A high-security EEG-based login system with RSVP stimuli and dry electrodes

Yiyu Chen, Ayalneh Dessalegn Atnafu, Isabella Schlattner, Wendimagegn Tariku Weldtsadik, Myung-Cheol Roh, Hyoung Joong Kim, Seong-Whan Lee, Benjamin Blankertz, Siamac Fazli

Abstract—Lately, EEG-based authentication has received considerable attention from the scientific community. However, the limited usability of wet EEG electrodes as well as low accuracy for large numbers of users have so far prevented this new technology to become commonplace. In this study a novel EEG-based authentication system is presented, which is based on the RSVP paradigm and uses a knowledge-based approach for authentication. 29 subjects’ data were recorded and analyzed with wet EEG electrodes as well as dry ones. A true acceptance rate of 100% can be reached for all subjects with an average required login time of 13.5 s for wet and 27.0 s for dry electrodes. Average false acceptance rates for the dry electrode setup were estimated to be 3.33 x 10^-7.

Index Terms—biometrics, EEG, brain-computer interfaces, computer security, authentication, RSVP, dry electrodes, ERP

I. INTRODUCTION

ELECTRONIC authentication is a process of establishing confidence in user identities electronically presented to an information system [1]. Conventional authentication systems are based on passwords and user input on a keyboard, touch screen or with a mouse. More recent approaches employ eye tracking or biometric mechanisms.

To overcome some of the shortcomings of these classic authentication techniques, novel authentication approaches are being explored that rely on the acquisition of neural activity, see, for example, [2], [3], [4], [5], [6], [7], [8], [9].

Non-invasive measurements of neural activity are most commonly performed by Electroencephalography (EEG) [10], Magnetoencephalography (MEG) [11], Near-Infrared Spectroscopy (NIRS) [12], [13] and functional Magnetic Resonance Imaging (fMRI) [14], [15]. Due to the ease-of-use, portability, relative low cost and its high temporal resolution EEG presents the most viable imaging technology for biometric authentication purposes.

To date, a number of different EEG-based authentication techniques have been proposed. Most of them rely on the finding that EEG characteristics are unique for every person [16], and can therefore be used as a biometric feature for identification. Since the first EEG-based biometric system was proposed in the late 90s [17], a whole range of subject-specific EEG features have been examined for a variety of conditions, such as eyes-open/eyes closed [17], [18], visual stimulus presentations [2], [7], motor imagery [4], [19], word generation [4] and imagined speech [20], among others.

Poulos et al. [17] used spectral features in combination with an autoregressive (AR) model of resting state EEG and employed learning vector quantification (LVQ) to achieve an identification accuracy of up to 84%. Paranjape et al. used single-channel EEG, recorded during a simple eyes-closed/eyes-open task for the identification of individual subjects and achieved an out-of-sample accuracy of over 80% [18]. Palaniappan and Raveendran [2] used visual evoked potentials (VEP) to identify users. They developed a task in which pictures of gray-scale objects were presented and analyzed corresponding power changes within the gamma range, which lead to an average classification accuracy of 95% across subjects. Marcel and Millán [4] proposed an authentication system, which is based on inherent features and used Gaussian mixture models. Subjects were cued to perform motor imagery tasks and word generation. Their results show that motor imagery is the more appropriate mental task for authentication and also that classification performance may degrade across multiple sessions. Rocca et al. [8] suggested an inheritance-based authentication method using PhysioNet data collected from 108 subjects with eye-open and eye-closed resting conditions. The authors estimate the robustness of functional connectivity measures between EEG sensors and use it as a feature for biometric recognition. Yeom et al. [7] designed an authentication system, where EEG responses to self-face and non-self-face visual stimuli are analyzed by means of non-linear support-vector machines (SVMs). Their average performance is reported as 86.1% across 10 subjects. For a more detailed review on recent EEG-based biometric user recognition approaches we would like to refer the interested reader to [21].

To date, most of the neural activity-based authentication systems employ gel-based EEG electrodes. While gel-based
electrodes offer the best signal quality, they are also inconvenient to use in daily life as they need to be set-up and leave the hair full of gel. For the authentication process which should take less than a minute, this is not feasible. To this end, dry and water-based electrodes can offer a viable solution. While dry electrode technology has first been proposed in the late 60’s and early 70’s [22], [23], it has only recently been validated in online BCI paradigms and thus gained more attention from the scientific community [24], [25], [26]. Besides, recent developments for deployable setups use electrodes that are placed hardly visible in-ear [27], [28], on the ear [29] or, as printed electrode arrays, around the ears [30].

Early stage authentication systems, based on commercially available dry electrodes and wireless EEG, have recently been reported [31], [6]. These existing dry-electrode authentication systems (and most others) are based on oscillatory features. From the neuroscientific and Brain-Computer Interface (BCI) literature it is known that oscillatory features are highly subject-dependent and that so-called BCI illiteracy occurs in approximately 15-30% of subjects [32]. In short, BCI illiteracy means that subjects are not able to change their oscillatory brain activity willingly upon command. As a result, at least some subjects will not be able to use these types of systems meaningfully.

An Event-Related Potential (ERP) is the electrical brain activity, as measured by EEG, that is time-locked to some event. Typically this event is an external sensory stimulus, but an ERP can also result from the execution of a motor, cognitive or psychophysiological task [33]. ERPs reflect, e.g., sensory perception and cognitive processing and they are modulated by various factors, e.g., user states like attention. In contrast to oscillatory features, ERP-based signals are known to have a much better reproducibility across subjects and can be employed in spelling devices [34], [35], [36].

In this paper we propose a novel EEG-based authentication system which is based on ERPs that are elicited by a rapid serial visual presentation (RSVP) paradigm [37]. In the RSVP paradigm, different symbols are presented one-by-one in a serial manner and in the same location of display, exploiting only the foveal visual field. The RSVP paradigm allows to present a large number of stimuli within a short time. Despite of the fast stimulus presentation, the RSVP elicits strong ERPs [36]. As a result the RSVP paradigm can lead to high spelling speeds [36] or short login times as we will show.

In contrast to most previous approaches, which extract specific inference factors from the EEG, here we provide a knowledge-based approach: users choose a combination of three images as their personal password. A large number of images, which include the password images, are then presented in rapid succession at a single central location. Selective attention of the images which comprise the password leads to an enhancement of event-related potential (ERP) components which can be robustly detected with the help of machine learning techniques in the EEG. The usability of this system is furthermore validated with dry electrodes.

For a detailed discussion on alternative methods, that do not rely on EEG as well as specific advantages of the proposed authentication system are discussed in Sec. V.
C. Paradigm

1) Key Idea: In order to introduce the main idea, we would like to explain the system in terms of a potential deployment (see Fig. 2): During registration, the subject chooses a username as well as a combination of three symbols as a password (see also Fig. 3A). Password symbols are flashed in rapid succession together with other non-password symbols in randomized order (Fig. 3C) and a subject-dependent model is estimated from the EEG data. During the login phase first the username is entered, then the same password and non-password symbols are flashed in rapid succession. The model that was derived from the registration phase is applied to the EEG data and the authentication is performed.

2) Detailed experimental paradigm: A total of six runs was performed. At the beginning of each run, subjects were instructed to select three password symbols (PS) from a collection of 260 pictures, which would subsequently act as their password. For the stimuli, a standardized set of 260 pictures was used [39], which are loosely based on Snodgrass and Vanderwart’s object set [40] (see part A of Fig. 3). The password symbols were selected by clicking the icon with a mouse (see part A of Fig. 3). Once three icons were chosen, the selected PS were displayed at the center of the screen for 15 s and the subject was asked to memorize them (part B of Fig. 3). 22 non-target symbols were randomly sampled from the data base and kept constant during each run.

During the RSVP phase, 25 different icons (including the three PS as well as 22 unrelated icons) were presented in RSVP with an inter-stimulus interval (ISI) of 200 ms. They were presented in bursts of 150 presentations (lasting 30s). Each burst contained a variable number of target symbols, which the user was instructed to count. After each burst a short break occurred. The break consisted of two phases: during the first phase, the subject was asked to type the number of targets (i.e. PS) he was able to count (see Fig. 3D). His actual input was not used in any further analysis, but merely to ensure his attention to the task. In the second phase, the target stimuli and a countdown of 15 s were shown (Fig. 3B). The countdown indicated the start of a new presentation period. This procedure was repeated four times until a total of 600 trials, consisting of 72 targets and 528 non-targets were presented. The sequence of stimuli was designed such that there was always a minimum gap of two other icons between two same icons.

Please note, that for deployment, only one run would be sufficient for registration. However, as mentioned earlier a total of six runs was performed. The additional runs were included to examine various scenarios, as described below.

a) Runs 1-3: Login with novel and specific symbols: During the first run, subjects had to choose three icons as their password. In the second run they had to choose three different icons. The icons of the first run were blocked during the selection process. Similarly, in run 3, icons of runs 1 & 2 could not be chosen. If an icon had been selected in any of the previous runs, it was not part of the stimulus set in the current run. Runs 2 and 3 were performed to simulate a password change on the user side.

b) Run 4: Influence of old password symbols: A different setting was chosen in run 4 in order to investigate whether old PS would still elicit specific ERPs that the login system might mistake as the actual PS. This issue is relevant for the case of changing passwords, as well as in situations when trying to resist a forced login. To this end, the old PS of run 1 were introduced as nontargets in run 4 to be tested against the newly chosen PS.

c) Runs 5-6: Categories vs. specific symbols: In the last two runs, we introduced some variability in the target symbols in order to challenge the permissiveness of the proposed login
system. As in runs 1-4, the participants were asked to choose three PS from a collection of 33 options. But this time, each chosen symbol was taken as a representative for a category (such as ‘car’ or ‘elephant’). During stimulus presentation, the symbol that represented the category in the selection process was exchanged with other examples from this category. Three examples per category were used during stimulus presentation. If an elephant was chosen as one PS, also pictures of other elephants were considered as (target) PS. These symbols will be called categorical symbols in the remainder of this article. The average ratio of targets was the same in runs 5 and 6, i.e. 12%. By using categorical symbols instead of specific symbols the size and complexity of the stimulus set is increased and thereby the system security.

Note, that in runs 1-4, the data base contained only one symbol from each category. However, there were some pictures that were quite similar, e.g., tiger, leopard, lion, fox, see Fig. 3A.

The comparatively large number of runs can lead to exhaustion and/or boredom on the subject side. To counter these effects, distractors were included in the experimental paradigm. After each run, short movie clips were shown to the subject (Fig. 3E). Furthermore, the subjects were required to perform six mathematical tasks, which consisted of addition and subtraction (Fig. 3F). Prior to the six actual runs a short test run was performed, such that participants could become familiar with the task.

For the dry electrode study, only the first three runs were performed, since we were interested to show that the basic setup is viable for dry electrodes. Also the movie clips were not shown during dry measurements. As a result the total recording time for the dry electrode measurements was reduced to approx. 12 minutes. Three subjects performed the experiment twice, once with each electrode configuration, while all others participated in only one experiment.

III. DATA ANALYSIS

For the ERP analysis, the EEG data was low-pass filtered by a Chebyshev digital filter with a passband of 40 Hz and a stopband at 49 Hz. The data was then downsampled to...
100 Hz from 1000 Hz (and 1200 Hz dry electrode data) by averaging consecutive blocks of 10 (of 12) samples. The data was epoched to a range of -200 ms to 1000 ms with respect to stimulus onset. A baseline correction was performed on the pre-stimulus interval from -200 ms to 0 ms. Epochs containing strong eye movements were detected and rejected using a minmax criterion of 75 µV on the channels F9, Fz, F10 and Fp2 for wet electrode, and 150 µV on all channels for dry electrode data.

Features were calculated from 28 wet (16 dry) electrode channels by averaging voltages in nine non-overlapping time-windows with a width of 100 ms, starting from 100 ms to 1000 ms with respect to stimulus onset. This resulted in 28 × 9 = 252 (dry: 16 × 9 = 144) dimensional feature vectors. Channels F9,10 and Fp2 were removed in the wet electrode setup, while none of the dry electrodes were removed.

Topographical maps of significant features were calculated by point-biserial correlation coefficients [41]. The point-biserial correlation coefficient is a special case of the Pearson product-moment correlation coefficient, and measures the association of a binary random variable (in this case 'target symbols' and 'non-target symbols') to a continuous random variable (here channel-wise ERP data). It is defined as:

\[
r_{pb} = \frac{M_1 - M_0}{s_n} \sqrt{\frac{n_1 n_0}{n^2}}
\]

with

\[
s_n = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2
\]

where \(M_1\) and \(M_0\) are the mean values of the data points in groups 1 and 0, respectively, \(n_1/n\) the number of examples in groups 1 and 0 and \(n\) the total sample size. Using Fisher's transformation the correlations were transformed into unit variance z-scores for each subject \(j\) [42], and grand average z-scores were obtained by a weighted sum of individual z-scores over all subjects:

\[
z_j = \tanh^{-1}(r_{j}) / \sqrt{m_j - 3}
\]

where \(m_j\) is the sample size of subject \(j\), and \(N = 16\) the total number of subjects. \(p\)-values for the hypothesis of zero correlation in the grand average were computed by means of a two-sided z-test. All reported \(p\)-values were Bonferroni-corrected to account for multiple hypothesis testing [43].

Classification of ERP components was performed with regularized linear discriminant analysis (RLDA), also known as shrinkage LDA [44], [45]. Let \(x_1, \ldots, x_n \in \mathbb{R}^d\) be \(n\) feature vectors, and \(\hat{\mu}\) and \(\hat{\Sigma}\) be the unbiased estimator of the mean and covariance matrix, respectively. If the dimensionality of the data is high with respect to the number of data points, the estimation of the covariance matrix can be imprecise and lead to a systematic bias [46], [47]. Shrinkage (i.e. regularization) of the estimated covariance matrix is a way of correcting this systematic bias:

\[
\hat{\Sigma}(\gamma) := (1 - \gamma)\hat{\mu} + \gamma\nu I,
\]

where \(\gamma \in [0, 1]\) is the shrinkage parameter, and \(\nu\) being the average eigenvalue of the estimated covariance matrix (\(\nu = \text{Tr}(\hat{\Sigma})/d\) with \(d\) being the dimensionality of the data) and \(I\) the identity matrix.

For shrinkage LDA, an analytical solution is available [48]. The analytical solution penalizes large sample-to-sample variance of entries in the empirical covariance, leading to stronger shrinkage. Let the \(i\)-th element of vectors \(x_k\) and \(\hat{\mu}\) be denoted as \((x_k)_i\) and \((\hat{\mu})_i\), respectively, and \(s_{ij}\) be the \(i\)-th row and \(j\)-th column of \(\Sigma\) and, define

\[
z_{ij} = ((x_k)_i - (\hat{\mu})_i)(x_k)_j - (\hat{\mu})_j,
\]

then the optimal shrinkage parameter \(\gamma^*\) in terms of generalization error can be calculated by

\[
\gamma^* = \frac{n}{(n - 1)^2} \cdot \frac{\sum_{i,j=1}^{d} \text{var}(z_{ij}(\hat{k}))}{\sum_{i,j} s_{ij}^2 + \sum_{i} (s_{ii} - \nu)^2}.
\]

While this parameter could in principle also be found by nested cross-validation, the analytical solution is computationally less expensive and was used in our analysis.

In order to investigate the minimum time required for a potential login with a predefined desired true acceptance rate (TAR), cross-validation was performed. Please note that three different targets were present in each run (with 22 non-targets; see Sec. II-C for more details). Furthermore, targets changed during individual runs (similarly for non-targets). However, all data was grouped into target and non-target trials, regardless of the actual symbol and run. During the cross-validation procedure, 21 target trials with corresponding non-target trials, i.e. 7 sequences consisting of 175 trials (21 targets and 154 nontargets), were left out for testing, while all other data was used for classifier training. These left out trials were fed to the classifier and the resulting output was written into a matrix. The trials were arranged in the 21 × 8 matrix such that the first column consisted of all 21 targets, while the other 7 columns consisted of 147 non-targets. Note, that the outputs of 7 nontargets were discarded. This arrangement of the classifier outputs as a matrix was chosen to analyze the recognition performance depending on the length of the symbol sequence that is presented for the login process. The true labels of each row are therefore \(C = [1 0 0 0 0 0 0 0]\). Only if the predicted trials matched the true labels of \(C\) exactly, the login attempt was granted.

To examine the trade-off of TAR vs. login time, the classification outputs of \(n\) sequences (i.e. \(n\) matrix rows) were averaged, where \(n\) was varied from 1 to 21. The number of required sequences (and thus login time) was calculated for three predefined TARs, namely 90%, 99% and 100%.

Due to the unbalanced number of trials per class, an appropriate loss function needs to be chosen, which considers the different class prior probabilities [49]. A typical choice is the area under the receiver operator characteristic (ROC) curve [50], which we have employed in the single trial classification analysis.

The false acceptance rate (FAR) of an authentication system is defined as the fraction of the number of successful authentications by impostors divided by the total number of impostor authentication attempts. Let us consider the following situation: person B is trying to gain access to the account of person A. To simulate this situation, the EEG data of person
B is classified by the model, which was derived from the data of person A. We examine two cases.

Case 1: Person B does not know the password of person A. We therefore map a permuted version of person B’s classes to those of subject A. In practical terms this means person B randomly selects three symbols as his password symbols. The class permutation was performed 50 times and sequence data used for the login procedure were randomly sampled 100 times. Thus a total of 5000 login attempts were simulated for person B trying to enter the account of person A. Not only person B, but all other subjects were also considered as potential attackers for the account of person A, resulting in $15 \times 5000 = 75000$ attempts in total. This analysis was repeated for each subject. The same number of sequences $\alpha$, which were estimated through the predefined TAR levels, were used for averaging.

Case 2: Person B knows the username and password of person A. We apply the same procedure as for case 1, except that now no permutation of classes is performed. As a result, $15 \times 100 = 1500$ attempts are simulated in total.

IV. RESULTS

A. ERP

The grand average ERPs for normal and categorical data can be obtained from Figure 4. The top row shows ERP timecourses, where the green line represents the average target response, and the gray line the non-target response. For both conditions electrode P7 shows a N2 peak at 350 ms, P3a peaks at 540-550 ms in channel Cz and P3b peaks at 630 ms in both depicted channels (Cz and P7). The first two rows of scalp maps indicate the average target and non-target amplitudes for selected intervals. These intervals can be obtained from the gray shaded areas in the ERP timecourses. In the interval 300-420 ms, the N2 can be seen in the occipital area. The third row of scalp maps as well as the plots underneath show the distribution of the sgn($r^2$) as a measure of the correlation between the target and non-target classes. The correlations in occipito-parietal areas are higher for specific symbols in the time interval between 480-580 ms, while correlations in fronto-central area are higher for categorical symbols in the time interval from 580 to 680 ms. Both conditions show a late negative peak around 800 ms.

Classwise averaged timecourses of grand average ERPs with the dry electrode configuration can be seen in Figure 5.

B. Including the old password as a distractor

As described in Sec. II-C the PS of run 1 were included as non-targets in run 4. Figure 6 shows the time evolution of statistical significance for the target symbols versus non-targets (top) and for the old target symbols of run 1 versus non-targets (bottom). As can be seen, new target symbols show highly significant differences as compared to non-targets in time windows starting from 400 ms and ending at 700 ms. Old target symbols do not show any significant differences to non-target symbols.

C. Single-trial classification

Single trial classification results of the specific symbol condition can be seen in Table I. As mentioned before, area of ROC is used as a loss function here. The first two columns show classification accuracy for wet electrode data with the full 28 channel setup and a 16 channel setup, respectively. The last column shows single-trial accuracy for 16 dry electrodes. Both 16 channel setups contain the same channels. Average accuracy across 16 subjects for the 28- and 16-channel wet configurations are $87.8\pm5.1 \%$ and $85.9\pm5.0 \%$. For the dry electrode setup, the average accuracy is $78.2\pm5.7 \%$. A paired, two-sided sign test with the hypothesis that the difference between the matched samples for the 16 wet and 16 dry electrode setups comes from a distribution whose median is zero, yielded significantly lower accuracy for the dry electrode setup ($p < 0.001$).

Categorical symbols scored an average accuracy of $82.62\pm4.78 \%$ across 16 participants with the full, wet channel setup (not shown in Table I). The result of a paired two-sided sign test shows significantly lower accuracy for the categorical symbol condition compared to the specific symbol condition ($p < 0.05$).

D. Estimation of required login time

The estimated login times were calculated as a function of required accuracy. Table II summarizes the results for the specific-symbol condition with wet and dry electrode data. If the required accuracy of the system is set to 100%, the average login time for the wet electrode setup is $10.7\pm4.6$ s and $27.0\pm11.7$ s for the dry electrode setup.

E. False acceptance rate

Results for the simulations of impostor login attempts, when the impostor does not know the password of the user (case 1) can be obtained from Table III. False acceptance rates are given for each subject and for three levels of true acceptance
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIFS.2016.2577551, IEEE Transactions on Information Forensics and Security

Fig. 4. Grand average ERP analysis of specific- and categorical-symbol conditions. The top row shows timecourses of two EEG electrodes, namely Cz and P7. Rows 2 and 3 show scalp maps of time-averaged amplitudes. Below the distribution of the sgn($r^2$) as a measure of the correlation between the target and non-target classes is shown as time-averaged scalp maps (row 4) and timecourse (row 5).

rate. The average FARs across subjects are $1.69 \cdot 10^{-4}$, $4.00 \cdot 10^{-5}$ and $3.33 \cdot 10^{-5}$ for TAR being 90%, 99% and 100%, respectively. The system design is such that the probability of guessing another subject’s password symbols correctly is $p = \frac{25}{27} \times \frac{2}{23} \times \frac{1}{23} \approx 4.35 \cdot 10^{-4}$. One sample t-tests with the hypothesis of equal means show that FAR values obtained from our simulation are significantly lower for all three tested TAR values ($p = 3.8 \cdot 10^{-6}$, $p = 1.7 \cdot 10^{-14}$ and $4.1 \cdot 10^{-15}$). The reason for this is the reduced classification accuracy due to the subject-to-subject transfer of a given classifier. If the impostor knows the user’s password symbols (case 2), the average FARs across subjects are $7.8 \pm 8.2$, $15.6 \pm 8.4$ and $19.7 \pm 10.1$ for TAR being 90%, 99% and 100%, respectively. While these numbers may seem high at first sight, please note that the impostor is in the possession of the username and password. Since the impostor knows the password symbols, these symbols will elicit an ERP with a strong deflection when displayed during the RSVP procedure. The neural signature of the ERP is only slightly different for every person and the system is designed to discriminate between inter-subject differences of ERPs. Please note, that a combination with an inherence-based approach would reduce FAR rates for impostors with password knowledge (see also Sec. VI, where this is discussed further).

V. DISCUSSION

We proposed a knowledge-based authentication approach that relies on the acquisition of brain signals via EEG. We demonstrated that the actual login time is relatively short for the established high security level even when employing novel dry electrode EEG caps which provide lower signal quality. However, the additional effort of requiring an EEG headset and mounting it needs to be considered. In order to trade-off the pros and cons of the proposed brain-based authentication, we review existing approaches in more detail.

A. Existing authentication systems

Existing authentication techniques can be categorized into three factors: knowledge factor, possession factor and inher-
Fig. 6. Top and bottom rows show grand average statistical significance of differences between password-symbols (PS) and non-target (NT) symbols for wet and dry electrodes, respectively. The middle compares old target symbols (OPS) of an earlier password to other non-targets (i.e. data from run 4, see Sec. II-C for more details).

TABLE II
shows the minimum number of targets that need to be averaged (#) and the login time (T) required to achieve a true acceptance rate (TAR) of 90%, 99% and 100%. Results are given for each subject and their mean for the wet (left) and dry electrode configurations (right).

<table>
<thead>
<tr>
<th></th>
<th>wet electrodes</th>
<th></th>
<th>dry electrodes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>90%</td>
<td>99%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zk</td>
<td>3</td>
<td>5.0</td>
<td>11.7</td>
<td>9</td>
</tr>
<tr>
<td>eal</td>
<td>2</td>
<td>3.3</td>
<td>5.0</td>
<td>5</td>
</tr>
<tr>
<td>lh</td>
<td>2</td>
<td>3.3</td>
<td>6.7</td>
<td>6</td>
</tr>
<tr>
<td>xk</td>
<td>2</td>
<td>3.3</td>
<td>5.3</td>
<td>6</td>
</tr>
<tr>
<td>jw</td>
<td>3</td>
<td>5.0</td>
<td>15.0</td>
<td>14</td>
</tr>
<tr>
<td>ad</td>
<td>3</td>
<td>5.0</td>
<td>8.3</td>
<td>8</td>
</tr>
<tr>
<td>yy</td>
<td>2</td>
<td>3.3</td>
<td>5.0</td>
<td>4</td>
</tr>
<tr>
<td>er</td>
<td>2</td>
<td>3.3</td>
<td>6.7</td>
<td>5</td>
</tr>
<tr>
<td>vi</td>
<td>1</td>
<td>1.7</td>
<td>3.3</td>
<td>3</td>
</tr>
<tr>
<td>se</td>
<td>2</td>
<td>3.3</td>
<td>5.8</td>
<td>8</td>
</tr>
<tr>
<td>hw</td>
<td>2</td>
<td>3.3</td>
<td>8.3</td>
<td>8</td>
</tr>
<tr>
<td>th</td>
<td>2</td>
<td>3.3</td>
<td>8.3</td>
<td>8</td>
</tr>
<tr>
<td>hj</td>
<td>1</td>
<td>1.7</td>
<td>3.3</td>
<td>3</td>
</tr>
<tr>
<td>tu</td>
<td>1</td>
<td>1.7</td>
<td>5.0</td>
<td>5</td>
</tr>
<tr>
<td>gz</td>
<td>1</td>
<td>1.7</td>
<td>5.0</td>
<td>4</td>
</tr>
<tr>
<td>sk</td>
<td>1</td>
<td>1.7</td>
<td>6.7</td>
<td>7</td>
</tr>
<tr>
<td>mean</td>
<td>1.8</td>
<td>3.0</td>
<td>4.3</td>
<td>7.2</td>
</tr>
</tbody>
</table>
TABLE I
SINGLE TRIAL CLASSIFICATION ACCURACY FOR WET AND DRY ELECTRODES

<table>
<thead>
<tr>
<th>ID</th>
<th>wet</th>
<th>dry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>28 ch</td>
<td>16 ch</td>
</tr>
<tr>
<td>zk</td>
<td>78.8</td>
<td>80.4</td>
</tr>
<tr>
<td>eal</td>
<td>90.2</td>
<td>87.6</td>
</tr>
<tr>
<td>lh</td>
<td>86.9</td>
<td>83.9</td>
</tr>
<tr>
<td>xk</td>
<td>84.4</td>
<td>84.2</td>
</tr>
<tr>
<td>jw</td>
<td>78.0</td>
<td>75.9</td>
</tr>
<tr>
<td>ad</td>
<td>85.6</td>
<td>81.6</td>
</tr>
<tr>
<td>yy</td>
<td>89.6</td>
<td>86.7</td>
</tr>
<tr>
<td>er</td>
<td>88.7</td>
<td>84.2</td>
</tr>
<tr>
<td>vi</td>
<td>95.8</td>
<td>95.4</td>
</tr>
<tr>
<td>sc</td>
<td>83.7</td>
<td>82.5</td>
</tr>
<tr>
<td>hw</td>
<td>86.6</td>
<td>84.9</td>
</tr>
<tr>
<td>th</td>
<td>88.0</td>
<td>85.2</td>
</tr>
<tr>
<td>hj</td>
<td>95.4</td>
<td>94.0</td>
</tr>
<tr>
<td>tu</td>
<td>89.0</td>
<td>87.2</td>
</tr>
<tr>
<td>gz</td>
<td>94.3</td>
<td>91.1</td>
</tr>
<tr>
<td>sk</td>
<td>90.0</td>
<td>89.1</td>
</tr>
</tbody>
</table>

mean 87.8 ± 5.1 85.9 ± 5.0 mean 77.5 ± 5.9

TABLE III
SHOWS FALSE ACCEPTANCE RATE FOR THREE TAR LEVELS. A TOTAL OF 75000 IMPOSTOR LOGIN ATTEMPTS WERE SIMULATED FOR EACH SUBJECT. DRY EEG ELECTRODE DATA WAS USED FOR SIMULATIONS.

<table>
<thead>
<tr>
<th>ID</th>
<th>False acceptance rate for TAR=90%</th>
<th>TAR=95%</th>
<th>TAR=100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>zk</td>
<td>2.93e-04</td>
<td>9.33e-05</td>
<td>4.00e-05</td>
</tr>
<tr>
<td>eal</td>
<td>4.00e-05</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>lh</td>
<td>1.47e-04</td>
<td>6.67e-05</td>
<td>2.67e-05</td>
</tr>
<tr>
<td>jak</td>
<td>1.33e-04</td>
<td>1.33e-05</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>ocb</td>
<td>1.33e-05</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>rsv</td>
<td>1.60e-04</td>
<td>0.00e+00</td>
<td>1.33e-05</td>
</tr>
<tr>
<td>sn</td>
<td>6.67e-05</td>
<td>1.33e-05</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>jeg</td>
<td>2.40e-04</td>
<td>5.33e-05</td>
<td>9.33e-05</td>
</tr>
<tr>
<td>fat</td>
<td>1.33e-05</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>hjh</td>
<td>2.93e-04</td>
<td>1.33e-05</td>
<td>4.00e-05</td>
</tr>
<tr>
<td>sd</td>
<td>4.00e-04</td>
<td>1.07e-04</td>
<td>8.00e-05</td>
</tr>
<tr>
<td>nsk</td>
<td>4.00e-04</td>
<td>1.87e-04</td>
<td>1.87e-04</td>
</tr>
<tr>
<td>bv</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>dk</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>mg</td>
<td>6.67e-05</td>
<td>0.00e+00</td>
<td>0.00e+00</td>
</tr>
<tr>
<td>jys</td>
<td>4.00e-04</td>
<td>9.33e-05</td>
<td>5.33e-05</td>
</tr>
</tbody>
</table>

mean 1.69e-04 4.00e-05 3.33e-05

enue factor [51]. Password-based authentication is a knowledge factor authentication that is most widely used in computer systems. Recent major leaks from giant IT companies like Twitter, LinkedIn, Dropbox and Yahoo are an indicator that password-based authentication systems do not provide sufficient security [52]. The major attacks on password-based authentication systems are grouped into 6 types. They are brute-force, dictionary, guessing, spyware, social engineering and shoulder surfing [53], [54]. One of the basic rules in text-based password security to protect against guessing and dictionary attacks is that passwords should be random and strong, but this makes it difficult for the user to remember. The conventional way of measuring the strength of a password has been improved in a number of ways. Malone and Maher [55] proposed a method to identify dangerously common passwords among different standard statistics. Komanduri et al. [56] invented a novel way of identifying weak passwords by predicting the next character that the user will type. The authors indicate that their method creates fewer weak passwords than conventional character composition methods. On the other hand, Schechter et al. [57] claim passwords made of simple words are not always weak. The strength of passwords depends on their popularity. The authors propose an oracle that can identify popular (undesirable) passwords using the count-min sketch method [58]. In addition, Florêncio et al. [59] argue that the common way of ruling out weak passwords and password re-use are suboptimal. They propose an optimal solution that allows the re-use of passwords by grouping accounts.

The difficulty of recalling strong text passwords by users lead researchers to look for better ways of representing passwords. Since the 19th century, psychologists have hypothesized that visual information is easier to recall than verbal or textual information [60], [61]. The main reason for the superiority of visual memory is that pictures are more easily associated with previous knowledge about the object they represent [62], [63]. Two decades ago, the first picture-based password system was proposed by Blonder [64]. Since then, a variety of graphical password systems have been introduced to improve security and usability [65]. Although graphical passwords have improved memorability and security as compared to text-based passwords, they are still vulnerable against shoulder surfing attacks. Several techniques have been suggested to overcome shoulder surfing on graphical password systems. Suo [66] designed a system that blurs all parts of the image except a small portion, which he termed decoy region. In this system the user only needs to respond ‘yes’ or ‘no’ by using mouse clicks to indicate whether his password appears in the decoy region. Sobrado et al. [67] and Wiedenbeck [68] suggested a convex hull-based shoulder surfing protection scheme for graphical passwords. The convex hull is an imaginary polygon, formed by connecting the password images, which are scattered among a large number of other pictures. This method avoids the need of revealing the password images directly, since the user is instructed to click anywhere within the convex hull. Shoulder surfing can also be prevented by changing the password input mechanism from mouse, keyboard or touch screen to biometric based mechanisms [69], [70]. Kumar et al. [71] suggest to use an eye-tracking system with an on-
screen keyboard to input the password. This makes it difficult for a peeper to identify the user’s password. Later Wu et al. [72] designed a method that combines the Convex Hull click method [67], [68] with the eye-tracker input method [71] in order to provide better security as compared to the individual techniques.

Osborn et al. [73] introduced a novel graphical password system, where the user chooses a number of picture-based categories as his password. During the login process, a set of pictures is shown to the user on a grid. Each picture belongs to a predefined category (such as car, house, flower, etc.) and appears with an associated letter. The user will then find the pictures, which belong to his password categories and type the corresponding letters. The pictures as well as the associated letters will be different for each login process. Matching two pictures of the same category is a computationally expensive task, rendering brute-force attacks useless. Furthermore, the security level of the system can be controlled by adjusting the total number of categories as well as the number of displayed pictures during the login process.

Biometric authentication has several advantages over the traditional knowledge-based and possession-based authentication techniques, however most biometric-based authentication require special hardware to capture biometric data, which results in increased implementation costs. Some of the biometric authentication techniques found today are based on fingerprint scans [74], hand geometry [75], [76], iris scans [77], [78], face recognition [79], [80] and keystroke dynamics [81], [82].

B. The proposed brain-based authentication

The results indicate that our knowledge-based authentication approach is appropriate for high security EEG-based login technology. Attack modes, such as shoulder surfing or eye-tracking are impossible to apply. While previous approaches focused on inheritance-based features, a knowledge-based approach offers a number of benefits. One of them is that the required TAR levels can be adjusted to the particular application. For example, if only low-security access (e.g. TAR ≥ 90%) is required, the average login time with wet electrodes is only 3.0 s. If needed, the system can be set to 100% authentication accuracy for every subject by averaging more trials and thereby increasing the signal-to-noise ratio. However, this happens at the expense of increased login time. Even under maximum security (i.e. TAR = 100%) situations an average required login time of 10.7 s seems viable (see Tab. II for more results). Furthermore, we would like to point out that login times could be reduced by so-called early stopping methods, which adaptively stop the stimulus presentation sequence when enough evidence for a preset accuracy is reached, see [83] for an overview. Finally, it is an important fact that no subjects were excluded from the analysis which indicates that our system is applicable to the general public.

C. Stimulus complexity and ERP morphology

In an earlier study, the RSVP paradigm was employed for a spelling application [36]. The stimuli consisted of individual letters and ERPs showed much earlier positive deflections, when compared to our study. In this study we compared the effect of specific as well as categorical symbols. The results of these two studies indicate that the delay of the largest positive ERP deflection is positively correlated with the stimulus complexity. In a future study, we are planning to look further into this interesting finding. By deepening our knowledge in this domain it may be possible to not only increase our neurophysiological understanding, but to also find optimal parameters for RSVP-based applications [36], [84].

D. Simulation of an abduction scenario

If a given person is abducted and forced to login, an inherence-based biometric can be conducted forcefully. To find out whether it is possible to suppress the password symbols, we asked the subjects to think of a new password and tested whether the old password symbols would still elicit a discriminative response. The analysis did not show a discriminative response for old password symbols (see middle row of Fig. 6). Therefore, we conclude that it is possible to effectively ‘hide’ the password, if needed.

VI. Conclusion

We tested the proposed system with commercially available dry electrodes. While wet EEG electrodes generally have better signal quality as compared to dry ones, wet electrodes need to be prepared by inserting a conductive gel, which can require up to 30 minutes for a whole head setup. In addition, the hair needs to be washed after the experiment (i.e. login process). As a result, wet electrodes are not feasible for authentication systems. Our analysis shows that while single-trial classification is significantly lower for dry electrodes as compared to wet ones, they show sufficient accuracy for high security applications. We conclude that the use of dry electrodes in a commercial authentication system is possible and furthermore highlight their practicality. A whole range of novel EEG-based technologies are currently being developed. Among those are commercial products, but also many research-based prototypes. They include mobile, low-cost and also dry EEG solutions. In our study, the application of the dry electrodes took 2 minutes on average. Future mobile and dry EEG technology will reduce this time even further and could presumably be shortened to 15 s with next generation EEG headsets.

While most previous approaches are based on inheritance factors, EEG-based biometrics can also serve as a knowledge-based or hybrid authentication system [3], [9]. As a next step, we plan to extend our approach to a one-step, two-factor authentication system, where inheritance and knowledge-based factors are combined. An inheritance-based authentication system could be applied to identify the user and be combined with the knowledge-based approach we introduced here by employing advanced multi-modal fusion techniques, which have recently been developed [85], [86].

In our analysis we calculated the number of required sequence repetitions to achieve three levels of TAR, namely 90%, 99% and 100%. As described earlier these results
were obtained through cross-validation with data of the same session. However, earlier studies observed the degradation of classification performance, when a so-called session-to-session transfer is performed, i.e. when the classifier of an earlier session is transferred to a newer session [4]. While we have not estimated this type of degradation here, we would like to mention that this degradation could in principle be counterbalanced by raising the number of averaged sequences, which would in turn lead to increased login times. In addition, recent technological advances in BCI have shown that session-to-session transfers are possible for expert users with only minimal increased loss [87]. These type of techniques alleviate the need for recalibrating the system. In addition, a recent in-depth study on the stability of EEG-based features for biometrics came to the conclusion, that in fact EEG signals contain discriminative information which are stable across time [88].

Another current limitation of the presented technology is the need of calibration data on the subject side. Subject-independent decoding [89], [90], [91], [92], [93] has recently been addressed in BCI research and allows any user to start a real-time BCI feedback session instantaneously. In the future, we plan to adopt these type of techniques for the presented paradigm and thereby hope to further increase the usability of the proposed system significantly.

APPENDIX A. SUPPLEMENTARY MATERIAL

A video of the stimulus presentation is available.

REFERENCES


Yiyou Chen received her B.S. degree in Computer and Communication Engineering from Korea University, Seoul, Korea, in 2014. Since 2014, she has enrolled in the integrated Master’s and Ph.D. course at the Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea. Her current research interests include Brain Computer Interfaces, Machine Learning, decision making and Cognitive Neuroscience.

Ayalneh Dessalegn Atnafu received his B.Tech in Electrical Engineering from Defense University College, Debre zait, Ethiopia, in 2003, MSc in Computer Engineering from Addis Ababa University, Addis Ababa, Ethiopia, in 2008. He is currently a PhD candidate in Information Security at Korea University, since 2012. His research interests include Digital forensics, information hiding, Authentication and Information Security Management.

Isabella Schlattner is in the integrated Master’s and Ph.D. program at the Department of Brain and Cognitive Engineering at Korea University. She received her Bachelor’s degree in Biology at the Freie University Berlin, Germany. Her current research interests include cognitive processes, especially intelligence and creativity and their underlying brain-based mechanisms. Furthermore, she is interested in brain imaging techniques, such as EEG, NIRS as well as fMRI.

Wendimagegn Tariku Weldtsadik received his BSc in Mathematics from Addis Ababa University, Ethiopia, in 1998. MSc in Computer Science from Addis Ababa University, Ethiopia, in 2009. He is currently a PhD candidate in Information Security at Korea University, since 2013. His research interests include Face Recognition, information hiding, Authentication and EEG signal analysis.

Myung-Cheol Roh received his B.S. degree in Computer Engineering from Kangwon University, Chun-Choen, Korea, in 2001, and the MS and PhD degrees in Computer Science and Engineering from Korea University, Seoul, Korea, in 2003 and 2008. Currently, he is working as a managing researcher at S1, Seoul, Korea. He won the best paper award of the 25th annual paper competition which is supervised by the Korea Information Science Society and is sponsored by Microsoft in 2006. He worked at the Center for Vision, Speech and Signal Processing in the University of Surrey, UK, as a collaborator researcher in 2004 and at the Robotics Institute in Carnegie Mellon University, US, as a researcher from 2008 to 2012. His present research interests include face alignment, face and gesture recognition, robot vision and the pattern recognition related fields.

Hyoung Joong Kim received his B.S., M.S., and Ph.D. degrees from Seoul National University, Korea, in 1978, 1986, and 1989, respectively. He joined the faculty of Kangwon National University, Korea, in 1989. He is currently a Professor of Korea University, Korea. He invented fast lossless compression algorithm using reversible data hiding technique. He also invented image hash algorithm which is robust against histogram equalization attack. He published numerous papers including more than 40 peer-reviewed journal papers. He served Guest Editor of several journals including IEEE Transactions on Circuits and Systems for Video Technology. He is a Vice Editor-in-Chief of the LNCS Transactions on Data Hiding and Multimedia Security. He was a prime investigator of several national R&D projects. His main research interests include high-performance computing and multimedia computing.
Seong-Whan Lee received the B.S. degree in computer science and statistics from Seoul National University, Korea, in 1984, and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology (KAIST), Seoul, Korea, in 1986 and 1989, respectively. Currently, he is the Hyundai-Kia Motor Chair Professor at Korea University, Seoul, where he is the Head of the Department of Brain and Cognitive Engineering. His research interests include artificial intelligence, pattern recognition, and brain engineering. He is a Fellow of the IEEE, the IAPR, and the Korean Academy of Science and Technology.

Benjamin Blankertz received his PhD in mathematics in 1998 and pursued several studies in music cognition. He started Brain-Computer Interface research in 2000 and became chair for Neurotechnology at Technische Universität Berlin in 2012. The Berlin BCI group is known for innovative machine learning approaches in the field of BCI and the development of novel experimental paradigms. This includes, the transfer of BCI technology from the lab to real world applications.

Siamac Fazli received his B.Sc. Physics degree from the University of Exeter in 2002, his M.Sc. in Medical Neurosciences from the Humboldt University Berlin in 2004 and his Ph.D. from the Berlin Institute of Technology in 2011. From 2011-2013 he worked as a Postdoc researcher at the Berlin Institute of Technology for the Bernstein Focus Neurotechnology. Since 2013 he works as an Assistant Professor at Korea University. His current research interests include neuroscience, machine learning, multi-modal neuroimaging and brain-computer interfacing.