" [17]. However, our mechanism has the following unique advantages (i) eliminating the requirement of any additional server/components, (ii) never presenting the real user credentials to the attackers and, (iii) making password cracking impossible without physical access to the targeted machine. The scheme runs internally in the server without requiring any changes to the user interfaces, clients and/or experiences. A more detailed discussion of related literature is presented in the next section.

One additional contribution our scheme provides is that it imposes risks to any adversary who obtains a file of leaked usernames and passwords, causing mistrust within the market for such files, and rendering their use risky for many parties. This is because the unique property of our scheme of having the username and password file look identical to the file generated by the traditional authentication scheme. This property benefits not only the early adopters of the scheme, but the overall security of other (non-adopting) systems. This is one of the distinguishing features of using ersatzpassword in comparison to Honeywords [13], PolyPasswordHasher [5], SAuth [15] and others.

2 Background

2.1 Passwords

There have been many high profile incidents involving the leak of hashed passwords files [10]. Users are still using poor passwords, even with the existence of passwords policies that try to guide users towards choosing more secure passwords. This can be seen in the analysis of more than 70 million users' passwords [3]. Bonneau et al. presented an extensive comparative analysis of many authentication schemes replacing passwords [4]. However, passwords will remain in use for many users because of their convenience, ease of use, and ease of deployment.

2.2 Password-Related Threats

The convenient and versatile use of passwords comes with its own challenges. We define password-related threats as the attacks adversaries can launch to retrieve one or more valid passwords of current legitimate users of the systems. These host-based[1] threats can be grouped into three main categories.

Technical Threats

There are two sub-categories of technical threats associated with the use of passwords; server-side and client-side. Any piece of malware and/or key logger that can be installed at the user's machine to exfiltrate the user's password is a threat to any password-based authentication system. At the server side, adversaries can obtain the file of stored password information and then impersonate the system user using the stolen passwords. Strong host security is needed to protect the client and server systems, but there are multiple opportunities for an attacker to capture a copy of the stored password information.

A computer system needs to save an "authenticator" for every user during user enrollment that is used to verify the identity claim during the login phase. Current computer systems store a salted cryptographic hash (H) of the password along with the username. In a system with n users, we have the pairs $(u_1, H(p_1))$, $(u_2, H(p_2))$, ..., $(u_n, H(p_n))$, where u_i is the username of user i and p_i is the password of user i.[2] An attacker who steals this list can launch an offline attack to recover the hashed passwords using some dictionary and replicating the hashing algorithm used. Many tools already exist to automate an attack, such as John the Ripper[3]. There have been many attempts to address this challenge, usually falling into one of three major approaches; (i) significantly increasing the resources needed to match a password, (ii) strengthening user passwords to make their recovery process unlikely as they will be unlikely to be found in a dictionary, and (iii) instrumenting passwords files with fake decoy passwords triggering an alarm when used indicating that the password file has been attacked.

The development of password hashing algorithms from crypt to bcrypt, scrypt, and others is mainly driven by the goal of increasing the resources needed to crack the users' passwords. The introduction of *private* salts [14] was also intended to increase the work required for cracking the password files. In addition, increasing the number of rounds these algorithms apply to a password is a parallel approach to increasing the work factor.

Cappos and Torres proposed "PolyPasswordHasher" [5] as a scheme to protect passwords from offline dictionary attacks. Their scheme additionally protects passwords with a secret share obtained using Shamir Secret Sharing scheme [21]. The secret is saved in memory and used to verify passwords. One of the

[1] We are ignoring network snooping and other such remote mechanisms as our attention is directed only at securing host-based password databases.

[2] Salts, as an additional item in many systems, are described later.

[3] http://www.openwall.com/john/

limitations of their scheme is that it requires additional fields in the password file specifying which share to use. Also, if an attacker obtains a single access to the system memory they can steal the secret.

Deception has been used to address the threats associated with cracking password files. One approach is to inject fake accounts with passwords into the password file. Another approach is to place decoy password files in the system luring the attackers to access them believing they are the real files. Schemes such as *Honeywords* [13] are intended to confuse the attacker by presenting him with many passwords associated with a single username, where all of them are fake except one.

Procedural Threats

Password-recovery procedures associated with password-based authentication systems are sometime exploited to override current user passwords [20].

User-Centric Threats

Threats such as phishing, shoulder-surfing, password re-use, and others can be used to undermine the security of password-based authentication systems. Our approach does not address these issues, but can be used in conjunction with well-established approaches to minimize these risks. For instance, a filter applied at password enrollment can prevent password reuse.

2.3 Injecting Deceit

Deception has been used in computing since at least the 1980s [24, 22]. The prefix "honey" has been used to refer to a wide range of techniques that incorporate the act of deceit. The fundamental idea behind the use of the word "honey" is for those techniques to work they need to entice attackers to interact with them, i.e., fall for the bait: "honey." The term honeytokens was proposed by Spitzner [23] to refer to honeypots but at a smaller granularity. Yuill et al used the term *honeyfiles* to refer to files that have enticing names distributed in the system that act as a beaconing mechanism when accessed [29]. HoneyGen was also used to refer to tools that are used to generate honeytokens [1].

The use of deceit has been used to address some of the limitations associated with the use of passwords. Yue and Wang proposed a scheme named BogusBiter that shows when users fall for phishing by submitting fake, i.e. deceitful credentials, [28]. Rivest and Jules also proposed augmenting the password database in Unix with negative information such that cracked password files can be detected [13]. Bojinov et al. proposed Kamouflage, a scheme that is intended to protect the list of passwords used by a user and saved locally by a password manager [2]. Their scheme hides the real list with a a set of "fake" lists. Kontaxis et al. proposed an authentication scheme (SAuth) that requires each user's login attempt to be vouched by another service provider where an attacker cannot impersonate a user by simply obtaining the password for one web site [15]. They

use deception in their scheme as a way to address the common behavior of passwords reuse across multiple service providers. We use deceitful passwords in our scheme, referred to as *ersatzpasswords*. We discuss this in further details in the next section.

3 Technical Specification

3.1 Background

A number of cryptographic functions have been used in computer systems to protect passwords, including crypt, bcrypt, and scrypt. As discussed earlier, part of the motivation to develop additional algorithms is to make the cracking process of stolen password hashes files a resource-intensive process. Our scheme works with any of these underlying functions; we will denote the function used as (**H**). In later discussion we will use bcrypt to give a concrete example, but without any loss of generality.

Throughout this section we will assume the following format of the stored password file. For each user (i) in the system we have the following triplet, at a minimum, ($u_i,\ s_i,\ \alpha_i$) saved in the password file:

- Username (u_i).
- Multibyte (multi character) public salt (s_i).
- The hash of the user's password p_i as $\alpha_i = H(p_i \| s_i)$.

In addition, we will use a hardware-specific function denoted as ($\mathbb{HDF}$). This can be implemented as a physically unclonable function (PUF) [25], a hardware security module (HSM) [18] with a unique key, or any other mechanism of equivalent general functionality.

Our goal to enhance the security of the storage of passwords in 3 ways: (i) require the process of computing the hash of the password to require access to a hardware dependent function, thereby thwarting offline cracking of stolen password files, (ii) when an adversary attempts to crack the password file he will be presented with a fake password that can trigger an alarm at the server when used, and (iii) maintain the same appearance and format of the password file while implementing the new scheme. The final property is essential to the success of the deceptive process of injecting "fake" passwords. Unlike the Honeyword scheme, which mixes real passwords with fake ones, our scheme eliminates the ability of an adversary to obtain the real password (without physical access to the targeted machine during the cracking process) and seamlessly presents a fake password during an offline cracking process.

3.2 One-time Initialization

The initialization steps in our scheme are performed in two stages; *system-side* initialization and *user-specific* initialization. The former makes all the

users' saved, hashed, passwords machine-dependent – applying the hardware-dependent function as follows. The hardware-dependent function, $\mathbb{HDF}$ is applied to each stored password hash α_i and is then fed to the same hashing function, $\mathbf{H}$, with the original salt, s_i. After that, the output is stored in the password file replacing the old stored value. This system-wide initialization will have each user password stored in the file as the following

$$\beta_i = \mathbf{H}(\mathbb{HDF}(\alpha_i) \| s_i)$$

If an adversary obtains this file and tries to crack any user passwords, the probability that he will get any apparent match is negligible, even if a user password is from a standard dictionary. The cracking software will be searching its dictionary for a password equal to $p_i{}' = \mathbb{HDF}(\alpha_i)$ and when hashed will give β_i. An adversary with knowledge of the scheme cannot distinguish between a password file that was computed using our scheme or using the traditional scheme. Even under a stronger assumption, where the adversary knows that the file has been computed using the new scheme, the attacker gains no advantage as he cannot crack the user passwords without access to hardware used to compute the function $\mathbb{HDF}$. In the case where the attacker is an insider, any extensive use of the $\mathbb{HDF}$ can be easily noticed with a clear spike in API usage.

To incorporate the additional deceptive alarm component into our scheme — returning an "ersatzpassword" when the adversary attempts cracking the password file — we need to involve each user in a seamless fashion during any normal user authentication. This process requires the user to enter her password, which is a natural step during any authentication, (because the password is not actually stored or recoverable) and can be done during the first login process after the system wide initialization.

When the user attempts the first login after the initialization of our system, the password is checked using the original hash function to see if it matches. If so, the scheme will recompute the stored password value β_i as follows. The hardware-dependent function will be applied to the actual password p_i and then an ersatzpassword (p^*) will be chosen – we will discuss the use, choice and characteristics of ersatzpassword later in this paper. A new user-specific salt is then selected, to be used when computing the function $\mathbf{H}$, to satisfy the following property; $[\ s_i{}' = \mathbb{HDF}(p_i) \ \oplus \ p^*\ ]$. The scheme will take the first 128-bits of the result, assuming we are using a function $\mathbf{H}$ such as *bcrypt* that uses 128-bit salts, as the new salt overwriting the existing salt s_i.

We note that the ersatzpassword password length can be, at maximum, as long as the salt. In the current implementation of the *bcrypt* function, widely adopted to implement the hash function $\mathbf{H}$, the salt is 128-bit long. This gives us an ersatzpassword of up to 16 characters long. This does not impact the plausibility feature of the ersatzpassword, which will be discussed below. In the largest user passwords study analyzing more than 70 million real user passwords, Bonneau reports that users tend to use passwords with 6-8 characters [3]. If the ersatzpassword is shorter than the salt, the above computation will result in having the salt include some of the output of the $\mathbb{HDF}$ function. This does

not affect the security of the system as such output does not leak any useful information about the real password even to someone who has knowledge of the scheme and the length of the ersatzpassword $p*$.

To compute the stored value β our scheme calculates the following:

$$\beta_i = \mathbf{H}[\ (\mathbb{HDF}(p_i) \oplus s_i{}')\ \|\ s_i{}'\ ]$$

If the output of the $\mathbb{HDF}$ is longer than the salt, we address this as follows. We divide this output into chucks of length equal to the salt length. After that, we XOR these chunks together and then XOR the result with the salt $s_i{}'$. Finally, this becomes the input to the hash function $\mathbf{H}$ along with the concatenated salt.

The stored value in the password file will be in the same format used in traditional schemes. When an adversary tries to crack the users' passwords file, he will try to find a password $p_i{}'$ that when hashed using $\mathbf{H}$ will give β_i. In our scheme, we compute beta in a format equivalent to the traditional password storage where the password is p^*, i.e. $\beta = \mathbf{H}(p^* \ \|\ s_i{}')$. As a result, an attacker who is launching a dictionary attack against a stolen passwords file will likely find a result identifying p^* as the user password, which is the *ersatzpassword* injected in the system. When the adversary uses this password to login, an internal alarm will be triggered alerting the administrator that someone exfiltrated and attempted to crack the user passwords file.

3.3 Login

There are three main cases of login in our scheme: successful login, when the user enters the correct user/password pair; malicious login, when the adversary uses an ersatzpassword; and error login, when the username/password pair does not match anything. In this section we discuss how to evaluate the login request, in the presented order, and determine a login decision.

When the user i wants to login she presents the username and password $\bar{p}$ to the authentication server. The system identifies the username record and obtains the stored value β_i and the salt s_i associated with it. The scheme computes

$$\beta_i{}' = \mathbf{H}[\ (\mathbb{HDF}(\bar{p}) \oplus s_i)\ \|\ s_i\ ]$$

and checks whether $\beta_i{}'$ equals β_i, and if so the user is successfully authenticated.

If the authentication fails, the scheme checks whether the password presented is the *ersatzpassword*. This is done by computing

$$\beta_i{}'' = \mathbf{H}[\ \bar{p}\ \|\ s_i\ ]$$

and checking whether this equals β_i. If they are equal, this indicates that someone is trying to impersonate the user after cracking the passwords file and an internal alarm is triggered.

If the two values are not equal, this can be treated as an erroneous login. The system's policy for erroneous login can then be applied.

3.4 Password Administration

3.4.1 Password Change

The user's password change requests can be treated exactly as a new password. The only difference from traditional password schemes is that our approach mandates the generation of a new salt that satisfies the property discussed above, the XOR operation between the salt and the output of applying $\mathbb{HDF}$ on the password gives an ersatzpassword.

3.4.2 Backup

One of the major factors that hinders the use of hardware-dependent functions is the fact that the system catastrophically fails in the rare case where the hardware associated with the $\mathbb{HDF}$ fails or is no longer available. Thus, we outline a secure backup feature that can be used to recover the system in such a failure scenario. This process utilizes public-key encryption and is initialized by generating a suitably strong public/private key pair. The private key is never used in normal operation and can be stored in a secure vault offline. It is only needed in the recovery process. The public key is used during the system wide initialization process and during the process of password change.

When the system is initialized to adopt the new authentication scheme, all the current username u_i, password hash α_i and salt s_i triplets are encrypted using the public key and stored as a backup. In addition, whenever a user changes her password, the new value $\alpha_i{}'$ (the new hash value resulting from the new password using the traditional hash) is computed and the new triple overwrites or is appended to the backup log, along with the u_i and s_i values. As a result, the backup file with have the following list (u_i, s_i, α_i), for every user i in the system, encrypted under the public key.

If a recovery is needed after failure, the private key is fetched and used to recover the log file, which is then used to restore a traditional version of the password file. That file can be instantiated on new hardware, with a new $\mathbb{HDF}$, and users can be forced to reset their passwords — leading to transition to our new scheme as they do so.

It worth noting that decrypting the backup file using means of brute-force should not be practical. Even if the adversary, hypothetically, manages to recover the information in the backup file the resultant password security is at least as strong as the currently deployed schemes. The cost in storage and computation to build the recovery log is minimal.

3.4.3 Previous Passwords Storage

It is common for many authentication server to store previously used users' passwords to prevent users from recycling them [9]. This can put users at risk when such files are compromised. Although users are not using these passwords to login, they can be used to impersonate users at other websites. If systems

need to store these passwords nevertheless, our scheme provides an additional advantage over traditional methods of securely storing these passwords.

As our scheme saves the user passwords in a machine-dependent format, using the function $\mathbb{HDF}$, we can have some assurance that this password cannot be cracked offline without physical access to the target machine. Later, when attempting to store the previous passwords used in the system, we can save the passwords using the $\mathbb{HDF}$ function.

3.4.4 Fail-Safe Procedure

We finally point out that in addition to the backup mechanism discussed above to recover the system in the rare case of $\mathbb{HDF}$ function failure, our scheme comes with an intrinsic fail-safe procedure. In this case, we can use the traditional authentication method to check the passwords, comparing $\mathbf{H}(p_i \mid\ s_i)$ with the stored value β_i, where the effective user password becomes the ersatzpasswords.

4 ErsatzPasswords – The Use of Deception

The scheme presented in this paper provides the guarantee that stored users passwords cannot be cracked without physical access to the hardware-dependent function ($\mathbb{HDF}$). With the increased complexity of computer systems and targeted attacks computer systems are still vulnerable to security compromise and the list of stored passwords can be stolen. In addition, the latest Verizon Data Breach Investigation Report (DBIR) shows that about 50% of attacks thwarting authentication mechanisms take months or longer to be discovered. Even worse, 88% of these attacks are discovered by external parties. Integrating deceptive passwords in the design of our scheme addresses these two issues.

When attackers obtain the stolen credentials passwords and apply the cracking process, we could design our scheme to negatively respond to this activity as in [7]. This allows an attacker, who obtains this file, to notice such behavior and simply look for other vulnerabilities to exploit. Instead, the scheme is designed to present an attacker with plausible deceptive passwords leading him to believe that he successfully cracked the passwords file. When a login is attempted using the deceptive passwords, system defenders will be immediately alerted to two facts: (i) that the login credentials database was leaked; and (ii) that an attacker is currently trying to impersonate the system's users to gain access. This design enables system defenders to use internal controls for detecting credentials' database leakages, and for alerting them of an ongoing attack before it succeeds.

4.1 ErsatzPasswords Generation

The process of generating an ersatzpassword for each user account can be formalized as follows. Let $Gen(u_i, p_i)$ be the function that takes the user's username and password and outputs the selected ersatzpassword. This function should

provide two essential properties; plausibility and non-deducibility. The former ensures that an ersatzpassword generated by $Gen()$ is plausible to an adversary as a real user password. The latter provides the guarantee that even when an adversary knows the scheme, he cannot deduce any information from the ersatzpassword about the real user password. We define these two properties more formally later in the paper. We want this function to be randomized and to give us an ersatzpassword every time we use it. Of course, the generated ersatzpassword should have the properties discussed later in this paper. We present below several constructions of how to realize this function and discuss the advantages and disadvantages of each construction.

4.1.1 Total Password Replacement

When $Gen()$ receives the user password it can generate the ersatzpassword using the following procedure. For every character in the user password, replace it with a randomly chosen character from the same category (alphabetical with alphabetical, a digit with a digit, and a special character with a special character). After this replacement process, a cyclic shift is applied to the password by a random number of positions to generate the ersatzpassword.

We note that this process reveals two properties of the real password to an adversary when he views the ersatzpassword: the password's length and its character composition. In this case adversaries can use probabilistic context-free replacement to significantly narrow down the space of possible user passwords using knowledge of the ersatzpassword [27]. One of the potential ways to overcome this is to randomly truncate or append some random characters to generate the ersatzpassword.

4.1.2 List-Based

One of the most straightforward ways of generating the ersatzpassword using $Gen()$ is to randomly choose a word from an internal dictionary of candidates. This realization of $Gen()$ has two major limitations: the generation of ersatzpassword is not influenced by user-specific information and the existence of such a list in the system can affect the stealthiness of the deceptive component (the existence of the list is a sign that such a scheme is currently being used by the system). The former limitation is not as significant because the attacker never sees the "real" users' passwords. The advantage of using such method is the ability to have high degree of plausibility of the ersatzpasswords. We can compile a list of the some of the previously leaked passwords used by real users and use them as our ersatz passwords.

4.1.3 Grammar-Based Methods

Bojinov et. al propose a new method of generating plausible user passwords in [2] extending the work of Ross et. al. in [19] and Weir in [27]. Their method is similar to our *total password replacement* method, however they tokenize the

password representing distinct syntactic elements. For example, the password *"wtyy234ou*"* has the following token sequence $W_1 = \{wtyy\}|D_2 = \{234\}|W_3 = \{ou\} \mid S_4 = \{*\} \mid$. When generating the ersatzpassword, each token will be replaced with another token, of the same length, from a dictionary.

The main drawback of this method is that it leaks the type, number, and length of tokens of the original password. We address this concern by enhancing their implementation of $Gen()$ as follows. After tokenizing the password, we perform the following:

- We can randomly append or delete k tokens. For example, let say we add token S_5 to the above password.
- After that, we can randomly shuffle the order of these tokens. In the above example, the shuffle can give us the following order $W_3 \mid S_5 \mid D_2 \mid S_4 \mid W_1$.
- Finally, we randomly choose a word from a dictionary that matches each token. The chosen token can have length that is different from the original token. In our example, let's say that W_3 = "*abc*", S_5 = "!", D_2 = "10", S_4 = " + " and W_1 = "*test*".

Using the grammar-based method with our modification can generate the following ersatzpassword *"abc!10+test"*.

4.1.4 Using User Input

Our discussion so far assumes that the scheme can work without any interaction with the system users. However, we note that ersatzpassword can be constructed with implicit or explicit user input. Many authentication servers save previously used user passwords in the system preventing users from recycling their old password when their current password expires. This implicit user input, previously chosen user passwords, can be used as the ersatzpassword for this user account. With explicit user input, the system can prompt the user to enter another password during registration and use this as the ersatzpassword password.

The main advantage of using implicit user input is ensuring a high degree of plausibly, discussed later, of the ersatzpassword as this has been previously used as a real password. However, this method suffers from two major disadvantages. First, if an adversary cracks the password file and recovers the ersatzpassword, this might put the user in danger as users are known to reuse passwords across multiple sites [8]. Second, this has the potential of signaling a false alarm in the case where the user forgets and uses his previous password to login.

Explicit user input requires some changes to the user interfaces. More importantly, users are likely to pay less attention, choosing very guessable passwords and/or confusing the ersatz passwords with their real ones leading to the problem of false alarms. In addition, users may provide an additional, ersatzpassword that is closely related to their real password, e.g. by appending a number or a character to their real password to create the ersatzpassword.

A combination of several of these methods may be the best approach.

4.2 ErsatzPasswords Properties

Incorporating deception in this scheme actively feeds an adversary cracking a stolen password file with some ersatzpasswords chosen to trigger internal alarms when used. These passwords should have the following properties to ensure their effectiveness.

4.2.1 Plausibility

When an adversary is cracking a password file, these words will present themselves as a successful outcome, i.e. when hashed along with the salt they will match the stored hashed user password in the traditional way. For the effectiveness of the scheme, these need to be plausible user passwords. Thus, some dictionary and generation algorithms should be present to produce plausible ersatzpasswords (so their generation is random subject to plausibility rather than in absolute terms).

We can define a plausible generator function $Gen()$ more formally as using the following game:

- The adversary views many runs of the function $p^* = Gen(u, p)$ where she can choose the values of u and/or p (p^* is the ersatzpassword).
- $\bar{p^*} = Gen(u, \bar{p})$ is computed and $\bar{p^*}$ and $\bar{p}$ are presented to the adversary.
- The adversary outputs (1) if he believes $\bar{p^*}$ is the ersatzpassword with probability Pr.

We say that $Gen()$ is a plausible function if the probability for adversary success is one-half. In other words, the adversary cannot do better than random guessing which of the two passwords is the ersatzpassword. That is, $Pr = 1/2+\epsilon$ where ϵ is negligible.

4.2.2 Typo-Resilience

When the user is typing her real password, she may make a mistake by mistyping some characters. The ersatzpassword associated with the account should have enough edit distance from the actual password to ensure that an alarm is not triggered by mistake. As the real user password is present when selecting which ersatzpassword to use, the server can easily compute an edit distance to ensure that the user does not mistype the ersatzpassword during the login process.

4.2.3 Non-Deducibility

It is essential for the ersatzpassword to not reveal any useful information about the real user password. Even though we do not actively give adversaries the ersatzpasswords, we store them with the same level of protection used to store current real users' passwords. We define the function $Gen()$ to provide non-deducibility using the following game:

- The adversary views many runs of the function $p^* = Gen(u, p)$ where she can choose the values of u and/or p (p^* is the ersatzpassword).
- The adversary chooses two passwords p_1 and p_2, and send then to function $Gen()$.
- $Gen()$ flips a coin and computes $p^* = Gen(u, p_1)$ if she gets heads or $p^* = Gen(u, p_2)$ otherwise. p^* is then presented to the adversary.
- The adversary outputs (1) if she thinks $p^* = Gen(u, p_1)$ and (0) otherwise with probability Pr.

We say that $Gen()$ is a non-deducible function if the probability for adversary success is half. In other words, the adversary cannot do better than random guessing which of the two passwords was used to generate p^*. That is, $Pr = 1/2 + \epsilon$ where ϵ is negligible.

4.2.4 Policy Adherence

It is essential that ersatzpasswords adhere to any system-wide policy of how users' password should appear. For example, some restrictions can be imposed on the length, format and composition of user passwords. An adversary who sees any password violating the system's policy can detect that this cannot be a real password as the system would not have accepted it. In addition, some websites mandate that user password cannot be dictionary words. In these cases, using a password list as the method to generate ersatzpassword can be challenging as it is not trivial to come up with a long list satisfying each server's policy. In addition, any change to the policy would require recomputing the list again. However, the use of grammar-based approaches, similar to the one illustrated above, can be much simpler as grammar can become part of the input of the generator function $Gen()$.

4.2.5 Crackable

Part of the plausibility aspect of our scheme deciding whether all ersatzpasswords should be crackable or not. Generally, this should not be the case. Many current systems add more stringent requirements of password choice to high privileged users. When they become easily crackable, this might increase adversary suspicion. In addition, it would also look suspicious if all user passwords were crackable. It might be wise to use some randomly-generated ersatzpasswords within a system to enhance the scheme's plausibility.

5 Implementation and Analysis

In this section, we describe the implementation details of the proposed system. A preliminary evaluation is also presented followed by a discussion driven by the observed results.

5.1 Implementation Details

We implemented the proposed scheme by modifying the authentication mechanism in an FreeBSD operating system. The *pam_unix* Pluggable Authentication Module (PAM), which handles the user authentication process, is modified to incorporate the proposed system. The design decision is driven by the simplicity of PAM modules as well as the preservation of expected behavior during user authentication. The effectiveness of the deception relies on the fact that the user authentication system appears no different than standard FreeBSD user authentication.

The proposed system relies on two key components: the hardware dependent function $\mathbb{HDF}$ and the ersatzpassword generation function $Gen()$. We used the basic Yubico's YubiHSM, a USB hardware security module, as our $\mathbb{HDF}$. Specifically, $\mathbb{HDF}$ is a HMAC-SHA1 with a fixed secret key (k) internally stored inside the HSM:

$$\mathbb{HDF}(p) := \text{HMAC-SHA1}_k(p)$$

For the ersatzpassword generation, $Gen()$, we implemented the List-Based approach described in section 4.1.2. This choice was mainly driven by the fact that we can pre-select ersatzpasswords and have more accurate measurements. The code can be easily modified to choose any ersatzpassword generation algorithm. As a proof of concept, we used a list of six-character dictionary words as our ersatzpasswords from [6]. A password is selected from the dictionary of 15,788 and used as the ersatzpassword during user account initialization.

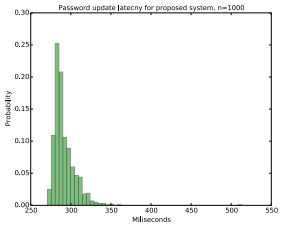
5.2 Analysis

We analyze two authentication processes when comparing our implementation of the new authentication scheme and the standard FreeBSD authentication. First, we compare the latency for adding a new user into the system and the latency for authenticating a valid user. Second, the storage of cryptographic hashes of the user's password must appear and behave as in a typical FreeBSD operating system. In addition to maintaining the fidelity for accurate user authentication user password hashes must also work with conventional password cracking tools such as John the Ripper [4] to ensure the plausibility of the ersatzpasswords. We conducted our analysis on a FreeBSD virtual machine with a single core clocked at 2.7 Ghz.

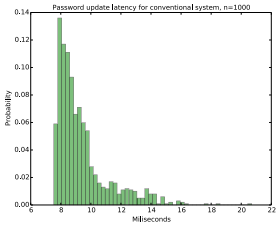
5.2.1 Password Update and Authentication Latency

To evaluate the performance of our authentication module, we compare the latency with the standard `pam_unix` module found in FreeBSD. Two measurements are considered: the latency to update an existing password and the latency to authenticate a user. The password is fixed to "password" for all experiments. Additionally, the authentication evaluation also considers the latency of using "ersatz" for the ersatzpassword. The evaluation consists of running the

[4] http://www.openwall.com/john/



(a) Distribution of password update latency in our system.



(b) Distribution of password update latency in FreeBSD.

Figure 1: Comparison of update latency in the modified and standard FreeBSD.

`pam_chauthtok` and `pam_authenticate` as found in `passwd` and `login`. Password update and authentication latencies are sampled 1000 times independently on an idle FreeBSD virtual machine.

As shown in figure 1, the median latency time to update a user's password for the proposed ersatz system is 287.3 ms while the latency on a standard FreeBSD system is 8.8 ms. These results indicate that further optimization is needed to reduce the latency for the our module to match the expected behavior of the standard FreeBSD `pam_unix` module.

A similar pattern is observed when comparing authentication latency. Figure 2 illustrates the latencies in system response observed when providing a valid password and an ersatzpassword in our system in comparison with the latency in system response when providing a valid password in a conventional FreeBSD system. Note that the latencies difference compared between our system and the conventional system are similar to the password update latency. The median system latency for authentication in our system is 277.76 ms when providing the correct password and 281.95 ms when providing the ersatzpassword. The system's latency for authenticating a valid user on a standard FreeBSD system is 5.14 ms.

We note that there are a number of reasons for the observed performance difference. The YubiHSM APIs are written in python and the implementation of our scheme is written in C as a modified `pam_unix` module. A call from C to Python has an impact on the system performance. To validate our concern,

we used `pmcstat`[5] to profile our modified `pam_unix` module. The results from `pmcstat` showed that the largest bottleneck is found in the `libpython2.7.so` library. Specifically, the bottleneck is `PyEval_FrameEx` which interprets and executes bytecodes from a given frame. Another bottleneck is `PyObject_Malloc` which is indirectly called when converting a C string to a Python string. Such conversion is needed in our modified `pam_unix` module when initializing the YubiHSM and generating a salt or the hash.

Another reason is that we are using a virtual machine and FreeBSD in that environment does not have a direct access to the USB port to call the YubiHSM APIs. A third reason for the performance difference may be the fact that we are using a basic $\mathbb{HDF}$ function, namely the YubiHSM, which is not optimized for performance. A built-in device rather than a USB device might also provide a speed improvement. We believe that a combination of optimizations might bring the times close enough that it would not be obvious to an observer what might be in use on the system. If that is not a consideration, the additional latency of the current, unoptimized implementation would be clearly insignificant in normal operation.

5.2.2 Crackable Ersatz Hashes

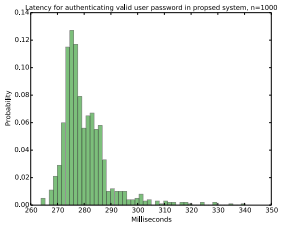
To demonstrate that the our scheme produces crackable hashes, we generated 1000 hashes with the real passwords of `password1, password2,` $\cdots$`, password1000` and ersatzpasswords randomly selected from our list of six-character dictionary words. We ran *John the Ripper* on all 1000 hashes created by our scheme to crack the generated hashes. John the Ripper successfully cracked all 1000 hashes and retrieved all the ersatzpasswords.

One interesting observation is that if two users select the same exact password and if the ersatz password selected is less than the length of the salt, then some of the bits in the salt are the same between both users. Such an anomaly would be unlikely a conventional `master.passwd` file and may raise suspicion of the deception. This can be mitigated by properly generating ersatzpasswords to avoid such situations.

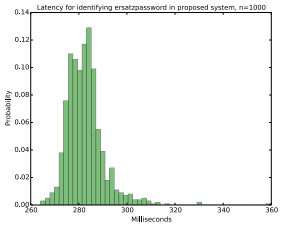
6 Conclusion

Passwords have been widely regarded as one of the weak points of securing any digital system. They come with their inherent weaknesses in how they are chosen, stored, memorized and managed. In this paper, we presented a scheme that address the wide-spread threat of stealing hashed password files and cracking them offline to impersonate user accounts to further infiltrate computer systems. Our scheme makes it impossible for an adversary to recover user passwords from their hashed format without physical access to the targeted machine. We show how can we instantaneously protect any system with the involvement of its user. Furthermore, we discussed how we can deceive an

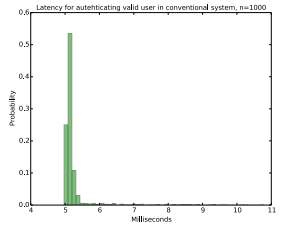
[5] `https://wiki.freebsd.org/PmcTools`



(a) Distribution of user authentication latency in our system using a valid password.



(b) Distribution of user authentication latency in our system using an ersatzpassword.



(c) Distribution of user authentication latency in FreeBSD.

Figure 2: User authentication latency in modified and standard FreeBSD.

attacker who steals the hashed users' passwords file by presenting him with ersatzpasswords that work as "decoy" passwords that trigger an alarm when used to access the system. We discussed how to generate these passwords and their properties. Finally, we implemented our scheme discussing the design decisions and the performance analysis. Our goal is with the deployment of our scheme, we can end the possibility of cracking user passwords and, at the same time, detect any exfiltration and cracking attempt on users' hashed password file.

7 Acknowledgments

The authors would like to extends their thanks to Jeff Avery, Dan Trinkle and Keith Watson for their time, discussion, and ideas they provided. Portions of this work were supported by National Science Foundation Grants CPS-1329979, Science and Technology Center CCF-0939370; by an NPRP grant from the Qatar National Research Fund; NGCRC grant; and by sponsors of the Center for Education and Research in Information Assurance and Security. The statements made herein are solely the responsibility of the authors.

8 Availability

A fully implemented `pam_unix` module can be obtained from `https://github.com/cngutierr/ErsatzPassword`

References

[1] Bercovitch, M., Renford, M., Hasson, L., Shabtai, A., Rokach, L., and Elovici, Y. HoneyGen: An automated honeytokens generator. In *Intelligence and Security Informatics (ISI), 2011 IEEE International Conference on* (2011), IEEE, pp. 131–136.

[2] Bojinov, Hristo and Bursztein, Elie and Boyen, Xavier and Boneh, D. Kamouflage : Loss-Resistant Password Management. In *Proceedings of the 15th European conference on Research in computer security* (2010), Springer-Verlag, pp. 286—-302.

[3] Bonneau, J. The science of guessing: Analyzing an anonymized corpus of 70 million passwords. In *Proceedings - IEEE Symposium on Security and Privacy* (2012), pp. 538–552.

[4] Bonneau, J., Herley, C., Van Oorschot, P. C., and Stajano, F. The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *Proceedings - IEEE Symposium on Security and Privacy* (2012), pp. 553–567.

[5] Cappos, J., and Torres, S. PolyPasswordHasher: Protecting Passwords In The Event Of A Password File Disclosure. Tech. rep., 2014.

[6] Chew, J. Common Six-Letter Words.

[7] Cvrcek, D. Hardware Scrambling - No More Password Leaks. Tech. rep., 2014.

[8] Das, A., Bonneau, J., Caesar, M., Borisov, N., and Wang, X. The Tangled Web of Password Reuse. In *NDSS '14: The 2014 Network and Distributed System Security Symposium* (San Diego, CA, USA, 2014).

[9] Defense Information Systems Agency (DISA) for the Department of Defense (DOD). Application security and development: Security technical implementation guide (STIG). Tech. rep.

[10] DeLuca, M., and Pepitone, J. eBay Warns Customers to Change Passwords After Database Hacked, 2014.

[11] Gaylord, C. LinkedIn, Last.fm, now Yahoo? Don't ignore news of a password breach.

[12] Gross, D. 50 million compromised in Evernote hack, Mar. 2013.

[13] Juels, A., and Rivest, R. L. Honeywords: making password-cracking detectable. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security* (2013), ACM, pp. 145–160.

[14] Klein, D. V. Foiling the Cracker; A Survey of, and Improvements to Unix Password Security. In *14th DoE Computer Security Group* (May 1991).

[15] Kontaxis, G., Athanasopoulos, E., Portokalidis, G., and Keromytis, A. D. SAuth: protecting user accounts from password database leaks. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security* (2013), ACM, pp. 187–198.

[16] Perlroth, N. Hackers in China Attacked The Times for Last 4 Months.

[17] Rao, S. Data and system security with failwords, 2005.

[18] Requirements, S. Hardware Security Module (HSM). *Security* (2009), 1–26.

[19] Ross, B., Jackson, C., Miyake, N., Boneh, D., and Mitchell, J. C. Stronger password authentication using browser extensions. In *Proceedings of the 14th conference on USENIX Security Symposium - Volume 14* (2005), p. 2.

[20] Schechter, S., Brush, A. J. B., and Egelman, S. It's no secret Measuring the security and reliability of authentication via 'secret' questions. In *Proceedings - IEEE Symposium on Security and Privacy* (2009), pp. 375–390.

[21] Shamir, A. How to share a secret. *Commun. ACM 22*, 11 (Nov. 1979), 612–613.

[22] Spafford, E. More than Passive Defense, 2011.

[23] Spitzner, L. Honeytokens: The other honeypot, 2003.

[24] Stoll, C. P. The Cuckoo's Egg: Tracing a Spy Through the Maze of Computer Espionage, 1989.

[25] Suh, G. E., and Devadas, S. Physical unclonable functions for device authentication and secret key generation. In *Proceedings - Design Automation Conference* (2007), pp. 9–14.

[26] Weir, M., Aggarwal, S., Collins, M., and Stern, H. Testing metrics for password creation policies by attacking large sets of revealed passwords. In *Proceedings of the 17th ACM conference on Computer and communications security* (2010), ACM, pp. 162–175.

[27] Weir, M., Aggarwal, S., De Medeiros, B., and Glodek, B. Password cracking using probabilistic context-free grammars. In *Proceedings - IEEE Symposium on Security and Privacy* (2009), pp. 391–405.

[28] Yue, C., and Wang, H. BogusBiter: A transparent protection against phishing attacks. *ACM Transactions on Internet Technology (TOIT) 10*, 2 (2010), 6.

[29] Yuill, J., Zappe, M., Denning, D., and Feer, F. Honeyfiles: deceptive files for intrusion detection. In *Information Assurance Workshop, 2004. Proceedings from the Fifth Annual IEEE SMC* (2004), IEEE, pp. 116–122.