# Private Authentication Keys Based On Wearable Device EEG Recordings

Hongxu Yang
Electrical Engineering Department
Eindhoven University of Technology
Email:h.yang@tue.nl

Vojkan Mihajlović
Wearable Health Solutions (WHS)
imec the Netherlands / Holst Centre
Email:vojkan.mihajlovic@imec-nl.nl

Tanya Ignatenko
Electrical Engineering Department
Eindhoven University of Technology
Email:t.ignatenko@tue.nl

***Abstract*—In this paper, we study an Electroencephalography (EEG) based biometric authentication system with privacy protection. We use motor imagery EEG, recorded using a wearable wireless device, as our biometric modality. To obtain EEG-based authentication keys we employ the fuzzy-commitment like scheme with soft-information at the decoder, see Ignatenko and Willems [2014]. In this work we study the effect of multi-level quantization together with binary encoding of EEG biometric at the encoder on the system performance, when EEG feature vectors have limited length. We demonstrate our findings on an experimental EEG dataset of ten healthy subjects.**

## I. Introduction

Access control systems based on biometric authentication gain popularity, since unlike passwords biometrics have strong bond with their user. Traditional biometric systems are based on voice, face, iris or fingerprint modalities. These biometric modalities, however, suffer from the problem that they can be acquired relatively easy and replicated by attackers, aiming at impersonating legitimate users. Therefore recently biometric research focused on finding alternative modalities that incorporate liveness detection property and are hard to obtain without a cooperating user. Physiological signals such as electroencephalography (EEG), electrocardiography (ECG), etc. became potential candidates for such biometric modalities.

Application of EEG signals in biometric systems is a relatively new research field that started in this century with initial work of Paranjape et al. [1]. Research in this field was mostly focused on assessing suitability of different EEG modalities and signal analysis methods to be used for authentication. In [1] the use of EEG patterns specific for eyes open/eyes closed conditions was explored. There, by applying autoregressive (AR) modeling on 8 channel EEG data and discriminant function analysis (DFA), classification accuracy over 80% was achieved. Palaniappan et al. [2] used EEG patterns generated by visual stimulation - a modality called Visual Evoked Potential (VEP); while Marcel and Millan [3] used EEG patterns associated with mental task stimulation. Combining AR modeling with kNN and Gaussian mixture modeling with MAP detector, respectively, they could achieve classification accuracy above 90%. Later works explored potential limitations of EEG biometrics. Brigham et al. [4] tested performance of the EEG biometric system over time. There EEG signals for imagery tasks were recorded for six subjects over four days. The EEG biometric system was created based on the AR modeling and support vector machines (SVM). Performance of the resulting system varied from 78.6% to 99.8%, showcasing difficulty to build a robust EEG-based system. Abdullah et al. [5] explored the effect of minimizing the number of channels on the classification performance. When using 2 to 4 channels on the eyes open/eyes closed modality, the highest accuracy achieved was 81%. Recently, Campisi et al. [6] proposed a new method, called Eigenbrain, that used VEP to realize an EEG-based authentication system. There the EEG signals were modeled using principal component analysis (PCA) applied to the EEG power spectrum density (PSD) over 19 electrodes, and classification was done using linear discriminant analysis (LDA). The achieved identification rate for a database of 60 participants was 99%, while equal error rate (EER) was 15%.

A vital aspect that has to be addresses in biometric authentication system design is biometric privacy. Since biometrics are unique identifiers of human beings, they cannot be revoked if compromised. Moreover, physiological signals also contain sensitive medical information. To address these problems, biometric cryptosystems based on key-binding and key-generation are often used, see e.g. [7] and [8]. Studies of EEG biometrics protection are however quite limited. Early work of Maiorana et al. [9] presented application of turbo codes for EEG secret generation. Their subsequent work [10] presented a case study on 40 subjects using eyes closed EEG patterns recorded over 19 electrodes. They achieved False Rejection Rate (FRR) of 9.5% and False Acceptance Rate (FAR) of 6.4% with 39 bits security, using turbo code of rate $1/15$.

*Our Contribution:* We concentrate on the EEG-based authentication system with privacy protection. We make use of fuzzy commitment with multi-level quantization at the encoder and soft information at the decoder [11]. One of the problems with physiological signals is that the size of the corresponding biometric feature vectors with independent components is relatively small to apply error-correcting codes (EEC) that are essential building blocks of biometric cryptosystems. Therefore we study the effect of using encoding techniques that introduce redundancy in the enrollment biometrics. To test our system, we use EEG data corresponding to motor imaginary activities. Our EEG database contains data recorded in multiple sessions during a week time-period using the wearable wireless EEG headset designed at imec. Thus our setting corresponds to the realistic authentication scenario that takes into account EEG

ISBN 978-0-9928626-7-1 

time variability as well as practical system usability.

## II. Biometric Privacy-Preserving Systems

Here we first present generic construction for biometric systems that creates unique reproducible secret keys for biometric observations. Such secret keys and the corresponding systems lay the basis for privacy-preserving biometric authentication. In these systems during enrollment, biometric sequence is observed and a secret key together with the helper data are produced. During authentication, a noisy observation of biometric is taken and the corresponding secret key is estimated based on this biometric observation and the helper data. Finally, biometric secret keys are hashed using cryptographic one-way functions and authentication decision is made by comparing the hashed versions of the biometric secret keys produced during enrollment and authentication phases.

Further we present a fuzzy commitment scheme, which is a particular realization of such generic biometric system, and its adaptation for continuous biometric sources.

### A. Generic Biometric System Based on Key-Binding

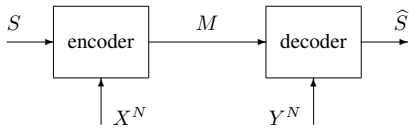


Fig. 1. Model for a biometric system based on key-binding.

Consider first a Gaussian biometric source. Let $x^N = (x_1, x_2, \ldots, x_N)$ and $y^N = (y_1, y_2, \ldots, y_N)$ be enrollment and authentication biometric sequences, respectively, composed of $N$ real-valued components. These sequences are produced by a Gaussian biometric source $\{G_\rho(x,y), x \in \mathbf{R}, y \in \mathbf{R}\}$, and thus the probability density corresponding to sequence pairs $(X^N, Y^N)$ is given by[1]

$$p_{X^N Y^N}(x^N, y^N) = \prod_{n=1}^{N} G_\rho(x_n, y_n), \text{ where} \quad (1)$$

$$G_\rho(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right),$$

for $x \in \mathbf{R}, y \in \mathbf{R}$ and correlation coefficient $|\rho| < 1$.

Here the pairs $\{(X_n, Y_n), n = 1, 2, \ldots, N\}$ are independent of each other and identically distributed (i.i.d.). The quality of biometric observations is characterized by the signal-to-noise ratio (SNR) expressed in terms of correlation coefficient $\rho$ as

$$\text{SNR} = \rho^2/(1-\rho^2). \quad (2)$$

Consider now a biometric system based on key-binding, see Fig. 1. During enrollment, a secret key $S \in \{1, 2, \ldots, |\mathcal{S}|\}$ is chosen uniformly at random and independently of biometrics:

$$\Pr\{S = s\} = 1/|\mathcal{S}|, \text{ for all } s \in \{1, 2, \ldots, |\mathcal{S}|\}. \quad (3)$$

An encoder observes the biometric enrollment sequence $X^N$ and secret $S$, and produces the helper data $M \in \{1, \ldots, |\mathcal{M}|\}$ as $M = e(S, X^N)$, where $e(\cdot,\cdot)$ is the encoder mapping.

[1]Scaling can always be applied to obtain unit variance for $X$ and $Y$.

The helper data produced by the encoder are communicated to the decoder and assumed to be public. The decoder in its turn also observes the biometric authentication sequence $Y^N$, and forms an estimate $\widehat{S} \in \{1, 2, \ldots, |\mathcal{S}|\}$ of the chosen secret based on available information, i.e., $\widehat{S} = d(M, Y^N)$, where $d(\cdot,\cdot)$ is the decoder mapping.

The goal of these biometric systems is to reliably share a secret key such that the amount of information the helper data provides about the secret (secrecy leakage) is negligible, while amount of information the helper data provides about the biometric enrollment sequence (privacy leakage) is minimal. We define secret-key and privacy-leakage rate pairs $(R_s, L_p)$ with $R_s \geq 0$ to be achievable if for all $\delta > 0$ and all $N$ large enough, there exist encoders and decoders such that

$$\begin{aligned} \Pr\{\widehat{S} \neq S\} &\leq \delta, \text{ (reliability)} \\ \log_2 |\mathcal{S}|/N &\geq R_s - \delta, \text{ (secret-key rate)} \\ I(S;M)/N &\leq \delta, \text{ (secrecy)} \\ I(X^N;M)/N &\leq L_p + \delta. \text{ (privacy)} \end{aligned} \quad (4)$$

In [12] the region $\mathcal{R}_\rho$ of all achievable secret-key and privacy-leakage rate pairs for key-binding for Gaussian biometric sources was characterized and the corresponding result is stated in the following theorem.

*Theorem 1:* [Key-binding based on Gaussian sources, [12]]

$$\begin{aligned} \mathcal{R}_\rho = \{(R_s, L_p) \;:\; & 0 \leq R_s \leq \frac{1}{2}\log_2\left(\frac{1}{\alpha\rho^2+1-\rho^2}\right), \\ & L_p \geq \frac{1}{2}\log_2\left(\frac{\alpha\rho^2+1-\rho^2}{\alpha}\right), \\ & \text{for } 0 < \alpha \leq 1\}. \end{aligned} \quad (5)$$

Related to this region, the relation between a given privacy leakage and the corresponding maximum achievable secret-key rate is given by the following rate-leakage function

$$R_\rho(L) \triangleq \max_{(R_s, L_p) \in \mathcal{R}_\rho} R_s = \frac{1}{2}\log_2\left(1 + \text{SNR}\frac{(2^{2L_p}-1)}{2^{2L_p}}\right). \quad (6)$$

The secrecy capacity of such biometric system is given by

$$\lim_{L_p \to \infty} R_\rho(L_p) = \frac{1}{2}\log_2(1+\text{SNR}) = I(X;Y). \quad (7)$$

### B. Fuzzy Commitment Based Schemes

A fuzzy commitment scheme [7] is a realization of generic biometric system with key-binding designed for binary biometric sequences. In fuzzy commitment during the enrollment a uniformly selected binary secret key $S^K \in \{0,1\}^K$ is encoded into a codeword $C^N$ using an error-correcting code of rate $R$. To produce the helper data, the codeword is masked with the biometric enrollment sequence $\underline{X}^N \in \{0,1\}^N$, thus the encoding function becomes

$$M^N = enc(S^K) \oplus \underline{X}^N = C^N \oplus \underline{X}^N, \quad (8)$$

where $\oplus$ denotes a modulo-2 addition. During authentication the decoder observes the authentication sequence $\underline{Y}^N \in \{0,1\}^N$ and subtracts it modulo-2 from the helper data. The

ISBN 978-0-9928626-7-1 © EURASIP 2017

resulting codeword corrupted with biometric measurement noise, given by $\underline{X}^N \oplus \underline{Y}^N$, is decoded to the closest codeword and subsequently to the secret key estimate $\widehat{S}^K$. Thus the decoding function becomes

$$\widehat{S}^K = dec(M^N \oplus \underline{Y}^N) = dec(C^N \oplus (\underline{X}^N \oplus \underline{Y}^N)). \quad (9)$$

Achievable region for fuzzy commitment scheme in case of a (virtual) memoryless binary symmetric channel (BSC) between enrollment and authentication measurements and uniformly distributed $X$ was characterized in e.g. [13]:

$$\begin{aligned}\mathcal{R}_{FC} = \{(R_s, L_p) \quad : \quad & 0 \le R_s \le 1 - h(p), \\ & L_p \ge 1 - R_s, \\ & \text{for } 0 < p \le 1/2\}, \end{aligned} \quad (10)$$

where $p$ is a crossover probability of the BSC and $h(\cdot)$ is a binary entropy function: $h(p) \triangleq -p\log_2 p - (1-p)\log_2(1-p)$.

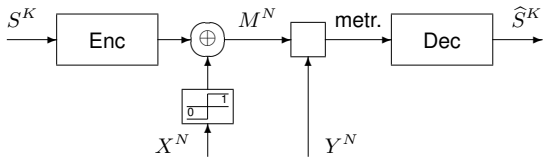


Fig. 2. Modified fuzzy commitment with quantization at the encoder only.

Performance of fuzzy commitment schemes is optimal for rates at the capacity, i.e., $R_s = 1 - h(p)$, thus requiring capacity achieving codes for their implementation. Note that, in general, biometric sequences are non-binary, and therefore, to apply fuzzy commitment, quantization has to be done at both encoder and decoder side. However, quantization at the decoder results in the performance loss, since decoding is done using hard decision. Therefore in this work we consider a modified version of fuzzy commitment, see Fig. 2, where quantization is only done at the encoder [11]. Moreover, we apply multi-level quantization at the encoder [14] in order to be close to the system secrecy capacity $I(X;Y)$.

## III. Coding for EEG Biometrics

In this section we present the building blocks for our EEG-based authentication system depicted in Fig. 2.

### A. Modeling EEG

Consider EEG data signals corresponding to two motor imaginary (MI) activities acquired using $C$ channels. The EEG measurements are filtered using common average reference (CAR) filter to remove average potential of all electrodes, and are further normalized, resulting into our raw EEG data.

In oder to obtain enrollment and authentication EEG sequences, we combine information from raw EEG data and EEG data transformed using common spatial pattern (CSP) filter. CSP is a method widely applied in motor imagery based brain-computer interface (BCI) systems [15]. The method transforms the multi-channel EEG data into a low-dimensional subspace, which maximizes the difference between two different MI activities. Since in biometric systems the goal is people recognition, we consider three different CSP transformations in order to obtain the most discriminating information, i.e.,

- PCSP: CSP generated based on the data of each subject separately. It relates to individual MI activities;
- GCSP: CSP generated based on the data coming from all subjects together. It relates to generic MI activities;
- SCSP: CSP defined as GCSP-PCSP. This transformation should capture individual variations in MI activities.

Both CSP-filtered and raw EEG data for each of $C$ channels are further modeled as autoregressive (AR) process of order $O$. The resulting feature vectors are represented by reflection coefficients (RC) estimated using Burg's method [16].

Note that the resulting feature vectors are typically high dimensional and correlated. Therefore, we apply independent component analysis (ICA) followed by Fisher discriminant analysis (FDA) to obtain biometric sequences composed of independent discriminating components.

### B. Multi-Level Quantization

Recall that fuzzy commitment requires uniform biometric input distribution to guarantee optimal performance. Therefore in this work we consider equiprobable quantization scheme, proposed in [14]. Consider a zero-mean unit-variance Gaussian distribution. The equiprobable quantizer $Q_J(\cdot)$ with quantization level $\log_2 J$ for this Gaussian distribution is given by

$$\begin{aligned} Q_J(x) = j, \qquad & \text{for } q_{j-1} < x \le q_j, \\ \text{where } q_j \text{ is such that } \quad & \textstyle\int_{-\infty}^{q_j} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{j}{J}. \end{aligned} \quad (11)$$

The quantization level $\log_2 J$ is selected such that the resulting recognition performance achieved with quantized data remains close to the one achieved with original data.

### C. Binary Encoding

After the quantization step, we obtain a quantization index that has to be further encoded into a binary string. We consider two types of encodings: Gray and Gray-unary encodings.

Gray encoding ensures that two consecutive integers are assigned to the codewords that differ by only one bit. In general, the Hamming distance between the codewords of the Gray code is not equivalent to the distance in the integer domain. Gray-unary encoding represents an integer $j, 0 \le j \le J$, as $J - j$ zeros followed by $j$ ones. Just like Gray encoding two adjacent codewords differ in one bit, but it does preserve the integer distance. Gray-unary encoding, however, introduces redundancy in the codewords.

### D. Error-Correction and Log-Likelihood Ratio

Consider now our modified fuzzy commitment scheme with soft input at the decoder. The encoder encodes a secret key $S^K$ using a turbo-code of rate $R$. Moreover, it observes biometric enrollment sequence $X^N$ and quantizes its each component $X_n, n = 1, 2, \ldots, N$ into a sequence $\underline{X}_n^{\mathcal{J}} = (X_{n,1}, X_{n,2}, \ldots, X_{n,\mathcal{J}})$, where $\mathcal{J}$ takes on a value of $\log_2 J$ or $J$ for Gray and Gray-unary encoding, respectively. The helper data is then formed as $C^{N\mathcal{J}} \oplus \underline{X}^{N\mathcal{J}}$. The decoder observes the helper data $M^{N\mathcal{J}}$ and the authentication sequence

ISBN 978-0-9928626-7-1 © EURASIP 2017

$Y^N$ and uses an BCJR decoding [17] to reconstruct $C^{N\mathcal{J}}$. The decoding step requires log-likelihood ratios (LLR).

Consider a Gray-unary encoding. Given helper data symbol $m_{n,i}$ and observed biometric symbol $y_n$, LLR is given by

$$LLR_{n,i} = \log_2 \frac{\Pr(m_{n,i}, y_n | c_{n,i} = 0)}{\Pr(m_{n,i}, y_n | c_{n,i} = 1)}. \quad (12)$$

Using Bayes rule and the fact that biometric observations are independent of the secrets and thus codewords, the probability $\Pr(m_{n,i}, y_n | c_{n,i} = 1)$ can be rewritten as

$$\begin{aligned}\Pr(m_{n,i}, y_n | c_{n,i} = 1) \\ &= \Pr(m_{n,i} | c_{n,i} = 1) \Pr(y_n | m_{n,i}, c_{n,i} = 1) \\ &= \Pr(\underline{x}_{n,i} \oplus 1 | c_{n,i} = 1) \Pr(y_n | \underline{x}_{n,i}) = \Pr(y_n, \underline{x}_{n,i}), \quad (13)\end{aligned}$$

and we need to consider probability of $\underline{x}_{n,i} = 0$, if $m_{n,i} = 1$ and $\underline{x}_{n,i} = 1$, if $m_{n,i} = 0$. Note that $\underline{x}_{n,i} = 0$ implies that $x_n \leq q_{J+1-i}$, where $q_{J+1-i}$ is a quantization boundary. Therefore, we can write

$$\begin{aligned}\Pr(x_{n,i} = 0 | y_n) &= \Pr(x_n \leq q_{J+1-i} | y_n) \\ &= \int_{-\infty}^{q_{J+1-i}} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(x-\rho y_n)^2}{2(1-\rho^2)}\right) dx \\ &= 1 - Q_f\left(\frac{q_{J+1-i} - \rho y_n}{\sqrt{1-\rho^2}}\right), \quad (14)\end{aligned}$$

where $Q_f(\cdot)$ is a Q-function. Combining (12)-(14) and analogous expressions for $\Pr(m_{n,i}, y_n | c_{n,i} = 0)$ and $\Pr(x_{n,i} = 1 | y_n)$, LLRs can be written as

$$LLR_{n,i} = (-1)^{m_{n,i}} \log_2 \left[\frac{1 - Q_f\left(\frac{q_{J+1-i} - \rho y_n}{\sqrt{1-\rho^2}}\right)}{Q_f\left(\frac{q_{J+1-i} - \rho y_n}{\sqrt{1-\rho^2}}\right)}\right]. \quad (15)$$

For the Gray encoding, see [14], LLR are given by

$$\begin{aligned}LLR_{n,i} &= (-1)^{m_{n,i}} \\ &\cdot \log_2 \left[\frac{1 - \sum_{j=1}^{2^{i-1}} (-1)^{j+1} Q_f\left(\frac{q_{(2j-1)2^{J-i}} - \rho y_n}{\sqrt{1-\rho^2}}\right)}{\sum_{j=1}^{2^{i-1}} (-1)^{j+1} Q_f\left(\frac{q_{(2j-1)2^{J-i}} - \rho y_n}{\sqrt{1-\rho^2}}\right)}\right]. \quad (16)\end{aligned}$$

## IV. Experimental Results

### A. Experimental Set-Up

EEG data was collected using imec's wireless EEG headset with 4 active dry EEG channels. The headset comprised 6 reusable silver-silver chloride (Ag/AgCl) electrodes (BIOPAC EL120) with 4 measurement electrodes placed at positions Cz, C3, C4 and Pz of the International 10-20 electrode positioning system, and ground and reference electrodes positioned at left and right mastoid, respectively. In-house developed EEG data acquisition software with a tailor made application that generate visual cues was used in the data collection.

The data acquisition protocol was approved by internal review committee at imec. Ten participants (all males; mean age: 30), recruited for the study, were healthy and had no history of neurological disorders. To avoid variability in primary cortex activation patterns, selected participants had right side laterality. Participants were seated comfortably, facing the computer screen about 1 meter away with both arms resting on the table. All the experiments were performed under laboratory conditions at constant room temperature and light ambiance.

During the user trials, EEG was recorded during left arm, right arm, right foot and tongue movements. In our experiments we only used the motor imaginary EEG recordings corresponding to left and right arm movements. Before the measurements were initiated, the participants were asked to hold down the left 'CONTROL' key (using the left arm index finger) and the number '0' key on the number-pad (using the right arm index finger). The measurements were divided into 4 sessions ($S_1 - S_4$) with 100 trials each. Sessions $S_1$ and $S_2$ were recorded on the same day; $S_3$ was recorded one to three days after $S_1$ and $S_2$; and $S_4$ was recorded one to three days after $S_3$. The total period of recording process covered about one week. Four different cues were generated in random order. Each trial was divided into four segments: the cue phase, the motor imagery phase, relax phase and activity phase. The cue phase lasted 3 seconds, following the instruction. At the end of cue phase, each participant was asked to imagine the movement, marking the initiation of the motor imagery phase. No physical movement was performed during this phase of the trial. The motor imagery phase lasted 3 seconds followed by a short period of relaxation. The final phase (activity phase) was marked by a command asking the participant to perform the activity as instructed during the cue phase of the trial. This phase lasted 3 seconds, the end of which marked the end of one trial. Each trial lasted 11 seconds.

### B. EEG Feature Vector Parameter Selection

In our experiments we selected SCSP filter and band pass filter to select [8,50] Hz of EEG rythm, see Fig. 3. We also used 18 RC in AR process per each of 4 channels. The number of coefficients was selected based on Akaike information criterion (AIC). Moreover, 150 ICA components were used as input to FDA. The resulting EEG biometric sequences, used as input in the privacy-preserving system, had 9 components (limited by the number of participants). The resulting performance of EEG based authentication system without privacy protection is shown in Fig. 4. The achieved average EER was 0.93% calculated using 4-fold cross-validation.

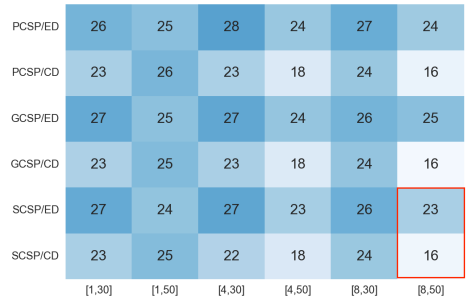


Fig. 3. RC performance (EER) for different CSP filters and frequency ranges, based on cosine (CD) and Eucledian distance (ED).

ISBN 978-0-9928626-7-1 © EURASIP 2017

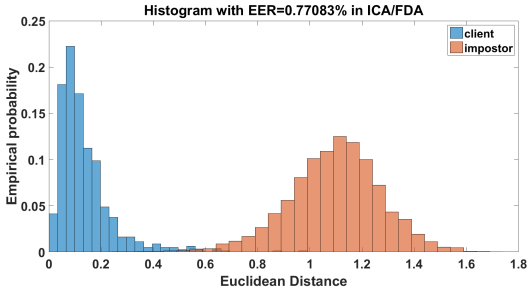


Fig. 4. Distribution of the Eucleadian distance for genuine and imposter comparisons in the unprotected system.

### C. Experimental Results

The EEG sequences used as input into the privacy-preserving system have only 9 components. We could have used 150 ICA components, however, ICA components have low correlation, viz. $\rho = 0.2$, resulting in enrollment binary sequences with too high error-rates. The components after ICA/FDA transformation, on the other hand, have high correlation, viz. $\rho = 0.85$, corresponding to SNR$=$ 4dB, allowing to extract 9 secret bits with negligible error probability.

Fig. 5 shows the simulation results for the EEG system with Gray encoding, where the performance is evaluated using word error rate (WER), an equivalent of FRR. The simulations are done for different source quality (SNR) and quantization levels. From these results, we see that only 4 bits secret keys can be achieved with FRR=2.4% and FAR=87% using the code rate of $1/5$, corresponding to the privacy leakage of $0.8$ bits per source symbol. Clearly we have two problems: on one hand, the source quality is relatively low here and, on the other hand, the length of the EEG sequences is too low to use more complex codes or lower code rates.

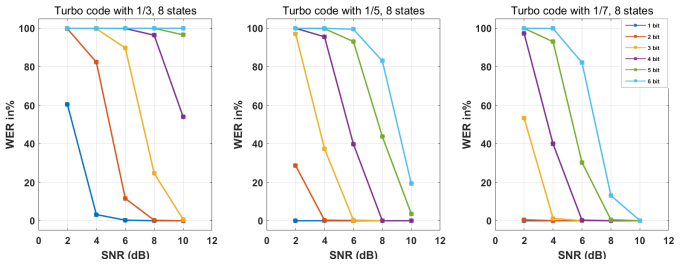


Fig. 5. Simulation results for the EEG-based authentication system with privacy protection implemented using Gray encoding.

To overcome these problems, we use Gray-unary encoding. This allows us to increase quantization level, which in turn increases the error rate. However, since Gray-unary encoding introduces redundancy, the overall error-rate of the resulting EEG sequences does not increase. Redundancy, however, has an effect that the secrecy leakage, $L_s$, becomes non-zero [13], and thus the effective key size reduces. The performance results for this setting are summarized in Table I. Here we used a turbo code with 64 trellis states. We see that here we can achieve FAR$=$ $1.83\%$ and FRR=$1.875\%$ with effective key size of 21 bits and privacy leakage of $0.116$ per source symbol. Note that the secret key size reported there is higher than secrecy capacity. This can be explained by the observation that the reported FRR is not negligible. Nevertheless, key size estimates might still be optimistic, since privacy and secrecy leakage estimates, see [13], are lower bounds.

TABLE I
PERFORMANCE OF THE EEG-BASED AUTHENTICATION SYSTEM WITH PRIVACY PROTECTION IMPLEMENTED USING GRAY-UNARY ENCODING.

| Key (bits) | FAR (%) | FRR (%) | $J$ | $R_s$ | $L_p$ | $L_s$ | $H(S^N\|M^N)$ (bits) |
|---|---|---|---|---|---|---|---|
| 40 | 4.57 | 1.875 | 28 | 1/5 | 0.113 | 0.134 | 13 |
| 55 | 1.93 | 2.13 | 36 | 1/5 | 0.090 | 0.143 | 16 |
| 70 | 1.76 | 2.13 | 44 | 1/5 | 0.076 | 0.148 | 18 |
| 80 | 3.83 | 1.5 | 70 | 1/7 | 0.057 | 0.111 | 18 |
| 110 | 1.83 | 1.875 | 95 | 1/7 | 0.043 | 0.116 | 21 |
| 140 | 2.458 | 2.375 | 128 | 1/7 | 0.033 | 0.120 | 22 |

## V. CONCLUSIONS

In this paper we studied biometric authentication with privacy protection based on EEG patterns, related to motor imagery activities. We considered a realistic scenario where EEG recording for 10 healthy subjects were performed with a wireless device developed at imec, and data was collected in four sessions separated in time. We demonstrated that we could achieve recognition performance of around EER=$1.87\%$ with effective key size of 21 bits and privacy leakage of $0.116$.

## REFERENCES

[1] R. Paranjape and et al., "The electroencephalogram as a biometric," in *Can. Conf. on Elec. and Comp. Eng.*, 2001.
[2] R. Palaniappan and D. P. Mandic, "Biometrics from brain electrical activity: A machine learning approach," in *IEEE Trans. on Patt. Analys. and Mach. Intellig.*, 2007.
[3] S. Marcel and J. d. R. Millán, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptiation," in *IEEE Trans. on Patt. Analys. and Mach. Intellig.*, 2007.
[4] K. Brigham and B. V. Kumar, "Subject identification from electroencephalogram (EEG) signals during imagined speech," in *Biometric Theory App. and Syst. (BTAS)*, 2010.
[5] M. K. Abdullah and et al., "Analysis of the EEG signal for a practical biometric system," *Int. Jour. of Med., Health, Biomed., Bioeng. and Pharm. Eng.*, 2010.
[6] E. Maiorana, D. La Rocca, and P. Campisi, "EEG-based biometric recognition using EigenBrain," in *Multimed. and Expo Work.*, 2015.
[7] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *6th ACM Conf. on Comp. and Comm. Sec.*, 1999.
[8] T. Ignatenko and F. Willems, "Biometric systems: Privacy and secrecy aspects," in *IEEE Trans. on Inf. For. and Sec.*, 2009.
[9] E. Maiorana, D. Blasi, and P. Campisi, "Biometric template protection using turbo codes and modulation constellations," *IEEE Work. on Inf. Forens.and Sec. (WIFS)*, 2012.
[10] E. Maiorana, D. La Rocca, and P. Campisi, "Cognitive biometric cryptosystems a case study on EEG," in *Int. Conf. on Sys., Sig. and Image Proc. (IWSSIP)*, 10-12 Sept. 2015.
[11] T. Ignatenko and F. Willems, "Privacy-leakage codes for biometric authentication systems," in *IEEE Int. Conf. on Acous., Speech and Sig.Proc. (ICASSP)*, 2014.
[12] F. Willems and T. Ignatenko, "Quantization effects in biometric systems," in *Inf. Theory and its App. (ITA), San Diego, USA*, 2009.
[13] T. Ignatenko and F. Willems, "Information leakage in fuzzy commitment schemes," *IEEE Trans. on Inf. Forens. and Sec.*, vol. 5, no. 2, 2010.
[14] C. Ye, A. Reznik, and Y. Shah, "Extracting secrecy from jointly gaussian random variables," in *IEEE Int.Symp. on Inf. Theory*, 2006.
[15] H. Lu and et al., "Regulairzed common spatial pattern with aggregation for EEG classification in small-sample setting," in *IEEE Trans. on Biomed. Eng.*, 2010.
[16] M. De Hoon and et al., "Why yule-walker should not be used for autoregressive modelling," in *Ann. of Nuc. Energy, Vol. 23, Iss. 15*, 1996.
[17] S. Lin and D. J. Costello, *Error Control Coding, 2nd Ed. Pearson*, 2004.

ISBN 978-0-9928626-7-1 © EURASIP 2017