Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in DISCRETE MATHEMATICS

Electronic Notes in Discrete Mathematics 58 (2017) 79–86

www.elsevier.com/locate/endm

EEG time series learning and classification using a hybrid forecasting model calibrated with GVNS

V. N. Coelho[a,b, 1,2], I. M. Coelho[a,b,c], B. N. Coelho[a,b], M. J. F. Souza[d], F. G. Guimarães.[e], E. J. da S. Luz[d], A. C. Barbosa[a], M. N. Coelho[a,b], G. G. Netto[a], R. C. Costa [a], A. A. Pinto[a], A. de P. Figueiredo[a], M. E. V. Elias [a], D. C. O. G. Filho [a] and T. A. Oliveira[a]

[a] *Grupo da Causa Humana, Ouro Preto, Brazil*

[b] *Instituto de Pesquisa e Desenvolvimento de Tecnologias, Ouro Preto, Brazil*

[c] *Dpt. of Computer Science at Universidade Estadual do Rio de Janeiro*

[d] *Dpt. of Computer Science at Universidade Federal de Ouro Preto*

[e] *Dpt. of Electrical Engineering at Universidade Federal de Minas Gerais*

Abstract

Brain activity can be seen as a time series, in particular, electroencephalogram (EEG) can measure it over a specific time period. In this regard, brain fingerprinting can be subjected to be learned by machine learning techniques. These models have been advocated as EEG-based biometric systems. In this study, we apply a recent Hybrid Focasting Model, which calibrates its *if-then* fuzzy rules with a hybrid GVNS metaheuristic algorithm, in order to learn those patterns. Due to the stochasticity of the VNS procedure, models with different characteristics can be generated for each individual. Some EEG recordings from 109 volunteers, measured using a 64-channels EEGs, with 160 HZ of sampling rate, are used as cases of study. Different forecasting models are calibrated with the GVNS and used for the classification purpose. New rules for classifying the individuals using forecasting

http://dx.doi.org/10.1016/j.endm.2017.03.011

models are introduced. Computational results indicate that the proposed strategy can be improved and embedded in the future biometric systems.

Keywords: Electroencephalogram, Brain fingerprinting, Biometrics, Variable Neighborhood Search, Forecasting and Time series.

1 Introduction

"Every living organism on our planet is surrounded with an energy in form of the signal environment" [8]. This statement reinforces that signals are everywhere. Recently, brain activity for biometric systems has been the focus of different researches [3]. In particular, with the rise of big time series data, novel machine learning techniques are being developed. Those generated with Electroencephalography (EEG) are generally measured with high sampling rate (usually from 100Hz to more than 1KHz).

One advantage of understanding brain signals using EEG and Magnetoencephalography (MEG) is that those techniques are a noninvasive way to look into our brains. As described by Farwell et al. [2], the term "brain fingerprinting" arises from an analogy to fingerprints that has several facets. In particular, they highlighted this term in order to match information from the crime scene with information stored in the brain of the subject. This and other works in the literature have been investigating the concealed information test or guilty knowledge test, usually, used for detecting concealed information since Lykken [9], around the fifties. This present study uses the term Brain Fingerprinting as a way to characterize individual patterns that might be learned through machine learning techniques, coming up with what has been called EEG-Based Biometric systems [3].

We propose the use of a novel Hybrid Forecasting Model (HFM), recently introduced by Coelho et al. [1], in order to learn EEG patterns from the electrodes signals of each volunteer. A hybrid metaheuristic calibration algorithm generates model's *if-then* fuzzy rules and weights, using the General Variable Neighborhood Search (GVNS) procedure [5]. In particular, different models are generated using an automatic learning framework that uses an expert input selection strategy. The latter is done by the use of Neighborhood Structures (NS) that change model's input during the learning phase.

[1] The authors acknowledge Brazilian institutes, foundations, universities and citizens, as well a great vibration of the universe, for supporting the development of this research.

[2] Email:vncoelho@gmail.com

The proposed strategy aims at checking the forecasting models potential for classifying individuals only by looking at its raw EEG time series. As mentioned by Varner et al. [10], evidences are starting to point out that EEG signals carry individual specific traits that might be induced due to genetic factors. The dataset used here is public and freely available by PhysioNet [4], composed of 109 volunteers originally acquired using the 64 channels BCI2000 system.

2 EEG learning, forecasting and classification strategy

Let us consider a target EEG time series $ts^{EEG} = y_1, y_2, ..., y_t$, comprising a set of t observations from a single electrode. The goal is to train the HFM with sections of this time series and check its performance regarding previous unseen information, known as testing set. The most intuitive case seen in the literature is the one where parts of the training phase are also used in the classification phase. Despite having few practical applications, this case can be an useful strategy for verifying the consistence of the classification rules.

Here, we adapt the use of forecasting models, using their forecasting capability for classification purpose. A given HFM can be applied to learn and generate k steps-ahead, with k indicating the number of steps ahead to be predicted, namely Forecasting Horizon (FH). One-step ahead ($k = 1$) is usually the most precise type of forecasting, returning the lowest training errors. However, they might not learn special characteristics of the time series.

The core of our strategy is to apply the forecasting model for learning a $ts^{trainingEEG}$ and testing it with another one, defined as ts^{valEEG}. Regarding this topic, several Data Splitting techniques can be used for sampling parts of a given time series and generate those training, validation and testing sets (or even sets of them).

Considering data from a specific electrode, in our case, the time series is split into only two parts, training and testing. Thus, the forecasting models can applied for learning the first tr samples of a given EEG band ts^{EEG}, $ts^{trainingEEG} = y_1, ..., y_{tr}$, with $tr < t$. The remaining $val = t - tr$ samples can be used to check the accuracy of the model regarding unseen data. Thus, a finite sequence $ts^{valEEG} = \{\hat{y}_{tr+1}, ..., \hat{y}_t\}$ is the one that should be predicted.

According to a given forecast accuracy [6], the forecasting errors are calculated. Let consider a set of models $SM = \{HFM^m_{v_1}, ..., HFM^m_{v_i}, ..., HFM^m_{v_n}\}$, composed of m models trained only with data regarding an individual/volunteer v_i, with $i = [1, n]$. In the proposed classification rule, after training, each model HFM^m_v can be applied for forecasting unseen data from all different volunteers. The model with the lowest error is assigned and labeled to the unseen time series. For simplicity, the classification rule will be explained considering a single model for each volunteer ($m = 1$).

Following this reasoning, the forecasting model HFM_{v_i} can be applied to learn different parts from the EEG electrodes of v_i. Later on, the unseen parts of the data are applied to be predicted using all SM available models. However, it is expected that a model HFM_{v_j} from an individual j applied to learn unseen parts of v_i (with $v_j \neq v_i$) will return higher forecasting errors than HFM_{v_i}. This is confirmed by this study and happens because the model HFM_{v_j} was not trained to forecast patterns from individual v_i. Thus, the HFM_{v_j} might obtain worse results and, consequently, higher forecasting errors.

The objective of our analysis is to check if the time series from v_i was better learned, providing lower errors, when predicted using the forecasting model HFM_{v_i}.

3 HFM metaheuristic training algorithm

A HFM solution comprises different fuzzy rules and can be represented as a matrix $s = [Y]$, being Y a matrix $4 \times |L|$, where $L = \{l_1, l_2, ..., l_z\}$ is the lags vector, containing the inputs used by the forecasting model. Each column of matrix Y, which is a respective index of vector L, contains two different fuzzy rules and its respective weights. A solution example can be seen in Figure 2.

Initial values for the fuzzy rules are generated using the procedure described in Coelho et al. [1]. The strategy consists in calculating the EEG time series mean values and applying a normal distribution for generating each rule and weight. In this study, model's inputs are pick at random, being the oldest input limited to 3% of the available samples for each time series that is being learned.

Quick training phases are studied in this current paper, up to 5 seconds per model. The GVNS procedure proceeds with the refinement phase, respecting this time limit. The basic idea of the GVNS is depicted in Figure 1, exemplifying a minimization problem with several NS. The procedure started refining a solution with the Variable Neighborhood Descent, in our case, by combining two Random Descent methods. When the procedure finishes, it may have found a local optimum and a shaking procedure usually jumps to a worse solution and the search with VND starts all over again.

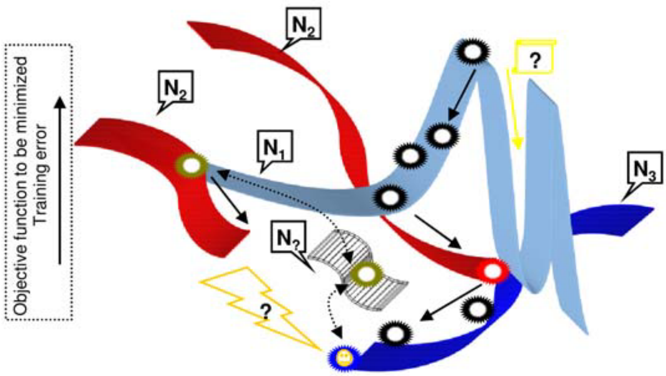


Fig. 1. VNS multi-neighborhood searching strategy

3.1 Neighborhood structures

For exploring the solution space of the problem and train each forecasting model, two basic neighborhood structures ($NS^{addX}(\cdot)$ and $NS^{changeLags}(\cdot)$) are used. These NS were also used as a perturbation strategy, being this mechanism a key strategy to diversify the search in the solution space.

Neighborhood $NS^{addX}(s, r, c, x)$ increment or decrement, with magnitude $x \in \mathbb{R}$, a rule or weight at position at row r and column c of a given solution s. Figure 2a exemplifies the use of the $NS^{addX}(\cdot)$ moves in a solution s, generating neighbors s' and s''. The move increments the weight of vector V, from the lag operator $K-2$, in ten units.

Neighborhood structure $NS^{changeLags}(s, l)$ has the property of increment or decrement, in one unit, the input lag of a given lag operator of the vector $L \in s$. Examples are given in Figure 2b.

Fig. 2. NS examples

$NS^{addX}(s, i = 2, j = 2, x = 10) \Rightarrow$

$$s' = \left[\begin{array}{c|ccc} & z(K-1) & z(K-2) & z(K-24) \\ A & 87 & 95 & 103 \\ V & 70 & \overrightarrow{80} \rightarrow 90 & 95 \\ B & 100 & 90 & 110 \\ W & 110 & 50 & 80 \end{array}\right]$$

(a) Application example of $NS^{addX}(\cdot)$ move

$NS^{addX}(s, l = 2, x = 1)$

$$s' = \left[\begin{array}{cc} z(K-1) & \overrightarrow{z(K-2)} \rightarrow z(K-1) \\ \ldots & \end{array}\right]$$

$NS^{addX}(s, l = 1, x = -1)$

$$s'' = \left[\begin{array}{c|cc} & \overrightarrow{z(K-1)} \rightarrow z(K-2) & z(K-2) \\ \ldots & & \end{array}\right]$$

(b) Application example of $NS^{changeLag}(\cdot)$ move

4 Computational experiments

The HFM, trained with the GGVNS algorithm, was implemented in C++ and can be find inside the optimization framework OptFrame 2.2 [3]. The tests were carried out on a OPTIPLEX 9010 Intel Core i7-3770, 8×3.40 GHZ with 32GB of RAM, with operating system Ubuntu 14.04, and compiled by g++ 4.8.4.

4.1 EEG time series used as cases of study

Some EEG recordings from 109 volunteers, measured using a 64-channels EEGs, BCI2000 system, with 160 HZ of sampling rate, are used for validating the proposal. Four out of 14 experimental runs available in the original dataset were used here. Data from up to 50 volunteers were used for analyzing the classification strategy, considering 3 distinct experimental runs, and one duplicated, (namely 1, 2, 4 and 5): two one-minute baseline runs (one with eyes open, one with eyes closed) and two two-minute runs (4 and 5) where the volunteers were subjected to a set of imagery and real movements, opening and closing their left or right fists.

4.2 HFM learning ability

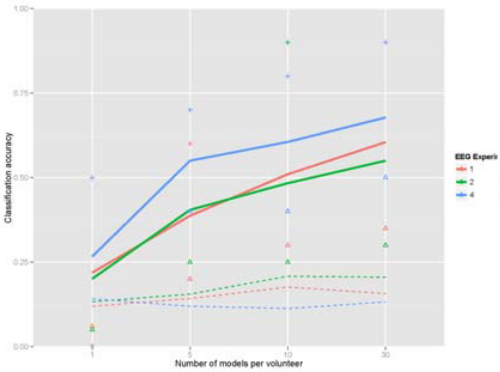
Two different analysis were done. In the first one, we decided to analyze the ability of the proposed strategy for classifying parts of a single run from each volunteer. Thus, 70% of the data (5320 and 10640 samples for one and two minutes experiments, respectively) was subjected to be learned by the HFM, with a 5 seconds calibration done by the GVNS. The second batch was done using two different experiment from the dataset, the models were applied to learn the whole time series from experiment run 4 and tested over the experiment run 5. We believe that these configurations are quite realistic and can be used in several real applications. However, since the data was not measured by us, we are still limited to the quality of the Physionet and their proposed protocol BCI2000. The batches of experiment were composed of around 120 different combinations of parameters, such as, different: forecasting horizons, number of analyzed volunteers (10, 20, 30 and 50) as well as the number of models per volunteer.

Figure 3a shows an interaction plot, indicating the accuracy (successes divided by the number of trials). The points in shapes of triangles and crosses

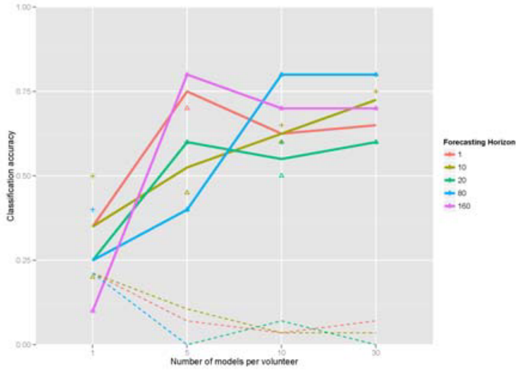
[3] Available at http://sourceforge.net/projects/optframe/

indicate respectively the minimum and maximum accuracy of each configuration. The dashed line indicates the standard deviation while the thicker line shows the average accuracy.

Fig. 3. Interactions plots and model classification accuracy



(a) Split EEG time series with random channels



(b) EEG time series with random channels and different experiments runs – 4 and 5

As suggested by Yang & Deravi [11], with experiment 4 (a motor imagery task), might better avoid the contamination of the EEG signal. If the classification strategy had been done at random, an average accuracy of around 5% (10%, 5%, 3% and 2% for 10, 20, 30 and 50 volunteers, respectively) might be expected. Furthermore, results indicate that more than one model per volunteer considerably increases the classification accuracy.

5 Final considerations and extensions

In this study, a novel EEG-based biometric system was designed using a forecasting model trained with a trajectory search metaheuristic algorithm, the VNS. Different EEG time series were trained and the initial results point out that the classification protocol seems to work. Other EEG datasets and new experiments, measured in different environments and considering other stimulus, will be addressed in a near future. Our future works will extend the proposal for handling signals from Functional Magnetic Resonance Imaging (fMRI), which is a powerful and versatile measurement technique. Other non-invasive that operates near quantum limited sensitivity [7] will be analyzed. Novel strategies and classification rules for forecasting models should also be developed and explored. The generation of a more diverse set of models is also a promising topic, such as a non-dominated set of trained networks.

References

[1] Coelho, V. N., I. M. Coelho, B. N. Coelho, A. J. R. Reis, R. Enayatifar, M. J. F. Souza and F. G. Guimarães, *A self-adaptive evolutionary fuzzy model for load forecasting problems on smart grid environment*, Applied Energy **169** (2016), pp. 567 – 584.

[2] Farwell, L. A., D. C. Richardson, G. M. Richardson and J. J. Furedy, *Brain fingerprinting classification concealed information test detects us navy military medical information with p300*, Frontiers in Neuroscience **8** (2014).

[3] Fraschini, M., A. Hillebrand, M. Demuru, L. Didaci and G. L. Marcialis, *An EEG-based biometric system using eigenvector centrality in resting state brain networks*, IEEE Signal Processing Letters **22** (2015), pp. 666–670.

[4] Goldberger, A. L., L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng and H. E. Stanley, *Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals*, Circulation **101** (2000), pp. e215–e220.

[5] Hansen, P., N. Mladenovic and J. A. M. Pérez, *Variable neighborhood search*, European Journal of Operational Research **191** (2008), pp. 593–595.

[6] Hyndman, R. J. and A. B. Koehler, *Another look at measures of forecast accuracy*, International Journal of Forecasting **22** (2006), pp. 679 – 688.

[7] Jensen, K., R. Budvytyte, R. A. Thomas, T. Wang, A. Fuchs, M. V. Balabas, G. Vasilakis, L. Mosgaard, T. Heimburg, S.-P. Olesen et al., *Non-invasive detection of animal nerve impulses with an atomic magnetometer operating near quantum limited sensitivity*, arXiv preprint arXiv:1601.03273 (2016).

[8] Lipkova, J. and J. Cechak, *Human electromagnetic emission in the ELF band*, Measurement Science Review **5** (2005), p. 29.

[9] Lykken, D. T., *The GSR in the detection of guilt*, Journal of Applied Psychology **43** (1959), p. 385.

[10] Varner, J. L., R. A. Potter and J. W. Rohrbaugh, *A procedure for automatic classification of EEG genetic variants*, in: *Engineering in Medicine and Biology Society*, 1991, pp. 451–452.

[11] Yang, S. and F. Deravi, *Quality filtering of EEG signals for enhanced biometric recognition*, in: *Biometrics Special Interest Group (BIOSIG)*, 2013, pp. 1–4.