A New Method for EEG-Based Concealed Information Test
Deng Wang, Duoqian Miao, and Gunnar Blohm

Abstract—Forensic electroencephalogram (EEG)-based lie detection has recently begun using the concealed information test (CIT) as a potentially more robust alternative to the classical comparative questions test. The main problem with using CIT is that it requires an objective and fast decision algorithm under the constraint of limited available information. In this study, we developed a simple and feasible hierarchical knowledge base construction and test method for efficient concealed information detection based on objective EEG measures. We describe how a hierarchical feature space was formed and which level of the feature space was sufficient to accurately predict concealed information from the raw EEG signal in a short time. A total of 11 subjects went through an autobiographical paradigm test. A high accuracy of 95.23% in recognizing concealed information with a single EEG electrode within about 20 seconds demonstrates effectiveness of the method.

Index Terms—Electroencephalogram (EEG), concealed information test (CIT), hierarchical knowledge base, feature extraction, EEG classification.

I. INTRODUCTION

R
cently, polygraphs based on central nervous system activity have been suggested to be more reliable for the detection of concealed information than those based on physiological measures [1]–[5]. In particular, electroencephalography (EEG) monitors brain activity through electrodes on the scalp, possesses a high temporal resolution and is considered to be relatively convenient, inexpensive, and harmless compared to other methods such as magnetoencephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and optical imaging [6]. In the last two decades, researchers have investigated the use of event-related changes in EEG (known as event-related potentials—ERPs) for concealed information detection [2]–[4], [7]–[11].

The traditional polygraph is based on a comparison of physiological responses (e.g., changes in blood pressure, pulse, respiration, and skin conductivity) to relevant and control questions [12]–[15]. A classical method is the control question test (CQT) which has been widely employed in criminal investigations, employee screening, and selection in several countries, most notably in North America and Israel [12]. However, criticisms of the CQT include nonsolid scientific principles and improper control questions which enhance the risk of stress-related false positive errors (e.g., innocent suspects classified as guilty) [12], [13]. An alternative technique, the concealed information test (CIT), also known as the guilty knowledge test [16], has recently drawn considerable attention among researchers. This test presents a set of question items to an examinee, which include one crime-related item (critical item) and several control items (noncritical items). Items are selected so that an innocent examinee (i.e., one who does not possess the information) would be unable to distinguish the critical item from the noncritical items [9], [11].

In this study, we used the CIT technique which relied on contrasting brain waves evoked by relevant and control stimuli, and developed a novel efficient (i.e., accuracy/time) EEG-based CIT using machine learning algorithms. Through EEG signal processing, we automatically detected brain waves corresponding to different mental activity patterns to uncover the critical item from noncritical items. Indeed, numerous studies have previously demonstrated that CIT based on brain signals can be very accurate [2], [3], [7]–[11]. Besides detection accuracy, it is important to develop fast algorithms that can be used in real-life CIT investigations, for a number of reasons. First, police work is typically carried out under the pressure of time [17]. Second, long time intervals may dissipate witness memory, particularly since identification accuracy decreases with time for eyewitness identification [18]. To our knowledge, no study to date has analyzed and described the accuracy/time trade-off for EEG-based CIT. Moreover, extracting relevant EEG features for classification is a crucial step in EEG analysis. Although there is a wealth of research and literature on EEG signal processing (see [19], [20] for reviews), analysis of EEG recordings from two or more channels often results in high-dimensional data vectors with very large amounts of spatiotemporal information. As a result of increases in dimensionality, processing time increases significantly. Indeed, many methods report good classification accuracy but in the expense of long processing times in order to optimize a test. In addition, the number of samples available for training classifiers is usually relatively small compared

Manuscript received May 03, 2012; revised November 09, 2012; accepted January 28, 2013. Date of publication February 01, 2013; date of current version February 14, 2013. This work was supported by the National Natural Science Foundation of China (60970061, 61273304), by the Fundamental Research Funds for the Central Universities, by CF1 (Canada), by NSERC (Canada), by the Boffert Fund (Queen’s University, Kingston, ON, Canada), and by ORF (Canada). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Z. Jane Wang.

D. Wang was with the Key Laboratory of Embedded System and Service Computing, Ministry of Education, and the Department of Computer Science and Technology, Tongji University, Shanghai 201804, China. He is now with the Centre for Neuroscience Studies, Queen’s University, Kingston, K7L 3N6, ON, Canada (e-mail: w_deng208@hotmail.com).

D. Miao is with the Key Laboratory of Embedded System and Service Computing, Ministry of Education, and the Department of Computer Science and Technology, Tongji University, Shanghai 201804, China (e-mail: dqmiao@tongji.edu.cn).

G. Blohm is with the Canadian Action and Perception Network (CAPnet) and the Centre for Neuroscience Studies, Queen’s University, Kingston, K7L 3N6, ON, Canada (e-mail: gunnar.blohm@queensu.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2013.2244884
to their dimensionality [21]. Therefore, the critical challenge is to map the high dimensional feature space into a lower dimensional space while maintaining relevant signal information, and still obtaining high classification accuracy. Inspired by human decision-making strategies, which attempt to meet criteria for adequacy rather than to identify an optimal solution [22]–[24], our method consists of a hierarchical knowledge construction algorithm and test algorithm for EEG-based CIT. This method provides an accurate, real-time algorithm that does not require an exhaustive global optimization search but nevertheless obtains high accuracy through supervised learning.

Here, our contributions mainly focus on tackling the questions of how to form an appropriate hierarchical feature space and determining the most sufficient level of this space to accurately detect the concealed information from the raw EEG signals. We employed the nonparametric weighted feature extraction (NWFE) [25], [26] method for EEG signal feature extraction. Then we developed hierarchical learning and test algorithms considering the computational costs—learning and test time. The k-nearest neighbor algorithm (k-NN), one of the simplest methods based on the closest training examples in the feature space, was implemented on the data as the classifier. We applied our method to experimental EEG data (mock attention-birthday which is similar to paradigms used in other CIT studies) and obtained a high accuracy (95.23%) in recognizing concealed information with a single electrode in about 20 seconds on an ordinary desktop computer demonstrating the effectiveness of the method.

II. MATERIALS AND METHODS

A. EEG Data Description

The data [29] used in this paper were recorded during an autobiographical paradigm test. Eleven healthy male volunteers between the ages of 22 and 35 years participated in this study. They were right handed and had normal or corrected-to-normal vision. They were not familiar with the scientific basis of the test and only had knowledge about how to perform the test. Each subject was asked to provide to the experimenter five numbers (all 4 digits long), one of them being their year of birth. The subjects did not reveal to the experimenter which one of the numbers corresponded to their birth date until the end of experiment. Each subject participated in 2 experimental runs, except for subject 11, who participated in 3 runs. For subject 1, 3, and 7, one run was discarded because of incorrect target stimulus counting (see below). Therefore, a total of 20 experimental runs were used in this study. In each run, each number was displayed to the subject randomly with thirty repetitions, resulting in a total of $5 \times 30 = 150$ stimuli. Each number was displayed for one second and between the numbers, the screen was blank for two seconds. The subjects did not respond to the items, but were instructed to count the number of times the target stimulus (year of birth) was presented (they were unaware that all stimuli were repeated 30 times in each run). EEG signals were recorded at frontal (Fz), central (Cz), and parietal (Pz) electrode positions of the 10–20 international electrode placement system. All electrodes were referenced to linked mastoids. Vertical EOG was also recorded for blink artifact detection. EEG signals were digitally sampled at 256 Hz.

B. Methodology

The method consists of EEG signal preprocessing, feature extraction, supervised learning for hierarchical knowledge base with a time cost ($\text{HKB}^T$) construction, and classification [30] (see Fig. 1). More specifically, raw EEG signals were first preprocessed. Then we determined the best features to be extracted as the eigenvectors with the largest eigenvalues using nonparametric weighted feature extraction (NWFE). The third stage included: (1) supervised learning of the discrimination rules and the corresponding background conditions (i.e., $k$-values) in a given learning time period, and (2) discrimination rules reorganization according to the validation test performance for constructing a hierarchical knowledge base ($\text{HKB}^T$). The final stage consisted of computing the test sample’s membership degree, computing the classification accuracy for recognized samples and calculating the rejection rate for unrecognized samples. At this stage, the test sample’s member degree was calculated by selecting one or several discrimination rules from the $\text{HKB}^T$ satisfying the minimal confidence level (MCL) demand at a given single sample average test time. The details of these methods are outlined below.

1) EEG Preprocessing and Feature Extraction: Several preprocessing operations were applied to the raw data before submitting the data to the feature extraction and learning algorithms. At first, all data were band-pass filtered offline using elliptic filters from 0.3 Hz to 35 Hz implemented by MATLAB software (The MathWorks, Natick, MA, USA). This is the frequency range that is typically used in P300-based concealed information detection studies [10]. To reduce the effect of large
amplitude outliers due to eye movements, muscle activity, or subject movement [31], a winsorization procedure [32], [33] was applied. To do so, the 5th percentile and the 95th percentile of the signals from each electrode were computed. Values below the 5th percentile or above the 95th percentile were replaced by the corresponding percentile values. The min–max normalization method was then applied to EEG signals to reduce variability. We applied the following equation to $X(c, s, t)$ where $X \in \mathbb{H}$ denotes a matrix that represents a segment of EEG series, where $C$, $S$, and $N$ denote the number of channels, number of measurement samples, and number of trials, respectively.

$$X' = \frac{X - \min(X)}{\max(X) - \min(X)}(U - L) + L.$$  

(1)

where $X'$ is the normalized data matrix, $X$ is the original data matrix and $U$ and $L$ are the upper and lower normalization bounds respectively. Equation (1) maps the data matrix into a bound between 0 and 1. We also compared the results for a bound between $-1$ and 1 (see Fig. 4). Finally, the continuous EEG signals were separated into epochs corresponding to number presentation. Each epoch was one second long; blank intervals were removed. Epochs containing eye blinks in the vertical EOG channel (i.e., potentials above 400 μV) were discarded.

Before classification, we applied a nonparametric weighted feature extraction (NWFE) [25], [26] on the previously preprocessed signals in order to extract the meaningful information. Linear discriminant analysis (LDA) is one of the classical methods for feature extraction and dimensionality reduction that transforms data into a new feature space in which the data has optimal discriminability. NWFE is an LDA-based nonparametric feature extraction method proposed by Kuo and Landgrebe [25]. Compared to parametric discriminant analyses, nonparametric scatter matrices are generally of full rank which provides the ability to specify the number of extracted features desired and to reduce the effect of the singularity problem. On the other hand, the nonparametric nature of scatter matrices reduces the effects of outliers and works well even for nonnormal datasets [25]. In particular, NWFE assigns greater weights to samples near the expected decision boundary. This method extracts fewer features to represent the EEG data distribution and generally performs well, e.g., in studies of EEG-based driver’s cognitive response classification [26], EEG-based driver’s cognitive state assessment [27], and EEG-based passenger’s motion sickness level classification [28]. However, to our knowledge, there is no relevant literature using NWFE for feature extraction in EEG-based concealed information tests.

The main idea of NWFE is to put different weights on every sample to compute the “weighted means”, and to compute the distance between samples and their weighted means as their “closeness” to a boundary, and then to define between-class and within-class scatter matrices nonparametrically which places large weights on the samples close to the boundary and de-emphasizes those samples far from the boundary. In our experimental data, there are two classes, EEG recordings related to birthday and nonbirthday numbers. The nonparametric within-class scatter matrix $S'_{nW}$ and between-class scatter matrix $S'_{bW}$ for L classes was defined as follows:

$$S'_{nW} = \sum_{i=1}^{L} P_i \sum_{k=1}^{N_i} \frac{\lambda_{i,j}^{(i)}}{N_i} \left(x_{k}^{(i)} - M_j \left(x_{k}^{(i)}\right)\right) \times \left(x_{k}^{(i)} - M_j \left(x_{k}^{(i)}\right)\right)^T,$$  

(2)

$$S'_{bW} = \sum_{i=1}^{L} P_i \sum_{j=1}^{L-1} \sum_{i=1}^{N_j} \lambda_{i,j}^{(i)} \frac{x_{k}^{(i)} - M_j \left(x_{k}^{(i)}\right)}{N_k} \times \left(x_{k}^{(i)} - M_j \left(x_{k}^{(i)}\right)\right) \times \left(x_{k}^{(i)} - M_j \left(x_{k}^{(i)}\right)\right)^T,$$  

(3)

where $P_i$ denotes the prior probability of class $i$, $N_i$ is training sample size of class $i$, and $x_{k}^{(i)}$ refers to the $k$th sample from class $i$. $\lambda_{i,j}^{(i)}$ denotes the scatter matrix weight which is a function of $x_{k}^{(i)}$ and $M_j \left(x_{k}^{(i)}\right)$:

$$\lambda_{i,j}^{(i)} = \frac{\text{dist} \left( x_{k}^{(i)}, M_j \left(x_{k}^{(i)}\right) \right)^{-1}}{\sum_{t=1}^{N_j} \text{dist} \left( x_{t}^{(i)}, M_j \left(x_{t}^{(i)}\right) \right)^{-1}},$$  

(4)

where $\text{dist}(a, b)$ denotes the Euclidean distance from $a$ to $b$. If the distance between $x_{k}^{(i)}$ and $M_j \left(x_{k}^{(i)}\right)$ is small, then its weight will be close to 1; if the distance is large, $\lambda_{i,j}^{(i)}$ will be close to 0. The sum of $\lambda_{i,j}^{(i)}$ for class $i$ is 1. $M_j \left(x_{k}^{(i)}\right)$ denotes the weighted mean of $x_{k}^{(i)}$ in class $j$ and is defined as

$$M_j \left(x_{k}^{(i)}\right) = \frac{\sum_{l=1}^{N_j} u_{kl}^{(i,j)} x_{l}^{(i)}}{\sum_{l=1}^{N_j} u_{kl}^{(i,j)}},$$  

(5)

where

$$u_{kl}^{(i,j)} = \frac{\text{dist} \left( x_{k}^{(i)}, x_{l}^{(i)} \right)^{-1}}{\sum_{t=1}^{N_j} \text{dist} \left( x_{t}^{(i)}, x_{l}^{(i)} \right)^{-1}}.$$  

(6)

is the weight of a sample $x_{k}^{(i)}$ in class $i$ corresponding to $x_{l}^{(i)}$ in class $j$. Similarly, if the distance between $x_{k}^{(i)}$ and $x_{l}^{(i)}$ is small, then its weight $u_{kl}^{(i,j)}$ will be close to 1; if the distance is large, $u_{kl}^{(i,j)}$ will be close to 0. Again, the sum of $u_{kl}^{(i,j)}$ for $M_j \left(x_{k}^{(i)}\right)$ is 1.

The extracted features (eigen vectors with largest eigenvalues) were determined by following matrix:

$$\left(S'_{nW} \right)^{-1} S'_{bW}.$$  

(7)

2) Hierarchical Knowledge Base With Time Constraint (HKB$^T$) Construction: We added a time limitation for the HKB construction based on our previous work [34].

Definition 1: Discrimination rule with time. A discrimination rule DR$^T$ is a 4-tuple defined as follows:

$$\text{DR}^T = \{F, K, R, T\}.$$  

(8)
where $F$ is the set of features, $K$ is the number of nearest neighbors corresponding to the largest $R$-value as determined through $k$-NN classification, $R \in [-1, 1]$ represents the agreement when using Cohen’s kappa [35] under $F$ and $K$, and $T$ is the elapsed time when the discrimination rule is used.

Definition 2: Hierarchical knowledge base with time. Let $\text{DR}_T^T$ be a discrimination rule. A hierarchical knowledge base $\text{HKB}^T$ is defined as a sequence:

$$\text{HKB}^T = \{\ldots, \text{DR}_i^T, \text{DR}_j^T, \ldots\}$$  \hspace{1cm} (9)$$

such that for $1 \leq i, j \leq n$, $R_i \geq R_j$, and $T_i \leq T_j$. $n$ is the number of discrimination rules.

From the above definitions, we can see that a hierarchical knowledge base with time is composed of all discrimination rules sorted by accuracy and time.

The construction algorithm is shown as follows. The Euclidean distance was used as the distance metric in the $k$-NN.

During the validation learning, the kappa coefficient was used to record the best ‘$K$’ corresponding to the highest kappa value. Note that the goal of this algorithm was not to learn a functional relationship, but a set of discrimination rules. Then, a hierarchical knowledge base $\text{HKB}^T$ was constructed by ranking the discrimination rules according to the kappa agreement ‘$R$’.

3) Classification Test: Since human decision-making in real life is always limited by available information, available time, and the information-processing ability of the mind, decisions inevitably require approximations, i.e., the decision-making is not done in full confidence. In order to quantify the degree of this confidence, we use a minimal confidence level (MCL) measure [34] in our classification test algorithm. The higher the MCL, the stricter the requirements for accepting a result as being correct, and vice versa. In order to evaluate the learning design, we compare the real sample label of an unseen sample to the classification result of our algorithm based on the best feature level ‘$F$’.

Algorithm 1: Supervised learning for $\text{HKB}^T$

Input: A learning set (LS) containing labeled samples; Total learning time (LT).
Output: A hierarchical knowledge base $\text{HKB}^T$.
begin
$\text{HKB}^T \leftarrow \emptyset$; $\text{SLT} \leftarrow \text{LT} / \# \text{ of samples in LS}$; Divide LS into $m$ subsets; for each subset $\text{LS}_m$ do
\{VTS\} $\leftarrow \text{LS}_m$; /* Initialize a validation test set. */ \{VLS\} $\leftarrow \text{LS} - \text{LS}_m$; /* Initialize a validation learning set. */ $F \leftarrow \emptyset$; /* Initialize a validation feature set. */ $i \leftarrow 1$; repeat
$F \leftarrow \{\text{VLS}\}$; /* Select a column data from VLS to $F$. */ $\text{ Tic;}$ /* Timing begin. */ $[l, \ldots] \leftarrow \text{KNN}(\text{VTS}, F)$; /* Classifies each row of the data in VTS into one of the groups in $F$ using k-NN classifier, $k$ from 1 to (\# of samples in VLS)/4. */ $R \leftarrow \text{Kappa} \{\text{labels of VTS}\}$; /* Calculate agreement using Cohen’s kappa. */ Save $K$ value corresponding to $R_{\text{max}}$. /* $T \leftarrow \text{Tic}$; /* Timing up, and save the elapsed time. */ $\text{DR}_{\text{best}} = <F, K, R, T>$; $i \leftarrow i + 1$; until ($F = \text{VLS}$) or ($T > \text{SLT}$); end
$\text{HKB}^T = \text{Sort}(\text{DR}_{\text{best}})$; /* Sort all $\text{DR}$ to $\text{HKB}^T$ in descending order of $R$ (primary key) and $T$ (secondary key). */ return $\text{HKB}^T$.
end

In Algorithm 2, for each test sample from the test set, we recursively classify the sample into the majority class and compute its membership using a $k$-NN classifier at a given single sample average test time. In each loop, the parameter $K$ and the “nearest” feature samples come from $\text{HKB}^T$ which is selected using a up-bottom search strategy. Finally, a label is assigned to the test sample satisfying the minimal confidence level (MCL) demand in the given single sample test time. Furthermore, the algorithm computes the classification accuracy for all recognized samples. For those test samples whose memberships are below the given MCL (the samples are called unrecognized samples), the algorithm computes the rejection rate.

In summary, a nonparametric weighted feature extraction method was employed for EEG signal feature extraction. Then through supervised learning, the discrimination rules along with the corresponding $k$-values were stored, and through ranking the discrimination rules according to validation tests, a $\text{HKB}^T$ was constructed (see Algorithm 1). Finally, given an unseen EEG sequence, the algorithm chose one or several
Fig. 2. Grand means of event-related potentials for critical items (solid lines) and noncritical items (dotted lines) on Fz, Cz, and Pz electrodes. The critical waveforms represent an average of 30 stimulus presentations, and the noncritical waveforms represent 120 stimulus presentations for each experimental run.

Fig. 3. Classification accuracy (black crosses) and runtimes (gray upward-pointing triangles) obtained for all possible combinations of channels across all subjects, different normalization techniques, and different MCL values (0 0.2 0.4 0.6 0.8 1).

Fig. 4. Classification accuracy obtained for different normalization techniques across different MCL values (0 0.2 0.4 0.6 0.8 1) from electrode Pz. Every trial was then winsorized for each subject. With a 90% Winsorization, the highest and the lowest 7–19 trials as outliers were replaced.

discrimination rules from the II1Kβ^2 using a top-down search strategy and calculated classification accuracy and rejection rate (see Algorithm 2). Classifiers used in the above two algorithms are all k-NN, but in Algorithm 2 the degree of membership of test sample p is also estimated. A higher value of p indicates a higher confidence level for the estimated label.
III. EXPERIMENTAL RESULTS

If the subjects have the relevant information (critical item) stored in their brain and they focus on the information present, the specific ERP component (P300) is different for the critical item compared to the noncritical item [2], [3], [9]. Hence we expected that the responses to the birth date year, regarded as the critical item in CIT, would contain a P300 component, whereas other responses would not. We used the average overall signal sampled from the onset of the stimulus (time 0) to the stimulus offset, and we did obtain the P300 wave. Fig. 2 presents the average EEG signals at electrode sites Fz, Cz, and Pz for critical and noncritical items across all subjects (20 experimental runs). The critical items and noncritical items were presented 30 and 120 times respectively for each experimental run and the resulting averaged waveforms are presented. It can be seen that P300 responses for critical items are significantly higher than that for noncritical items within a time window from 300 to 800 ms (Fz, t = 2.550, p < 0.001; Cz, t = 3.926, p < 0.001; Pz, t = 4.690, p < 0.001).

In order to show the performance of the proposed method, a 5-fold cross-validation method was employed, that is, each dataset from each subject was randomly divided into five mutually exclusive subsets, with one of the subsets being used as the test set and the other four subsets merged to form the learning set. This procedure was repeated five times and the average classification accuracy was reported. All the experiments of this section were carried out individually for all eleven subjects.

In Fig. 3, we report the average accuracy and runtimes (obtained using an Intel(R) Core(TM)2 Quad CPU Q9550 @2.83 GHz, 3.25 GB RAM) using 5-fold cross-validation for all possible combinations of channels across eleven subjects. It is noted that as the number of channels increases, the runtimes increase sharply from around 60 seconds to 175 seconds (learning time for each sample SLT is set to 10 ms, and test time for each sample STT is set to 0.5 ms), but the achieved accuracies across four normalization and nonnormalization techniques increased only marginally (from around 85% to 88%). Since the electrode Pz performed the best, we select Pz as the target channel in the following results. Another consideration is that one electrode would take less time to set up, which would be more user-friendly compared to multiple electrodes.

Fig. 4 shows the average classification accuracy obtained from the 5-fold cross-validation. Four normalization techniques (’minmax[−1, 1]’, ’minmax[0, 1]’, ’x-score’, and ’log’ [36]) were used and compared to nonnormalized data in our study. We found that the accuracy of using ’minmax[0, 1]’ was clearly better than that of using the other normalization techniques for all subjects.

Fig. 5 shows an example of the accuracy of classification for all subjects and for the different minimal confidence levels (MCL) (SLT = 10 ms and STT = 0.5 ms). We observed that, for most subjects, the higher the MCL, the higher the rejection rate (the number of unrecognized samples increased); however, lower MCL was prone to misclassification. Therefore, the highest classification performance was obtained for medium MCLs. Based on this observation, we set MCL to 0.4 and extended the learning time (LT) and test time (TT) to ensure that all samples were recognized, i.e., rejection rate (RR) equals 0. The results of this analysis are reported in Table I using electrode Pz as the target channel. From the above experimental results, we can draw several observations. As LT increased, discrimination rules were obtained because more features were included in Algorithm 1. In Algorithm 2, the longer the TT, the lower the RR because seeking the optimized discrimination rules from HKB2 usually required more CPU time. For MCL, the higher the parameter value, the more rigorous each classification was. Under the same HKB2 and MCL, a short test time for single sample testing increased the probability of incorrect classification. Also, as the number of samples for each subject
TABLE I
SUMMARY OF FINAL FIVE-FOLD CROSS-VALIDATED PERFORMANCES OVER ALL ELEVEN SUBJECTS

<table>
<thead>
<tr>
<th>Subject</th>
<th>Total samples</th>
<th>Valid samples</th>
<th>Size of HKB</th>
<th>Actual SLT (ms)</th>
<th>Actual STT (ms)</th>
<th>Accuracy (%)</th>
<th>CT (s)</th>
<th>Runtimes (s)</th>
<th>Actual runtimes (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>150</td>
<td>144</td>
<td>64.6</td>
<td>18.884</td>
<td>0.257</td>
<td>93.5±4.66</td>
<td>2.189</td>
<td>12.225</td>
<td>13.532</td>
</tr>
<tr>
<td>S2</td>
<td>300</td>
<td>278</td>
<td>40</td>
<td>19.787</td>
<td>0.175</td>
<td>96.7±3.73</td>
<td>4.410</td>
<td>24.068</td>
<td>27.391</td>
</tr>
<tr>
<td>S3</td>
<td>300</td>
<td>245</td>
<td>41.2</td>
<td>18.416</td>
<td>0.249</td>
<td>96.7±2.74</td>
<td>3.620</td>
<td>19.984</td>
<td>22.547</td>
</tr>
<tr>
<td>S4</td>
<td>150</td>
<td>115</td>
<td>75.2</td>
<td>18.953</td>
<td>0.145</td>
<td>100.0±0.00</td>
<td>1.747</td>
<td>9.688</td>
<td>10.641</td>
</tr>
<tr>
<td>S5</td>
<td>150</td>
<td>145</td>
<td>65</td>
<td>18.938</td>
<td>0.214</td>
<td>97.2±1.54</td>
<td>2.203</td>
<td>12.062</td>
<td>13.234</td>
</tr>
<tr>
<td>S6</td>
<td>300</td>
<td>200</td>
<td>50</td>
<td>18.674</td>
<td>0.160</td>
<td>96.0±3.35</td>
<td>2.994</td>
<td>16.643</td>
<td>18.469</td>
</tr>
<tr>
<td>S7</td>
<td>300</td>
<td>223</td>
<td>45</td>
<td>18.072</td>
<td>0.277</td>
<td>98.1±1.02</td>
<td>3.236</td>
<td>17.948</td>
<td>20.125</td>
</tr>
<tr>
<td>S8</td>
<td>150</td>
<td>150</td>
<td>62</td>
<td>19.000</td>
<td>0.221</td>
<td>97.2±1.54</td>
<td>2.287</td>
<td>12.902</td>
<td>14.188</td>
</tr>
<tr>
<td>S9</td>
<td>300</td>
<td>242</td>
<td>44.8</td>
<td>19.343</td>
<td>0.258</td>
<td>95.0±4.32</td>
<td>3.757</td>
<td>20.631</td>
<td>23.172</td>
</tr>
<tr>
<td>S10</td>
<td>300</td>
<td>269</td>
<td>40</td>
<td>19.561</td>
<td>0.298</td>
<td>89.4±2.15</td>
<td>4.226</td>
<td>23.097</td>
<td>26.188</td>
</tr>
<tr>
<td>s11</td>
<td>450</td>
<td>384</td>
<td>25</td>
<td>17.277</td>
<td>0.247</td>
<td>87.3±2.20</td>
<td>5.327</td>
<td>29.554</td>
<td>35.687</td>
</tr>
</tbody>
</table>

Average                                18.810  0.227  95.2±2.48  3.272  18.073  20.470

*CT = Valid samples × (ms−1/m) × Actual SLT + Valid samples × 1/m × Actual STT, m=5.
*Runtimes = Preprocessing runtimes + CT*m.
*Actual runtimes: obtained using an Intel(R) Core(TM)2 Quad CPU Q9550 @ 2.83GHz, 3.25GB RAM.

increased, longer learning was required. We set LT and TT to a large number as the initial value in order to ensure that our algorithms had enough learning and test times (SLT = 20 ms and STT = 5 ms). Parameters Actual SLT and Actual STT denote the actual learning time per sample and actual test time per sample, respectively. Table I shows the Actual SLT was no more than 20 ms, but STT was far less than 5 ms. After careful analysis of algorithm 2, we discovered that the test procedure stopped as soon as condition \( T > STT \) or \( q > \text{size}(HKB^T,1) \) or \( d > \text{MCI} \) was true. We found that in most cases the test loop stopped when \( q > \text{MCI} \). The computation time (CT) represents the total time cost for each subject. As seen in Table I, subject 4 had the best classification accuracy (ACC = 100% ± 0.00%) with the largest HKB\(^T\) size (75.2) and subject 11 had the lowest accuracy (ACC = 87.37% ± 2.20%) with the lowest HKB\(^T\) size (25). We achieved an overall accuracy of 95.23% ± 2.48 ms with CT = 3.272 seconds. The accuracy could be improved by adding more channels, however, as shown in Fig. 3, the accuracy would not improve much, while the time costs would increase sharply.

IV. CONCLUSION

Exhaustive searches for global optimization are inappropriate most of the time, while, fast or real-time decision-making based on limited information could be more useful and necessary in forensic investigations. In this work, we developed a simple and feasible hierarchical knowledge base construction and test method for a concealed information test (CIT) based on objective EEG measures. Based on a speed-accuracy trade-off, discrimination rules were learned and structured into a knowledge base in a given learning time, then during the classification stage, the rules were selected to make a decision for a novel sample given a certain test time while ensuring certain confidence levels. The proposed method was successfully applied to an autobiographical concealed information test paradigm with high classification accuracies and low computation time that would be acceptable for real-time applications requiring only one electrode. The present study suggests that further development is worthwhile, and would provide assistance to forensic investigations in the future.

ACKNOWLEDGMENT

The authors would like to thank Dr. V. Abootalebi and the Research Center of Intelligent Signal Processing (RCISP), Iran, for providing the data. They also would like to thank Dr. A. Khan and A. Reichelt, and the anonymous reviewers for their constructive comments and suggestions.

REFERENCES


