2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Multi-factor EEG-based User Authentication

Tien Pham, Wanli Ma, Dat Tran, Phuoc Nguyen, and Dinh Phung

***Abstract*—Electroencephalography (EEG) signal has been used widely in health and medical fields. It is also used in brain-computer interface (BCI) systems for humans to continuously control mobile robots and wheelchairs. Recently, the research communities successfully explore the potential of using EEG as a new type of biometrics in user authentication. EEG-based user authentication systems have the combined advantages of both password-based and biometric-based authentication systems, yet without their drawbacks. In this paper, we propose to take the advantage of rich information, such as age and gender, carried by EEG signals for user authentication in multi-level security systems. Our experiments showed very promising results for the proposed multi-factor EEG-based authentication method.**

I. Introduction

Authentication plays a very important role in security systems; however, the current methods, such as password-based, token-based, and biometrics-based authentication, have been exposed their own security weaknesses. Password-based authentication is not immune from malicious attacks such as offline dictionary attack, popular password attack, exploiting user mistakes, and exploiting multiple password use [9]. Token-based authentication requires users always bringing and providing tokens when accessing the systems. Presenting a token, which is not a part of a human body, can cause inconvenient. Moreover, all the tokens require special readers and tokens can be physically stolen, be duplicated, as well as be hacked [7] [9]. Although biometric authentication can avoid some disadvantages of password-based and token-based authentication systems, the biometrics modalities have some drawbacks. Face, fingerprint, and iris information can be photographed. Voice could be recorded, and handwriting may be mimicked [6] [23]. Individuals can also be lost or changed their biometric characteristics such as finger or face. These disadvantages of the 3 current authentication methods require a better modality for security systems.

Tien Pham is with Faculty of Education, Science, Technology and Mathematics, University of Canberra, Australia (e-mail: Tien.Pham@Canberra.edu.au).

Wanli Ma is with Faculty of Education, Science, Technology and Mathematics, University of Canberra, Australia. He is also with Department of Computer Science, University of Houston Downtown, USA (e-mail:Wanli.Ma@Canberra.edu.au)

Dat Tran is with Faculty of Education, Science, Technology and Mathematics, University of Canberra, Australia. (e-mail:Dat.Tran@Canberra.edu.au)

Phuoc Nguyen is with Faculty of Education, Science, Technology and Mathematics, University of Canberra, Australia (e-mail: Phuoc.Nguyen@Canberra.edu.au)

Dinh Phung is with Faculty of Education, Science, Technology and Mathematics, University of Canberra, Australia (e-mail: Dinh.Phung@Canberra.edu.au)

Recently, human electroencephalography (EEG) signals emerge as a potential biometric modality with the advantages of being difficult (close to impossible) to fake, impossible to observe or intercept, unique, un-intrusive, and alive person recording is required [19] [21].

EEG signals were discovered in early 1900s, and they have been playing an important role in health and medical applications. Epileptic seizure detection is one of the most well-known applications. Another common usage of EEG signal in health is the study of sleep disorders. In additional, the relations between EEG signals and brain diseases have been investigated. Recording EEG signals is non-invasive with a portable device, so EEG is also widely used in Brain Computer Interface (BCI) which can provide a link between the human subject and the computer without physical contact [22].

Varieties of feature extraction methods and machine learning models have been applied to extract representing information from EEG signals for person identification and verification. Manhattan distances on Auto Regression (AR) coefficients with PCA were used to compute thresholds to determine test patterns of clients or impostors in 2 stages [17]. Fisher's Linear Discriminant (FLD) was first employed in [3] to reduce the dimensions of AR and power spectrum density (PSD) feature vectors. In [9], the authors tried to analyse EEG signals for person authentication based on an ARMA (Auto-Regressive and Moving Average) model while three sets of features are extracted 6th order autoregressive (AR) coefficients, power spectral density, and total power were applied in [5]. Multi-sphere Support vector data description (MSSVDD) is used in [13]. In addition, MSSVDD is also used with universal background model (UBM) in [14]

Most of the current studies only focus on analyzing users' characteristics or "EEG password" while other factors, such as gender and age information of the person, have not been investigated and used to improve the accuracy and strengthen the security of the system. In addition, an authentication system may need different credentials for different levels of security depending on zones and resources.

According to a comprehensive study on cybercrime in 2013 of United Nations [20], cybercrime perpetrators are overwhelmingly male. This study also confirms that cybercrime perpetrators are most commonly aged between 18 and 30 years. In addition, in [15] the authors had good result in age and gender classification rates. In this paper, we propose to take the advantage of rich information, such as age and gender, carried by EEG signals for user authentication in multi-level security systems where users are asked to provide EEG signals and then not only users' characteristics but also users' gender and age information are used to authenticate

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

that user in the first, the second, and the third layer, respectively.

The rest of the paper is organized as follows. In Section 2, we study the using EEG for multi-factor user authentication. Section 3 describes EEG features. Section 4 describes Small Sphere Two Large Margins Support Vector Data Description (SS2LM-SVDD) modelling technique and summarizes hypothesis testing. Experiments and results are presented in Section 5. We conclude the paper with a discussion and our future work in Section 6.

II. Multi-factor EEG-Based User Authentication System

While the types of authentication have their own shortcomings as discussed previously, EEG emerges as a potential modality for authentication, because of the following advantages, yet without shortcomings of the conventional types:

1. EEG is confidential because it corresponds to a secret mental task which cannot be observed;
2. EEG signals are very difficult to mimic because the signals of similar mental tasks are person dependent;
3. It is almost impossible to steal because the brain activity is sensitive to the stress and the mood of the person. An aggressor cannot force the person to reproduce the same signals while he or she is under stress [22];
4. EEG signals, by nature, require alive person to produce [4].

Moreover, other useful information can be extracted from EEG signal, for example, age and gender [15-16]. However, using gender and age information extracted from EEG signals to improve the performance of EEG based authentication has not been studied. We propose a multi-factor authentication system using EEG signals to take the advantage of rich information, such as age and gender, carried by EEG signals as illustrated in Fig.1, Fig.2, and Fig.3 below.

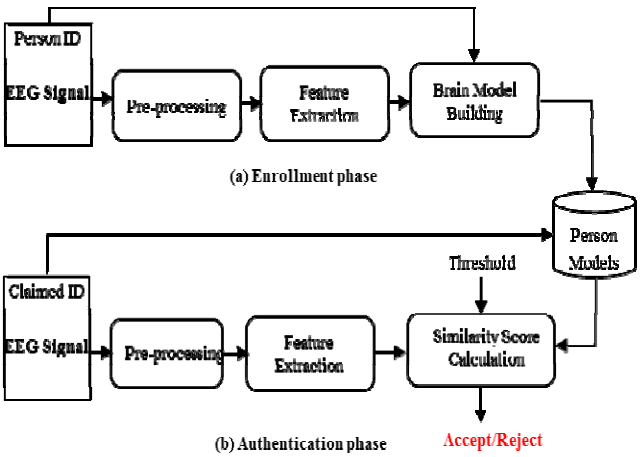

Fig.1. Typical EEG-based user authentication diagram

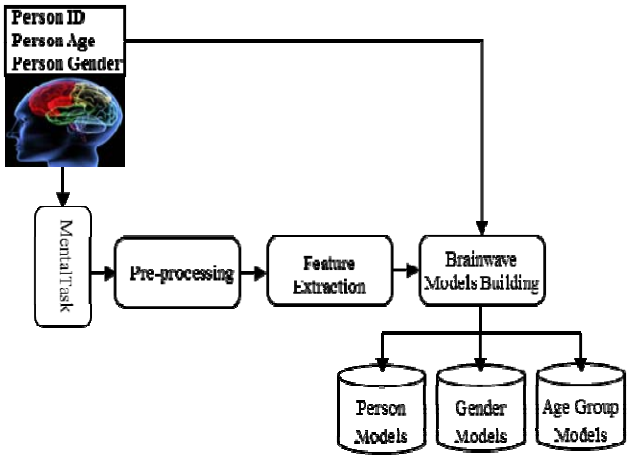

Fig.2. Enrolment phase of the proposed multi-factor EEG-based user authentication system

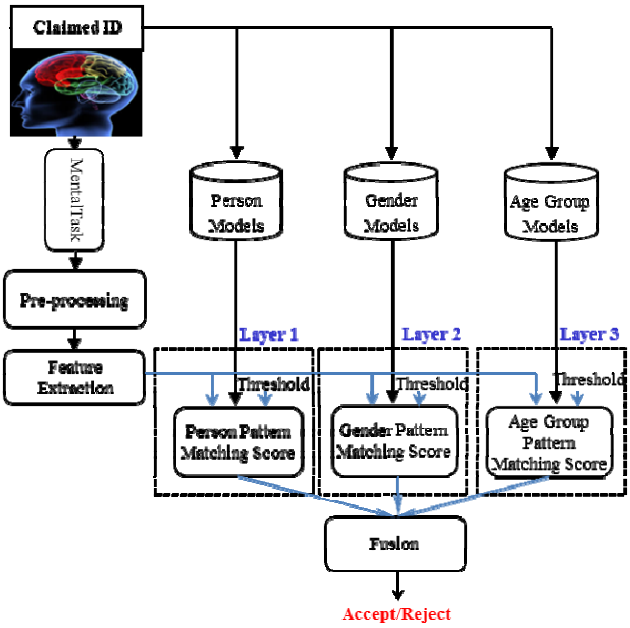

Fig.3. Verification phase of the proposed multi-factor EEG-based user authentication system

We regard the system as authentication by *something a user thinks*. An EEG based authentication system has two phases: enrolment and verification. In the enrolment phase, a user firstly has to provide his or her age and gender information. After that, the user is asked to do some tasks, for example, imagining moving a hand, a foot, a finger or just relax, and EEG signals associated with the tasks are recorded. For authentication purposes, the mental tasks themselves are also a part of the credentials and could not be seen by any third party. After collecting the data, the EEG signals of the tasks corresponding to the user are pre-processed, and the representing features are extracted. These features are used to train the person model, gender model, and age group model for this person, which are securely kept in a database.

In the verification phase, when a user wants to access a security system, he or she has to provide EEG signal, which

carries gender and age group information, by repeating the tasks which he or she did in the enrolment phase. The input EEG data is processed in the same way as seen in the enrolment phase. The obtained features are then fed into the classifier as testing data. Firstly, testing data is calculated for matching scores to the person model of the person who he or she claims to be. After that, authentication system continues to calculate matching scores of testing data to the gender model of the individual who he or she claims to be in the second layer. Finally, testing data is calculated for matching scores to the age group model of the claimed user in the third layer. These results are fused at the decision level. A user is verified when his or her testing data is matched in all the layers of the security system (AND case).

III. EEG FEATURES

The spectral power in 2 Hz frequency bins from 1 to 30 Hz was computed for each channel. The central frequency of each bin was an integer. The relative power, which is the owner in a specific frequency, divided by the total power in all frequency bins from 1 to 30 Hz, together with the total power were also used. In addition, 11 AR coefficients of the 11th-order AR model and 3 Hjorth parameters (activity, mobility and complexity) were extracted for each electrode.

A. Autoregressive (AR) features

Autoregressive model can be used for a single-channel EEG signal. It is a simple linear prediction formula that best describes the signal generation system. Each sample $s(n)$ in an AR model is considered to be linearly related with respect to a number of its previous samples [22]:

$$s(n) = -\sum_{k=1}^{p} a_k s(n-l) + x(n) \qquad (1)$$

where a_k, $k = 1, 2, \ldots, p$ are the linear parameters, n denotes the discrete sample time, and $x(n)$ is the noise input. The linear parameters of different EEG channel were taken as the features.

B. Power spectral density (PSD) features

Power spectral density (PSD) of 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 (ACS)

$$\emptyset(\omega) = \sum_{k=-\infty}^{\infty} r(k) e^{-i\omega k} \qquad (2)$$

where the auto covariance sequence $r(k)$ is defined as

$$r(k) = E\{s(t)s^*(t-k)\} \qquad (3)$$

and $s(t)$ is the discrete time signal $\{s(t); t = 0, \pm 1, \pm 2, \ldots\}$ assumed to be a sequence of random variables with zero mean.

In this paper, the Welch's method using periodogram is used for estimating the power of a signal at different frequencies. The Welch's method can reduce noise but also reduce the frequency resolution as compare to the standard Bartlett's method, which is desirable for this experiment.

IV. MODELLING TECHNIQUE

A. Small Sphere Two Large Margins Support Vector Data Description (SS2LM-SVDD)

The Small Sphere Two Large Margins Support Vector Data Description (SS2LM-SVDD) [24] aims to construct an optimal hyper-sphere in feature space to include EEG data of a person and exclude the EEG data of other persons. Two margins will be determined to reduce both false acceptance and false rejection rates at the same time. These two margins are proportional to the adjustable parameter δ . This parameter depends on the proportion of EEG data of the claimed person and EEG data of other persons, i.e. impostors. Consider the training set $x_1, x_2, \ldots, x_{m_1+m_2}$ where first m_1 data points are labeled +1 and belong to the claimed person and the remaining m_2 data points are labeled –1 and belong to impostors. Let us also denote the label of data point x_i by $y_i \left(i = 1..\overline{m_1 + m_2} \right)$. It is obviously that $y_i = 1$ ($i = \overline{1..m_1}$) and $y_i = -1$ ($i = \overline{m_1 + 1..s}$) where $s = m_1 + m_2$.

The optimization problem is as follows

$$\min_{R,c,\xi,\rho} R^2 - \nu\rho^2 + \frac{1}{\nu_1 m_1}\sum_{i=1}^{m_1}\xi_i + \frac{1}{\nu_2 m_2}\sum_{i=m_1+1}^{s}\xi_i \qquad (4)$$

$$\text{s.t. } \left\|\phi(x_i) - c\right\|^2 \le R^2 - \delta\rho^2 + \xi_i, i = \overline{1..m_1} \qquad (5)$$

$$\left\|\phi(x_i) - c\right\|^2 \ge R^2 + \rho^2 - \xi_i, i = \overline{m_1 + 1..s} \qquad (6)$$

$$\xi_i \ge 0, i = \overline{1..s}$$

where R and c are radius and center of the optimal hyper sphere, respectively, $\xi = [\xi_1, \xi_2, \ldots, \xi_s]^T \in R^s$ are slack variables, ρ is outside margin (distance from abnormal data to decision boundary), $\delta\,(0 \le \delta \le \nu)$ is the ratio between outside margin and inside margin. Hence $\delta\rho$ can be considered as inside margin (distance from normal data to decision boundary).

Through doing minimization of the objective function in (4), the minimal radius R and two maximal margins ρ and $\delta\rho$ will be determined simultaneously. Combining (5) and (6) gives

$$y_i \left\|\phi(x_i) - c\right\|^2 \le y_i R^2 - z_i\rho^2 + \xi_i, i = \overline{1..s} \qquad (7)$$

$$\text{where } z_i = \frac{1}{2}\left[(1 - y_i) + (1 + y_i)\delta\right], i = \overline{1..s} \qquad (8)$$

To derive the solution of the above optimization problem, the following Lagrange function is introduced

$$L(R, c, \rho, \xi, \alpha, \beta) = R^2 - \nu\rho^2 + \frac{1}{\nu_1 m_1}\sum_{i=1}^{m_1}\xi_i + \frac{1}{\nu_2 m_2}\sum_{i=m_1+1}^{s}\xi_i$$

$$+ \sum_{i=1}^{s}\alpha_i\left(y_i\left\|\phi(x_i) - c\right\|^2 - y_i R^2 + z_i\rho^2 - \xi_i\right) - \sum_{i=1}^{s}\beta_i\xi_i \qquad (9)$$

Maximizing this function gives the dual form:

$$\min_{\alpha} \sum_{i=1}^{s}\sum_{j=1}^{s}\alpha_i\alpha_j y_i y_j K(x_i,x_j) - \sum_{i=1}^{s}\alpha_i y_i K(x_i,x_i)$$
$$\text{s.t. } \sum_{i=1}^{s}\alpha_i y_i = 1; \sum_{i=1}^{s}\alpha_i = \frac{\nu+1-\delta}{\delta+1}; 0 \le \alpha_i \le \frac{1}{\nu_1 m_1},$$
$$\overline{i=1..m_1}; 0 \le \alpha_i \le \frac{1}{\nu_2 m_2}, i = \overline{m_1+1..s} \quad (10)$$

where $K(x, x')$ is a kernel function.

The quadratic optimization problem (10) shows that the SS2LM-SVDD model can be solved using $\nu - SVM$ package in LIBSVM.

The radius R and ρ are calculated as follows

$$R^2 = \frac{1}{n_1}P_1 \text{ and } \rho^2 = \frac{1}{n_2}P_2 - \frac{1}{n_1}P_1 \quad (11)$$

where $n_1 = |SV_p|$ and $n_2 = |SV_n|$

$$SV_p = \left\{ i : 1 \le i \le m_1 \text{ and } 0<\alpha_i < \frac{1}{\nu_1 m_1} \right\}$$
$$SV_n = \left\{ i : m_1 < i \le s \text{ and } 0<\alpha_i < \frac{1}{\nu_2 m_2} \right\} \quad (12)$$

P_1 and P_2 can be computed as:

$$P_1 = \sum_{i\in SV_p}\|\phi(x_i)-c\|^2 = \sum_{i\in SV_p}\left(K(x_i,x_i) + \|c\|^2 - 2\sum_{k=1}^{s} y_k\alpha_k K(x_k,x_i)\right)$$
$$P_2 = \sum_{i\in SV_n}\|\phi(x_i)-c\|^2 = \sum_{i\in SV_n}\left(K(x_i,x_i) + \|c\|^2 - 2\sum_{k=1}^{s} y_k\alpha_k K(x_k,x_i)\right)$$
$$\|c\|^2 = \sum_{i=1}^{s}\sum_{j=1}^{s} y_i y_j \alpha_i\alpha_j K(x_i,x_j) \quad (13)$$

For a new pattern x for authentication we calculate the distance between $\Phi(x)$ and center c of the hyper-sphere and then classify x as claimed data if this distance is less than radius R and as impostor data if otherwise. The decision function is of the following form:

$$f(x) = sign\left(R^2 - \|\phi(x)-c\|^2\right)$$
$$= sign\left(R^2 - \|c\|^2 - K(x,x) + 2\sum_{i=1}^{s}\alpha_i y_i K(x_i,x)\right) \quad (14)$$

B. Hypothesis testing

The verification task can be stated as a hypothesis testing between the two hypotheses: the input is from the hypothesis person (H0), or not from the hypothesis person (H1).

Let λ_0 be the claimed person model and λ_1 be a model representing all other possible people, i.e. impostors. For a given input x and a claimed identity, the choice is between the hypothesis H0: x is from the claimed person λ_0, and the alternative hypothesis H1: x is from the impostors λ_1. A claimed persons' score $L(x)$ is computed to reject or accept the person claim satisfying the following rules

$$L(x) = \begin{cases} \ge \theta_L \text{ accept} \\ < \theta_L \text{reject} \end{cases} \quad (15)$$

where θ_L are the decision threshold.

Let x be an EEG feature vector, the probability of x belonging to the class y is defined as $P(x|\theta_y) = ce^{yf(x)}$ where c is normalization factor and $f(x)$ is from (14).

If $x_1,..,x_k$ is a sequence of independent identical density (iid) feature vectors of class y, the probability of $x_1,..,x_k$ belonging to the class y in the AND case is:

$$P(x_1,\dots,x_k|\theta_y) = \prod_{i=1}^{k} ce^{yf(x_i)} = c'e^{\sum_{i=1}^{k} f(x_i)} \quad (16)$$

Then the score $L(x)$ in (15) for SS2LM-SVDD will become

$$L_{AND}(x) = P(x_1,\dots,x_k|\theta_y) = c'e^{\sum_{i=1}^{k} f(x_i)} \quad (17)$$

$$L'_{AND}(x) = \sum_{i=1}^{k} f(x_i) \quad (18)$$

V. EXPERIMENT AND RESULTS

A. Data set

The dataset used in this research is Australian EEG Database, which consists of EEG recordings of 40 patients at the John Hunter Hospital [10], over an 11-years period. There are 20 males and 20 females and their ages are between 19 and 69. The EEGs were recorded using 23 electrodes followed the standard International System 10-20 electrode placements. The recordings were sampled at 167 Hz for about 20 minutes in the resting state with eyes open and eyes closed.

TABLE I
DATA DESCRIPTION

Dataset	#persons	#tasks	#trials	#sessions	Length(s)
Australian EEG	40	free	1	1	1200

B. Feature extraction

In the Australian EEG dataset, epochs of 15 seconds were split for training and testing from 8 channels F3, F4, C3, C4, P3, P4, O1, and O2 on the frontal, central, parietal, and occipital locations. The channel signal in each epoch were used to extract features, and these features were merged together to make a single feature vector. Choosing those electrodes is suggested by [15], where the authors had good result in age and gender classification rates.

The autoregressive (AR) linear parameters and power spectral density (PSD) components from these signals were extracted as features. In details, the power spectral density (PSD) in 2 Hz frequency bins from 1 to 30 Hz was estimated. The Welch's averaged modified periodogram method [18] was used for spectral estimation.

In AR model, each sample is considered linearly related with a number of its previous samples. The AR model has the

advantage of low complexity and has been used for person identification and authentication [12]. Burg's lattice-based method was used with the AR model of order 11^{th}. In addition, relative powers in different frequency bands and Hjorth parameters were also extracted.

Cybercrime perpetrators are overwhelmingly male, and they are most commonly aged between 18 and 30 years [20]. In addition, in [15] the authors had good age group classification rate when they used age range from 19 to 34 for the class of young people. Therefore, in this experiment the feature vectors were labelled for 2 age groups and 2 gender groups which are male, female, young, and older. The young age range is 19-34, and the older age range is above 34. There are 20 subjects in each male and female class while young class has 12 subjects and older class has 28 people.

C. *Results*

The SS2LM-SVDD method was used to train person EEG models and gender models. Experiments were conducted using 5-fold cross validation training, and the best parameters found were used to train models on the whole training set and test on a separate test set. The RBF kernel function $K(x_i, x_j) = e^{-\gamma\|x_i - x_j\|^2}$ was used. The parameter γ was searched in $\{2^k: k = -4, -3, \ldots, 1\}$. The parameter δ will be searched in grid $\{0.1k\nu : k = 0..10\}$, ν is in $\{0.1k, 0.01k\}$ where k is an integer number ranging from 1 to 9. The best parameter is $\gamma = 0.8$ for each person in the Australian EEG dataset.

Tables II and III present the confusion matrices of gender classification and age group classification respectively of the feature sets from the Australian EEG dataset. These matrices were calculated from the confusion matrix of 2-class classification experiment in the test phase by summing the predictions over the desired classes. These gender and age group classification rates are with decision threshold=0.

TABLE II
CONFUSION MATRIX OF GENDER CLASSIFICATION IN TEST PHASE IN AUSTRALIA EEG DATASET

Classified as →	Male	Female	Accuracy
Male	745	11	97.1%
Female	26	804	

TABLE III
CONFUSION MATRIX OF AGE GROUP CLASSIFICATION IN TEST PHASE IN AUSTRALIA EEG DATASET

Classified as →	Young	Older	Accuracy
Young	512	5	96.7%
Older	18	1051	

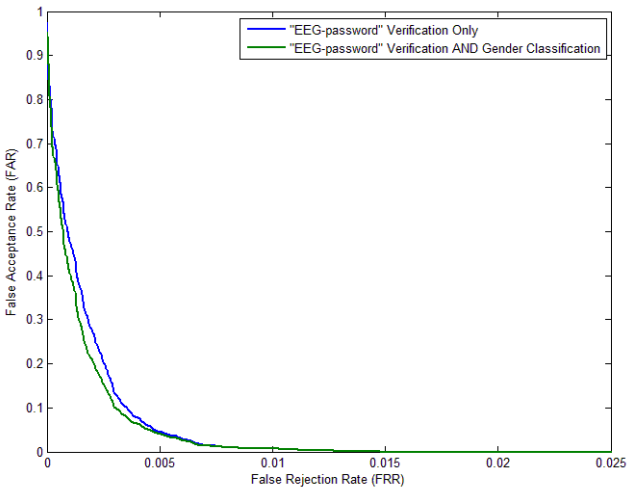

Fig.4. DET curves of single and multil-factor EEG-based user authentication using users‘ characteristic and gender information

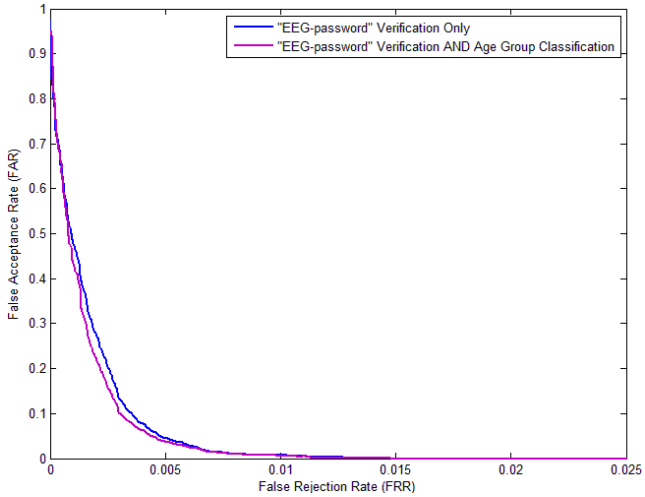

Fig.5. DET curves of single and multil-factor EEG-based user authentication using users‘ characteristic and age group information

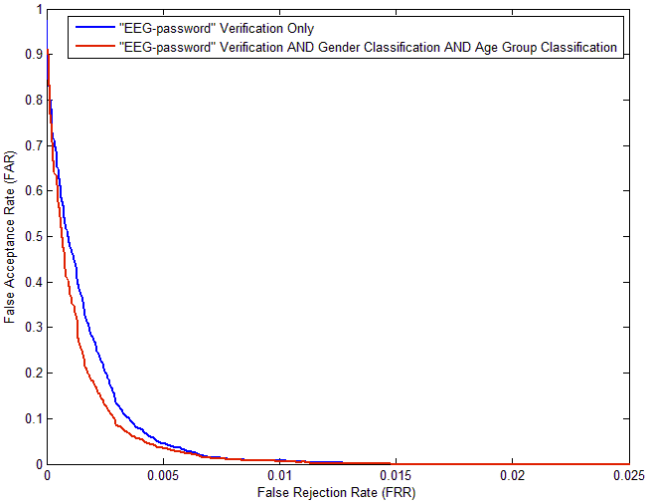

Fig.6. DET curves of single and multi-factor EEG-based user authentication using users‘ characteristic, gender, and age group information

Fig. 4, Fig.5, and Fig.6 illustrate the False Rejection Rate (FFR) and False Acceptance Rate (FAR) when using single and different multi-factor authentication method by using combination of EEG-based users' characteristic, gender and age group information. A DET curve, which is a plot of FAR on y-axis versus FRR on x-axis, is considered as a means of representing performance on detection tasks that involve a trade-off of error types [1]. Therefore, the above DET curves confirm that errors are significantly reduced when EEG-based multi-factor are used for user authentication instead of single factor. As a result, it is much more difficult for an imposter to access system when multiple matched EEG-based authentication policy is applied. It is also seen that the number of EEG-based factors used for authentication provides different accuracy; therefore, it can be adjusted flexibly for different level security of multiple level security systems, depending on zones and resources. To sum up, multi-factor EEG-based authentication is useful and suitable for security systems.

VI. Discussion and Future Work

Using EEG signals for authentication has the advantages of both password based and biometric based authentication approaches, yet without their drawbacks. Firstly, EEG signals are biometric information of individuals. Secondly, brain patterns correspond to particular tasks, and they be regarded as individualized passwords. As the result, EEG based authentication can overcome the disadvantages of password based and conventional biometric based authentication.

In addition, EEG signals carry rich personal information, such as gender and age etc., which can be exploited to implement a multilevel security system. EEG based authentication provides multilevel security systems by using an improved authentication mechanism with mental tasks, age, and gender information combination.

In the near future, we will experiment our proposed method EEG based authentication on other large datasets. Other useful information which can be extracted from EEG signal will also be investigated to enhance the EEG based person authentication system.

References

[1] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. "The det curve in assessment of detection task performance." *Technical report, DTIC Document*, 1997.

[2] A. Vallabhaneni, T. Wang, and B. He, "Brain—Computer Interface," *Neural Engineering*, pp.85–121, 2005.

[3] A. Yazdani, A. Roodaki, S. Rezatofighi, K. Misaghian, and S.K. Setarehdan, "Fisher linear discriminant based person identification using visual evoked potentials," *ICSP 2008*. pp.1677–1680, 2008.

[4] B. Allison, "Trends in BCI research: progress today, backlash tomorrow?," *The ACM Magazine for Students*, pp. 18-22, 2011.

[5] C. Ashby, A. Bhatia, F. Tenore, and J. Vogelstein, "Low-cost electroencephalogram (EEG) based authentication," *Neural Engineering (NER)*, pp. 442–445, 2011.

[6] C.He, H. Chen, and Z.Wang, "Hashing the MAR Coefficients From EEG Data For Person Authentication," *ICASSP 2009. IEEE International Conference on,* pp. 1445 – 1448, 2009.

[7] C. Rathgeb, A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," *EURASIP Journal on Information Security*, 2011.

[8] J. Hu, "Biometric System based on EEG Signals by feature combination," *ICMTMA*, pp. 752 – 755, 2010.

[9] L. Brown, *Computer Security: Principles and Practice*, William Stallings, 2008.

[10] M. Hunter, R. Smith, W. Hyslop, O. Rosso, R. Gerlach, J. Rostas, D. Williams, and F. Henskens, "The australian eeg database," *Clinical EEG and neuroscience*, vol. 36, no. 2, pp. 76–81, 2005.

[11] M. Poulos, M. Rangoussi, and N. Alexandris, "Neural network based person identification using EEG features," *ICASSP'99.Proceedings, 1999 IEEE International Conference On*, pp. 1117–1120, 1999.

[12] M. Poulos, M. Rangoussi, N. Alexandris, and Evangelou, "A person identification from the EEG using nonlinear signal classification," *Methods of information in medicine 41*, vol. 41, pp. 64–75, 2002.

[13] P. Nguyen, D. Tran, T. Le, and T. Hoang, "Multi-sphere support vector data description for brain-computer interface," *Communications and Electronics (ICCE), 2012 Fourth International Conference on*, pp. 318 – 321, 2012.

[14] P. Nguyen, D. Tran, T. Le, X. Huang, and W. Ma, "EEG-Based Person Verification Using Multi-Sphere SVDD and UBM," *17th PAKDD*, pp. 289-300, 2013

[15] P. Nguyen, D. Tran, T. Le, X. Huang, and W. Ma, "Age and Gender Classification Using EEG Paralinguistic Features," *the 6th International IEEE EMBS Conference on Neural Engineering*, 2013, in press.

[16] P. Nguyen, D. Tran, X. Huang, T.Vo, D.Phung and W. Ma, "EEG-Based Age and Gender Recognition Using Tensor Decomposition and Speech Features," *ICONIP (2)*, pp. 632-639, 2013.

[17] R. Palaniappan, "Two-stage biometric authentication method using thought activity brain waves," *Int Journal of Neural Systems,*vol.18, 2008.

[18] P. Welch, "The use of Fast Fourier Transform for the estimation of power spectra: a method based on time averaging over short, modified periodogram", *IEEE Trans Audio Electroacoustics*, pp. 70–73, 1967.

[19] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles. "The electroencephalo-gram as a biometric." *In Electrical and Computer Engineering, 2001. Canadian Conference on*, vol. 2, pp. 1363-1366, 2001.

[20] S. Malby, R. Mace, A. Holterhof, C. Brown, S. Kascherus, and E. Ignatuschtschenko. "Comprehensive Study on Cybercrime," *Available*: http://www.unodc.org/documents/organized-crime/UNODC_CCPCJ_EG.4_2013/CYBERCRIME_STUDY_210213.pdf

[21] S. Marcel, J.R. Millán, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on 29*, pp. 743–752, 2007.

[22] S. Sanei, J. Chambers, *EEG signal processing*, Wiley-Interscience, 2007.

[23] V. Matyás□, Z. R□iha, "Security of biometric authentication systems, " *CISIM*, pp.18-28, 2010.

[24] T. Le, D. Tran, W. Ma, and D. Sharma, "An Optimal Sphere and Two Large Margins Approach for Novelty Detection," *in Proceedings of WCCI (IJCNN)*, pp. 909-914, 2010.

[25] W. Ma, J. Campell, D. Tran, and D. Kleeman, "Password Entropy and Password Quality," *NSS2010*, pp. 583 – 587.