

A Novel Hybrid Swarm Algorithm for P300-Based BCI Channel Selection

Víctor Martínez-Cagigal, Eduardo Santamaría-Vázquez,
and Roberto Hornero

Abstract

Channel selection procedures are essential to reduce the curse of dimensionality in Brain-Computer Interface systems. However, these selection is not trivial, due to the fact that there are 2^{N_c} possible subsets for an N_c channel cap. The aim of this study is to propose a novel multi-objective hybrid algorithm to simultaneously: (i) reduce the required number of channels and (ii) increase the accuracy of the system. The method, which integrates novel concepts based on dedicated searching and deterministic initialization, returns a set of pareto-optimal channel sets. Tested with 4 healthy subjects, the results show that the proposed algorithm is able to reach higher accuracies (97.00%) than the classic MOPSO (96.60%), the common 8-channel set (95.25%) and the full set of 16 channels (96.00%). Moreover, these accuracies have been obtained using less number of channels, making the proposed method suitable for its application in BCI systems.

Keywords

Brain-computer interface • Multi-objective optimization
Swarm intelligence • Electroencephalography
Channel selection

1 Introduction

Brain-Computer Interfaces (BCI) have proven to be able to establish an effective communication system that allows users to control applications using their own brain signals

[1]. Due to the non-invasiveness, portability and low cost characteristics of the electroencephalography (EEG), brain signals are usually registered by placing several electrodes (i.e., channels) on the users' scalp [1]. In order to identify the users' intentions in real time, it is essential to employ a recognizable control signal, such as the P300 evoked potentials. These potentials are positive peaks of the EEG mainly produced in the parietal cortex in response to infrequent and particularly significant stimuli at about 300 ms after their onset [1]. The most common setup, known as P300 Speller, allows users to spell words or select certain commands [2]. The user just need to focus attention on one of the character cells of a displayed matrix, while its rows and columns are randomly flashing. Whenever the target's row or column are intensified, a P300 potential is produced in the user's scalp. Thus, the desired character can be determined by computing the intersection where those P300 responses were found [2].

However, the inter-session variability and the low signal-to-noise-ratio that are present in these event-related responses make it difficult to obtain a reliable P300 potential. Thus, it is necessary to compute an average of several sequences, which may produce an over-fitting of the classifier, resulting in a spoiled system performance [3]. The curse of dimensionality may be reduced by using a channel selection procedure, which also reduces the power consumption in wireless EEG caps, increases the users' comfort and assures suitable performances [3]. Nevertheless, this selection is not trivial, owing to the fact that there are 2^N possible combinations for an N -channel cap, making the exhaustive search intractable [3]. For this reason, most P300-based studies omit this stage and use a combination of 8 typical channels in parietal and occipital positions as a general rule of thumb [4]. Nonetheless, due to the intrinsic inter-subject variability of the EEG, an optimization is required for leveraging the system performance.

In this regard, metaheuristics based on evolutionary computation have demonstrated excellent performances solving complex optimization problems. Even though

V. Martínez-Cagigal (✉) · E. Santamaría-Vázquez · R. Hornero
Biomedical Engineering Group, University of Valladolid,
Valladolid, Spain
e-mail: victor.martinez@gib.tel.uva.es

several methods have been successfully applied to P300-based BCI systems, most of them have used single-objective strategies, ignoring the trade-off between the final number of selected channels and the system performance [5, 6], or merging both trade-off objectives using aggregation approaches [7–9]. However, a practical BCI channel selection algorithm should simultaneously optimize a two-fold purpose: (i) to maximize the performance of the system, and (ii) to minimize the required number of channels. Multi-objective optimization algorithms, such as Non Sorting Genetic Algorithm II (NSGA-II) [7, 10] or Multi-Objective Particle Swarm Optimization (MOPSO) [10] have been applied in this regard. Although all of them have been proved to be suitable to its application in this field, their inter-trial variability, as well as their lack of deterministic approaches, hinder their full adaptation to binary BCI systems.

In this study, a novel multi-objective hybrid algorithm that merges the key aspects of MOPSO and forward selection (FS) is proposed for selecting the optimal channel sets in BCI applications. The method, which performs a dedicated local search over each channel, provides a set of Pareto optimal solutions that minimizes the error of the system and the required number of channels.

2 Subjects

The subject pool was composed of 4 male healthy subjects (mean of 26.25 ± 5.19 years) that were asked to spell a total of 200 characters with the P300 Speller in 2 sessions [2] (i.e., half for training and half for testing). EEG signals were recorded using a 16-channel cap with a g.USBamp amplifier (g.Tec, *Guger Technologies*, Austria). Sampling rate was fixed at 256 Hz and bandpass (0.1–60 Hz), notch (50 Hz) and common average reference filters were applied.

3 Methods

In order to evaluate the usefulness of our proposed algorithm for BCI systems, the method has been compared with the traditional MOPSO, described in [11].

3.1 Processing Pipeline

The signal processing pipeline is detailed in the Fig. 1a. In the feature extraction stage (i) a 0–700 ms window from the stimuli onset was selected; and a (ii) sub-sampling to 20 Hz was computed. Then, the multi-objective algorithm is applied to the training subset, returning sets of optimal

combinations of channels. Finally, these sets are evaluated in the testing dataset using a Linear Discriminant Analysis (LDA) classifier and final accuracies are calculated. This pipeline is repeated 20 times in order to discard stochastic effects.

3.2 Two-Fold Objective

Multi-objective metaheuristics involve the optimization of multiple conflicting objective functions at the same time. In this case, our two-fold objective is:

$$\min F(\mathbf{x}) = \begin{cases} f_1(\mathbf{x}) = 1 - AUC(\mathbf{x}) \\ f_2(\mathbf{x}) = \sum \mathbf{x} \end{cases}, \quad (1)$$

where \mathbf{x} denotes a solution (i.e., particle position, a specific set of channels where $\mathbf{x}_c = \{0, 1\}$ with $c = 1, \dots, N_c$) and AUC denotes the area under ROC curve, derived from a 5-fold cross-validated LDA that is trained and tested with the same solution \mathbf{x} . Therefore, $f_1(\mathbf{x})$ involves the minimization of the error rate, whereas $f_2(\mathbf{x})$ the minimization of the required number of channels.

3.3 δ MOPSO/FS Algorithm

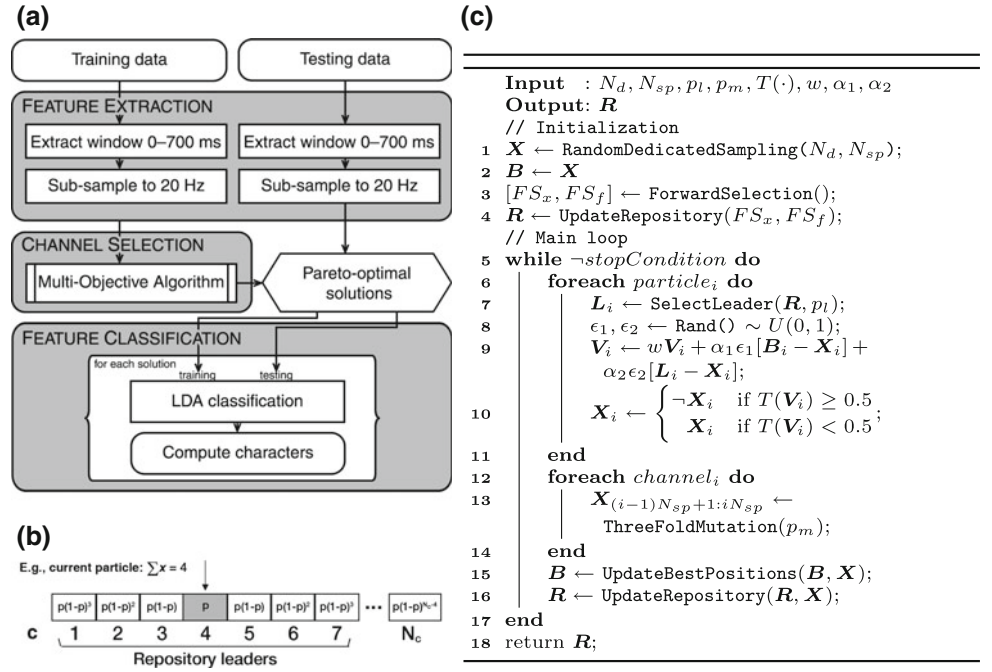
The proposed algorithm, dedicated MOPSO with FS (δ MOPSO/FS), whose pseudo-code is detailed in Fig. 1c, was developed to overcome the inter-trial variability and the lack of search depth that binary MOPSO experiments when is applied to BCI systems [10]. In order to achieve this objective, the method provides a set of novel concepts, such as (i) deterministic initialization, (ii) dedicated particle subgroups for each channel, (iii) leader selection based on binary tournament, and (iv) three-fold mutation, which are detailed below.

Deterministic initialization Forward selection is applied in order to reduce the inter-trial variability due to stochastic effects. Starting from an empty set, the algorithm tests each channel for its inclusion based on the metric $f_1(\mathbf{x})$. Then, the repository is filled with the non-dominated solutions.

Dedicated particles In order to perform a depth local search and favor the convergence, MOPSO/FS dedicates subgroups of N_{sp} particles focused on each possible number of channels $c \in 1, \dots, N_c$. Thus, each particle's position is randomly initialized as long as $\sum \mathbf{x} = c$, where c is the number of channels that belongs to its subgroup.

Leader selection Each particle should point to a repository leader that has the same number of channels than the particle's subgroup. Thus, the leader selection is based on binary tournament odds: each particle selects its

Fig. 1 **a** Detailed signal processing pipeline. **b** Leader selection based on distances. Each particle should select its corresponding leader with probability p , whereas the rest have decreasing probabilities as distance increases. **c** Pseudo-code of the proposed MOPSO/FS algorithm



corresponding leader with probability p ; and the rest of them with probabilities $p(1-p)^d$, where d is the distance from its leader. The procedure is illustrated in Fig. 1b.

Position updating Due to the dichotomous nature of the problem, once the velocity of each particle is calculated, the positions are updated on a transfer function basis. If the transformed velocity in each dimension $T(v_i)$ exceeds a threshold t , the position is inverted, otherwise is maintained. The transfer function is v-shaped: $T(v) = |v/\sqrt{v^2+1}|$, based on [12].

Three-fold mutation The mutation operator is similar to [11], but applied to each subgroup of particles. Therefore, each subgroup is divided in three parts, in which: (1) no mutation is applied, (2) uniform mutation with probability P_m is applied, and (3) non-uniform mutation is applied (i.e., probability $P_n = 1 - [(gen - 1)/\max_{gen}]^5$).

Repository update The repository (i.e., the set of Pareto optimal solutions) is updated in each generation of the algorithm as follows: (i) the current population X and the repository R are merged into a new population R_n , (ii) fitness is calculated in R_n according to Eq. (1), and (iii) non-dominated solutions of R_n are stored into the new repository. Due to the discrete nature of the channel selection problem, the repository keeps a maximum of N_c solutions, which provides a range of combinations to choose from, depending on the number of channels that the user would want to use.

4 Results

The proposed method has been compared with a traditional binary MOPSO, described in [11]. Both of them have run a fixed number of 500 generations in order to facilitate the comparison between them. Optimal training phase Pareto Fronts of both methods, as well as testing final accuracies of the optimal channel sets, are depicted in Fig. 2. Moreover, the highest reached accuracies using MOPSO/FS, MOPSO, the classical 8-set [4] and the 16-channel full set are shown in Fig. 3a. Finally, the Fig. 3b displays the averaged number of times that each channel has been selected as a Pareto Optimal solution across subjects.

5 Discussion and Conclusion

Although multi-objective metaheuristics have been proved to be suitable for the channel selection procedure in BCI systems, there is still room for improvement. Their lack of deterministic approaches cause the algorithms to suffer from stochastic effects, making necessary the computation of several runs in order to reach the optimal solutions [7, 10]. As can be observed in Fig. 2, not only δ MOPSO/FS reaches more optimal Pareto Fronts than MOPSO, but also its convergence is faster (mean of 96.36 generations for δ MOPSO/FS, and 417.70 for δ MOPSO). In addition, the

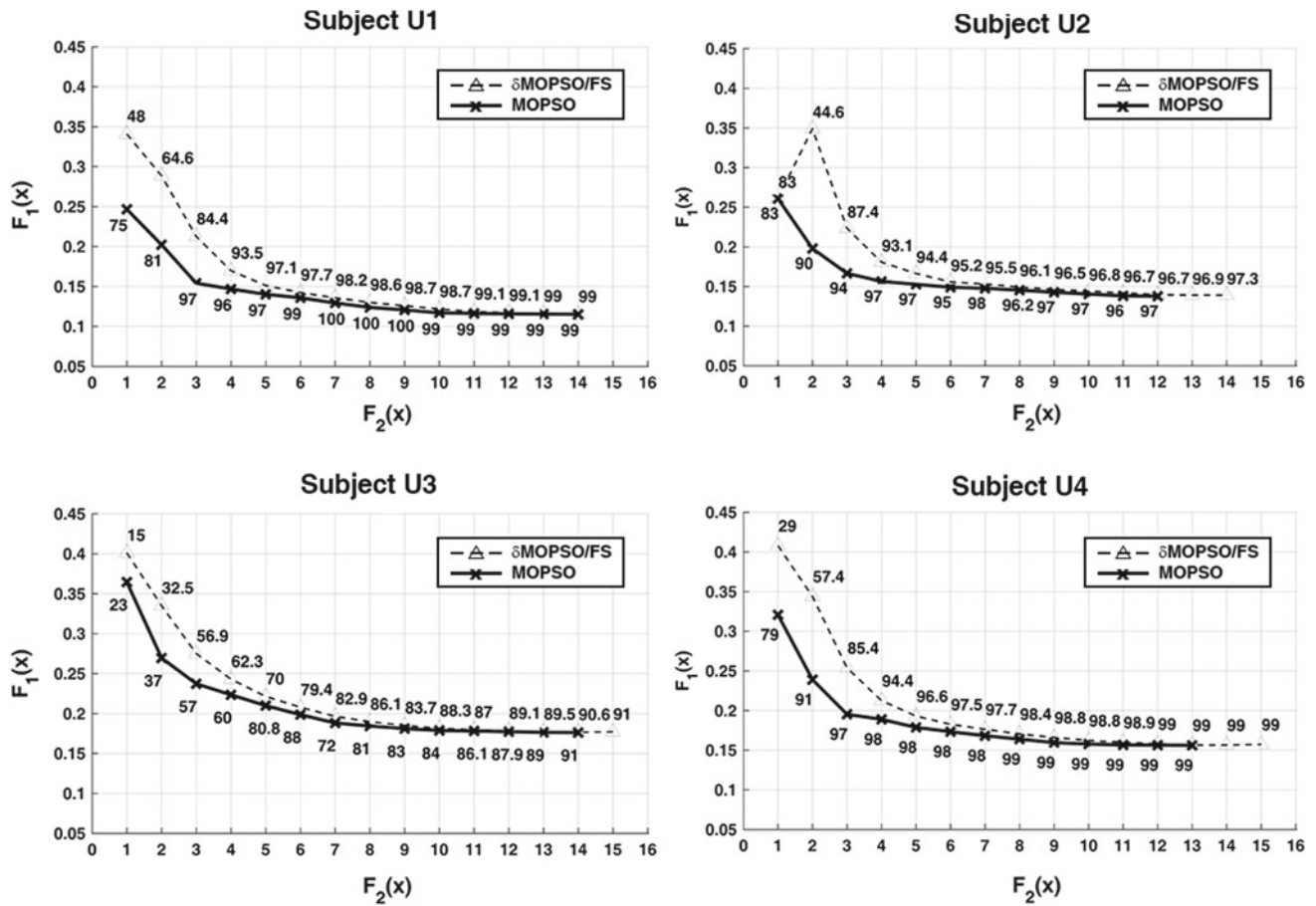


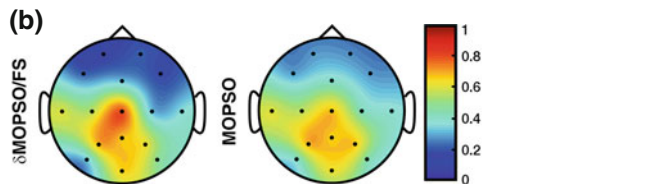
Fig. 2 Optimal Pareto solutions for δ MOPSO/FS (solid line) and MOPSO (dashed line) for each user. The curves, composed by the optimal solutions returned by both methods, depict the trade-off

between both objectives in training data. Final testing accuracies are also show next to each solution

(a)

	δ MOPSO/FS		MOPSO		8-set		16-set	
	Acc.	Ch.	Acc.	Ch.	Acc.	Ch.	Acc.	Ch.
U1	100.0 \pm 0.0	7	99.1 \pm 0.3	11	99.0	99.0	99.0	99.0
U2	98.0 \pm 0.0	7	97.3 \pm 0.5	14	94.0	97.0	97.0	97.0
U3	91.0 \pm 0.0	14	91.0 \pm 0.0	15	90.0	89.0	89.0	89.0
U4	99.0 \pm 0.0	8	99.0 \pm 0.0	12	98.0	99.0	99.0	99.0

Fig. 3 a Highest reached accuracies and their required number of channels using different approaches. **b** Averaged normalized channel ranks for the obtained Pareto optimal solutions (e.g., a value of 1



indicates that the channel is selected in all the solutions that belongs to the Pareto Front of every single user)

results indicate that the final δ MOPSO/FS accuracies are higher than that obtained with MOPSO and, furthermore, the solutions use less number of channels. These accuracies are also higher than that obtained using the common 8-set [4] or the entire full set of channels, which reinforces the idea that channel selection is beneficial for the system's performance. It is also noteworthy to mention that the standard deviation of the final accuracies across trials for δ MOPSO/FS is null, which means that every single trial has converged to the same set of solutions. This fact, in conjunction with the rapid

convergence of the algorithm, demonstrates that δ MOPSO/FS has successfully avoided the stochastic effects and thus, it can assure the identification of global optima in a single run, saving a high amount of computation time. Moreover, the Fig. 3b shows that there are certain channels that have been repeatedly selected along the Pareto Fronts of the subject pool. These channels are mainly distributed over the parietal and occipital regions, which reinforces the study of Krusienski et al. [4], who stated that the P300 potentials are mainly generated in those positions. It can also be

noticed that the dispersion of the selected channels by MOPSO is smoother than by δ MOPSO/FS, which indicates again that δ MOPSO/FS successfully finds the same global optima in each run.

Even though this new technique has been proved to be suitable for use in P300-based BCI systems channel selection, we can point out several limitations. Firstly, the algorithm requires several hyperparameters to be fixed, whose optimization lies on the user experience. Moreover, both MOPSO and δ MOPSO/FS use a transfer function for adapting them into binary-based approaches. In order to overcome this limitations, we contemplate the following future research lines: (i) to implement a dynamic fixation of the hyperparameters, and (ii) to apply these novel concepts to binary objected algorithms, such as the ones that are based on genetic algorithms.

In conclusion, a novel swarm-based algorithm has been proposed for selecting the optimal channel set in BCI applications. The proposed algorithm, δ MOPSO/FS, has been tested with 4 healthy subjects and compared with the traditional binary MOPSO. Results show that δ MOPSO/FS is not only able to converge more faster than MOPSO, but also to reach higher accuracies (mean of 97.00%) than that obtained by using MOPSO (mean of 96.60%), the common 8-channel set (mean of 95.25%) and the full set of 16 channels (mean of 96.00%). Moreover, these accuracies are obtained using less number of channels than MOPSO, approximately the half of the full set. For these reasons, we conclude that δ MOPSO/FS is suitable for use in P300-based BCI systems channel selection procedures.

Acknowledgements This study was partially funded by projects TEC2014-53196-R of ‘Ministerio of Economía y Competitividad’ and FEDER, the project “Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer” (Inter-regional cooperation program VA Spain-Portugal POCTEP 2014–202) of the European Commission and FEDER, and

project VA037U16 of the ‘Junta de Castilla y León’ and FEDER. V. Martínez-Cagigal was in receipt of a PIF-UVa grant of the University of Valladolid. The authors declare no conflict of interest.

References

1. Wolpaw, J.R., Birbamer, N., *et al.*: Brain-computer interfaces for communication and control. *Clin. Neurophysiol.*, 113(6), pp. 767–791 (2002).
2. Farwell, L.A., Donchin, E.: Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroen. Clin. Neuro.*, 70(6), pp. 510–523 (1988).
3. Cecotti, H., Rivet, B., *et al.*: A robust sensor-selection method for P300 brain-computer interfaces. *J. Neural Eng.*, 8(1), p. 016001 (2011).
4. Krusienski, D.J., Sellers, E.W., *et al.*: Toward enhanced P300 speller performance. *J. Neurosci. Meth.*, 167(1), pp. 15–21 (2008).
5. Jin, J., Allison, B.Z., *et al.*: P300 Chinese input system based on Bayesian LDA. *Biomed. Tech.*, 55(1), pp. 5–18 (2010).
6. Perseh, B., Sharafat, A.R.: An Efficient P300-based BCI Using Wavelet Features and IBPSO-based Channel Selection. *J. Med. Signals Sens.*, 2(3), pp. 128–143 (2012).
7. Kee, C., Ponnambalam, S.G., *et al.*: Multi-objective genetic algorithm as channel selection method for P300 and motor imagery data set. *Neurocomputing*, 161, pp. 120–131 (2015).
8. Martínez-Cagigal, V., Hornero, R.: P300-Based Brain-Computer Interface Channel Selection using Swarm Intelligence. *Rev. Iberoam. Autom. In.*, 14(4), pp. 372–383 (2017).
9. Martínez-Cagigal, V., Hornero, R.: A Binary Bees Algorithm for P300-Based Brain-Computer Interfaces Channel Selection. In: *Advances in Computational Intelligence, LNCS, IWANN2017*, pp. 453–463. Cádiz, Spain (2017).
10. Martínez-Cagigal, V., Hornero, R.: Multi-Objective Optimization for P300-Based Channel Selection. In: *Proceedings of the 9th CEA symposium*, pp. 73–78. Barcelona, Spain (2017).
11. Sierra, M.R., Coello, C.A.: Improving PSO-Based Multi-objective Optimization Using Crowding, Mutation and E-Dominance. *Lect. Notes Comput. Sc.*, 3410, pp. 505–519 (2005).
12. Mirjalili, S., Lewis, A.: S-shaped versus V-shaped transfer functions for binary Particle Swarm Optimization. *Swarm Evol. Comput.*, 9, pp. 1–14 (2013).