2014 IEEE Congress on Evolutionary Computation (CEC)
July 6-11, 2014, Beijing, China

Selecting the Optimal EEG Electrode Positions for a Cognitive Task using an Artificial Bee Colony with Adaptive Scale Factor Optimization Algorithm

Shreyasi Datta[1], Pratyusha Rakshit[2], Amit Konar[3]
Department of Electronics and Telecommunication Engineering
Jadavpur University
Kolkata, India
[1]shreyasidatta@gmail.com, [2]pratyushar1@gmail.com, [3]konaramit@yahoo.co.in

Atulya K. Nagar
Department of Mathematics and Computer Science
Liverpool Hope University
United Kingdom
nagara@hope.ac.uk

***Abstract*—The present work introduces a proposed Artificial Bee Colony with Adaptive Scale Factor (ABC-ASF) optimization algorithm-based optimal electrode selection strategy from which the acquired EEG signals enlighten the major brain activities involved in a cognitive task. In ABC-ASF, the scale factor for mutation in traditional Artificial Bee Colony is self adapted by learning from the previous experiences. Experimental results obtained from the real framework of estimating optimal electrodes indicate that the proposed algorithm outperforms other state-of-art techniques with respect to computational accuracy and run-time complexity.**

Keywords—artificial bee colony; electroencephalogram; independent component analysis; self adaptation

I. Introduction

Brain computer interface (BCI) technology [1] has revolutionized rehabilitation, assistance and treatment of neuro-motor, sensory-motor and cognitive disabilities. BCIs have been implemented using a number of different techniques that include invasive brain implants, partially invasive procedures like electrocorticography and non-invasive techniques like processing of brain signals that are captured through Electroencephalogram (EEG), magnetoencephalogram (MEG) or functional magnetic resonance imaging (fMRI). EEG has proved to be the most popularly used non-invasive interface because of its superior temporal resolution, easy acquisition, simple processing, portability, cost-effectiveness and freedom from exposure of the subject to intense magnetic fields. Several instances of EEG based BCI research has been found in literature [2-7].

EEG signals are generated from independent and localized sources through neuronal firing in the outer cortex of the brain and are recorded using electrodes placed on various locations on the scalp [8]. At each electrode the obtained EEG signal is a mixture of the effects of the signals from a number of sources. EEG for a task may be recorded using a number of electrodes placed at different locations; however, it may happen that signals from two or more EEG electrodes convey redundant information and hence only one out of these is to be considered. The selection of the most relevant electrodes for which non-overlapping information is obtained without any information being lost is necessary for faster signal processing in the real time scenario.

The present work proposes a novel evolutionary technique of EEG electrode selection specific to a particular cognitive activity. EEG signals available on the scalp are termed as sink signals henceforth, which are acquired using 14 scalp electrodes. The corresponding set of 14 source signals is estimated using Independent Component Analysis (ICA) [8-9]. The source signals as well as the sink signals (i.e., EEG signals) are represented using a feature space consisting of four standard EEG features, namely, Adaptive Autoregressive parameters [10], Hjorth parameters [11], Power Spectral Density [12-13] and Wavelet Coefficients [14]. The most important electrodes are found out using a proposed Artificial Bee Colony with Adaptive Scale Factor (ABC-ASF) algorithm with an aim to maximize the correlation between the set of source features for original and selected sink signals while minimizing the mutual information among each of the selected sink feature pairs.

In ABC-ASF, the scale factor used for food source (candidate solution) mutation is gradually self-adapted by learning from their previous experiences in producing improved locations of food sources. Specifically, at each generation, a set of scale factor values will be separately assigned to each individual bee (corresponding to a food source) of the current generation according to the selection probabilities learned from their previous generations. A comparison of ABC-ASF with some standard evolutionary algorithms like Artificial Bee Colony (ABC) [15], Differential Evolution (DE) [16] and one variant of ABC, known as Self-Adaptive Artificial Bee Colony (SAABC) [17] for solution of the present problem is provided.

The rest of the paper is structured as follows. The standard principles and methods are explained in Section II. An overview of the ABC algorithm is presented in Section III while the proposed ABC-ASF algorithm is described in Section IV. The experiments and results are explained in Section V. Finally, in Section VI conclusions are drawn.

II. PRINCIPLES AND METHODS

This section provides a schematic overview of the proposed scheme. Fig. 1 illustrates the basic steps of the present work.

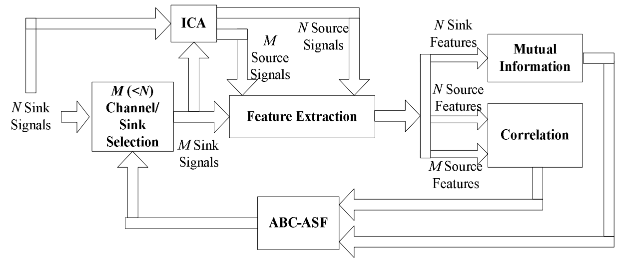


Fig. 1. The overview of the proposed scheme

A. Independent Component Analysis for EEG Source Estimation

Estimation of the EEG sources from the signals collected at the scalp electrodes has long been tried to solve through a wide variety of techniques [18-19]. Independent Component Analysis (ICA) [8-9], [20-21] is a method that can estimate the original signals from their linear transformation provided that the transformed data is non-gaussian and the estimated signals are independent.

Suppose there are N signal sources $x_i(t)$ within the brain giving rise to an equal number of sinks $y_i(t)$ on the scalp as shown in Fig. 2, for i,j=[1, N]. Let the time series representing the source as well as the sink signals be discretized with the sampling frequency $1/T$ at n integer points k=[1, n]. Assuming negligible propagation delay between the sources and the sinks, the linear transformation relating the discrete source and the sink EEG signals can be described by (1), where $x_i(kT)$ denotes the i-th source and $y_i(kT)$ denotes the i-th sink, for i=[1, N], while a_{ij} for i,j=[1, N] denote parameters that establish the relation between the above mentioned signals and depend on the distance between the sources and the sinks.

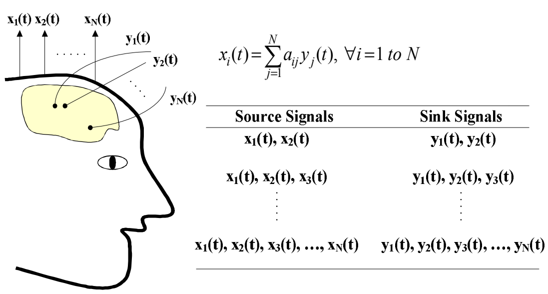


Fig. 2. Distribution of the N source $x_i(t)$ - sink $y_j(t)$ pairs and the linear relation between them in terms of the transformation parameters a_{ij} for i,j=[1, N]

$$
\begin{aligned}
y_1(kT) &= a_{11}x_1(kT) + a_{12}x_2(kT) + \ldots + a_{1N}x_N(kT) \\
y_2(kT) &= a_{21}x_1(kT) + a_{22}x_2(kT) + \ldots + a_{2N}x_N(kT) \\
&\vdots \\
&\vdots \\
y_N(kT) &= a_{N1}x_1(kT) + a_{N2}x_2(kT) + \ldots + a_{NN}x_N(kT)
\end{aligned}
\tag{1}
$$

A matrix form representation of (1) is presented in (2) where $\boldsymbol{A}$ with elements a_{ij} is called the 'mixing matrix', $x=[x_1(t),\ x_2(t),\ldots\ x_n(t)]$ and $y=[y_1(t),\ y_2(t),\ldots\ y_n(t)]$ for a particular time index.

$$y = \mathbf{A}x \tag{2}$$

ICA involves estimation of the mixing matrix **A** assuming the sources to be independent, computation of its inverse **W** (the 'demixing' matrix) and evaluating the sources from (3).

$$x = \mathbf{W}y \tag{3}$$

ICA estimation is done through the maximization of non-gaussianity that can be measured through kurtosis, negentropy, mutual information or maximum likelihood. In the present work ICA is performed on the acquired EEG data using the Infomax algorithm [20] that maximizes the entropy of the estimated source signals.

B. Feature Extraction

EEG signals of the sinks as well as the sources have been represented using four standard features used in BCI problems—Adaptive Autoregressive (AAR) Parameters [10], Hjorth Parameters [11], Power Spectral Density (PSD) [12-13] and Wavelet Coefficients [14].

C. Selection of best Electrodes/Sinks by Evolutionary Algorithm

In the present context, the objective function formulation to select M best representative sinks from the feature space of N source-sink signals by means of evolutionary algorithm is inspired by two crucial observations:

a) The Mutual Information (MI) between the features two selected sinks must be minimized.

According to information theory, mutual information (MI) measures the independence of random variables [22-23]. From the point of view of the present problem, in order to maximize the separability of information between two selected sinks such that only maximally distinct sinks are obtained, the mutual information between these sink features has to be minimized. For two sink feature spaces A_1 and A_2 mutual information $MI(A_1;A_2)$ is computed using (4), where, average information or entropy $H(A_1)$ is the uncertainty in the features of the first sink before observing the second sink and conditional entropy $H(A_1|A_2)$ is the uncertainty in the features of the first sink after observing the second sink.

If there are n samples in an observation, then $H(A_1)$ and $H(A_1|A_2)$ are given by (5) and (6), where $P(a_{1i})$ is the probability of each sample in A_1, $P(a_{1i},a_{1j})$ is the joint probability of that in A_1 and A_2 and $P(a_{1i}|a_{2j})$ is the transition probability of that from A_1 to A_2 and all cases are considered equally likely.

$$MI(A_1;A_2) = H(A_1) - H(A_1 \mid A_2) \tag{4}$$

$$H(A_1) = -\sum_{i=1}^{n} P(a_{1i}) \times \log_2 P(a_{1i}) \quad (5)$$

$$H(A_1 \mid A_2) = -\sum_{i=1}^{n} P(a_{1i}, a_{2j}) \times \log_2 P(a_{1i} \mid a_{2j}) \quad (6)$$

b) The correlation between the features of the source signals estimated from the selected sink signals and that of the original number of sink signals must be maximized.

Another factor is the linear correlation coefficient [24] between the set of reduced source features and the original number of source features. In statistics, the linear correlation coefficient in the range [-1, 1] between two random variables is a measure of the strength and the direction of a linear relationship between them. +1 and -1 values respectively stand for perfect positive and perfect negative correlations where as the absence of any correlation between the variables produce a correlation coefficient having near zero value. The correlation coefficient between two source features B_1 and B_2 each of length n is computed using (7).

$$R = \frac{n\sum B_1 B_2 - (\sum B_1)(\sum B_2)}{\sqrt{n(\sum B_1)^2 - (\sum B_2)^2}\sqrt{n(\sum B_2)^2 - (\sum B_1)^2}} \quad (7)$$

A larger value of correlation coefficient between the sets of source features of original and selected sink signals indicate an efficient representation of the original number of sink signals with the reduced number of sinks signals. In other words, on successful selection of the electrodes, the absolute value of the correlation between these two above mentioned entities must be close to unity (indicating the selected and the original set of source features covary perfectly, either positively or negatively) while a low absolute value of the same concludes high degree of disassociation (between the features of selected and original group of source signals).

With these two considerations, an objective function formulated in (8) needs to be minimized with the evolutionary algorithm. The minimization is effected over all the subjects under consideration. Here *S, N* and *M* denote the total number of subjects, original number of sinks and selected number of sinks respectively. $MI(A_j, A_k)$ denotes the mutual information between the features of the j-th and k-th selected sink signals while $R(B_j, B_k)$ denotes the correlation coefficient between features of the source signals corresponding to the j-th selected sink and k-th original sink. λ denotes a constant which is set in a manner so as to have all terms in the right hand side of (8) in the same order of magnitude. In our experiments λ is selected to be 20.

$$f = \sum_{S} (\lambda \sum_{j=1}^{M} \sum_{\substack{k=1 \\ k \neq j}}^{M} MI(A_j, A_k) - \sum_{j=1}^{M} \sum_{k=1}^{N} | R(B_j, B_k) |) \quad (8)$$

III. An Overview of Artificial Bee Colony (ABC) Algorithm

In ABC [15], the nectar amount of a food source corresponds to the fitness of the associated solution. The number of employed bees and onlooker bees is equal to the number of solutions in the population.

A. Initialization

ABC starts with a population of *NP*, *D*-dimensional food sources (candidate solutions) $\vec{X}_i(t) = \{x_{i,1}(t), x_{i,2}(t), x_{i,3}(t), \ldots x_{i,D}(t)\}$ at generation t=0 for i= [1, *NP*] within the search range $[\vec{X}_{\min}, \vec{X}_{\max}]$ where $\vec{X}_{\min} = \{x_{\min-1}, x_{\min-2}, \ldots, x_{\min-D}\}$ and $\vec{X}_{\max} = \{x_{\max-1}, x_{\max-2}, \ldots, x_{\max-D}\}$. Each food source $\vec{X}_i(t)$ is assigned a fitness value (nectar amount) $fit(\vec{X}_i(t))$.

B. Employed Bee Phase

An employed bee produces a modification $\vec{X}'_i(t) = \{x_{i,1}(t), \ldots x'_{i,j}(t), \ldots x_{i,D}(t)\}$ on the food source position in her memory $\vec{X}_i(t) = \{x_{i,1}(t), \ldots x_{i,j}(t), \ldots x_{i,D}(t)\}$ depending on the local information as stated by (9) and tests $fit(\vec{X}'_i(t))$.

$$x'_{i,j}(t) = x_{i,j}(t) + F \times (x_{i,j}(t) - x_{k,j}(t)) \quad (9)$$

Here F is the scale factor in [-1, 1], j is a randomly chosen index from [1, *D*] and k is any number between 1 to *NP* but not equal to i. Provided that $fit(\vec{X}'_i(t)) > fit(\vec{X}_i(t))$, the bee memorizes $\vec{X}'_i(t)$ and forgets $\vec{X}_i(t)$.

C. Calculation of Probability of Food Source Selection

The probability of each food source $\vec{X}_i(t)$ to be selected by the onlooker bee is proportional to the nectar amount $fit(\vec{X}_i(t))$ of $\vec{X}_i(t)$ and is given by (10).

$$prob(i) = \frac{fit(\vec{X}_i(t))}{\sum_{j=1}^{NP} fit(\vec{X}_j(t))} \quad (10)$$

D. Onlooker Bee Phase

An onlooker bee then selects a food source $\vec{X}_i(t)$ depending on the associated probability, $prob(i)$ as calculated by (10). After that, as in case of employed bee, onlooker bee produces a modification on the food source in her memory as described in section III-B and checks the nectar amount of the candidate source. Providing that its fitness is better than that of the previous one, bee remembers the new position and forgets the old one.

E. Scout Bee Phase

In the ABC algorithm, if a food source cannot be improved further through a predefined number of cycles called 'limit', the food source is abandoned and is reinstated by the scouts by randomly producing a position.

After each evolution, we repeat from step B until the termination condition is satisfied.

IV. Proposed Artificial Bee Colony with Adaptive Scale Factor (ABC-ASF)

In ABC-ASF, a candidate pool of scale factors, denoted as $\{F_1, F_2, \ldots, F_n\}$ is maintained. Prior to the mutation operation (as in (9)) corresponding to each food source, a scale factor F_i is selected from the candidate pool according to its success probabilities gradually learned from past experience to generate improved neighborhood food source locations. The more successfully one scale factor F_i behaves in the previous generations to produce quality food source positions, the more probably F_i will be selected for mutation operation in the current generation. It is realized in ABC as follows.

1. Let $p_{l,t}$ denotes the probability of selecting the scale factor F_l from the candidate pool $\{F_1, F_2, \ldots, F_n\}$ at generation t for $l = [1, n]$.
2. At t=0, $p_{l,t}$ is initialized with $1/n$ signifying that all the scale factors in the candidate pool have the equal probability of selection.
3. The number of neighborhood food sources successfully replacing the original food sources after mutation with F_l (by employed/onlooker bee) at generation t is recorded as $s_{l,t}$. The value is stored in a *success memory*.
4. The number of neighborhood food sources discarded with respect to the original food sources after mutation with F_l (by employed/onlooker bee) at generation t is recorded as $f_{l,t}$. The value is stored in a *failure memory*.
5. Steps 3 and 4 are repeated for a fixed number of generations, known as *Learning Period LP*. It is pictorially represented in Fig. 3(a) for success and failure memory of dimension $LP \times n$.
6. When $t>LP$, steps 3 and 4 are also performed. However, the rows corresponding to the rows s_{t-LP} and f_{t-LP} will be removed from the respective memories leaving room for the newly available $s_{l,t}$ and $f_{l,t}$. It is elaborated in Fig. 3(b).
7. When $t>LP$, the probabilities of choosing F_l will be updated at each subsequent generation based on the success and failure memories as follows

$$p_{l,t} = S_{l,t} \Big/ \sum_{l=1}^{10} S_{l,t}\,, \quad l = [1,10], t > LP \tag{11}$$

$$S_{l,t} = \frac{\sum_{g=t-LP}^{t-1} s_{l,g}}{\sum_{g=t-LP}^{t-1} s_{l,g} + \sum_{g=t-LP}^{t-1} s_{l,g}} + \varepsilon \tag{12}$$

where $S_{k,t}$ represents the success rate of the neighborhood sources generated by F_l within the previous LP generations with respect to generation t. The small constant value ε =0.01 is used to avoid the possible null success rates. Obviously, a high value of the success rate for the F_l within the previous LP generations leads to the larger probability of applying it for mutation at the current generation.

index	F_1	F_2		F_n
1	$s_{1,t-LP}$	$s_{2,t-LP}$	....	$s_{n,t-LP}$
2	$s_{1,t-LP+1}$	$s_{2,t-LP+1}$	....	$s_{n,t-LP+1}$
:	:	:	....	:
LP	$s_{1,t-1}$	$s_{2,t-1}$	....	$s_{n,t-1}$

Success Memory

index	F_1	F_2		F_n
1	$f_{1,t-LP}$	$f_{2,t-LP}$	....	$f_{n,t-LP}$
2	$f_{1,t-LP+1}$	$f_{2,t-LP+1}$	....	$f_{n,t-LP+1}$
:	:	:	....	:
LP	$f_{1,t-1}$	$f_{2,t-1}$	....	$f_{n,t-1}$

Failure Memory

(a)

$s_{1,t-LP}$	....	$s_{n,t-LP}$
:	....	:
$s_{1,t-1}$	....	$s_{n,t-1}$

$s_{1,t-LP+1}$	....	$s_{n,t-LP+1}$
:	....	:
$s_{1,t}$	....	$s_{n,t}$

$s_{1,t-LP+2}$	....	$s_{n,t-LP+2}$
:	....	:
$s_{1,t+1}$	....	$s_{n,t+1}$

(b)

Fig. 3. (a) Success Memory and Failure Memory (b) Progress of Success Memory

The pseudo code of ABC-ASF is given below.

Procedure ABC-ASF

Begin

1. Initialize a population P_t of NP, D-dimensional food sources $\vec{X}_i(t)$ at generation t=0 with $trial_i$=0 and evaluate $fit(\vec{X}_i(t))$ for i= [1, NP].

2. While termination condition is not reached **do begin**

- 2.1. **If** $t>LP$ **Then** calculate $p_{l,t}$ using (11) and (12) and remove $s_{l,t-LP}$ and $f_{l,t-LP}$ from success and failure memories respectively for l= [1, n].
  End If

//Employed Bee Phase

- 2.2. Using stochastic universal sampling, select a scale factor F_l from the candidate pool $\{F_1, F_2, \ldots, F_n\}$ corresponding to each food source $\vec{X}_i(t)$ for i= [1, NP].
- 2.3. Produce a new food source $\vec{X}'_i(t)$ using F_l following (9) for i= [1, NP].
- 2.4. Evaluate $fit(\vec{X}'_i(t))$ for i= [1, NP].
- 2.5. **If** $fit(\vec{X}'_i(t)) > fit(\vec{X}_i(t))$
  Then $\vec{X}_i(t+1) \leftarrow \vec{X}'_i(t)$; $trial_i \leftarrow 0$; $s_{l,t} \leftarrow s_{l,t}+1$;
  Else $trial_i = trial_i + 1$; $f_{l,t} \leftarrow f_{l,t}+1$;
  End If
  Repeat the step for i= [1, NP].

//Onlooker Bee Phase

- 2.6. Select a food source $\vec{X}_i(t) \in P_t$ based on its probability of selection $prob(i)$ for calculated using (10) for i= [1, NP].
- 2.7. Repeat from 2.1 to 2.5.
- 2.8. Reinitialize the food source with highest *trial* value exceeding "limit" by the **scout bee**.
- 2.9. $t \leftarrow t+1$.

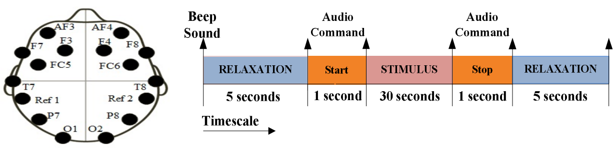
End While

End

V. Experiments and Results

A. EEG Acquisition

EEG channel selection can be done in response to different types of external stimuli. In the present work we have undertaken two case studies — 1) an audio-visual stimulus that consisted of a length of video to be watched along with audio and 2) according to a visual stimulus the subject has to move his hand in right or left directions. Experimental data is acquired in the laboratory from 5 subjects, 2 male and 3 female, in the age group 25±5 years, with their consent. EEG is acquired using a 14 channel Emotiv headset [25] which has a sampling rate 128Hz. The placement of electrodes follows the standard 10/20 system of electrode placement [26] and is shown in Fig. 4(a). Each experiment was conducted while the subject sat comfortably during EEG recording and responded according to the stimulus. Periods of relaxation were included in between the several instances of stimuli for EEG recording. A sample data acquisition queue for an experiment is illustrated in Fig. 4(b).



(a) (b)

Fig. 4. (a) Electrode Placement (b) Queue for Stimulus Presentation

B. Pre-processing and Source Signal Estimation

From the study of the EEG power spectrum that spans different frequency bands [27], it is seen that the stimuli produce significant changes in the 4-30 Hz range. To extract the EEG signals in the desired frequency range and thereby eliminate the other frequencies, an Elliptical Band pass filter of order 6 with 1dB pass-band ripple and 50dB stop-band ripple in the bandwidth 4-30Hz has been used. EEG source signals are extracted using the ICA algorithm based on Infomax principle [20] from the EEG sink signals selected in each iteration of the evolutionary process.

C. Feature Extraction

Four standard features have been extracted from the source as well as the sink signals. AAR parameters of order 6 have been extracted using Kalman Filtering [28] as the estimation algorithm. The update coefficient of the AAR estimation has been taken as 0.0085. PSD has been computed with Welch method [29] with 50% overlap between the signal segments using a Hamming Window at the integer frequency points between 4-30Hz. For extracting the wavelet coefficients Daubechies order 4 mother wavelet has been used. For each source/sink EEG signal an instance of duration 30 seconds has been considered such that the dimensions of the four feature spaces are 6, 3, 27 and 965 respectively for each EEG channel. The proposed ABC-ASF has been tested with all the possible combinations of these four features after normalization with their respective maximum values.

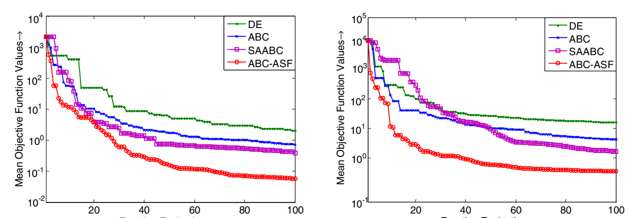
D. Selection of best electrodes/ Sink signals

For selecting the best electrodes using ABC-ASF, a candidate solution is encoded as a food source of dimension M ($<N$) representing the indices of the M selected electrodes out of the N. In each iteration of ABC-ASF the process of ICA for source signal estimation and feature extraction from the selected sink signals and corresponding source signals have been performed to evaluate the objective function in (8).

E. Experimental Results

The performance of the proposed ABC-ASF is examined here with respect to minimizing the objective in (8). Here, we compare ABC-ASF with traditional ABC [15], Differential Evolution (DE/current-to-best/1) [16] and Self Adaptive Artificial Bee Colony (SAABC) [17] algorithms. For each algorithm, the population size is kept at 50. We employ the best parametric set-up for all these competitor algorithms as prescribed in their respective sources. For the proposed ABC-ASF algorithm, we have selected the size of the candidate pool of scale factors n= 10 and learning period LP= 50.

We plot the mean objective function values taken over 50 runs versus Function Evaluations (FEs) in Fig. 5, and observe that ABC-ASF outperforms all other algorithms in terms of the relative speed of convergence and quality of solution. In Tables I and II a comparison of ABC-ASF with the other three algorithms has been made while varying the number of selected electrodes M for the respective case studies considering the combination of all four features. The mean and standard deviation (in parenthesis) of the cost functions for 50 independent runs are provided. The statistical significance levels (SS) of the difference of the means of best two algorithms using t-test are reported, with maximum FEs as 500. Here "+" indicates that the t value of 49 degrees of freedom is significant at a 0.05 level of significance by two-tailed test, whereas "–" means the difference of mean is not statistically significant, and "NA" stands for not applicable, covering cases for which two or more algorithms achieve the best accuracy results. The best algorithm is marked in bold. In Case Study I, ABC-ASF outperforms its competitors in 7 cases out of 10 in a statistically significant manner. In two cases (M=10 and M=12) SAABC has better results than ABC-ASF, being the second best algorithm. In Case Study II, ABC-ASF outperforms its competitors in all of the 10 cases in a statistically significant manner.



(a) (b)

Fig. 5. Plots of Mean Objective Function Values vs. FEs (a) Case Study I and (b) Case Study II

With a view to find the smallest number of electrodes necessary in the recognition of a task from EEG signals, from the observations of the mean values of the cost functions in Table I and Table II, it is found that for Case Study 1 a minimum of 5 electrodes provides the best results with ABC-ASF while 4 electrodes are sufficient for Case Study 2. Tables III and IV are constructed to focus on the performance of ABC-ASF with the best solutions (5 electrodes for Case Study I and 4 for Case Study II) and different combinations of signal features. In Tables III and IV, the mean values of the total sum of Correlation Coefficient over all pairs of selected and original source features over all subjects, that of Mutual Information over all pairs of selected and original sink features over all subjects and the Cost Function, along with the corresponding standard deviation values (within parenthesis) for 50 independent runs for each combination of feature sets have been reported for Case Study I and Case Study II respectively, taking maximum FEs as 500.

TABLE I. PERFORMANCE OF THE PROPOSED ABC-ASF WITH OTHER STATE-OF-ART ALGORITHMS FOR CASE STUDY-1

M	ABC-ASF	SAABC	ABC	DE	SS
4	**0.0539 (0.001)**	0.0838 (0.019)	0.1656 (0.030	0.3516 (0.054)	+
5	**0.0119 (0.033)**	0.0781 (0.044)	0.2629 (0.049	0.3804 (0.057)	+
6	**0.0758 (0.042)**	0.1523 (0.058)	0.1621 (0.073	0.5497 (0.135)	+
7	**0.1299 (0.072)**	0.2289 (0.095)	0.3112 (0.103	0.2858 (0.147)	+
8	**0.3371 (0.095)**	**0.3371 (0.095)**	0.5285 (0.159	0.5678 (0.166)	NA
9	**0.4505 (0.140)**	0.4693 (0.172)	0.5852 (0.195	0.6019 (0.309)	+
10	0.5383 (0.196)	**0.5307 (0.160)**	0.6540 (0.219	0.7537 (0.320)	–
11	**0.5688 (0.205)**	0.6892 (0.231)	0.7572 (0.326	0.8258 (0.360)	+
12	0.7791 (0.347)	**0.7481 (0.248)**	0.8308 (0.347	0.9133 (0.361)	–
13	0.7942 (0.341)	0.9171 (0.364)	0.9340 (0.377	0.9961 (0.384)	+

TABLE II. PERFORMANCE OF THE PROPOSED ABC-ASF WITH OTHER STATE-OF-ART ALGORITHMS FOR CASE STUDY-2

M	ABC-ASF	SAABC	ABC	DE	SS
4	**0.0254 (0.044)**	0.0285 (0.051)	0.0780 (0.069)	0.1135 (0.072)	+
5	**0.0275 (0.055)**	0.1015 (0.060)	0.1260 (0.090)	0.1369 (0.141)	+
6	**0.0369 (0.082)**	0.2227 (0.092)	0.3137 (0.094)	0.3374 (0.162)	+
7	**0.0777 (0.094)**	0.3883 (0.095)	0.4375 (0.102)	0.5243 (0.164)	+
8	**0.2215 (0.125)**	0.5058 (0.129)	0.5245 (0.181)	0.6337 (0.187)	+
9	**0.2536 (0.175)**	0.5429 (0.184)	0.6402 (0.202)	0.6517 (0.239)	+
10	**0.5558 (0.216)**	0.5945 (0.227)	0.7246 (0.258)	0.7325 (0.262)	+
11	**0.5648 (0.242)**	0.6061 (0.277)	0.7307 (0.279)	0.7657 (0.301)	+
12	**0.6587 (0.251)**	0.6793 (0.283)	0.7660 (0.311)	0.7675 (0.329)	+
13	**0.7471 (0.294)**	0.7601 (0.343)	0.7719 (0.354)	0.7764 (0.355)	+

To view the brain activation areas for the selected electrodes, the independent components or sources obtained by using ICA are illustrated in Fig. 6 and Fig. 7 respectively for Case Study I and II. In each of these figures, (a) provides the independent components estimated for all the 14 electrodes, while (b) provides the independent components obtained for the selected electrodes. In Case Study I an audio-visual stimulus is presented, and hence the selected electrodes lie in the occipital, temporal, frontal and prefrontal regions [30] that is evident from the selected electrodes as well as the corresponding independent components as shown in Fig. 6. In Fig. 6(b) it is clearly visible that ABC-ASF provides the best selection such that five completely distinct source activations are obtained in comparison with the other algorithms. Similar observations are found for Case Study II in Fig. 7. Here the selected electrodes span the occipital, parietal and motor cortex regions having accordance with the active regions during motor imagination [30-31] with visual stimulus.

TABLE III. PERFORMANCE OF THE PROPOSED ABC-ASF ALGORITHM FOR CASE STUDY-1 WITH NO. OF SELECTED ELECTRODES=5

Features	Correlation Coefficient	Mutual Information	Cost Function
AR +Hjorth	275.273 (0.068)	13.7654 (0.014)	0.0349 (0.029)
Hjorth +PSD	283.086 (0.002)	14.1557 (0.029)	0.0268 (0.038)
AR +PSD	288.666 (0.046)	14.4347 (0.034)	0.0276 (0.047)
Wavelet +Hjorth	290.229 (0.060)	14.5127 (0.047)	0.0253 (0.084)
Wavelet +AR	301.698 (0.049)	15.0862 (0.083)	0.0255 (0.064)
Wavelet +PSD	315.597 (0.098)	15.7806 (0.097)	0.0147 (0.090)
AR +Hjorth +PSD	293.294 (0.078)	14.6653 (0.106)	0.0128 (0.044)
Wavelet +AR +Hjorth	317.362 (0.007)	15.8695 (0.025)	0.0291 (0.072)
Wavelet +Hjorth +PSD	322.546 (0.037)	16.1286 (0.031)	0.0256 (0.074)
Wavelet +AR +PSD	324.637 (0.045)	16.2325 (0.092)	0.0138 (0.028)
Wavelet +AR +PSD +Hjorth	331.524 (0.083)	16.5767 (0.094)	0.0119 (0.033)

TABLE IV. PERFORMANCE OF THE PROPOSED ABC-ASF ALGORITHM FOR CASE STUDY-2 WITH NO. OF SELECTED ELECTRODES=4

Features	Correlation Coefficient	Mutual Information	Cost Function
AR +Hjorth	163.155 (0.003)	8.1591 (0.005)	0.0285 (0.001)
Hjorth +PSD	172.837 (0.007)	8.6536 (0.010)	0.0235 (0.004)
AR +PSD	179.377 (0.018)	8.9701 (0.012)	0.0250 (0.012)
Wavelet +Hjorth	180.559 (0.059)	9.0290 (0.017)	0.0216 (0.014)
Wavelet +AR	182.980 (0.084)	9.1503 (0.030)	0.0256 (0.019)
Wavelet +PSD	204.040 (0.057)	10.2031 (0.035)	0.0214 (0.023)
AR +Hjorth +PSD	184.170 (0.029)	9.2097 (0.038)	0.0238 (0.024)
Wavelet +AR +Hjorth	183.236 (0.082)	9.1629 (0.045)	0.0218 (0.027)
Wavelet +Hjorth +PSD	197.773 (0.048)	9.8901 (0.047)	0.0281 (0.028)
Wavelet +AR +PSD	204.093 (0.045)	10.2061 (0.070)	0.0299 (0.028)
Wavelet +AR +PSD +Hjorth	217.310 (0.069)	10.8668 (0.070)	0.0254 (0.044)

VI. CONCLUSIONS AND DISCUSSIONS

The paper proposes a novel evolutionary approach for estimating the best electrode positions to be selected for EEG analysis of a particular cognitive task. The work involves source signal estimation from acquired EEG signals by ICA, feature extraction to represent EEG signals by four standard features and using an evolutionary technique to determine the best electrodes by minimizing the difference of mutual information between the EEG features of each pair of the selected electrodes and the correlation coefficient between the source signal features estimated from the selected electrodes and that of the original electrodes. The novelty in the proposed evolutionary approach is the use of a self adapted mutation scale factor in traditional Artificial Bee Colony algorithm. The proposed strategy when compared with other three traditional optimization procedures yields better results in two case studies in the present context. Future works in this direction include determination of minimum number of EEG electrodes specific to a large number of well defined activities using the proposed scheme.

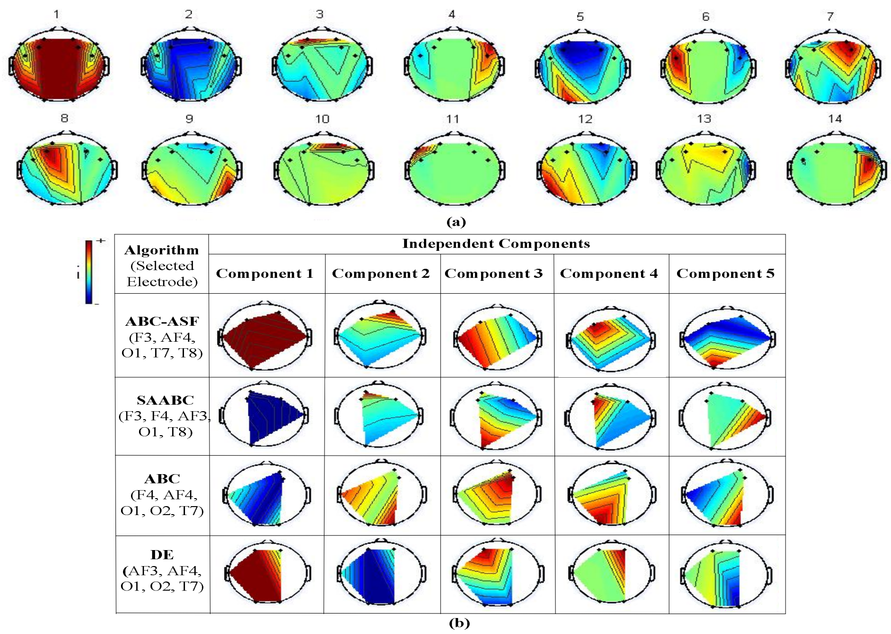


Fig. 6. Independent Components for Case Study I from Subject 1 (a) for 14 electrodes (b) For reduced number of electrodes

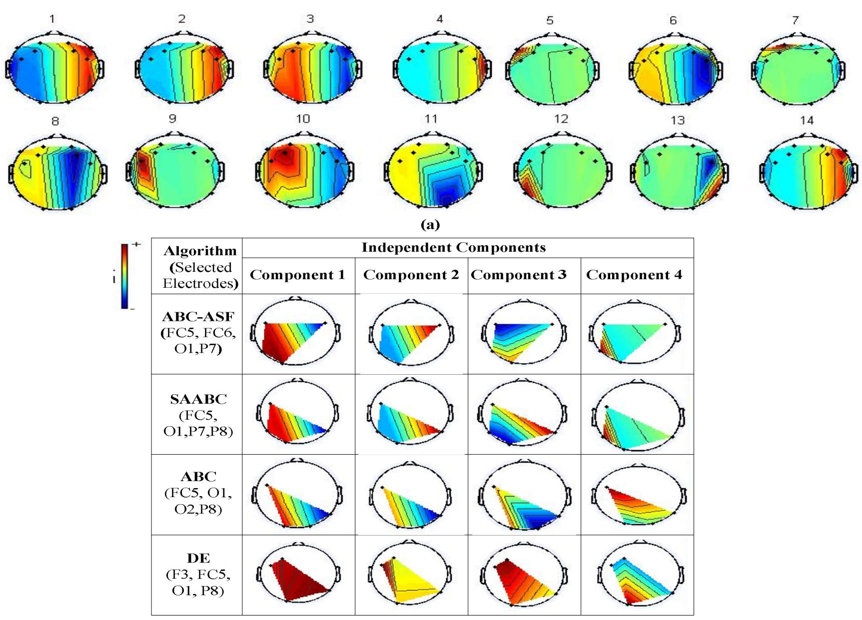


Fig. 7. Independent Components for Case Study II from Subject 1 (a) for 14 electrodes (b) For reduced number of electrodes. The vertical colour bar indicates intensity/activation levels in Fig. 6 and Fig. 7. The electrode locations are given by black points.

ACKNOWLEDGMENT

Funding by Council of Scientific and Industrial Research (CSIR), India (for awarding Senior Research Fellowship to the second author) and funding by University Grants Commission, India, University of Potential Excellence Program (UGC-UPE) (Phase II) in Cognitive Science, Jadavpur University are greatly acknowledged for the present work.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," Clinical neurophysiology, vol. 113, no. 6, pp.767-791, 2002.

[2] A. Vuckovic, V. Radivojevic, A. C. Chen, and D. Popovic, "Automatic recognition of alertness and drowsiness from EEG by an artificial neural network," Medical Engineering & Physics, vol. 24, no. 5, pp. 349-360, 2002.

[3] A. Gevins, and M. E. Smith, "Detecting transient cognitive impairment with EEG pattern recognition methods," Aviation, space, and environmental medicine, vol. 70, no. 10, pp. 1018-1024, 1999.

[4] G. Pfurtscheller, and C. Neuper, "Motor imagery and direct brain-computer communication," Proceedings of the IEEE, vol. 89, no. 7, pp. 1123-1134, 2001.

[5] C. Guger, W. Harkam, C. Hertnaes, and G. Pfurtscheller, "Prosthetic control by an EEG-based brain-computer interface (BCI)," in Proc. aaate 5th european conference for the advancement of assistive technology, pp. 3-6, November 1999.

[6] K. Schaaff and T. Schultz, "Towards an EEG-based emotion recognizer for humanoid robots," in The 18th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2009, pp. 792-796, September 2009.

[7] A. Khasnobish, A. Konar, D. N. Tibarewala, S. Bhattacharyya, and R. Janarthanan,"Object Shape Recognition from EEG Signals during Tactile and Visual Exploration," in Pattern Recognition and Machine Intelligence, Springer Berlin Heidelberg, pp. 459-464, 2013.

[8] L. Zhukov, D. Weinstein, and C. Johnson, "Independent component analysis for EEG source localization," Engineering in Medicine and Biology Magazine, IEEE, vol. 19, no. 3, pp. 87-96, 2000.

[9] S. Makeig, A. J. Bell, T. P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," Advances in neural information processing systems, pp. 145-151, 1996.

[10] G. Pfurtscheller, C. Neuper, A. Schlogl, and K. Lugger, "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters," IEEE Transactions on Rehabilitation Engineering, vol. 6, no. 3, pp. 316-325, 1998.

[11] C. Vidaurre, N. Krämer, B. Blankertz, and A. Schlögl, "Time domain parameters as a feature for EEG-based brain–computer interfaces," Neural Networks, vol. 22, no. 9, pp. 1313-1319, 2009.

[12] F. Cona, M. Zavaglia, L. Astolfi, F. Babiloni, and M. Ursino, "Changes in EEG power spectral density and cortical connectivity in healthy and tetraplegic patients during a motor imagery task," Computational intelligence and neuroscience, 2009.

[13] J. F. D. Saa, and M.S. Gutierrez, "EEG Signal Classification Using Power Spectral Features and linear Discriminant Analysis: A Brain Computer Interface Application," 2010.

[14] P. Jahankhani, V. Kodogiannis, K. Revett, "EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks," in IEEE John Vincent Atanasoff International Symposium on Modern Computing, pp.120-124, 2006.

[15] B. Basturk, D. Karaboga, "An artificial bee colony (ABC) algorithm for numeric function optimization," In: Proceedings of the IEEE Swarm Intelligence Symposium 2006, Indianapolis, Indiana, USA, 12-14 May 2006.

[16] R. Storn, K. V. Price, and J. Lampinen, "Differential Evolution–A Practical Approach to Global Optimization", Berlin, Germany: Springer- Verlag, 2005.

[17] W. Gu, M. Yin, and C. Wang, "Self Adaptive Artificial Bee Colony for Global Numerical Optimization," IERI Procedia 1, pp. 59-65, 2012.

[18] R. D. Pascual-Marqui, "Review of methods for solving the EEG inverse problem," International journal of bioelectromagnetism, vol. 1, no.1, pp. 75-86, 1999.

[19] R. Grech, T. Cassar, J. Muscat, K. Camilleri, S. Fabri, M. Zervakis, ... and B. Vanrumste, "Review on solving the inverse problem in EEG source analysis," Journal of neuroengineering and rehabilitation, vol. 5, no. 1, pp. 25, 2008.

[20] A. J. Bell, and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," Neural computation, vol. 7, no. 6, pp. 1129-1159, 1995.

[21] A. Hyvärinen, and E. Oja, "Independent component analysis: algorithms and applications," Neural networks, vol. 13, no. 4, pp. 411-430, 2000.

[22] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," Physical Review E, 69(6), 066138, 2004.

[23] R. Steuer, J. Kurths, C. O. Daub, J. Weise, and J. Selbig, "The mutual information: detecting and evaluating dependencies between variables," Bioinformatics, 18(suppl 2), S231-S240, 2002.

[24] N. G. Das, Statistical Methods,TataMcGrawHill, 2008.

[25] http://www.emotiv.com/eeg/

[26] G. Dornhege, Towards Brain-Computer Interfacing, MIT Press, 2007.

[27] M. Teplan, "Fundamentals of EEG Measurement," J. Measurement Sc. Review, vol. 2, no. 2, 2002.

[28] A. Schlögl, "The electroencephalogram and the adaptive autoregressive model: theory and applications," Germany: Shaker, 2000.

[29] A. Alkan, and A. S. Yilmaz, "Frequency domain analysis of power system transients using Welch and Yule–Walker AR methods," Energy conversion and management, vol. 48, no. 7, pp. 2129-2135, 2007.

[30] S. Sanei and J. A. Chambers, EEG signal processing, 2008, Wiley. com.

[31] G. Buccino, F. Binkofski, G. R. Fink, L. Fadiga, L. Fogassi, V. Gallese, and H. J. Freund, "Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study," European journal of neuroscience, vol. 13, no. 2, pp. 400-404, 2001.