Fighting Coercion Attacks in Key Generation using Skin Conductance

Payas Gupta
School of Information Systems
Singapore Management University
payas.gupta.2008@phdis.smu.edu.sg

Debin Gao
School of Information Systems
Singapore Management University
dbgao@smu.edu.sg

Abstract

Many techniques have been proposed to generate keys including text passwords, graphical passwords, biometric data and etc. Most of these techniques are not resistant to coercion attacks in which the user is forcefully asked by an attacker to generate the key to gain access to the system or to decrypt the encrypted file. We present a novel approach in generating cryptographic keys to fight against coercion attacks. Our novel technique incorporates the user's emotional status, which changes when the user is under coercion, into the key generation through measurements of the user's skin conductance. We present a model that generates cryptographic keys with one's voice and skin conductance. In order to explore more, a preliminary user study with 39 subjects was done which shows that our approach has moderate false-positive and false-negative rates. We also present the attacker's strategy in guessing the cryptographic keys, and show that the resulting change in the password space under such attacks is small.

1 Introduction

Many techniques have been proposed to generate strong cryptographic keys for secure communication and authentication. Some of these techniques, e.g., those using biometrics [15, 24, 27, 28, 35], offer desirable security properties including ease of use, unforgettability, unforgeability (to some extent), high entropy and etc. However, most of these schemes are not resistant to coercion attacks in which the user is forcefully asked by an attacker to reveal the key [32]. When the user's life is threatened by an attacker, one would have to surrender the key, and the system will be compromised despite all the security properties described above. In this paper, we present a novel approach to protection against coercion attacks in generating keys.

For a cryptographic key generation technique to be coercion attack resistant, it is required that when the user is under coercion, he/she will have no way of generating the key, or the key generated will never be the same as the one generated when he/she is not being coerced. If this requirement is met, then an adversary would not apply any threat to him/her because the adversary understands that the user would not be able to generate the key when he is threatened to do so. Here we assume that the coercion resistance property is publicly known to everyone, including the attackers; otherwise it might lead to a dangerous situation for the user, a problem we do not address in this paper.

To show how desirable it is to have a coercion-resistant cryptographic key generation technique, here we list a few scenarios in which such a technique could be useful:

- Bank's vault and safe: According to statistics released by the FBI [17], there were $1,094$ reported robberies (out of which 58 cases were of vault/safe robberies) of commercial banks between July 1, 2009 and September 30, 2009 totaling more than \$9.4 million. If such systems are used to fight against these attacks, then managers will never be forced to open the vault.
- Cockpit doors on airliners: The hijackers of the September 11, 2001 use the fueled aircraft as a missile to destroy ground targets. If the cockpit doors on airliners are well equipped with coercion resisted techniques, then hijackers can never force a flight attendant to open the door.
- Secret/capability holders in a war: secret and capability holders would not be forced to reveal the secret or use the capability.

In this paper, we explore the incorporation of user's emotional status (through the measure of skin conductance) into the process of key generation to achieve coercion resistance. We demonstrate this possibility by incorporating skin conductance into a previously proposed key generation technique using biometrics [24] (see Figure 1).

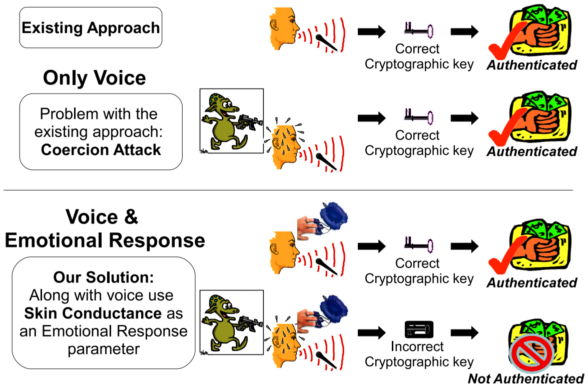

Existing Approach
Only Voice
Correct Cryptographic key
Authenticated
Problem with the existing approach: Coercion Attack
Correct Cryptographic key
Authenticated
Voice & Emotional Response
Correct Cryptographic key
Authenticated
Our Solution: Along with voice use Skin Conductance as an Emotional Response parameter
Incorrect Cryptographic key
Not Authenticated

Figure 1: Coercion attacks in key generation

Incorporating skin conductance information into key generation is nontrivial. First, the fact that a change in a user's emotional status leads to changes in a user's skin conductance does not necessarily mean that our proposed technique is coercion resistant. If known patterns exist in such changes, an attacker might be able to guess the skin conductance of the user when he is not nervous by, e.g., flipping a few bits of the feature key (see Section 4) generated from the skin conductance of the user when he is nervous. We analyze this attack and its consequences, and show that the reduction in password space is small.

Second, we hope that the key generation algorithm will take in the least amount of user specific information except the live data collected when it is used. This is because the key generation algorithm might be executed from the client's machine, and the inputs to the algorithm could potentially be retrieved by the attacker during a coercion attack. However, when dealing with biometrics data, removing such user specific information from the inputs of the algorithm is not plausible, as different people have different sets of consistent and inconsistent biometric features. The algorithm would have too high false-negative rates without this additional user specific information. We propose using only user-specific feature lookup tables which contain valid key shares or garbage. We also analyze conceivable attacks that result from our proposal.

Third, it is nontrivial how a user study can be performed to evaluate our technique. We need to collect biometric data corresponding to different emotional states of real human beings. Efforts in this area are more demanding than traditional efforts to get pattern recognition data [31]. To analyze the effectiveness of our proposal, we perform a user study to see how one's skin conductance changes when he/she is being coerced. This is used to evaluate the false-positive and false-negative rates of our model, and to analyze the attacker's strategy in guessing the cryptographic key. With 39 participants in our user study, we find that our technique enjoys moderate false-positive and false-negative rates in key generation. Furthermore, we find that the reduction in the password space for an informed attacker is small.

The rest of the paper is organized as follows. In Section 2, we discuss some state-of-the-art approaches in cryptographic key generation and recognition of emotional response. Background knowledge about the chosen biometrics and fingerprint are discussed in Section 3. In Section 4, we present the details of our approach in key generation using skin conductance and voice. The user study and results are presented in Sections 5 and Section 6 respectively. We conclude in Section 7 with some plausible future work.

2 Related Work

In this section we review some of the techniques and methodologies used to generate cryptographic keys from biometrics and some previous work on the emotion recognition schemes using physiological signals.

Many key generation techniques from biometrics, e.g., voice, iris, face, fingerprints, keystroke dynamics, and etc., have been proposed in the last decade [15, 24, 27, 28, 35]. The pioneer work in cryptographic key generation from behavioral biometrics uses keystroke dynamics of a user while typing the password [25]. The features of interest are the duration of keystrokes and the latency between each pair of keystrokes. The generated cryptographic key is called the hardened password. However the password generated is not very long and is susceptible to brute-force attacks [25]. Another method using secret sharing was proposed to generate the biometric key from voice [24]. The distinguishing biometric features are selected based on the separation between the authentic and the imposter data, and then binarized by some thresholds. However, this method is not resistant to coercion attacks (which our proposed model trying to target), as the attacker can force the user to speak out the password in a normal way. We will discuss key generation approach from voice in more detail in the formal framework of our model (see Section 3).

Another work on key generation from voice uses phonemes instead of words, as it is possible to generate larger keys with shorter sequences [15]. Using the information of the voice model and the phoneme information of the segments, a set of features are created to train an SVM (Support Vector Machine) that could generate a cryptographic key. False-positives and entropy of the system were not demonstrated, which does not give a clear picture of the security of the scheme.

There are many risk and security concerns over biometric systems [32, 33, 40]. Some of the threat models include fake biometrics at the sensor, tampering with the stored templates, coercion attacks. Biometrics liveness detection is proposed to thwart fake biometrics attacks, e.g., by using perspiration in the skin [1] or blood flow [22]. However, no previous work has been proposed to resist coercion attacks in generating cryptographic keys using biometrics. There have been suggestions like panic alarm or duress code to fight against coercion attacks, but they are different from what we are proposing here because in previous schemes users *choose* not to generate the key but to send a signal to authorities without catching the adversary's attention, whereas in our scheme we require that users simply will not be able to generate the key. It is clear that our scheme offers much stronger security properties.

Previous work also shows that emotion recognition using physiological signals, affects from speech, and facial expressions have various success rates between 60% and 98% [31]. Although many techniques have been proposed for emotion recognition [31, 20, 29, 21], none has looked into the incorporation of emotional status into key generation as what we propose in this paper.

3 Background

In this section, we present some background knowledge of voice and skin conductance, and discuss why in future an addition of fingerprint in our model would be better as an authentication measure for the protection against coercion attack. We also discuss the reasons for the selection of these features and the advantages over others in terms of acceptability, feasibility and usability.

3.1 Why Skin Conductance?

An emotion is a mental and physiological state associated with a wide variety of feelings, thoughts, and behavior. Emotions are subjective experiences, often associated with mood, temperament, personality, and disposition [11]. This emotional behavioral change is the key component in our model in fighting against coercion attack. Several physiological peripheral activities have been found to be related to emotional processing of situations. Many physiological parameters were studied for emotion recognition, e.g., heart beat rate [3] (HR), skin conductance [23] (SC), EMG (Electromyography) signals, ECG (Electrocardiography) signals, body temperature, BVP (Blood Volume Pulse) signals, and etc., among which HR and SC are especially attractive due to their strong association with behavioral activation system (BAS) and behavioral inhibition system (BIS) respectively [14].

SC is the change in the electrical properties of an individual person's skin caused by an interaction between environmental events and the individual psychological state. Human skin is a good conductor of electricity and when subject to a weak electrical current, a change in the skin conductance level occurs [42]. We chose SC over HR for the following reasons.

1. The skin conductance is one of the fastest responding measures of stress response [16]. It is one of the most robust and non-invasive physiological measures of autonomic nervous system activity [7]. Researchers have linked skin conductance response to stress and autonomic nervous system arousal [37].

2. The change in HR not only accounts for stress but for many other reasons, including jogging or doing some heavy work load. SC, on the other hand, has been shown to be a promising measure in experimental studies [36] for its reliability.

3. According to [41], HR is also impacted when stress levels rise but the shifts take a bit of time to happen and by the time the changes are noticeable the triggering stimulus is long past, whereas SC responses are rapid and easy to measure.

4. HR is not suitable to our model due to prevailing feasibility issues. HR can be measured using an Electrocardiogram (ECG) machine or a stethoscope. Using an ECG machine is impractical because it is very cumbersome due to many (at least three) electrodes required and installation costs [6]. Stethoscope is not good either because different placements of the stethoscope could lead to high FTC rate (failure to capture rate) [30].

5. Using SC has an extra advantage as it can be measured simultaneously while fingerprints are being scanned. This ensures that SC is measured from the authentic person (more on this in the coming subsection). The wide acceptance of finger scanning [18, 39] also suggest that SC measurement would have the potential to gain user acceptance.

There are some limitations of using skin conductance as with any other biometric. Some skin lotions can be used to manipulate the skin conductance level. In a test done by [34], the usage of specific solutions produced significant increase in skin water content, and was indicated by increase in skin conductance level. According to the product after the application of the cream by EncoSkin, skin moisture level can be significantly increased which can be monitored by skin conductance [12].

3.2 Why Voice?

Voice has been used previously to generate cryptographic keys [15, 24]. Voice as a biometric is desirable for generating keys for two important reasons. First, it is the most familiar way of communication, which makes it ideal for many applications. Second, voice is a dynamic biometric and is not static like iris or fingerprint. A user can have different keys for different accounts by just changing the password (what to pronounce) or the vocalization of the same password (how to pronounce) to generate different cryptographic keys. In an event of key compromise a new cryptographic key can be easily generated. Note that voice has a potential disadvantage when used in fighting against coercion, namely that the attacker may blame the user for intentionally pronouncing the wrong password. We demonstrated our technique with voice; however, our scheme is not limited to using voice, other biometric can be used as well.

3.3 Why Fingerprint?

A potential threat to our biometric system is to use spoken password from the genuine user *(under stress)* and SC responses from another person *(normal emotional state)*. To ensure that SC is not unforgeable, one can make use of a device to collect fingerprint and skin conductance of the user at the same time so that the fingerprint of the user can be checked and mapped to his/her skin conductance signal. However, we did not demonstrate how to use this as a measure in our proposed model as this is not the contribution of this paper and is left for the future work.

4 Key Generation from Voice and Skin Conductance

In order to show how skin conductance can be used to fight against coercion attacks in cryptographic key generation, in this section, we present the details of a cryptographic key generation technique using voice and skin conductance. Note the criteria behind choosing skin conductance and voice in Section 3. Other biometrics in lieu of voice could be used as well. Our way of using voice is similar (with some differences) to an earlier proposal of generating cryptographic keys using voice [24]. Table 1 shows some notations used in the rest of this paper.

4.1 An Overview

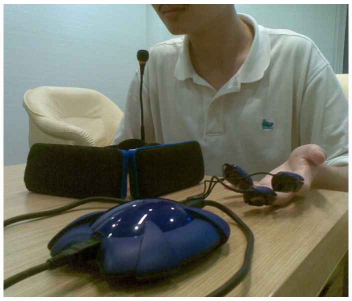


Figure 2: Input devices

Inputs to our model include the voice captured when the user utter the password into the microphone and the skin conductance measured. Figure 2 shows the input devices we used in our experimental setup. Output of our model is a cryptographic key generated.

In the first phase (Figure 3 (a)–(h)), features extracted from the spoken password are used to generate a sequence of frames $f_V(1), \ldots, f_V(n)$ (3 (c)), from which an optimal segmentation of s segments (component sounds) (3 (f)). The segmentation obtained are then mapped to the feature descriptor using a random α_V plane (3 (g)). Furthermore, features are also extracted from the SC sample and the corresponding feature descriptors are computed (3 (h)). These feature descriptors should be "sufficiently similar" for the same user and "sufficiently different" for different users. By the end of the first phase, we have feature descriptors for both voice and SC signal.

In the second phase (Figure 3 (i)–(l)), we perform lookup table generation and cryptographic key reconstruction. A total of N_V samples from voice and N_{SC} samples from SC are used to generate lookup tables T_V and T_{SC} . In cryptographic key reconstruction, feature keys are generated from the spoken password (m_V bits) and SC (m_{SC} bits). The two lookup tables generated and the features keys are then used to generate the cryptographic key.

In the next two subsections, we will present these two phases in more detail.

4.2 Phase I: Feature descriptors derivation

4.2.1 Feature descriptors from voice

In the last six decades, speech recognition and speaker recognition have advanced a lot [8]. A speaker recognition system usually has three modules: feature extraction, pattern matching and decision making, among which feature extraction is especially important to our research as it estimates a set of features from the speech signal that represent the speaker-specific information. These features should be consistent for each speaker and should not change over time. The way we extract these features and derive the feature descriptors is very similar to the previous approach [24], except that we use the

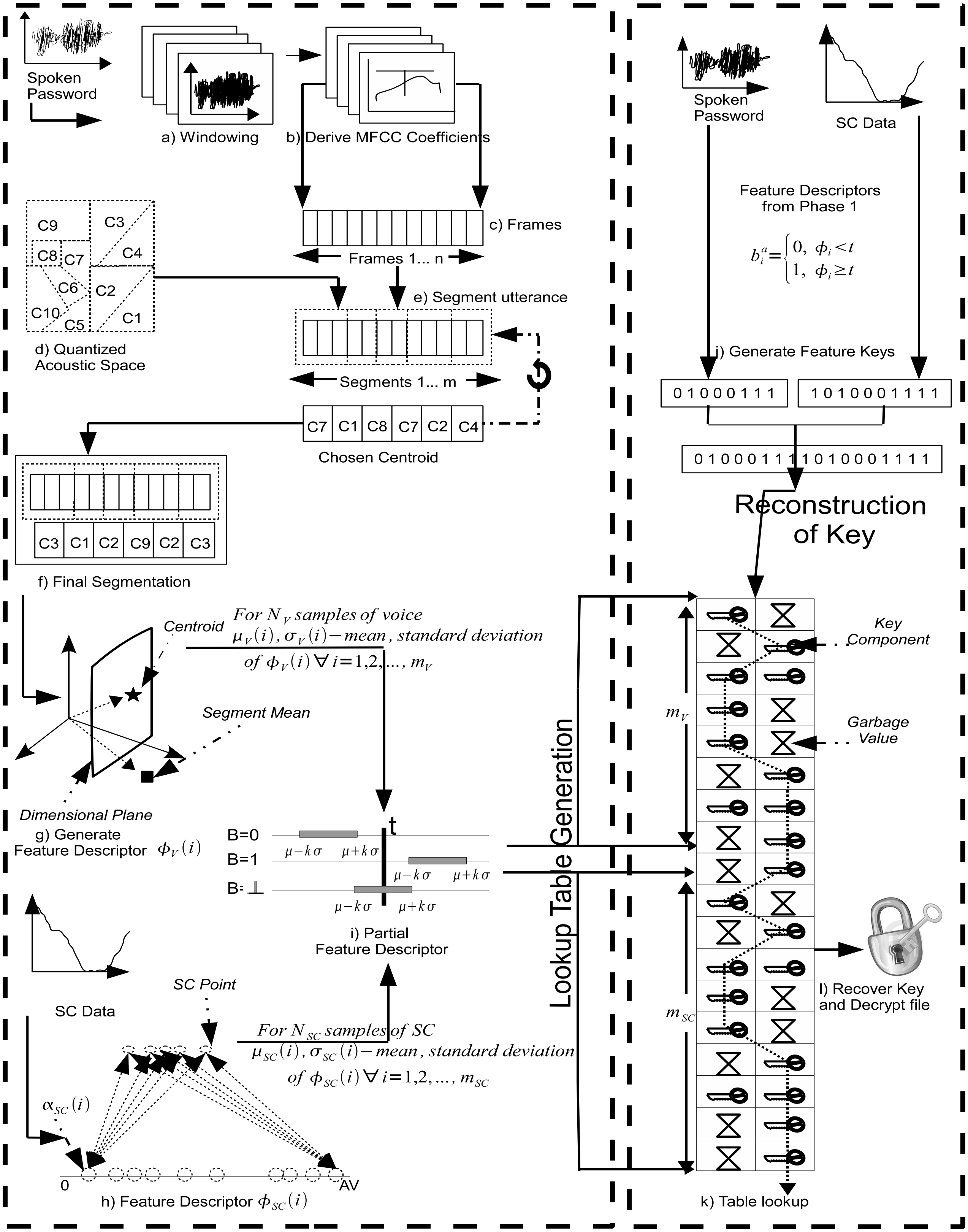

Spoken Password
a) Windowing
b) Derive MFCC Coefficients
c) Frames
Frames 1... n
C9
C3
C4
C8
C7
C6
C2
C10
C5
C1
d) Quantized Acoustic Space
e) Segment utterance
Segments 1... m
C7 C1 C8 C7 C2 C4
Chosen Centroid
C3 C1 C2 C9 C2 C3
f) Final Segmentation
Centroid
For N_V samples of voice
$\mu_V(i), \sigma_V(i) -$ mean, standard deviation
of $\phi_V(i) \forall i=1,2,\ldots,m_V$
Segment Mean
Dimensional Plane
g) Generate Feature Descriptor $\phi_V(i)$
t
B=0
B=1
B= ⊥
$\mu-k\sigma$
$\mu+k\sigma$
i) Partial Feature Descriptor
SC Data
SC Point
For N_{SC} samples of SC
$\mu_{SC}(i), \sigma_{SC}(i) -$ mean, standard deviation
of $\phi_{SC}(i) \forall i=1,2,\ldots,m_{SC}$
$\alpha_{SC}(i)$
0
AV
h) Feature Descriptor $\phi_{SC}(i)$
Spoken Password
SC Data
Feature Descriptors from Phase 1
$b_i^a = \begin{cases} 0, & \phi_i < t \\ 1, & \phi_i \geq t \end{cases}$
j) Generate Feature Keys
0 1 0 0 0 1 1 1
1 0 1 0 0 0 1 1 1 1
0 1 0 0 0 1 1 1 1 0 1 0 0 0 1 1 1 1
Reconstruction of Key
Lookup Table Generation
Key Component
Garbage Value
m_V
m_{SC}
l) Recover Key and Decrypt file
k) Table lookup

Figure 3: Design overview, refer to Section 4.2 for detailed description

General Notations		Notations related to Spoken Password		Notations related to Skin Conductance	
$\mathcal{K}$	cryptographic key	V	Voice	SC	Skin Conductance
C	a set of centroids	N_V	# samples in V during training	N_{SC}	# samples in SC during training
c	a centroid in C	f_V	frame vector	f_{SC}	vector containing sampled values of SC
m	$m = m_V + m_{SC}$	ϕ_V	feature descriptor	ϕ_{SC}	feature descriptor
		n	number of frames	ℓ	number of frames
		T_V	lookup table generated using V	T_{SC}	lookup table generated using SC
		m_V	total bits in a feature descriptor of V	m_{SC}	total bits in a feature descriptor of SC
		b_V	feature key using V	b_{SC}	feature key using SC
		s	number of segments		
		R	segment vector		

Table 1: Notations

Mel-frequency Cepstral Coefficients (MFCCs) instead of linear cepstrum [24]. MFCC has advantages over linear cepstrum that the frequency bands are equally spaced on the mel scale, which approximates the human auditory system's response more closely than the linearly-spaced frequency bands used in the linear cepstrum [13].

Associating centroids to the acoustic model We convert the raw speech signal into a sequence of acoustic feature vectors in terms of the Mel-frequency Cepstral Coefficients (MFCCs) [10]. In the next paragraph we provide a short description on the extraction of MFCC (see Figure 4).

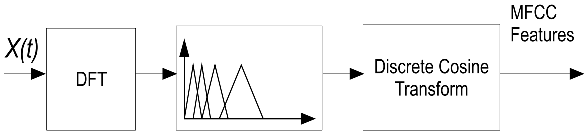


Figure 4: Block diagram of extracting MFCC

The voice signal is first divided into blocks of 20 to 30 msec (see Figure 3(a)), and Discrete Fourier Transform (DFT) is performed to obtain the frequency representation of each block. The neighboring frequencies in each block are grouped into bins of overlapping triangular bands of equal bandwidth. These bins are equally spaced on a Mel-scale instead of a normal scale as the lower frequencies are perceptually more important than the higher frequencies. The content of each band is now summed and the logarithmic of each sum is computed. To see this effect in time domain, Discrete Cosine Transform is applied to yield a "spectrum like" representation $\psi(t)$ that collectively make up an *MFC*, and $\psi(1), \ldots \psi(12)$ are called MFCC, where higher order coefficients are discarded. This vector is called a frame (f_V).

We run a sliding window of 30 msec over an utterance to obtain blocks 10 msec apart from one another, and extract the MFCC, $\langle \psi(1), \ldots \psi(12) \rangle$, for each block (see Figure 3(b)). n frames are obtained from utterance of the password (see Figure 3(c)). An acoustic model of vectors from a speaker-independent and text-independent database of voice signals is obtained, from which vector quantization is used to partition the acoustic model into clusters (see Figure 3(d)). A multivariate normal distribution for each cluster is generated, where each cluster is parameterized by the vector c of a component-wise means (called a centroid) and the covariance matrix Σ for the vectors in the cluster. The density function for this distribution is

$$P(c \mid x) = \frac{1}{(2\pi)^{\delta/2}\sqrt{det(\Sigma)}} e^{-(x-c)^T \Sigma^{-1} (x-c)/2}$$

where δ is the dimension of the vectors. We denote the set of centroids as C.

Segmentation of frames After getting the centroids from a speaker-independent database of voice signals, we try to obtain the transcription, i.e., the starts and ends, of the phonemes of an individual user's utterance.

To do this, we perform segmentation on the spoken password. Let $f_V(1), \ldots f_V(n)$ be the sequence of frames from the utterance, and $F(R_1), \ldots F(R_s)$ be the sequence of s segments (s is a constant and same for all users), where $F(R_i)$ is the i^{th} segment containing the sequence of frames $f_V(j), \ldots f_V(j')$ such that, $1 \leq j \leq j' \leq n$. Intuitively, each $F(R_i)$ corresponds to one "component sound" of the user's utterance.

We did this with an iterative approach (see algorithm 1). Ranges $R_1, \ldots, R_s$ are first initialized to be equally long. We then calculate the matching centroid c for a segment F(R), i.e., the one for which the likelihood of F(R) w.r.t. c is maximum. Dynamic programming is then used to determine a new segmentation for that frame sequence. This process is repeated until an optimal segmentation is obtained, which is mapped to the feature descriptor (see Figure 3(e,f)).

Feature descriptor Having derived a segmentation for a spoken password, we next define the feature descriptor (ϕ_V) of this segmentation that is typically the same when the same user speaks out the same utterance. To do this,

Algorithm 1 Spoken password segmentation

```
Segmentation $(f_V(1), \ldots, f_V(n), s)$
Score$'$ $\longleftarrow 0$
for $i = 1$ to $s$ do
    $R_i \longleftarrow \left( \left\lfloor \frac{(i-1) \times n}{s} \right\rfloor, \left\lfloor \frac{i \times n}{s} \right\rfloor \right)$
end for
repeat
    Score $\longleftarrow$ Score$'$
    for $i = 1$ to $s$ do
        while $\forall c \in C$ do
            $L(F(R_i)|c) \longleftarrow \prod_{j \in R_i} (f_V(\text{j})|c)$
        end while
        $c(R_i) \longleftarrow arg \max_{c \in C} \{L(F(R)|c)\}$
    end for
    let $\bigcup_{i=1}^{s} R_i' \longleftarrow [1, n]$
    Score$'$ $\longleftarrow \prod_{i=1}^{s} L(F(R_i'|c(R_i)))$
    $R_i \longleftarrow R_i'$
until Score$'$ - Score $< \Delta$
```

we use a fixed vector α_V, and define the i^{th} bit of the feature descriptor as (see Figure 3(g))

$$\phi_V(i) = \alpha_V.(\mu_V(R_i) - c(R_i)), \quad \forall \quad 1 \leq i \leq s$$

That is, we normalize $\mu_V(R_i)$ with $c(R_i)$ and let ϕ_V(i) be the linear combination of components in it as specified by α_V. This process results in a feature descriptor (ϕ_V), where N_V feature descriptors are then generated from N_V voice samples and used to generate a lookup table T_V (in Phase II).

4.2.2 Feature descriptor from skin conductance

When some external or internal stimuli occur that makes a person stressed, the skin becomes a better conductor of electricity. This conductance can be measured between two points on the body (e.g., two fingers) and the level of electrical conductance is called skin conductance. Since we want to detect changes in the emotional status of a person, we record skin conductance over a time period.

SC signal was measured with our device and sampled at a frequency of 30 samples per second. Let $f_{SC}(1), \ldots, f_{SC}(\ell)$ denote the sampled values obtained from the SC signal. We model the feature values into a feature descriptor (ϕ_{SC}) in a similar way as we did in the processing of voice. We choose a random vector α_{SC}= $[\alpha_{SC}(1), \alpha_{SC}(2), \ldots, \alpha_{SC}(m_{SC})]$ (m_{SC} is a constant), and use the Euclidean distance between all the points of the α_{SC} vector and f_{SC} to compute the distance measure M and henceforth the feature descriptor (ϕ_{SC}).

$$M(i,j) = \alpha_{SC}(i) \times f_{SC}(j) \quad \forall \quad 1 \leq i \leq m_{SC}, 1 \leq j \leq \ell$$

ϕ_{SC} is the mean of all the distance measures for each $\alpha_{SC}(i)$ values (see Figure 3(h)), i.e.,

$$\phi_{SC}(i) = \frac{1}{\ell} \sum_{j=1}^{\ell} M(i,j) \quad \forall \quad 1 \leq i \leq m_{SC}$$

Note that the upper bound of $\alpha_{SC}(i)$ needs to be carefully chosen to maintain a good entropy on the feature descriptor of different people. Also note that we do not store skin conductance information directly but rather the feature descriptor generated from the distance measure is stored (same as in the case of voice). N_{SC} feature descriptors are derived from N_{SC} SC samples and then are used to generate a lookup table T_{SC} (in Phase II).

4.3 Phase II: Lookup table and cryptographic key generation

We explain how we obtained the feature descriptors from voice and skin conductance in the previous subsection. Here, we will explain how we constructed lookup tables (training of the model) and obtained the cryptographic keys from the tables (usage of the model). The basic idea is that each entry of the lookup tables contains a share of the correct key or some garbage value, and the feature descriptor is used to determine the corresponding entry from the lookup table. In the end, the shares from the lookup tables are used to reconstruct the key.

4.3.1 Lookup table generation

Intuitively, if a feature descriptor is the same as the one recorded previously (i.e., in training), then the system should choose the correct key share from the lookup table, or the garbage otherwise. In order to tolerate some small deviation of a user's utterance and skin conductance, we calculate the mean (μ_{ϕ_V}(i), $\mu_{\phi_{SC}}$(i)) and standard deviation (σ_{ϕ_V}(i), $\sigma_{\phi_{SC}}$(i)) of each feature descriptor over N_V, N_{SC} training samples, and define the partial feature descriptors B_V, B_{SC} as

$$B_V(i) = \begin{cases} 0, \text{ if } \mu_{\phi_V}(\text{i}) + k\sigma_{\phi_V}(\text{i}) < t_V \\ 1, \text{ if } \mu_{\phi_V}(\text{i}) \text{ - } k\sigma_{\phi_V}(\text{i}) > t_V \\ \perp, \text{ otherwise} \end{cases} \quad \forall \ 1 \leq i \leq m_V$$

$$B_{SC}(i) = \begin{cases} 0, \text{ if } \mu_{\phi_{SC}}(\text{i}) + k\sigma_{\phi_{SC}}(\text{i}) < t_{SC} \\ 1, \text{ if } \mu_{\phi_{SC}}(\text{i}) \text{ - } k\sigma_{\phi_{SC}}(\text{i}) > t_{SC} \\ \perp, \text{ otherwise} \end{cases} \quad \forall \ 1 \leq i \leq m_{SC}$$

for some threshold t_V and t_{SC} respectively (see Figure 3(j)). This phase is the training phase in our model. Here k is a parameter to acquire a tradeoff between security and usability. With the increase in value of k, the user has better chance to generate the key successfully, but will hamper the security of the scheme. More precisely, the increase in the value of k will increase the false-positive rate and decrease the false-negative rate (as shown in our results in the evaluation Section 6).

The idea of defining the partial feature descriptor in this way is illustrated in Figure 5 (where the set $\{B, \mu, \sigma, t\}$ is replaced by $\{B_V, \mu_{\phi_V}, \sigma_{\phi_V}, t_V\}$ and $\{B_{SC}, \mu_{\phi_{SC}}, \sigma_{\phi_{SC}}, t_{SC}\}$ for voice and skin conductance respectively). If the i^{th} feature descriptor is consistently same i.e. $\mu(i) + k\sigma(i) < t$ (the first case in Figure 5), then there is a high probability that the value of the i^{th} feature descriptor will be less than t during key reconstruction. Therefore, we can let the cell $T(i, 0)$ of the lookup table contain a valid share of the key (and let $T(i, 1)$ contain random bits). If the i^{th} feature descriptor is consistently different, i.e. the value of the feature descriptor is unreliable (when compared to the threshold t as in the third case in Figure 5), we let both $T(i, 0)$ and $T(i, 1)$ contain valid shares (typically different). Unlike [24], lookup tables are not encrypted (for discussion on this, see section 4.4).

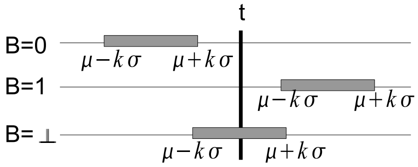


Figure 5: Definition of partial descriptor

Having valid shares in both $T(i, 0)$ and $T(i, 1)$ leads to different key shares used and consequently different keys being generated, which might not be desirable in systems that require a unique key. To solve this problem, a random cryptographic key $\mathcal{K}$ (unique for each user) is first generated, which is then encrypted with all possible valid keys (K_{H_i}) that can be derived from $<T_V \| T_{SC}>$. The key generation template therefore comprises of key $\mathcal{K}$ encrypted with $Z = |K_{H_i}|$ derived keys and the lookup tables $<T_V \| T_{SC}>$. Thus, the template = $<<T_V | T_{SC}>$, $<E_{K_{H_1}}(\mathcal{K} \| B), E_{K_{H_2}}(\mathcal{K} \| B), \ldots, E_{K_{H_Z}}(\mathcal{K} \| B)>>$, where $E_{\mathcal{K}_{H_i}}(msg)$ is a publicly known encryption algorithm and B is a unique string associated to each user which helps us to determine whether the decryption is correct or not in section 4.3.2.

4.3.2 Cryptographic key reconstruction

When a user tries to reconstruct the cryptographic key, he/she first presents his/her spoken password and the skin conductance. The model collect this information, extracts the features and generates the feature descriptors for both voice and the SC. Corresponding shares from the lookup tables are chosen based on the feature descriptors.

$$b_V(i) = \begin{cases} 0 & \text{if } \phi_V(i) < t_V \\ 1 & \text{otherwise} \end{cases} \quad \forall \quad 1 \le i \le m_V$$

$$b_{SC}(i) = \begin{cases} 0 & \text{if } \phi_{SC}(i) < t_{SC} \\ 1 & \text{otherwise} \end{cases} \quad \forall \quad 1 \le i \le m_{SC}$$

For example, if the feature descriptor $\phi_{SC}(i)$ is less than the threshold t_{SC}, then $b_{SC}(i) = 0$ and $T_{SC}(i, 0)$ is chosen from T_{SC} as a key share; otherwise $b_{SC}(i) = 1$ and $T_{SC}(i, 1)$ is chosen (see Figure 3(i)). b_V and b_{SC} are the feature keys and are obtained from voice and SC respectively.

A key K' is derived by concatenating the key shares (see Figure 3(k)). This derived key is then used to decrypt the $|K_{H_i}|$ encrypted keys stored in the template. If the decryption succeeds (by matching the released B and the stored B), then the key $\mathcal{K}$ is released.

$$\mathcal{K}_D = \begin{cases} D_{K'}(E_{K_{H_i}}(\mathcal{K} | B)), & if \quad K' = K_{H_i} \\ Random, & if \quad K' \neq K_{H_i} \end{cases}$$

where, $D_{K'}(msg)$ is a publicly known decryption algorithm.

4.4 Discussions

While we try to use the consistency of voice and skin conductance to generate the correct key only when it is the genuine user in the normal emotional state, the inconsistency of voice and skin conductance poses challenges, too. Voice produced and skin conductance measured of the genuine user in a non-stressed emotional status might change due to tiredness, illness, noise, and etc.

We used an error correction technique, in particular, hamming distance, to improve the usability of the scheme. ${}^mC_d$ different keys are derived from any freshly generated key K' obtained from the feature descriptors and T (similar to the one derived in section 4.3.2), which are d distance away from the derived key K'. All of these ${}^mC_d$ keys are then used to decrypt the encrypted keys before giving any negative answer to the user. If the decryption succeeds then the key $\mathcal{K}$ is released. For example, if $d = 2$ and length of the key is m, then ${}^mC_2$ different keys are derived. Thus, $|K_{H_i}| \times^m C_2$ decryptions are performed in attempting to recover $\mathcal{K}$.

Another issue concerns the privacy of the biometric data used. Ballard et al. propose using randomized biometric templates protected with low-entropy passwords to provide strong biometric privacy [4]. One can use this in conjunction with our model to provide both coercion resistance and biometric privacy. However, it is unclear whether the use of low-entropy passwords may have a negative impact on coercion resistance since, intuitively, an attacker may blame the user for providing the wrong low-entropy password in a coercion (similar problem discussed in section 3.2). We leave this as future work to develop a solution that satisfies both requirements.

5 Experimental Setup

We presented our design in generating a cryptographic key using voice and skin conductance in Section 4. It is important to test it out with real human beings to evaluate its performance. However, this is difficult as we need to find a way to make the participants feel stressed or nervous. It is clear that we cannot actually coerce them to do something by, e.g., putting a gun over their heads. Nevertheless, we performed case studies to induce stress on the participants and measure their voice and skin conductance. (IRB approval was obtained from our university before the user study.) We present the experimental setup in this section and the evaluation results and discussion in the next section.

5.1 Demographics

Since we were going to induce stress on the participants, we decided to concentrate on the younger generation (undergraduate and graduate students in the age from 18 to 30). We had altogether 43 participants, from which 4 participants detached the sensors from their fingers when they were nervous during the experiment. Therefore, we successfully performed our experiments on 39 participants, out of which 22 were male and 17 were female.

5.2 Experimental settings

Participants were asked to sit in a small office where the overhead fluorescent lights were turned off and a dim red incandescent lamp was turned on to reduce the possible electrical interference with the monitoring equipments. The room was air conditioned to approximately 72°F and humidity level was generally dry. This is done in accordance to the variation of skin conductance in different environmental conditions [36].

Skin conductance sensors[1] were attached to the three middle fingers of the participant to record SC (shown in Figure 2). The participant was also asked to keep her left hand (with sensors attached) as still as possible to avoid interference from the sensors. Fake heart rate tags were tied to the wrist, which gave an illusion of monitoring the heart rate.

Initially, there was an incomplete disclosure regarding the purpose and the steps of the study in order to ensure that the participant's responses will not be affected by her knowledge of the research.

5.3 Procedure

We ran two experiments (`e1` and `e2`). Each experiment consisted of two parts, where the first parts (`e1n` and `e2n`) were conducted when the participants were in a normal (calm) condition, and the second parts (`e1s` and `e2s`) were conducted when the participants were stressed.

We ran experiment `e1n` by

- showing nice (geographical) pictures one after another and short phrases (the spoken password embedded) which are related to the pictures, and asking the participant to read them out;
- showing fake visual heartbeats at a normal rate at the bottom of the screen and correspondingly playing heartbeats sound.

In order to capture the emotional responses in the stress scenario in `e1s`,

- a frightening horror movie was played, replacing the nice pictures;
- the rate of the heartbeats were gradually increased to induce more stress on the participant;
- the participant was asked to read out some short phrases at the end of each horror scene (rather than along with the video) to avoid distraction.

Similar studies have been performed previously to measure the stress level in users [26, 19].

In `e2`, we went a bit further to induce more stress on the participant. Figure 6 shows the change in skin conductance in response to different events in `e2`. During `e2`, the participant was asked to type a few sentences (e.g., "Work is much more fun than fun") shown to her in a fixed period of time. She was also warned (prior to the experiment) not to press the "ALT" key on the keyboard, as it would cause the computer program to crash and all data would be lost (event A). We then left the participant alone in the room to continue typing (event B). We configured the computer to restart after 3 minutes irrespective of whether the participant actually touched the "ALT" key or not. The computer would then boot from a USB drive into MS-DOS and display some error messages (event C). This completes the first part of `e2`, i.e., `e2n`.

Stress started to develop at this point in time as the participant believed that she had pressed the "ALT" key which caused data loss on the computer (event D). We purposely left the participant alone so that stress could develop further and she could not get immediate help to resolve the "problem". After that, the researcher entered the room and examined the keyboard and the computer (event E) and then accused the participant of her negligent act of pressing the "ALT" key (event F). This turned

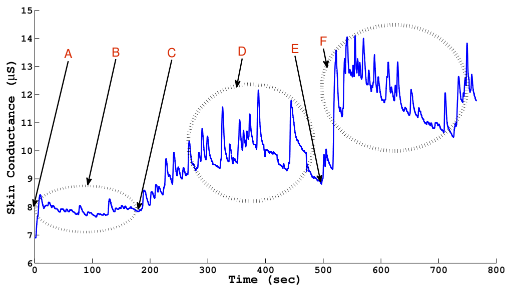


Figure 6: Change of skin conductance in e2

out to be successful in making the participant stressed as we observed that many participants were nervous at this point in time. Some kept saying "sorry"; some tried very hard to fix the "problem", and some started calling for help. There were also voluntary confession statements from the participants, e.g., "I hit the ALT key by mistake in place of typing the 'X' key", "It was a mistake from my side.".

5.4 Discussion

In this section, we discuss the difference of the emotional state of a user in real life and in our user study, and limitations of our experiment.

1. Training of the system
   - Real life: the user is in a (controlled) environment specified by our system, in which the stress level is low. This allows us to generate the lookup table for that particular user with the normal skin conductance level.
   - User study: the user is in exactly the (controlled) environment specified by our system, i.e., when watching a relaxation movie.
2. Trying to generate the cryptographic key; no coercion
   - Real life: a user could be in various emotional states, including being happy, sad, angry, etc.
   - User study: same as in training when the user is watching a relaxation movie. In this work, we only try to analyze how our system performs when users are calm and relaxed. It remains future work to analyze how it works when the user is in other emotional states. We do expect the false-negative rate to rise when the user is in other emotional states.
3. Trying to generate the cryptographic key; in coercion
   - Real life: a user can be forced/coerced in many different ways, e.g., a gun to the head, or a knife under the throat, etc.
   - User study: watching a horror movie and being forced to plead guilty (having damaged a notebook computer). We tried our best to approximate the real-life scenarios, but there is a limit we could go when doing this to real human beings (e.g., IRB restriction). However, we believe that what we did is a clever way of studying human behavior when being coerced.

Discussions above highlight some limitations of our scheme, e.g., we have not tested how it reacts to other emotional status (happy, sad, angry, etc.) and how skin conductance may change naturally (due to oily fingers, etc.). There are two other important limitations in the present study. First, our study does not test the repeatability of using our scheme, i.e., we did not ask the participants to come back and try again. The second limitation comes with the over-controlled environment, e.g., quiet office (because of the use of voice), controlled temperature and humidity [9](because of the use of skin conductance), and etc. It remains further work to test our scheme in different settings.

6 Evaluation and Discussion

In this section, we analyze the data collected in our user study. We first describe how we partition the data into different groups (e.g., for training and test purposes), see Section 6.1. We then present a series of analysis on the false-positive and false-negative rates (Section 6.2). Finally we show the change in the password space where an attacker has perfect knowledge of our design and the content stored. We show that this change in the password space in this worst case is small (Section 6.3).

6.1 Training and Testing Datasets

We have collected voice and skin conductance signals for 39 participants. For each participant, we have collected many samples of the signals when the participant is either calm or stressed. Table 2 shows the number of samples we collected in each experiment for each participant. Voice signals are typically 2 to 3 seconds long, while skin conductance signals are about 10 seconds long to avoid fluctuations.

Figure 7 shows how we obtain dataset to

- split original sample sets $\{\nu^{\mathsf{full}}_{\mathsf{e1}n}, \omega^{\mathsf{full}}_{\mathsf{e1}n}, \omega^{\mathsf{full}}_{\mathsf{e2}n}\}$ into two equal halves $\{\nu^{\mathsf{train}}_{\mathsf{e1}n}, \omega^{\mathsf{train}}_{\mathsf{e1}n}, \omega^{\mathsf{train}}_{\mathsf{e2}n}\}$ and $\{\nu^{\mathsf{test}}_{\mathsf{e1}n}, \omega^{\mathsf{test}}_{\mathsf{e1}n}, \omega^{\mathsf{test}}_{\mathsf{e2}n}\}$ to obtain datasets for training and testing (see the half circles);

Feature		e1n	e1s	e2n	e2s
Voice	# of samples	26	5	0	0
	Notation	$\nu^{\mathsf{full}}_{\mathsf{e1}n}$	$\nu^{\mathsf{full}}_{\mathsf{e1}s}$	-	-
SC	# of samples	26	60	18	60-80
	Notation	$\omega^{\mathsf{full}}_{\mathsf{e1}n}$	$\omega^{\mathsf{full}}_{\mathsf{e1}s}$	$\omega^{\mathsf{full}}_{\mathsf{e2}n}$	$\omega^{\mathsf{full}}_{\mathsf{e2}s}$

Table 2: Number of samples collected for each participant

- combine different voice samples and skin conductance samples to create new datasets to test our system (see circles in the middle column). $\{\nu^{\mathsf{train}}_{\mathsf{e1}n}$ & $\omega^{\mathsf{train}}_{\mathsf{e1}n}\}$, $\{\nu^{\mathsf{test}}_{\mathsf{e1}n}$ & $\omega^{\mathsf{test}}_{\mathsf{e1}n}\}$, $\{\nu^{\mathsf{train}}_{\mathsf{e1}n}$ & $\omega^{\mathsf{train}}_{\mathsf{e2}n}\}$, $\{\nu^{\mathsf{test}}_{\mathsf{e1}n}$ & $\omega^{\mathsf{test}}_{\mathsf{e2}n}\}$ are combined to create $\{\xi^{\mathsf{train}}_{\mathsf{e1}n}\}$, $\{\xi^{\mathsf{test}}_{\mathsf{e1}n}\}$, $\{\xi^{\mathsf{train}}_{\mathsf{e2}n}\}$, $\{\xi^{\mathsf{test}}_{\mathsf{e2}n}\}$ respectively.
- to obtain the stress dataset $\{\nu^{\mathsf{full}}_{\mathsf{e1}s}$ & $\omega^{\mathsf{full}}_{\mathsf{e1}s}\}$, $\{\nu^{\mathsf{full}}_{\mathsf{e1}s}$ & $\omega^{\mathsf{full}}_{\mathsf{e2}s}\}$ are combined to create $\{\xi^{\mathsf{full}}_{\mathsf{e1}s}\}$, $\{\xi^{\mathsf{full}}_{\mathsf{e2}s}\}$ respectively.

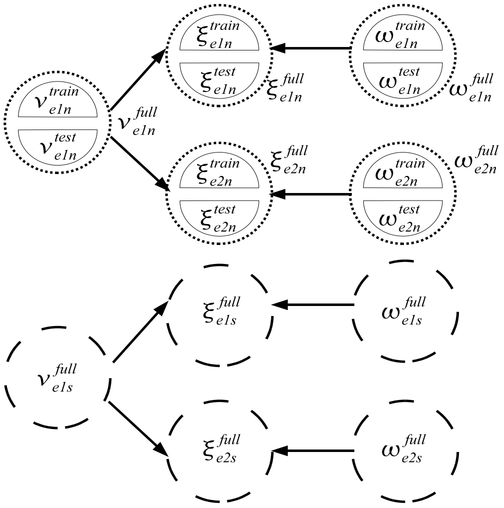


Figure 7: Splitting and combining datasets

Note that the voice and skin conductance samples that are combined together might not have been captured at exactly the same time. We allow a time gap because an attacker might record the voice of the victim to be used in conjunction with the skin conductance of the victim at a slightly different time. Both samples were captured in the same part of the experiment, though, i.e., both from e1s or both from e2s.

6.2 Accuracy of our model

The false-negative rate of our system is defined as the percentage of failed login attempts by a legitimate user with her cryptographic key generated, averaged over all users in a population A. Similarly, the false-positive rate is defined as the percentage of failed detection of attempts by illegitimate users or legitimate users in a stressful situation, averaged over all users in a population A.

Voice samples only We first evaluate the voice samples we collected in our experiments. The purpose is to check out the false-positive and false-negative rates, in an event if only voice samples are used to generate cryptographic keys. The system is trained with $\nu^{\mathsf{train}}_{\mathsf{e1n}}$ of user a_i, and is tested against $\nu^{\mathsf{full}}_{\mathsf{e1n}}$ of user a_j where $i \neq j$, $\forall\ j \in A$ to calculate the false-positive rates; and against $\nu^{\mathsf{test}}_{\mathsf{e1n}}$ of user a_i to calculate the false-negative rates. Results are averaged on all users in A. We try different random α_V vectors and choose the one that yields the smallest sum of the false-positive and false-negative rates. We try different settings of the hamming distance parameter d, and find that 2 gives a reasonable tradeoff between false-positive and false-negative rates. The false-positive and false-negative rates for different values of k are plotted in Figure 8.

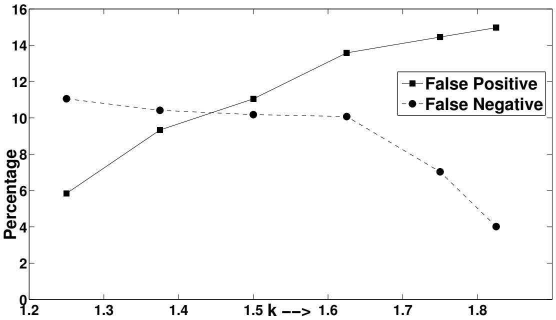
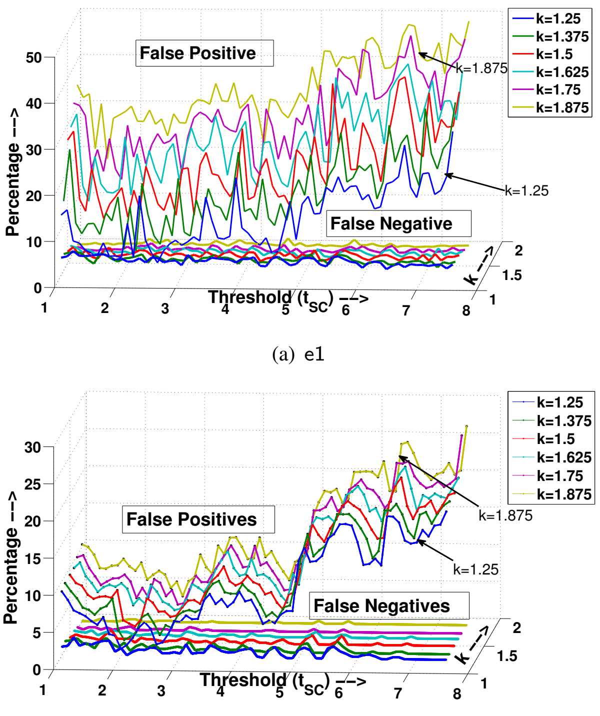


Figure 8: False-positive and false-negative rates for spoken passwords

Figure 8 shows that we manage to get a comparable accuracy with the previous work [24] in terms of the false-negative rate. False-positive rate was not reported in [24].

Skin conductance only Next, we evaluate the skin conductance samples to see how well they reflect the change in the participants' emotional status. We show the results in Figure 9(a) and Figure 9(b) for experiment e1 and e2, respectively. The different color lines denotes different 'k' values in Figure 9 and Figure 10. The system is trained with $\omega^{\mathsf{train}}_{\mathsf{e1n}}$ (and $\omega^{\mathsf{train}}_{\mathsf{e2n}}$, respectively) of user a_i, and is tested against the stressed full data set, $\omega^{\mathsf{full}}_{\mathsf{e1s}}$ (and $\omega^{\mathsf{full}}_{\mathsf{e2s}}$, respectively) of the same user a_i to calculate the false-positive rates; or against the normal test data set, $\omega^{\mathsf{test}}_{\mathsf{e1n}}$ (and $\omega^{\mathsf{test}}_{\mathsf{e2n}}$, respectively) of the same user a_i to calculate the false-negative rates. Results are averaged over all users in A.



(a) e1

(b) e2

Figure 9: False-positive and false-negative rates for skin conductance

Note that the false-positive and false-negative rates are higher for e1 in Figure 9(a). We believe, this is because of the reason that the intensity of some of the horror videos was not very high, which did not result in a noticeable change in the skin conductance for many users.

We can observe the tradeoff of various settings of k and the threshold from these figures. In general, this shows that whenever a user is under stress, her skin conductance can be used to differentiate between the two emotional state with good accuracy. For example in e2, when $k = 1.25$ and $t_{SC} = 2.1$, we obtained a false-positive rate of 3.2% and a false-negative rate of 2.2% (see Figure 9(b)). If we increase the value of k from 1.25 to 1.875 in both Figures 9(a) and 9(b), we could see a decrease in the false-negative rates (increasing usability) and increase in the false-positive rates (compromising with the security). We used the hamming distance parameter $d = 2$ in our setting.

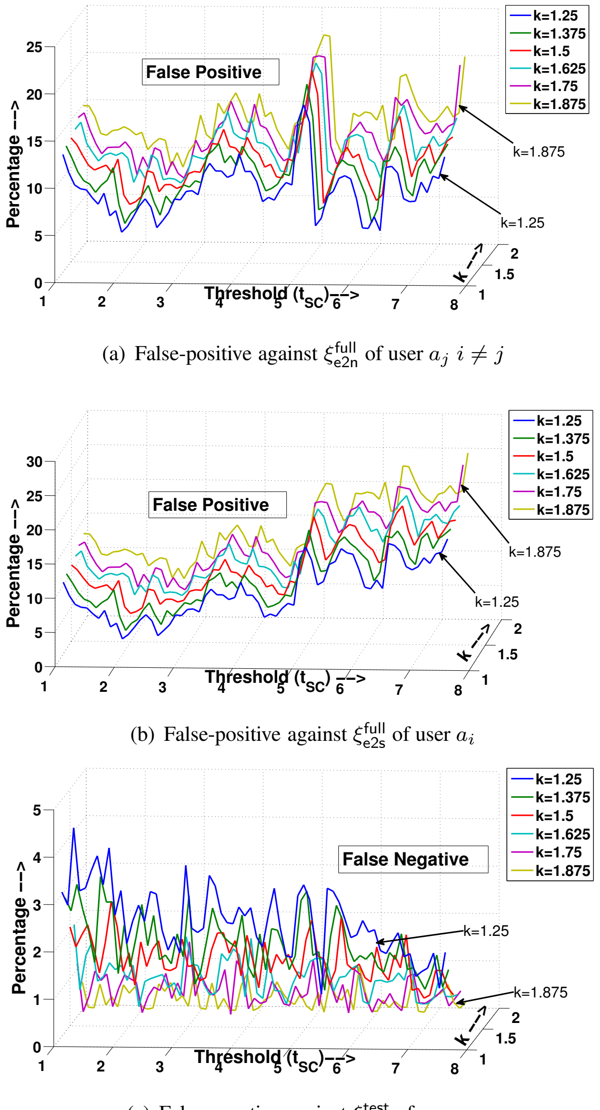
Voice combined with skin conductance Voice and skin conductance samples are combined as shown in Figure 7 to obtain the samples needed in this evaluation. We first train the system with $\xi^{\mathsf{train}}_{\mathsf{e2n}}$, and then evaluate the system against three different datasets to evaluate the false-positive and false-negative rates.

a $\xi^{\mathsf{full}}_{\mathsf{e2n}}$ of user a_j where $i \neq j, \forall\, j \in A$: when a different person tries to generate the key (Figure 10(a));

b $\xi^{\mathsf{full}}_{\mathsf{e2s}}$ of user a_i: when the same user tries to generate the key when she is being coerced (Figure 10(b));

c $\xi^{\mathsf{test}}_{\mathsf{e2n}}$ of user a_i: when the same user tries to generate the key when she is not being coerced (Figure 10(c)).

We evaluate the false-positive rates in the first two cases and the false-negative rates in the third case. Results are averaged over all users in A. We use a hamming distance parameter $d = 4$, and show the results in Figure 10.



(a) False-positive against $\xi^{\mathsf{full}}_{\mathsf{e2n}}$ of user a_j $i \neq j$

(b) False-positive against $\xi^{\mathsf{full}}_{\mathsf{e2s}}$ of user a_i

(c) False-negative against $\xi^{\mathsf{test}}_{\mathsf{e2n}}$ of user a_i

Figure 10: False-positive and false-negative rates for voice combined with skin conductance

These results show that generating cryptographic keys from voice and skin conductance is effective in fighting coercion attacks, as we observe false-positive rates between 6% to 15% for $1 \leq t_{SC} \leq 4$, which can also rise up to 22% for $t_{SC} \geq 5$. False-negative rates are between 0% and 4.5% for all values of t_{SC}. Further efforts are needed to reduce the false-positive and false-negative rates. Same as in the previous subsection, if we increase the value of k from 1.25 to 1.875, we could see a decrease in the false-negative rates and increase in the false-positive rates.

6.3 Change in password space

In this subsection, we discuss more advanced attacks on our system (if implemented) beside forcing the victim to obtain her spoken password and skin conductance. If such system is implemented, then we need to approximate the entropy in the worst case of these advanced attacks, in which the attacker makes use of the group information about the skin conductance and information stored in the key generation module.

The group information about skin conductance refers to the patterns observed in the change in the users' feature key generated from the skin conductance (b_{SC}) when they are coerced. An attacker could use this information to selectively modify the victims skin conductance feature key in order to improve the probability of generating the correct key. To know how we obtained the feature key (b_{SC}) for SC, see section 4.

Although we do not store any biometric information of the user directly on the device (see discussions in Section 4), we still need to store the lookup tables (T_V and T_{SC}) which are derived from the user specific data (e.g., feature descriptors). Although this table can be encrypted with a user password as discussed in previous work [24], however we try not to rely the security of our model on the secrecy of this table because we are dealing with coercion attacks. In the rest of this subsection, we assume that an attacker has perfect knowledge in both the group information about skin conductance and the lookup tables. We want to approximate the guessing entropy, i.e., the reduction in the password space for this more powerful attacker.

More precisely, we assume in the worst case that an attacker has access to

- the lookup tables T_V and T_{SC};
- the recorded spoken password of the user and the corresponding feature key $\{b_V(i)\}$;
- the recorded skin conductance when the user is stressed and the corresponding feature key $\{b^S_{SC}(i)\}$;
- the database D which contains the mapping of the SC feature keys when users are normal ($\{b^N_{SC}(i)\}$) to the scenario when they are stressed ($\{b^S_{SC}(i)\}$) for all users in a population A.

A sample database D for such mapping of SC is shown in Table 3 for $|A|$ users. Each row in the table is a record of the feature key of a user when she is normal and stressed, and the last column shows the index of the feature keys that had changed from b^N_{SC} to b^S_{SC}.

#	b^N_{SC}	b^S_{SC}	Flipped bits' pos.
1	011011011011	001101110011	2,4,5,7,9
2	010010010111	010100110110	4,5,7,12
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\|A\|$	010101001100	111111100110	1,3,5,7,9,11

Table 3: A sample Database D

The attacker's strategy would be to analyze D to learn patterns in which people's feature keys $\{b^N_{SC}\}$ changes to $\{b^S_{SC}\}$, e.g., whenever the i-th index of the feature key changes, the j-th one will change too.

These patterns can be easily learned by applying a well studied technique called association rule mining [2]. The attacker can then use these patterns to reduce the password space. Here, we use a simple example to demonstrate the idea.

We first represent the password space by a sequence of 0's (the corresponding index in $\{b^S_{SC}\}$ will definitely not change when a user's emotional status changes), 1's (the corresponding index in $\{b^S_{SC}\}$ will definitely change), and $*$'s (don't know), e.g., $[1, *, *]$ represents a password space in which only the first index of $\{b^S_{SC}\}$ will change, and therefore the password space is $2^2 = 4$. When the attacker makes use of a pattern learned, e.g., "the change of the first index of $\{b^S_{SC}\}$ implies the change of the second one", he can convert the password space from $[1, *, *]$ to $[1, 1, *]$, since the second index of the $\{b^S_{SC}\}$ will definitely change, too. With this, the password space reduces to $2^1 = 2$.

We present the detailed algorithm with an example in estimating this reduction in the password space in the Appendix A.

We constructed the database D with the skin conductance samples collected in our user study, mine all association rules, and then use the above algorithm to find out the change in the password space. Figure 11 shows the results for different settings of the threshold and minimum confidence in the association rule mining.
k is set to 1.25 in this experiment, and the minimum support is set to 30%. Note that the original password space is $2^{m_{SC}} = 2^{50}$. Although in the worst case the effective number of bits to represent the password space reduces

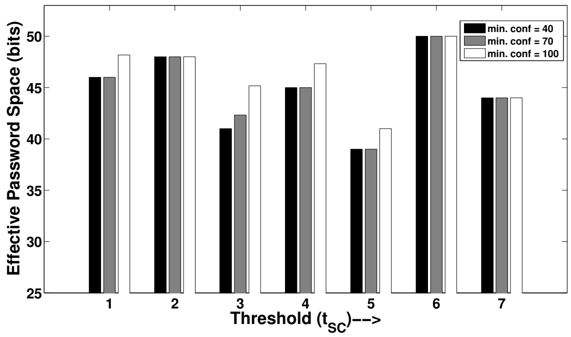


Figure 11: Password Space reduction

by roughly 20%, many settings of the threshold value result in only 10% reduction.

Another way to attack our system is to make the user take a sedative to relieve his/her anxiety before capturing SC. The attacker can then use this skin conductance to generate the key. We are trying to collaborate with medical practitioners and researchers to see the correlation between the two skin conductances, one under normal condition without taking any sedative and the other under coercion and having taken the sedative. For now this remains as a future work.

7 Conclusion and Future Work

In this paper we present a novel approach for fighting against coercion attacks in generating cryptographic keys using *skin conductance* (SC) of a person. In coercion attack, the attacker forces a user to grant him access to the system. SC was used to determine the person's overall arousal state i.e. (emotional status). The change in the emotional status of a person results in different keys. We discussed the reasons of adopting SC as an emotional response parameter and why it was preferred over other physiological signals like Electrocardiography, Electromyography, Heart Rate, respiration, skin temperature etc. In this paper, we have chosen skin conductance along with voice in generating cryptographic keys; however, one can choose any other biometric for e.g. iris, fingerprint, face etc. in lieu of voice. Cryptographic key is generated using lookup table method as discussed in [24].

In our knowledge the presented work is the first in fighting coercion attacks in generating cryptographic keys. We conducted two experiments in our user study and have shown some interesting results. The proposed model was tested with 39 user's voice and skin conductance data to compute the false-positive and false-negative rate. Furthermore our results showed that the cryptographic key generated in two different scenarios are different for the same person. This bolsters our heuristic to use skin conductance for fighting against coercion attacks. As both skin conductance and voice are not static biometrics, in some cases we obtained high false-negatives. We evaluated the security of the proposed model in terms of entropy and several threat models and discussed how difficult it is for an attacker, in an event when she has full information about the key generation module; the skin conductance of the victim in the stressful scenario; and the group information about the skin conductance.

Note that guessing entropy and guessing distance [5] might provide deeper insight in the security of our model. We leave it as our future work. In terms of feasibility, in future we will also like to see in some possibilities of building the system (may be a mobile device) with all three: voice, skin conductance and fingerprint extraction mechanism to authenticate to the system. Furthermore, we would like to look into other emotional responses like happy, joy, anger, sad etc., to make the claim of using SC in fighting coercion attacks stronger. This paper does not study the repeatability of the key using the proposed scheme and is left as a future work.

Acknowledgments We are deeply grateful to Lucas Ballard for comments and suggestions in key construction, as well as David Lo for helpful discussions on the analysis of the password space.

References

[1] ABHYANKAR, A., AND SCHUCKERS, S. Integrating a wavelet based perspiration liveness check with fingerprint recognition. *Pattern Recognition 42*, 3 (2009), 452–464.

[2] AGRAWAL, R., IMIELIŃSKI, T., AND SWAMI, A. Mining association rules between sets of items in large databases. In *SIGMOD '93: Proceedings of the 1993 ACM SIGMOD international conference on Management of data* (New York, NY, USA, 1993), ACM, pp. 207–216.

[3] ANTTONEN, J., AND SURAKKA, V. Emotions and heart rate while sitting on a chair. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2005), ACM, pp. 491–499.

[4] BALLARD, L., KAMARA, S., MONROSE, F., AND REITER, M. K. Towards practical biometric key generation with randomized biometric templates. In *CCS '08: Proceedings of the 15th ACM conference on Computer and communications security* (New York, NY, USA, 2008), ACM, pp. 235–244.

[5] BALLARD, L., KAMARA, S., AND REITER, M. K. The practical subtleties of biometric key generation. In *SS'08: Proceedings of the 17th conference on Security symposium* (Berkeley, CA, USA, 2008), USENIX Association, pp. 61–74.

[6] BIEL, L., PETTERSSON, O., PHILIPSON, L., AND WIDE, P. Ecg analysis: a new approach in human identification. *Instrumentation and Measurement, IEEE Transactions on 50*, 3 (Jun 2001), 808–812.

[7] CACIOPPO, J. T., AND TASSINARY, L. G. Inferring psychological significance from physiological signals. *American Psychologist 45*, 1 (Jan 1990), 16–28.

[8] Campbell, J. P. Speaker recognition: a tutorial. *Proceedings of the IEEE 85*, 9 (1997), 1437–1462.

[9] Conklin, J. E. Three Factors Affecting the General Level of Electrical Skin-Resistance. *The American Journal of Psychology 64*, 1 (Jan 1951), 78–86.

[10] Davis, S., and Mermelstein, P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing], IEEE Transactions on 28*, 4 (1980), 357–366.

[11] Ekman, P. *Basic Emotions*, vol. 476 of *Handbook of Cognition and Emotion*. T. Dalgleish and M. Power, John Wiley & Sons Ltd. Sussex UK, 1999.

[12] EncoSkin. ENCOLL's Collagen Technology, 2010. `http://www.encoll.com/SkinCare_and_Dietary_Products.htm`.

[13] Fang, Z., Guoliang, Z., and Zhanjiang, S. Comparison of different implementations of mfcc. *J. Comput. Sci. Technol. 16*, 6 (2001), 582–589.

[14] Fowles, D. C. The Three Arousal Model: Implications of Gray's Two-Factor Learning Theory for Heart Rate, Electrodermal Activity, and Psychopathy. *Psychophysiology 17*, 2 (1980), 87–104.

[15] García Perera, L., Nolazco Flores, J., and Mex Perera, C. Cryptographic-Speech-Key Generation Architecture Improvements. In *IbPRIA05* (2005), p. II:579.

[16] Helander, M. Applicability of drivers' electrodermal response to the design of the traffic environment. *Journal of Applied Psychology 63*, 4 (1978), 481–488.

[17] INVESTIGATION, F. B. O. Bank crime statistics (bcs) federal insured financial institutions july 1, 2009 september 30, 2009. `http://www.fbi.gov/publications/bcs/bcs2009/bank_crime_2009q3.htm`.

[18] Jain, A. K., Ross, A., and Pankanti, S. Biometrics: a tool for information security. *IEEE Transactions on Information Forensics and Security 1*, 2 (2006), 125–143.

[19] Johnson, K. J., and Fredrickson, B. L. We all look the same to me: Positive emotions eliminate the own-race bias in face recognition. *Psychological Science 16* (2005), 875–881.

[20] Kim, J. Bimodal Emotion Recognition using Speech and Physiological Changes. In *In M. Grimm, K. Kroschel (Ed.), Robust Speech Recognition and Understanding*. I-Tech Education and Publishing, Vienna, Austria, 2007, pp. 265–280.

[21] Kim1, K. H., Bang, S. W., and Kim, S. R. Emotion recognition system using short-term monitoring of physiological signals. *Medical and Biological Engineering and Computing 42*, 3 (May 2004), 419–427.

[22] Lapsley, P. D., Lee, J. A., Pare, Jr., D. F., and Hoffman, N. Anti-fraud biometric scanner that accurately detects blood flow. US Patent # 5737439, 1998.

[23] Lee, C. K., Yoo, S. K., Park, Y., Kim, N., Jeong, K., and Lee, B. Using Neural Network to Recognize Human Emotions from Heart Rate Variability and Skin Resistance. In *27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005* (2005), pp. 5523–5525.

[24] Monrose, F., Reiter, M. K., Li, Q., and Wetzel, S. Cryptographic Key Generation from Voice(Extended Abstract). In *SP '01: Proceedings of the 2001 IEEE Symposium on Security and Privacy* (Washington, DC, USA, 2001), IEEE Computer Society, p. 202.

[25] Monrose, F., Reiter, M. K., and Wetzel, S. Password hardening based on keystroke dynamics. *International Journal of Information Security 1*, 2 (2002), 69–83.

[26] Nachson, I., and Feldman, B. Psychological Stress Evaluator - Validity Study. *Crime and Social Deviance 7*, 2 (1979), 65–81.

[27] Nandakumar, K., Jain, A. K., and Pankanti, S. Fingerprint-Based Fuzzy Vault: Implementation and Performance. *IEEE Transactions on Information Forensics and Security 2*, 4 (2007), 744–757.

[28] Nandakumar, K., Nagar, A., and Jain, A. Hardening Fingerprint Fuzzy Vault Using Password. In *ICB07* (2007), Springer Berlin / Heidelberg, pp. 927–937.

[29] Nasoz, F., Alvarez, K., Lisetti, L., and Finkelstein, N. Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cognition, Technology and Work 6*, 1 (2004), 4–14.

[30] Phua, K., Chen, J., Dat, T. H., and Shue, L. Heart sound as a biometric. *Pattern Recogn. 41*, 3 (2008), 906–919.

[31] Picard, R. W., Vyzas, E., and Healey, J. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Transaction Pattern Analysis Matching Intelligence 23*, 10 (2001), 1175–1191.

[32] Prabhakar, S., Pankanti, S., and Jain, A. K. Biometric recognition: Security and privacy concerns. *IEEE Security and Privacy 1*, 2 (2003), 33–42.

[33] Ratha, N. K., Connell, J. H., and Bolle, R. M. An Analysis of Minutiae Matching Strength. In *AVBPA '01: Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication* (London, UK, 2001), Springer-Verlag, pp. 223–228.

[34] Sakai, K., and Quick, T. W. Moisturizing skin preparation, July 1988.

[35] santos A, M. F., aguilar A, J. F., and garcia A, J. O. Cryptographic key generation using handwritten signature. *Biometric Technologies for Human Identification III, Processing of SPIE* (2006).

[36] Schmidt, S., and Walach, H. Electrodermal Activity (EDA) - State of the Art Measurement and Techniques for Parapsychological Purposes. *Journal of Parapsychology 64* (June 2000), 139 – 163.

[37] Selye, H. *The Stress of Life*. McGraw-Hill, 1956, ch. 1-7.

[38] to the Wild Divine, J. *Skin conductance aquisition device, Lightstone*. `http://www.wilddivine.com/`.

[39] UIDIA. Unique Identification Authority of India Card Project-India. `http://www.uidaicards.com/`.

[40] Uludag, U., and Jain, A. Attacks on biometric systems: a case study in fingerprints. In *Proc. SPIE-EI 2004, Security, Seganography and Watermarking of Multimedia Contents VI* (2004), pp. 622 – 633.

[41] Valentine, D. Skin Conductance One Of The Fastest Ways To Test Stress, 2009. http://www.articlesbase.com/health-articles/skin-conductance-one-of-the-fastest-ways-to-test-stress-1464442.html, [Online; accessed 16-November-2009].

[42] Westerink, J. H. D. M., van den Broek, E. L., Schut, M. H., van Herk, J., and Tuinenbreije, K. *Computing Emotion Awareness Through Galvanic Skin Response and Facial Electromyography*, vol. 8 of *Philips Research Book Series*. Springer Netherlands, New York, December 2007.

A Guessing Entropy for Skin Conductance

Let R be the set of rules the attacker can use to reduce the password space from S to S'. So, for a rule R_i

$$antecedent(\mathsf{A}) \Rightarrow consequent(\mathsf{C})$$

such that, A=$[y_1, \ldots, y_{E_a}]$ and C=$[z_1, \ldots, z_{E_c}]$, where E_a are the elements in the antecedent and E_c in consequent. The process of calculating the new password space from a given one is shown in algorithm 2. S' indicates a lower bound for the password space which shows the minimum number of combinations an attacker needs to guess if he has a full knowledge of the mappings in the database.

Let Ψ denote the candidate set and Φ be the large itemset, Ψ^I and Φ^I are the two dimensional vectors derived from the rules $R_1, \ldots, R_I$. Each item (Ψ^I_J) in a Ψ^I is a vector of the form $[x_1, x_2, \ldots, x_{mSC}]$, $\forall$ $0 \leq J \leq L$, where $x_i \in (0, 1, *)$ and $L = |\Psi^I|$. Similarly, each item (Φ^I_J) in a Φ^I is also vector of the form $[x_1, x_2, \ldots, x_{mSC}]$, $\forall\ 0 \leq J \leq L$, where $x_i \in (0, 1, *)$ and $L = |\Phi^I|$.

Algorithm 2 Reduced Password Space for SC

```
PasswdSpace $(R)$
$\Phi^0_1 \longleftarrow [*, *, *, *, *, *, *, \ldots, *]$
$\mathsf{S} \longleftarrow 2^{m_{SC}}$
for I = 1 to $|R|$ do
    L $\longleftarrow$ length($\Phi^{I-1}$)
    $\Psi^I \longleftarrow$ NULL
    for J = 1 to L do
        if any $\left((\Phi^{I-1}_{J,y_1}, \Phi^{I-1}_{J,y_2}, \ldots, \Phi^{I-1}_{J,y_{E_a}}) == *\right)$ then
            $\Psi^I \longleftarrow \Psi^I \bigcup$ split($\Phi^{I-1}_J$)
        else
            $\Psi^I \longleftarrow \Psi^I \bigcup \Phi^{I-1}$
        end if
    end for
    $\Psi^I \longleftarrow$ unique $(\Psi^I)$
    cnt $\longleftarrow$ 1
    L $\longleftarrow$ length($\Psi^I$)
    for J = 1 to L do
        if $\left(\prod_{p=1}^{E_a} \Psi^I_{J,y_p} == 1\& \prod_{q=1}^{E_c} \Psi^I_{J,z_q} == 0\right)$ then
            delete $\left(\Psi^I_J\right)$
        else
            $\Phi^I_{cnt} \longleftarrow \Psi^I_J$
            cnt++
        end if
    end for
end for
$\mathsf{S}' \longleftarrow \Phi^{|R|}$
```

$*$ denotes *don't care* and can be assigned 0 or 1. The set of rules R obtained are passed to the algorithm 2 to generate S'. Φ^0_1 is initialized to $[* * * * * * * \ldots *]$ and S = $2^{m_{SC}}$. Below is the short description of the functions used in the algorithm.

- **length**(Ψ^I) - gives the total vectors in the candidate set Ψ^I i.e. $|\Psi^I|$.
- **any**$(\Phi^J_I(y_1, y_2, \ldots, y_{E_a}) == *)$ - a boolean function

$$= \begin{cases} \mathbf{1}, & (\Phi^I_{J,y_1} == *) \vee \ldots \vee (\Phi^I_{J,y_{E_a}}) == *) \\ \mathbf{0}, & \text{else} \end{cases}$$

- **split**(Φ^I_J) - this function generates a new candidate set Ψ^I from a large itemset Φ^{I-1} based on a rule R_I. It generates the vectors for Ψ^I *s.t.*
    - Mark y_i, if (Φ^{I-1}_{J,y_i} $==$ $*$), $\forall$ y_i $\in$ $[y_1, \ldots, y_{E_a}]$.
    - Generate all possible combination of the marked bits; which implies if total number of marked bits are mb then total possible combinations are 2^{mb}. For e.g. if Φ^I_J = [***1*1] and the rule R_I is $1 \Rightarrow 2$, then the result is $([11 * 1 * 1]\ [10 * 1 * 1]\ [01 * 1 * 1]\ [00 * 1 * 1])$
- **unique**(Ψ^I) - gives the unique vectors from Ψ^I.
- **delete**(Ψ^I_J) - delete Ψ^I_J from the candidate set Ψ^I.

During the Candidate Itemset Generation, a $*$ in the large itemset triggers a *split*; 1 and 0 indicates *do nothing*. However during the Large Itemset Generation a 1 in a candidate itemset triggers *add 1*; 0 indicates *do nothing*. During the whole procedure, each time one rule is used and the sets which does not comply with that rule are omitted to create the new set. The final password space is calculated by computing the total number of vectors which can be generated using $\Phi^{|R|}$, where $\Phi^{|R|}$ is the final large itemset generated from the rules $R_1, \ldots, R_{|R|}$.

An example shown in Table 4 with *5* elements, to how to generate the candidate itemset and the large itemset from *3* rules. The total number of guesses which an attacker needs to make is *14* which implies the effective number of bits in the new password space are *4*; original was *5*.

R		Candidate Itemset		Large Itemset
Initialization			Φ^0_1	$* * * * *$
$\mathbf{R_1}$ $1 \Rightarrow 3$	Ψ^1_1	$1 * * * *$	Φ^1_1	$1 * 1 * *$
	Ψ^1_2	$0 * * * *$	Φ^1_2	$0 * * * *$
$\mathbf{R_2}$ $(1, 2) \Rightarrow 5$	Ψ^2_1	$101 * *$	Φ^2_1	$101 * *$
	Ψ^2_1	$111 * *$	Φ^2_1	$111 * 1$
	Ψ^2_2	$01 * **$	Φ^2_2	$01 * **$
	Ψ^2_3	$00 * **$	Φ^2_3	$00 * **$
$\mathbf{R_3}$ $5 \Rightarrow 1$	Ψ^3_1	$101 * 0$	Φ^3_1	$101 * 0$
	Ψ^3_2	$101 * 1$	Φ^3_2	$101 * 1$
	Ψ^3_3	$111 * 1$	Φ^3_3	$111 * 1$
	Ψ^3_4	$01 * *1$	Φ^3_4	$01 * *0$
	Ψ^3_5	$01 * *0$	Φ^3_5	$00 * *0$
	Ψ^3_6	$00 * *1$		
	Ψ^3_7	$00 * *0$		

Table 4: Generating candidate set and large itemset

Notes

[1] We use a physiological data acquisition device called Lightstone from WildDivine [38].