Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 126 (2018) 703–712

Procedia Computer Science

www.elsevier.com/locate/procedia

22nd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Emotional Influences on Cryptographic Key Generation Systems using EEG signals

Dang Nguyen[a], Dat Tran[a], Dharmendra Sharma[a], Wanli Ma[a]

[a] *Faculty of Science and Technology*
University of Canberra, ACT 2601, Australia

Abstract

This paper presents a research conducted to verify the influences of emotion on electroencephalogram (EEG)-based cryptographic key generation system. Emotion, such as negative and positive feelings, involves in EEG signal, and hence it may influence on the system. This issue has not been analyzed. Using parametric spectral estimation technique for feature extraction, and devised a quantization technique for error correction, a key is generated from EEG data. These techniques are believed to be suitable for a noisy data such as EEG. For experiment implementation, we use the Database for Emotion Analysis using Physiological Signals (DEAP) dataset. The emotional dimensions used are valence (positive and negative) and arousal (calm and excited) that is divided each in two classes: low and high. For experimental methodology, we performed on two groups of subjects, valence and arousal. Experimental results show that emotion actually has impacts on the performance of the system.

Keywords: Knowledge discovery, Electroencephalogram, cryptographic key, emotion, valence, arousal, NIST.

1. Introduction

Biometric-based cryptographic key generation has received increasing attention from many researchers in the last few years. It is regarded as a data mining approach that exploits knowledge discovery techniques to derive biometric information for key generation purpose. This application has been widely used in security systems to avoid limitation of passwords. People can generally use their biometrics (voice, face, iris, and fingerprint) as good alternatives, or supplements, to PINs and passwords. It has been shown that electroencephalogram (EEG) perhaps provide a potential biometric modality in biometric-based key generation. EEG signals contain genetic information which means that there is a connection between genetic information and EEG of an individual[11]. Hence, EEG has been successfully used for person identification[15]. EEG is also changeable and random[20]. Moreover, EEG features are universal as all living and functional persons have recorded EEG signals[20]. Therefore, EEG can be useful for a cryptographic key generation system[18]. The use of brain waves patterns from EEG data as a new modality of biometrics for key

* Corresponding author. Tel.: +61405639877
E-mail address: Dang.van.Nguyen@canberra.edu.au

1877-0509
10.1016/j.procs.2018.08.004

generation has several advantages that that are impossible to be faked or compromised[11]: (a) EEG is confidential, because it corresponds to a secret mental task which cannot be observed. (b) EEG signal is very difficult to mimic, because mental tasks are person dependent. (c) EEG is brainwave so it is almost impossible to steal, because the brain activity is sensitive to the user's stress and mood. The user cannot be forced to reproduce the same EEG signal while he is under stress, and (d) EEG signals require alive person for recording by nature.

Emotion plays an important aspect in the interaction and communication between people. In particular, peoples moods may heavily influence their way of communicating, acting and productivity. For example, there are two car drivers, one being happy and the other being very mad. They will drive totally different. In the last few years, EEG-based emotion recognition has gained increasing attention. Emotion also plays a very important role in brain computer interface (BCI) systems, which will effectively improve the communication between human and machines[23]. Many other researchers mainly focus on two dimension scales on which emotions are mapped according to their valence (positive versus negative), and arousal (calm versus excited) on the field of EEG-based emotion recognition[6,9]. The recognizability of different emotions depends on how well the EEG features can be mapped onto chosen emotion representation. The emotion representation used is the two dimensional mapping with valence and arousal axes[6]. If valence is difficult to identity, it is able to exploit usable results on the arousal, and vice versa. Correct EEG-based emotion recognition is about 84.50% for the SVM method[14].

Accuracy is one of the crucial requirements of any biometrics-based key generation system, including EEG-based. Factors which may affect the accuracy of an EEG-based key generation system can be signal noises, feature extraction methods, and/or error-correction algorithms. Emotion may also have some effects on the performance as it has been shown that the negative/calm quadrant was a little underdeveloped compared to the other sections[6]. As a result, brain state changes from less to more ordered state, from more to less chaotic, or from more to less complexity during emotional recording of EEG signals.

In this study, therefore, we study the influences of emotion on EEG-based cryptographic key generation systems. There is little work on this issue. We firstly present a system for EEG-based key generation. The power spectral density estimate technique based on Autoregressive (AR) model from Burg's method for feature extraction is employed to obtain EEG features for EEG-based key generation. Then, we investigate the impacts of emotion by two factors: the success rates of system and the randomness of generated keys by through two dimensions (valence and arousal) because they are simple and suitable[6]. Two binary problems for emotion are used including the impact of low/high valence, and low/high arousal for the DEAP dataset[10]. The ratings of participants during the experiment were used to label the data. The ratings for each of the valence and arousal dimension were divided into two classes: low and high. Particularly, the 9-point rating scales were simply divided in the middle. If emotion has impacts on the system, the success rates of low valence (low arousal) group will be different from the high valence (high arousal) group.

2. EEG-based Cryptography Key Generation Method

Our method from EEG comprises of three steps: feature extraction, enrollment and key generation as shown in Figure 1[13].

2.1. Feature Extraction

First, a number of EEG channels is selected that aims to produce a cryptographic key that contained the possible maximum amount of biometric information achieved from subjects. For feature extraction, first a frequency sub-band is computed. Then, autoregressive (AR) spectral analysis with Burg method is used to obtain the Power Spectral Density (PSD) of the signal. AR process is chosen due to its ability to handle short segment of data (in our case, 1 second of data) while giving high frequency resolution, and smooth power spectra[21]. Autoregressive process is given by

$$x(i) = \sum_{k=1}^{d} a_k x(i-k) + e(n)$$

where $x(i)$ is the *ith* sampled data, d is the model order, a_k is the AR coefficients, and $e(i)$ is the prediction error term. Burg's method is used to estimate the AR coefficients, because this method is more accurate than it used the data directly, whereas inaccuracies occur when Yule-Walker equations were directly used, or error source like bias in the autocorrelation function estimate happen in other parameter estimation methods[21]. Power spectral density (PSD) of

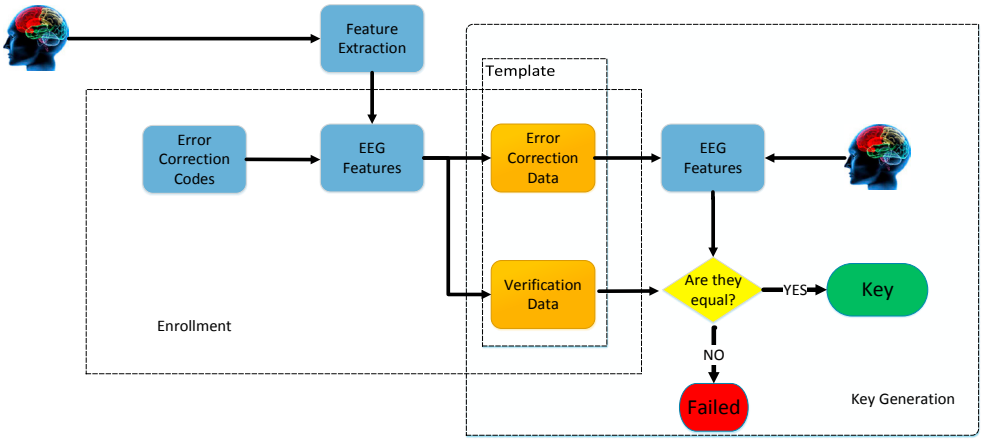


Fig. 1: EEG-based cryptographic key generation system [13]

a signal is a positive real function of a frequency variable associated with a stationary stochastic process. The PSD is defined as the discrete time Fourier transform (DTFT) of the covariance sequence [22]:

$\psi(\omega) = \sum_{k=-\infty}^{\infty} r(k)e^{-i\omega k}$

with the auto-covariance sequence $r(k) = E(y(t)y^*(t-k))$ and $y(t)$ is the discrete time signal assumed to be a sequence of random variables with zero mean.

After PSD estimate, the average power is computed to obtain a feature vector represented by:

$$\Phi = (\phi_1, ..., \phi_N) \tag{1}$$

2.2. Enrol

We use the notation $\|$ to refer to string concatenation, and $L[i]$ is the i^{th} element in the list L. π is notated as a password of an user and $x \xleftarrow{R} X$ be the uniform selection of x at random from a set X, and $x \leftarrow A$ is to show that x is an output of algorithm A. Let $[a, b]_k = \{a + ik : i \in [0, \lfloor (b - a/k) \rfloor]\}$. E and D are 192-bit AES encryption and decryption algorithms [17]. Four cryptographic hash functions are used: H_0 and H_1 are two functions that map a password in a set of passwords into two different elements in the space of 128 bits. H_{ver} is a SHA-192 hash function used to generate a token to check whether a generated key is correct, and H_{key} is a SHA-192 hash function to generate a cryptographic key [16]. We notate $B = \{\beta_1, \ldots, \beta_{M+1}\}$ to be a set of $(M + 1)$ biometric samples from a user, $\Delta = 1 + max_i(\delta_i)$, and $\Phi = \{\phi_1, \ldots, \phi_N\}$ be a set of N feature vectors extracted from B that will be presented in the next section.

The method is designed for legitimate users to generate cryptographic keys so that an adversary cannot learn how to measure biometric inputs, and recover the keys. In traditional biometric-based key generation, a template encodes a same set of features that are used to measure biometric inputs. However, in this proposal, we assign randomly EEG features to users, and encode the features so that adversaries cannot determine which features were originally used to generate a key. The advantages of this approach is to increase the adversary's attempts to guess for the correct key because he has to guess both the correct EEG sample and the set of features that were used. Moreover, this approach advances in decreasing equal error rates because it allows to measures EEG inputs more accurately by using features as a function of EEG sample. In addition, we added a random string to combine with a password to protect better templates rather than a password alone.

The method consists of two algorithms: Enrol (Algorithm 1) and KeyGen (Algorithm 3). The Enrol phase is a process of 3 steps: threshold estimation, error correction, key generation and template creation. We assume IV to be a random and secrete number of 128-bit length that can be generated by random number generators [4].

Threshold Estimation. First, a user provides M biometric samples $\phi_i(\beta_1), \ldots, \phi_i(\beta_M)$, $i = 1, 2, ..., N$ and a password π. The key generation system computes EEG features used for key generation. To determine thresholds to correct features of a user, we do the following steps (algorithm 2):

Algorithm 1 Enrol

Input: Password π, sample set $\{\beta_1, \ldots, \beta_M\}$, and feature set $\Phi_u = \{\phi_1, \ldots, \phi_N\}$ of a user u, and multi-thresholds $\delta_1, \ldots, \delta N$ with $\Delta = max_i(\delta_i)$
Output: Key K and template T

1. $L \leftarrow Permute\{1, ..., N\}$
2. $k_0 \leftarrow H_0(\pi \oplus IV)$, $k_1 \leftarrow H_1(\pi \oplus IV)$
3. For $j = 0$ to $\mid L \mid -1$
   (a) $i \leftarrow L[j]$
   (b) $\mu_i \leftarrow Mean(\phi_i(\beta_1), \ldots, \phi_i(\beta_M))$
   (c) $\alpha_i \leftarrow \lfloor \mu_i - \delta_i/2 \rfloor \bmod \delta_i$ if $\mu_i \geq \delta_i/2$. Otherwise, $\lfloor \mu_i + \delta_i/2 \rfloor$
   (d) $x_i \leftarrow \max(0, \lfloor \mu_i - \delta_i/2 \rfloor)$
   (e) $\rho_i \overset{R}{\leftarrow} [\alpha_i, \Delta]_{\delta_i}$
   (f) $C_j = (E^N_{k_0}(i), E^{\Delta}_{k_1}(\rho_i))$
   (g) $K_j = i \parallel x_i$
4. $K \leftarrow H_{key}(\pi \parallel K_0 \parallel \ldots \parallel K_{|\Psi|-1})$
5. $T \leftarrow (C, v) = ((C_0, \ldots, C_{|L|-1}), H_{ver}(\pi \parallel K_0 \parallel \ldots \parallel K_{|\Psi|-1}))$
6. Return K and T

Algorithm 2 Multi-Threshold Estimation

Input: Feature set $\Phi = \{\phi_1, \ldots, \phi_N\}$ and sample set $\{\beta_1, \ldots, \beta_M\}$ of a user u.
Output: Quantization thresholds $\delta_1, \ldots, \delta_N$

1. For $i = 1 \ldots N$
   (a) Sorting $\phi_i(\beta_1), \ldots, \phi_i(\beta_M)$ in ascending order
   (b) $\lambda_i = mean(\phi_i(\beta_1), \ldots, \phi_i(\beta_M))$
   (c) $t_i = max(\phi_i(\beta_M) - \lambda_i, \lambda_i - \phi_i(\beta_1))$
   (d) $\epsilon_i = \frac{1}{M-1} \sum_{k=1}^{M-1} (\phi_i(\beta_{k+1}) - \phi_i(\beta_k))$
   (e) $\delta_i = 2(t_i + \epsilon_i)$
2. Return $\delta_1, \ldots, \delta_N$

Error-Correction Codes. In this step, we correct EEG features using the vector quantization technique. Its purpose is to correct features into a single, repeatable value by partitioning them into intervals, and performed as follows (see algorithm 1, step 4(a)-4(d)). First, we compute μ_i as the mean of $\phi_i(\beta_1), \ldots, \phi_i(\beta_M)$ for each index $i \in L$. Then, the range $R_i = [0, r_i]$ of each feature ϕ_i is partitioned around μ_i into intervals ($[\alpha_i + k\delta_i, \alpha_i + (k+1)\delta_i, k \in [0, \lfloor (r_i/k) \rfloor] - 1$) of length δ_i by computing a lowest boundary:

$$\alpha_i = \begin{cases} \lfloor \mu_i - \delta_i/2 \rfloor \bmod \delta_i \text{if} \mu_i \geq \delta_i/2 \\ \lfloor \mu_i + \delta_i/2 \rfloor \text{if} \mu_i < \delta_i/2 \end{cases} \tag{2}$$

Finally, an offset ρ_i of quantization is randomly chosen in the range $[\alpha_i, \Delta]_{\delta_i}$ with $\Delta = \max(\delta_i)_i$, and $x_i = \max(0, \lfloor \mu_i - \delta_i/2 \rfloor)$ is computed as the border of partition that contains μ_i.

Key generation. This step is to generate a cryptographic key. A key is derived from a password π, the feature indexes, and the quantized feature by computing $K_j = L[j] \| x_{L[j]}$ with $j = 0, 1, ..., N-1$, and $K = H_{key}(\pi \parallel K_0 \parallel \ldots \parallel K_{|\Psi|-1})$. It means that K is the output of a hash function applied to the password, feature indexed and the lower boundary of partition.

Template creation. Our goal is to protect the feature indexes and the quantization information so that only a user who has π and the ability of reproduction EEG signal that is close to the samples in enrolment phase can regenerate the correct key (step 3(e+f) in algorithm 1). To do that, we first combine the password and IV to create two random and secrete values $k_0 \leftarrow H_0(\pi \oplus IV)$, $k_1 \leftarrow H_1(\pi \oplus IV)$. Next, for each $j \in [0, N-1]$, we select randomly $\rho_i \overset{R}{\leftarrow} [\alpha_i, \Delta]_{\delta_i}$ (with $i = L[j]$) to encode α_i to be a random integer that is smaller than the largest quantization width $\Delta = max_i(\delta_i)$. Then, we encrypt both i and ρ_i by computing $C_j = (E_{k_0}(i), E_{k_1}(\rho_i))$. Then, a template is:

$$T = (C, v) = ((C_0, \ldots, C_{|L|-1}), H_{ver}(\pi \parallel K_0 \parallel \ldots \parallel K_{|\Psi|-1})) \tag{3}$$

The value of v is used for only verification, but is different from the key K because of two different hash functions H_{ver} and H_{key}.

2.3. Key Generation

The input of the algorithm is a password π, the template $T = (C, v)$, and a new biometric sample β_{M+1}. This algorithm (see algorithm 3) consists of the following steps:

(a) Extracting features $\phi_1(\beta_{M+1}), ..., \phi_N(\beta_{M+1})$ from the sample β_{M+1}.

(b) Decrypting the vector C to reproduce the list L and the quantization offsets $\rho_{L[j]}$, $j = 0, ..., |C| - 1$. The values in list L are a list of feature indexes. After that, computing x_i to be the largest boundary of partition that is smaller than or equal to $\phi_i(\beta_{M+1})$ for $i = L[j]$, $j \in [0, |C| - 1]$, then letting $K_j = i \| x_i$ and concatenating these values to produce a temporary key (see Steps 2(a)-2(c)).

(c) Hashing the temporary key $H_{ver}(\pi \parallel K_0 \parallel ... \parallel K_j)$ and checking the result with v. If they are equal, the key $K = H_{ver}(\pi \parallel K_0 \parallel ... \parallel K_j)$ is output, otherwise the algorithm fails (see Steps 2(d) and 2(e)).

Algorithm 3 KeyGen

Input: Template $T = (C, v)$, passwords π, sample β_{M+1}, feature $\Phi_{M+1} = (\phi_{M+1,1}, ..., \phi_{M+1,N})$ of this sample, and multi-thresholds $\delta_1, ..., \delta_N$

Output: Key K or $\perp$

1. $k_0 \leftarrow H_0(\pi \oplus IV)$, $k_1 \leftarrow H_1(\pi \oplus IV)$
2. For $j = 0$ to $| C | -1$
   (a) $i \leftarrow D^N_{k_0}(C[j][0])$
   (b) $\alpha_i \leftarrow D^N_{k_1}(C[j][1])$s
   (c) $x_i \leftarrow \max_{x \in 0 \cup [\alpha_i, \phi_{M+1,i}(\beta_{M+1})]_{\delta_i}} x$
   (d) $K_j = i \parallel x_i$
   (e) if $H_{ver}(\pi \parallel K_0 \parallel ... \parallel K_j) = v$ then $K = H_{key}(\pi \parallel K_0 \parallel \ldots \parallel K_j)$
   (f) Return K
3. Return $\perp$

2.4. Security Analysis

To successfully unlock the cryptographic key (K) an adversary would require the password, the user biometric, and the template.

The compromise of any one of these factors only is not enough to aid an adversaries to regenerate the correct key K. The password or user biometric cannot be used only to regenerate the correct key without the stored template (contained either in a smart card or central database store). Alternatively even if the template is compromised by an adversary, what an adversary can do is to guess a biometric β', a password π', and run the algorithm 3 with T, and hope the output matches K. However, no information about the user biometric and the key K can be leaked. The template is derived from C and v which is generated independently from completely separate random process. C is produced from any cryptographically secure encryption algorithm whose outputs do not leak any information about its inputs because IV is sufficient secure[3], therefore it will not leak no information about the user biometric, and protect the template better than RBTs. Moreover, K and v produced from independently hash functions whose outputs are random, so there is no information about K to be leaked if v is compromised.

On the other hand, if the K is badly comprised, an adversary is unable to find a collision to discover the user biometric with knowledge of password and the template because of secure hash function and random permutation used. In this case, a new key can be issued, and the user will be required to provide a new biometric measurement to the system, and the old biometrics will be canceled. This method is flexible for EEG signals, but the new biometrics may be derived from the old template[2] for iris (cannot be changed) to generate a new key.

3. Experiments and Results

DEAP is a dataset for emotion analysis using EEG (Dataset for Emotion Analysis using Electroencephalogram, Physiological and Video Signals) which is an open database proposed by Koelstra et al.[10]. EEG signals of 32 participants were recorded while they watched 40 one-minute long excerpts of music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity, with ratings from 1 to 9 after watching videos. EEG was recorded with 32 electrodes, placing according to the international 10-20 system. Each electrode recorded 63s EEG signal, with 3s baseline signal before the trial. The data was recorded at a sampling rate of 512 Hz in two separate locations with 40 channels. Then, the data was preprocessed to a sample rate at 128Hz, segmented into 60 second trials and removed a 3 second pre-trial baseline. The dataset is summarized in Table 1.

Table 1: DEAP Dataset Descriptions

Number of subjects	Number of Channels	Number of trials	Number of sessions	Trial length (seconds)	Sampling (Hz)
32	32	40	1	60	128

We use the equal error rate (EER) when the false rejection rate (FRR) equals to the false acceptance rate (FAR) to evaluate this method. To compute the FRR we use the repeated leave-out-κ cross validation method. Given ν enrolment samples, we random choose $\nu - \kappa$ samples to create a training data for generating a key and a template in the enrolment phase. Then we use the remaining κ samples to create a testing data to regenerate the key with this template, and measure the number of samples that the key are not regenerated. We set the ratio of ν and κ to be 4:1. This process is repeated 10 times and averages across all 10 runs are used to compute FRR. To compute FAR, we use all of samples of a user from a training data to create a key and a template, and all samples from a testing data of remaining users to test the ability of forgers to regenerate the correct key.

Two binary problems for emotion are low/high arousal, and low/high valence. The ratings of participants during the experiment were used to label the data. The ratings for each of the valence and arousal dimension were divided into two classes: low and high. Particularly, the 9-point rating scales were simply divided in the middle. Experiments aim to achieve the highest accuracy based on two factors: a variation of AR order, and a number of channel selection.

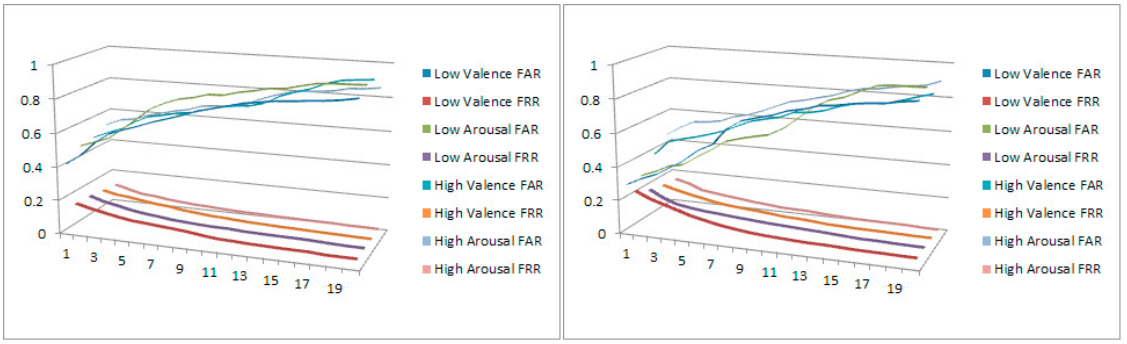
AR-based spectral analysis has an inconvenience that the order of AR model have to be estimated prior to the spectral analysis[5]. Among the criteria for selection of the optimal order of the AR model, the Akaike information criterion (AIC)[1] were evaluated in this work that has been regarded as one of the important breakthroughs in statistics

Table 2: Optimal Order of DEAP datasets. StD = Standard Deviation

	Valence		Arousal	
	High	Low	High	Low
Mean	27.15	26.28	24.6	27.68
Std	9.18	10.01	11.98	8

in the twentieth century[8]. These criteria measure the variance of the prediction error of the model in order to determine the order that minimizes the error without leading to unnecessary computation[7]. The technique used in this paper is similar to the approach presented in[5] in order to determine optimal AR model orders for spectral analysis of short segments of the time series including EEG. Based on the results shown in Table 2, we determined that AR model orders between 12 and 36 should be used.

For suitable channel selection to produce the possible maximum amount of biometric information achieved from all subjects, we do experiments and measure the FAR and FRR to determine the EERs as small as possible. Experimental results are presented in Figures 2 and 3. We choose two channels F3 and F4 to be the first because they are two most locations of use for emotion recognition from EEG signals[6]. For other selections, an easy way used is to select randomly EEG channels for 10 times, and measure the FARs and FRRs of each to obtain averages of these values as our results. As shown in Figure 2, with 2 and 4 channels selected, the FARs and FRRs do not match to get the EERs for four instances. Figure 3 shows that while there are half of selections for 8 channels that FARs and FRRs meet to obtain EERs, they meet together at all for selection of 16 channels. In addition, 16-channel selection provides the better performance as seen in Table 3, so it will be used from now. Our limitation is that this may not provide the best performance because of randomly selection, so in the future we want to do more experiments to exactly identify the channels.



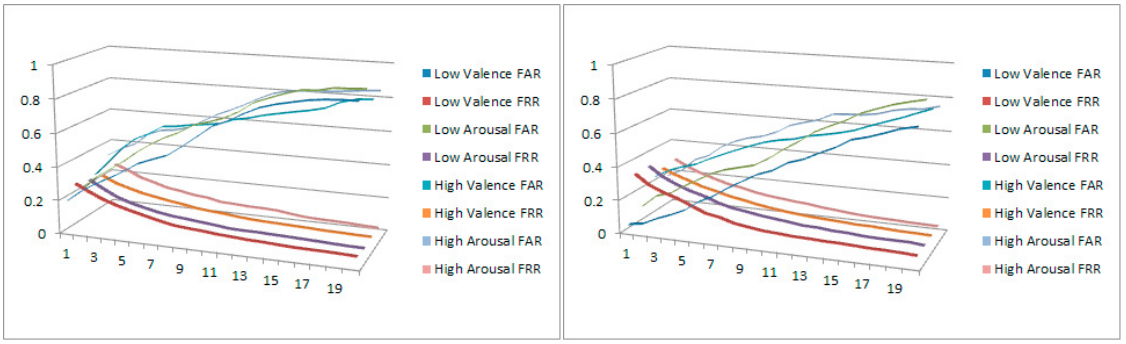
(a) (b)

Fig. 2: Average performance on error rates of the method: (a) 2 channels (F3 and F4), (b) 4 channels. The FARs increase all, while the FRRs decrease, and they do not match for these selections.

Table 3: Comparison of the accuracy between 8 and 16 channels of selection.

	8 Channels		16 Channels	
	Valence	Arousal	Valence	Arousal
High	-	-	73.92	75.56
Low	80.1	78.3	82.65	78.28

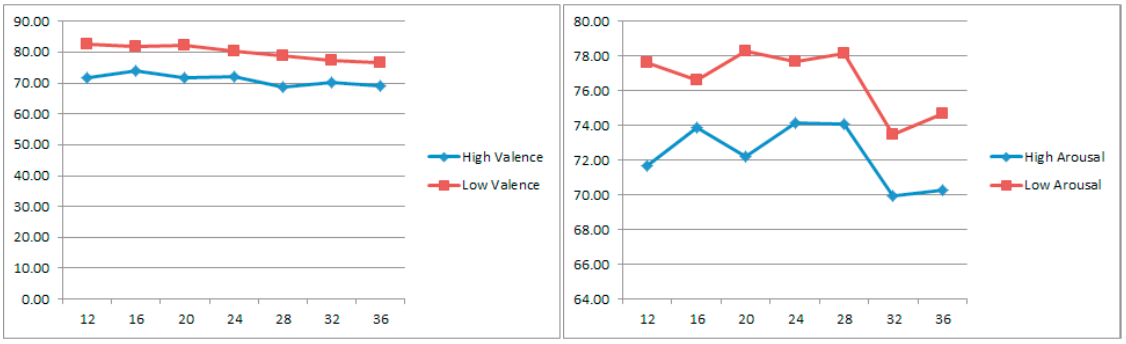
For analysis of the influence of emotional dimensions, experimental results are presented in Figures 4. We do investigation the variation of accuracy (complement of EER) by a change an order of AR model. We are not aiming to compare the accuracy on the same feature extraction method, but we can have a remark on which group of emotional dimension has a smaller error rates than others. As shown in the figure 4, the accuracy of the low groups are always



(a) (b)

Fig. 3: Average performance on error rates of the method: (a) 8 channels, (b) 16 channels. The FARs increase all, while the FRRs decrease. They match to get EERs all for a selection of 16 channels, but low groups only for 8 channels.

smaller than of the high group for all values of AR order. The difference fluctuates and peaks at 10.86% observed at order 12 for valence dimension, it is quite similar for the arousal dimension with the minimum of nearly 3% observed at order 16. Although it is not significant to see whether random difference between the two datasets or emotion has contributed to the performance of the system, the later case would be appropriate as the negative/calm quadrant was a little underdeveloped compared to the other sections[6]. As a result the negative/calm stimuli were less extreme, thus, it may help to increase the accuracy of the key generation systems.



(a) (b)

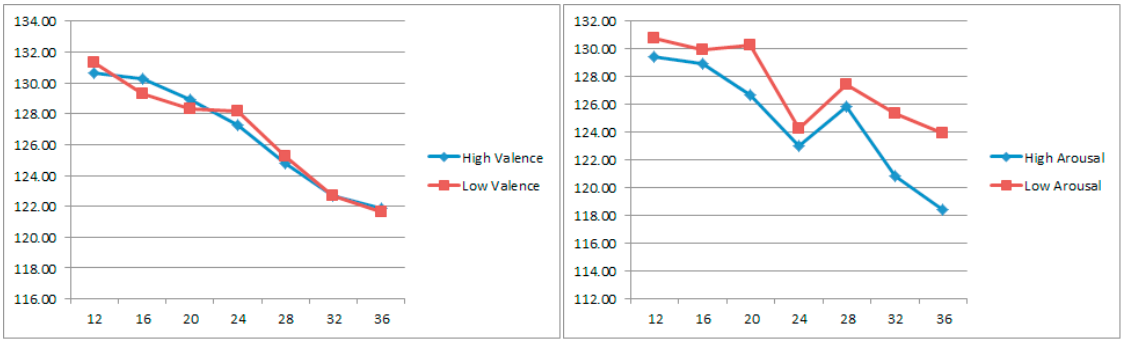
Fig. 4: Comparison of the method's performance for (a) valence, (b) arousal. The labels on the x-axis correspond to a change of AR order from 12 to 36.

Although the accuracy of method is not as high as the voice[12]. As summarized in Table 3, the success rate highest is 82.65%, while the voice reaches nearly 98%. We can improve the success rate by combining these emotional dimension. Accordingly, the error equal rate is close to that of the voice, that is, 97.88% versus 98% respectively (see Table 4).

Table 4: The accuracy for combined emotional dimensions

Order	12	16	20	24	28	32	36
High Valence, High Arousal	96.97	96.67	95.76	96.46	96.67	97.27	96.57
High Valence, Low Arousal	95.15	94.85	94.20	94.85	93.33	94.85	93.94
Low Valence, High Arousal	97.58	**97.88**	96.52	96.97	96.97	95.30	94.85
Low Valence, Low Arousal	95.05	94.24	94.97	94.55	94.30	94.85	94.85

We used a comprehensive battery of tests from NIST[19] to evaluate the randomness of keys generated. We firstly measure the average of key length for users in the dataset. As shown in figure 5, key length is selected to be 128 bits that is the largest for the EEG signal and suitably used for 128-bit AES application. The reason is that each channel generated a character containing at least 8 bits as seen in the step 3(g) on algorithm 1, and it can be the maximum of 256-bit keys obtained from the dataset. Then, for randomness tests, we use six out of fifteen tests in the NIST Test Suite as other tests requires more than 1000-bit length. We do not apply NIST test for the beta band because of shorter key length generated. Six testes include frequency, block frequency, runs, longest run, approximate entropy, and serial that has two tests. As shown in table 5, generated keys pass the most of six tests with high success rates, it means that the keys are good random. Moreover, the success rate of high group for both valence and arousal is similar the low group, and it shows that emotional dimensions does not has impacts on the randomness of keys obtained.



(a) (b)

Fig. 5: The variation of length of keys obtained: (a) valence, (b) arousal. The labels on the x-axis correspond to a change of AR order from 12 to 36.

Table 5: Results of NIST Test Suite for randomness. Values in table is percentage of passing rates

Emotional Dimension	Valence		Arousal	
	High	Low	High	Low
Frequency	100.00	100.00	96.00	100.00
Block Frequency	100.00	96.99	100.00	100.00
Runs	68.75	78.12	72.00	80.00
Longest Run	100.00	100.00	92.00	100.00
Approximate Entropy	100.00	100.00	100.00	100.00
Serial	96.88	93.75	92.00	92.00
Serial	96.88	93.75	96.00	92.00
Average Success Rate	**94.64**	**94.66**	**92.57**	**94.86**

Overall, we can see the changes on accuracy of EEG-based cryptographic key generation system on different emotional datasets, but they do not have impacts on the randomness of keys generated. This can contribute to EEG-based emotion recognition. For instance, if we have two undetermined groups containing negative and positive feelings. Then, by EERs evaluation of the key generation system, the negative group can be identified to be a smaller one of EERs.

4. Conclusion

We have explored that emotion does have impacts on the accuracy of EEG-based cryptographic key generation. An extensive analysis has been carried out to investigate, and the results show that, with suitable parameter selection, the success rate decrease for the negative and calm datasets if the keys obtained contains these kind of emotions. The

experimental results show that the accuracy of the system is at maximum of 97.88% for 16 channels selected from the DEAP dataset and observed at the AR order of 16 for the combined emotional dimension. This implies that key generation systems will not yield high performance on the users who have positive and excited emotions. This also have a contribution to EEG-based emotion recognition. In contrast, the randomness of keys cannot be affected by these particular type of emotions. There is a limitation of the study is that the accuracy for two emotional dimensions is low at 82.65% and there is a need of improvement. For further studies, we will conduct experiments on other feature extraction and channel selection methods, and on a larger scale of datasets to verify the method comprehensively, and improve the system accuracy.

References

1. H. Akaike, "A new look at the statistical model identification," *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974.
2. L. Ballard, S. Kamara, and M. K. Reiter, "The practical subtleties of biometric key generation." in *USENIX Security Symposium*, 2008, pp. 61–74.
3. E. Barker, "Nist special publication 800-57 part 1 revision 4 recommendation for key management–part 1: General," 2016.
4. E. Barker and J. Kelsey, "Recommendation for random number generation using deterministic random bit generators," *NIST Special Publication*, vol. 800, p. 90A, 2015.
5. A. Boardman, F. S. Schlindwein, and A. P. Rocha, "A study on the optimum order of autoregressive models for heart rate variability," *Physiological measurement*, vol. 23, no. 2, p. 325, 2002.
6. D. O. Bos *et al.*, "Eeg-based emotion recognition," *The Influence of Visual and Auditory Stimuli*, vol. 56, no. 3, pp. 1–17, 2006.
7. J. Carvalho, A. Rocha, I. Dos Santos, C. Itiki, L. Junqueira, and F. Nascimento, "Study on the optimal order for the auto-regressive time-frequency analysis of heart rate variability," in *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, vol. 3. IEEE, 2003, pp. 2621–2624.
8. J. Fan and Q. Yao, *Nonlinear time series: nonparametric and parametric methods*. Springer Science & Business Media, 2008.
9. R. Horlings, D. Datcu, and L. J. Rothkrantz, "Emotion recognition using brain activity," in *Proceedings of the 9th international conference on computer systems and technologies and workshop for PhD students in computing*. ACM, 2008, p. 6.
10. S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
11. S. Marcel and J. R. Del Millan, "Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 4, pp. 743–752, 2007.
12. F. Monrose, M. K. Reiter, Q. Li, and S. Wetzel, "Cryptographic key generation from voice," in *Security and Privacy, 2001. S&P 2001. Proceedings. 2001 IEEE Symposium on*. IEEE, 2001, pp. 202–213.
13. D. Nguyen, D. Tran, D. Sharma, and W. Ma, "On the study of eeg-based cryptographic key generation," *Procedia Computer Science*, vol. 112, pp. 936–945, 2017.
14. S. Paul, A. Mazumder, P. Ghosh, D. Tibarewala, and G. Vimalarani, "Eeg based emotion recognition system using mfdfa as feature extractor," in *Robotics, Automation, Control and Embedded Systems (RACE), 2015 International Conference on*. IEEE, 2015, pp. 1–5.
15. M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou, "Person identification based on parametric processing of the eeg," in *Electronics, Circuits and Systems, 1999. Proceedings of ICECS'99. The 6th IEEE International Conference on*, vol. 1. IEEE, 1999, pp. 283–286.
16. F. PUB, "Secure hash standard (shs)," 2015.
17. N. F. Pub, "197: Advanced encryption standard (aes)," *Federal Information Processing Standards Publication*, vol. 197, pp. 441–0311, 2001.
18. K. Ravi, R. Palaniappan, C. Eswaran, and S. Phon-Amnuaisuk, "Data encryption using event-related brain signals," in *Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on*, vol. 1. IEEE, 2007, pp. 540–544.
19. A. Rukhin, J. Soto, J. Nechvatal, E. Barker, S. Leigh, M. Levenson, D. Banks, A. Heckert, J. Dray, S. Vo *et al.*, "Statistical test suite for random and pseudorandom number generators for cryptographic applications, nist special publication," 2010.
20. S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
21. R. Shiavi, *Introduction to applied statistical signal analysis: Guide to biomedical and electrical engineering applications*. Academic Press, 2010.
22. P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson Prentice Hall Upper Saddle River, NJ, 2005, vol. 452.
23. N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from eeg signals using multidimensional information in emd domain," *BioMed research international*, vol. 2017, 2017.