A STUDY ON THE PARALLELIZATION OF MOEAS TO PREDICT THE PATIENT'S RESPONSE TO THE ONABOTULINUMTOXINA TREATMENT

Franklin Parrales Bravo
Alberto A. Del Barrio
José Luis Ayala

Dept. of Computer Architecture and Automation
Universidad Complutense de Madrid
Prof. José García Santesmases 9
Madrid, Spain
{fparrale,abarriog,jayala}@ucm.es

ABSTRACT

This work deals with the decrease of the computational cost in the task of feature weighting for the predictive models of the response to the treatment of migraine with OnabotulinumtoxinA (BoNT-A). More specifically, we consider the multiobjective evolutionary algorithms (MOEAs) that support parallelization. All this with the aim of improving the training times of predictive models of response to the treatment. The results obtained show that accuracies close to 84 % are obtained while training times are decreased from 8 to less than 2 hours when using 8 threads. All in all, this work remarkably reduces the feature weighting execution time in comparison with Simulated Annealing, while getting similar values of accuracy.

Keywords: MOEA, parallelism, feature weighting, Pareto front, migraine.

1 INTRODUCTION

One of the biggest problems faced by chronic diseases is their continuous treatment to mitigate or eliminate their symptoms. Chronic diseases such as Parkinson's or migraine are progressive and this entails a significant increase in personal, social and work disability (Parrales Bravo et al. 2017). Chronic migraine is defined as a headache occurring on 15 or more days per month for more than 3 months, and which has the features of a migraine headache on at least 8 days per month (IHS 2013). Globally, approximately 2% of the population experiences chronic migraine (Natoli et al. 2010).

OnabotulinumtoxinA (BoNT-A) has been an extended treatment for chronic migraine since its approval in 2010 by the Food and Drug Administration in the United States (FDA), having also shown a more sustained effect and better tolerability than topiramate in the few comparative studies performed (Mathew and Jaffri 2009, Cady et al. 2011). Today, it is known that 70-80% of patients with chronic migraine show an improvement with this treatment (improvement defined as a reduction in migraine attack frequency or days with attacks by at least 50% within 3 months, leading to a significantly improved functioning of the patients and their overall quality of life) (Cernuda-Morollón et al. 2015, Lipton et al. 2011, Oterino et al. 2011, Sandrini et al. 2011). Moreover, there is evidence that patients with chronic migraine who do not show the desired treatment response after the first cycle of BoNT-A treatment may indeed experience clinical improvement after one or two additional treatment cycles (Silberstein et al. 2015). However, in clinical

practice, around 20-30% of chronic migraineurs do not respond to BoNT-A (Silberstein et al. 2015). As has been mentioned in certain publications (Lovati and Giani 2017), it is very important to predict if the BoNT-A treatment will be effective in a patient.

In (Parrales et al. 2019) authors achieve a 85% of accuracy when predicting the response to the first and second stages of the migraine treatment with BoNT-A. Although this work provides a good basis for obtaining predictive models of the response to each stage of the treatment, it has some limitations: on the one hand, it does not allow a multiobjective optimization. With this approach it is possible to know what medical features are important for all prediction models instead of knowing the relevance in a specific prediction model. On the other hand, it has a high computational cost (Ram et al. 1996) as a consequence of the use of simulated annealing (SA) (Kirkpatrick et al. 1983), a single-solution heuristic, for the feature weighting task. Therefore, in this article, we intend to make a study of the parallelism of various algorithms that allow multiobjective optimization in order to diminish the computational cost while allowing multiobjective optimization when predicting the response to BoNT-A treatment. In this sense, we present a comparison in terms of time execution, number of threads and accuracy of each multiobjective algorithm performed. Results ranging from 80% to 85% of accuracy are obtained for every stage of treatment when applying *multiobjective evolutionary algorithms* (MOEAs) (Coello et al. 2007, Van Veldhuizen and Lamont 2000), which is comparable to the SA. However, the parallelization of the MOEAs reduces the runtime of the SA in the feature weighting task from more than 8 hours to less than 2 when 8 threads are employed.

The rest of the paper is organized as follows. Section 2 describes the work related with some techniques applied to migraine and other illnesses. In Section 3, our methodology for predicting treatment results is explained. Section 4 describes the experiments and comparisons between different algorithms and our solution. Finally, our conclusions and future lines of work are presented in Section 5.

2 RELATED WORK

The favorable response to treatment with BoNT-A has been analyzed by several studies. Although, conclusive results are not yet available for use in clinical practice. Possible predictors of a good response have been proposed: allodynia (painful hypersensitivity to superficial stimuli) (Mathew et al. 2008), the unilateral character of a migraine (Mathew et al. 2008, Lainez et al. 2006), associated migraine aura (visual, language, motor or sensory alterations occurring prior to pain) (Grogan et al. 2013), or the build-up time to maximum pain (shorter time, better response to BoNT-A) (Schulman et al. 2008). Pain directionality also seems to be a possible clinical predictor. This feature refers to whether the headache feels like it is exploding, imploding or ocular. The term exploding refers to when the discomfort is felt pushing from the inside out. Patients suffering from imploding or ocular pain tend to be relieved with the BoNT-A treatment than those with the exploding (Jakubowski et al. 2006). Pagola et al. (2014) studied a number of possible clinical predictive features in parallel, including unilateral location of headache, pericranial muscular tension, directionality of pain, duration of migraine history and medication overuse, comparing responders to BoNT-A treatment with non-responders, but no significant differences emerged.

In order to predict the response to every stage of the BoNT-A treatment, we can take use of the most common data mining algorithms used by the research community in many fields. For instance, according to Wu et al. (2008): C4.5, *k*-means, Support Vector Machines (SVM), Expectation-Maximization (EM) algorithm, PageRank, AdaBoost, *k*-NN, Naive Bayes, and CART. In the work presented by Parrales et al. (2019), a comparison between some classification algorithms was carried out. Authors found high accuracy values in the response prediction to BoNT-A treatment (close to 85% of accuracy) with the use of Random trees algorithm in combination with Simulated annealing (SA) on the clinical dataset. However, some computational drawbacks of the prior work are: (1) the time (Ram et al. 1996) that the computer requires to train the model when SA is applied for a high number of iterations; (2) SA is a single-solution heuristic (Kirk-

patrick et al. 1983). This approach presents some advantages like simplicity and a low number of function evaluations. However, it could be trapped in local optima with high probability (Mirjalili et al. 2016). In addition, information is not shared between the candidate solutions when the single solution approach is applied, and some aspects such as premature convergence, isolation of optima and search space biases must be treated with high care. (3) To manage clinical treatments with multiple stages (as in the case of BoNT-A treatment to chronic migraine), it is necessary to train a prediction model of each stage separately given the impossibility of addressing them in a unified way.

The methodology presented in this work takes advantage of the MOEAs (Coello et al. 2007, Van Veldhuizen and Lamont 2000) because they allow to handle multiple optimization problems while improving the time when training prediction models and handling population-based solutions instead of single-solutions. Additionally, this work incorporates some MOEAs that allow parallelism with the purpose of reducing time when training the treatment response prediction models.

3 METHODOLOGY

The issues involved in the MOEAs parallelization experiment when training the prediction models to the BoNT-A treatment response will be described in this section. Figure 1 presents the experiment workflow on which this paper is based. Firstly, a database is loaded with the medical records from the two participating hospitals. Secondly, the class attribute is defined by considering the retrospective data available. Thirdly, clinical features are categorized in order to work with homogeneous data. After that, some parallel multiobjective methods are applied to the feature weighting task. Finally, the Random tree classification algorithm is applied for improving the time used for training the prediction models and having a low computational cost.
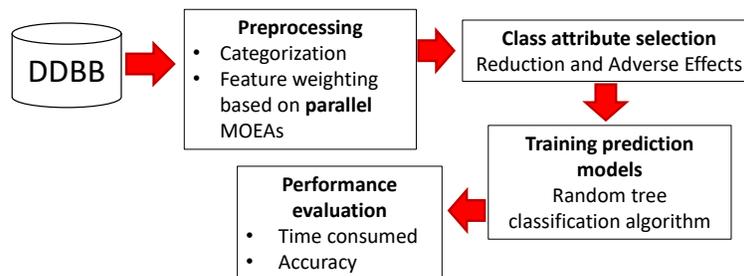
Figure 1: Experiment workflow.

3.1 Clinical Data

Retrospective medical data has been collected from various medical histories of patients under previous or current chronic migraine treatment with BoNT-A with follow-up at the Headache unit of two tertiary-level hospitals.

The number of patients collected from two hospitals were 173 (116 from *Hospital Clínico Universitario* in Valladolid and 57 from *Hospital Universitario de La Princesa*, in Madrid). A total of sixty-two baseline features have been categorized. Collected medical features were related to the following points: clinical pain features, demographic features of patients, comorbidities, tested and concomitant preventive drugs, pain impact measures, and available analytical parameters. The latter were obtained from blood tests recorded in the clinical history which were performed for other reasons in the 3 months prior to, or 3 months after, the first infiltration, and included hemogram and liver, renal, thyroid, ferric, vitamin B12, folic acid and

vitamin D profiles. The efficacy of BoNT-A was evaluated by comparing the baseline situation (before the first infiltration) and the situation after 12-16 weeks following each of the infiltrations, through the following parameters: number of days of pain per month, percentage reduction in days with pain, subjective intensity of pain, number of days of disability due to pain per month, drug consumption for pain and adverse effects of infiltration. Since this was a retrospective study, not all the data could be obtained for each patient in a systematic way.

In order to train the treatment response prediction models, clinical data need to be previously processed in order to achieve a high level of accuracy. In fact, some patients are *non-respondent*, while others respond after the ith session. With the purpose of predicting the outcome after the ith session, the outcome after the $(i-1)^{th}$ infiltration and the clinical data of the patient are considered. Nevertheless, problems like having a small set of patients with many features, still need to be solved. The incompleteness of data, continuous numeric values, and medically categorized data are difficulties to tackle in our medical dataset. All in all, it is hard to properly process all this information for infering prediction models with high accuracy through to data mining algorithms. Table 1 presents example of continuous and categorized data available in our clinical dataset.

Table 1: Example of continuous and categorized features available in the clinical data.

Toxin-age of onset (years)	Body mass index (kg/m^2)	Hemoglobin (g/dL)	Creatinine (mg/dL)	Platelets (u/mcL)	Reduction effects (1-4)
51	20.39	13.4	0.71	213000	4
49	26.5	14.2	0.55	252000	2
36	23.15	13.5	0.44	304000	3
26	17.7	13.1	0.66	218000	2
31	NA	14.8	0.71	327000	1
50	NA	16.2	0.74	327000	3

3.2 Preprocessing

3.2.1 Categorization Of Clinical Features

In order to improve prediction accuracy for the BoNT-A treatment, the heterogeneous data from the hospitals are first categorized. The method selected for the categorization of our medical data is based on the mean and standard deviation. Applying this method is possible to work with more homogeneous values.

The mean and standard deviation categorization type centers the intervals around the mean (μ), and defines subsequent intervals by adding or subtracting the standard deviation (σ). For instance, three categories are defined for each of our clinical features. The intervals *(V_{min}, $\mu - \sigma$), ($\mu - \sigma, \mu + \sigma$)* and *($\mu + \sigma$, V_{max})* are used to refer to value 1, value 2 and value 3, respectively. It should be noted that V_{min} and V_{max} are the minimum and maximum values of the data, respectively.

3.2.2 Feature Weighting

Our purpose is to find those weights that improve the representation of the numeric labels encoded by doctors for each stage. This problem is multiobjective because we need to find the optimal weights that improve the accuracy of all the predictive models of the treatment stages. Applying SA implementation of (De Vicente et al. 2000) for the feature weighting task does not simultaneously solve the minimization of the prediction error for all stages. For this reason, the use of MOEAs has been considered for the feature weighting task in this work. The objective of the MOEAs is to achieve a set of efficient solutions, not

dominated or Pareto optimal (Zitzler et al. 2000). These solutions are called Pareto Optimum, when there is no other solution that takes a lower value (in minimization problems) in some objective without causing a simultaneous increase in at least another. An important point to consider is that MOEAs handle a set of solutions (population) instead of a single solution as in the case of the SA. As a consequence of having more solutions, its computational cost is greater than algorithms with a single solution approach, specially when performing without parallelism (Durillo et al. 2008). The parallelization allows to distribute the computational load on different cores of the computer, making the execution of tasks efficiently. We can use parallel implementations of MOEAs in order to achieve faster execution of algorithms and a superior numerical performance (Alba and Tomassini 2002).

$$e_i = 100 - Acc_i , \quad \forall i \in [1, s]. \tag{1}$$

Regarding our problem, the objective is to minimize the error e_i for all the stages (s), described by Equation 1, where Acc_i refers to the accuracy percentage of the corresponding prediction model. In this paper we contemplate two stages, so there are two objectives to be minimized simultaneously, i.e. e_1 and e_2. The approach has been implemented using the MOEA framework presented in (Hadka, D. 2019). More specifically, we will focus on using those MOEAs that can be parallelized for diminishing the computational cost. Those selected algorithms are: GDE3 (Kukkonen and Lampinen 2005), PESA2 (Corne et al. 2001), SMPSO (Nebro et al. 2009), NSGA-II (Deb et al. 2002), NSGA-III (Deb and Jain 2014) and SPEA2 (Zitzler et al. 2001).

3.3 Class Attribute Selection

With the purpose of measuring how efficient a treatment stage has been, we need to define the *class attribute*. This refers to the clinical feature used to measure the effectiveness of treatment. HIT6 value (Yang et al. 2011), intensity, duration and frequency of attacks (Gasbarrini et al. 1998) are good candidates for class attributes according to doctors. However, the values of some of these features are not usually provided in our medical dataset. In fact, the HIT6 value being is missing in many of our clinical records. In consequence, a combination of the reduction (R) and the adverse (A) effects, which are those features more frequently found in our database, has been selected to define the class attribute. Reduction and adverse effects are defined with values directly provided by doctors.

These clinical features, *R* and *A*, are measurable values from an objective point of view based on definitions. *R* takes values from 1 to 4 according to the percentage of reduction of days of migraine. It takes the value of 1 when the percentage reduction of days of migraine is less than or equal to 25%, 2 for the interval between 25% and 49%, 3 for the interval between 50% and 74% and 4 when the percentage is greater than or equal to 75%. *A* is equal to 1 when there are no adverse effects, 2 when there are mild adverse effects (easily tolerated), 3 when there are moderate adverse effects (interfere with usual activities and may require suspension of treatment) and 4 when there are serious adverse effects (incapacitate or disable usual activities, and require suspension of treatment as well as medical intervention) (Parrales et al. 2019).

High levels of *R* indicate good treatment results, while high levels of *A* point to many adverse effects. With the purpose of obtaining a directly proportional feature, our class attribute (N_{AC}) has been determined by dividing *R* and *A*.

In this work, a similar approach as the one based on HIT6 (Yang et al. 2011) (two response categories: low and high) (Silberstein et al. 2015) has been considered for class attribute categorization, instead of the three categories (low, medium and high) used for the rest of the clinical features. Lower responses are labeled when the N_{AC} value falls into the *(V_{min}, cut-off point) interval, while high response labels are used for those*

values falling within the (cut-off point, V_{max}) interval. In this case, $V_{min} = 0.25$ occurs when $R = 1$ and $A = 4$, while $V_{max} = 4$ occurs when $R = 4$ and $A = 1$. We select a cut-off point of 1.40. The reason to use this value is the fact of trying to emulate the criterion used of the 30% decrease in the HIT6 value. The treatment is considered to be effective if the HIT6 is reduced in a 30% or more, according to the PREEMPT clinical trial Silberstein et al. (2015). In this way, values lower than 1.40 represent the 30% of the values that N_{AC} can take. Then, the low and high categories are defined with the intervals (0.25, 1.40) and (1.40, 4), respectively. Table 2 depicts an instance of the N_{AC} computation using different values provided by the hospitals.

Table 2: Class attribute categorization.

Reduction effects (R)	Adverse effects (A)	R/A	Categorized value
1	1	1	low
2	1	2	high
3	2	1.5	high
1	2	0.5	low

3.4 Training Prediction Models

With the purpose of predicting the treatment response to the different stages of BoNT-A treatment, several classification algorithms have been considered when building the prediction models because they identify categories for new records based on the previous data (training dataset) (Witten et al. 2016).

In the study performed by (Parrales et al. 2019), several classification algorithms were tested, the Random tree + SA being the <best combination with an accuracy close to 85% when predicting the response to first and second stage of treatment. Random tree is a non-deterministic algorithm that builds a tree considering K randomly chosen features for each node. Authors conclude that as a result of being a non-deterministic algorithm, Random tree was benefited with the use of SA. In contrast to the deterministic algorithms, it allowed a deeper exploration of the search space to avoid being trapped in a local minimum. For this reason, the Random tree algorithm will be considered in combination with the MOEAs algorithms when obtaining prediction models in our experiment.

4 EXPERIMENTAL

In this section, we have performed a comparison of execution times and accuracies between SA and the MOEAs described in Section 3 combined with the Random tree classifier algorithm for the feature weighting task. For completing the comparisons, we have also employed the SA algorithm used by Parrales et al. (2019) for the same task. It is important to note that that the SA implementation used in that article was not implemented with parallel execution support. The number of threads that has been considered has been: 1 (no parallelism), 2, 4, 6 and 8. The machine used to perform the experiments consists of an Intel Core i7-4790 CPU 3.60GHz (4 cores) processor and 16GB of RAM. The number of iterations of the experiment was established in 10^6 as in Parrales et al. (2019). The population size for MOEAs was established in 100. This has value has been selected in order to guarantee the diversity of solutions while avoiding a slow convergence of individuals (Chen et al. 2012), (Zitzler and Thiele 1999).

4.1 Runtime

Table 3 presents the execution time of the previously employed algorithms following the hour:minutes:seconds format. Since the SA method applied (De Vicente et al. 2000) does not perform a multiobjective optimization, we have divided the time taken to perform the feature weighting task into two operations, one for each stage of the BoNT-A treatment.

Table 3: Time performance of SA and MOEA parallel algorithms on the feature weighting task.

Methods	Number of threads				
	1	2	4	6	8
SPEA2	8:16:53	4:06:08	2:30:37	2:02:08	1:54:20
NSGAIII	8:12:20	4:05:14	2:28:52	2:02:17	1:55:32
NSGAII	8:12:04	4:05:12	2:28:48	2:02:05	1:54:35
SMPSO	8:13:02	4:06:16	2:29:04	2:07:27	1:59:15
PESA2	8:16:53	4:07:01	2:29:27	2:05:33	1:58:19
GDE3	8:13:02	4:06:08	2:29:45	2:02:39	1:54:38
SA-stage 1	4:13:05	NA	NA	NA	NA
SA-stage 2	4:32:31	NA	NA	NA	NA

According to the results, we can observe that parallel MOEAs executed on two or more threads have required less time than the SA algorithm. Parallel MOEAs are benefited from the use of more threads to distribute the computational load in the feature weighting task. However, it is necessary to note that the time difference between 6 and 8 threads is much smaller than the difference between 1, 2 and 4 threads. On the other hand, apparently, we can observe that MOEAs have a longer execution time than SA when only one thread is used. However, SA only performs the feature weighting task for a single stage. Therefore, the real total time employed by SA is the sum of the times of the first and second stages. This value is around 30 minutes higher than the employed by the MOEAs with 1 thread.

4.2 Accuracy

Table 4 presents the best accuracy values when predicting the treatment response to BoNT-A for the first and second stages. To present the results of this table, we have selected for each algorithm those non-dominated solutions that have the lowest errors e_1 and e_2 (first and second stages, respectively) as described in Section 3.2.2.

Table 4: Accuracy percentage of SA and parallel MOEAs in combination with Random tree algorithm.

Classification algorithm	First infiltration			Second infiltration		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
SPEA2	77.91%	0.75	0.81	79.06%	0.83	0.75
NSGAIII	81.39%	0.77	0.83	82.56%	0.81	0.83
NSGAII	82.56%	0.85	0.77	**84.88%**	0.87	0.81
SMPSO	79.06%	0.83	0.75	82.56%	0.81	0.83
PESA2	76.74%	0.82	0.71	**84.88%**	0.87	0.81
GDE3	82.56%	0.85	0.77	81.39%	0.77	0.83
SA	**84.88%**	0.87	0.81	81.39%	0.77	0.83
Mean	80.73%	0.82	0.78	82.39%	0.82	0.81

According to the results, we can observe that high values of accuracy, sensitivity and specificity are obtained both when the MOEA methods are applied and when SA is used. Sensitivity and specificity refer to the fraction of true positives and true negatives over predicted values, respectively. In our dataset, positive and negative values refers to 'high" and "low" responses, respectively. According to the results shown in this table, SA achieves the best accuracy (84.88%) for stage 1 while GDE3 and NSGAII achieve the best accuracy (84.88%) for stage 2. In all these three results, 0.87 and 0.81 were obtained as values of sensitivity and specificity, indicating a low number of false positives and false negatives.

To compare the runtime and the error obtained by each of the MOEAs contained in Table 4, Figure 2 is presented. Figures 2a and 2b depict the errors and runtimes produced during the feature weighting task for the first and second stages of the BoNT-A treatment, respectively. In addition, the solutions provided by SA have been considered in both figures, since they present the results of each stage separately. In the figures, the best points in terms of accuracy are marked with red circles. It is important to note how the charts show a better performance for MOEAs when using 6 and 8 threads than when using 1, 2 and 4 threads (both in runtime and accuracy), since the error for each stage decreases when each one is considered separately. In Figure 2a it can be observed that SA achieves the best accuracy (84.88%, i.e. an error of 15.12%) when only stage 1 is considered. However, SA implementation performed (De Vicente et al. 2000) does not support multiobjective optimization or parallelism, as it has been commented in Section 3.2.2. Thus, it takes more time in the feature weighting task (close to 4 hours) without being able to minimize errors for both stages at the same time, while the MOEAs are able to do this. One of them, GDE3, achieves an error of 17.44% for stage 1, but it gets an error of 18.60% for stage 2 (see Figure 2b), being surpassed by NSGAII and PESA2 in that stage. These last two obtain the best error minimizations for stage 2 (errors of 15.2%), but PESA2 is surpassed by SA in stage 1. In this stage, SA does not manage to minimize the error as much as GDE3 and NSGAII do, despite its low error achieved in stage 1. It must be noted that the SA solution does not even appear in Figure 2b because its long execution time.

Given that it is difficult for us to visualize which method maximizes the accuracy (i.e., minimize the error) for both stages simultaneously, Figure 3 is presented considering only the MOEAs. It is important to note that SA is not taken into account in Fig. 3 since it does not minimize both stages simultaneously. As can be seen in this figure, the best tradeoff is the one that minimizes the error for both stages, which is achieved with NSGAII when performing on 8 threads (marked with a red circle). However, as shown in Figure 2, the best tradeoff does not always imply to be the best in every stage.

(a) 1st stage

(b) 2nd stage
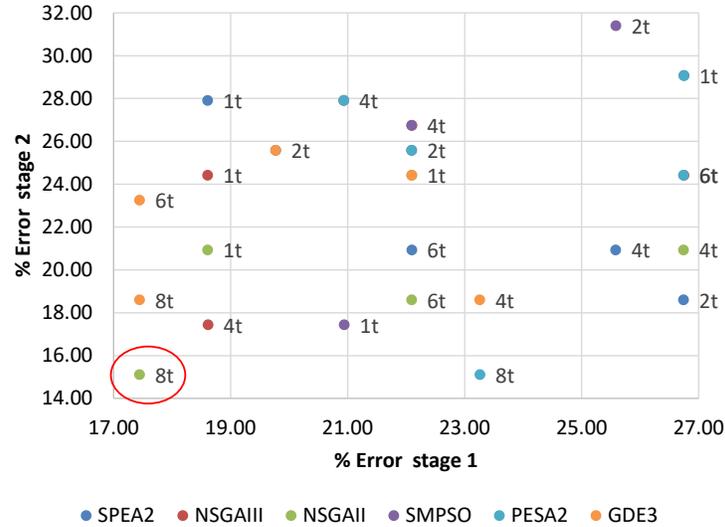
Figure 2: Time vs Error in 1st and 2nd stages

Figure 3: Best points for each thread setting and MOEA method.

5 CONCLUSIONS

This work has studied the improvements in terms of time when multiobjective evolutionary algorithms (MOEAs) are used for the task of feature weighting. Leveraging the use of multiple threads, our approach has needed less time to calculate those weights than a Simulated Annealing-based solution when carrying out the execution of the MOEAs. The accuracy values obtained by both approaches has resulted to be quite close, so we can conclude that MOEAs have provided an interesting advance to accelerate the weighting task in the migraine scenario.

Hence, given the remarkable reduction in execution time, in the future it could be beneficial to extend the use of the MOEAs for other stages of the treatment. Furthermore, given the runtime savings, we could increase the number of iterations to explore more solutions.

REFERENCES

Alba, E., and M. Tomassini. 2002. "Parallelism and evolutionary algorithms". *IEEE transactions on evolutionary computation* vol. 6 (5), pp. 443–462.

Cady, R. K., C. P. Schreiber, J. A. Porter, A. M. Blumenfeld, and K. U. Farmer. 2011. "A multi-center double-blind pilot comparison of OnabotulinumtoxinA and Topiramate for the prophylactic treatment of chronic migraine". *Headache* vol. 51 (1), pp. 21–32.

Cernuda-Morollón, E., C. Ramón, D. Larrosa, R. Alvarez, N. Riesco, and J. Pascual. 2015. "Long-term experience with onabotulinumtoxinA in the treatment of chronic migraine: What happens after one year?". *Cephalalgia* vol. 35 (10), pp. 864–868.

Chen, T., K. Tang, G. Chen, and X. Yao. 2012. "A large population size can be unhelpful in evolutionary algorithms". *Theoretical Computer Science* vol. 436, pp. 54–70.

Coello, C. A. C., G. B. Lamont, D. A. Van Veldhuizen et al. 2007. *Evolutionary algorithms for solving multi-objective problems*, Volume 5. Springer.

Corne, D. W., N. R. Jerram, J. D. Knowles, and M. J. Oates. 2001. "PESA-II: Region-based selection in evolutionary multiobjective optimization". In *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*, pp. 283–290. Morgan Kaufmann Publishers Inc.

De Vicente, J., J. Lanchares, and R. Hermida. 2000. "Adaptive FPGA placement by natural optimisation". In *Rapid System Prototyping, 2000. RSP 2000. Proceedings. 11th International Workshop on*, pp. 188–193. IEEE.

Deb, K., and H. Jain. 2014. "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints.". *IEEE Trans. Evolutionary Computation* vol. 18 (4), pp. 577–601.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. "A fast and elitist multiobjective genetic algorithm: NSGA-II". *IEEE transactions on evolutionary computation* vol. 6 (2), pp. 182–197.

Durillo, J. J., A. J. Nebro, F. Luna, and E. Alba. 2008. "A study of master-slave approaches to parallelize NSGA-II". In *2008 IEEE International Symposium on Parallel and Distributed Processing*, pp. 1–8. IEEE.

Gasbarrini, A., A. L. De, G. Fiore, M. Gambrielli, F. Franceschi, V. Ojetti, E. Torre, G. Gasbarrini, P. Pola, and M. Giacovazzo. 1998. "Beneficial effects of Helicobacter pylori eradication on migraine.". *Hepato-gastroenterology* vol. 45 (21), pp. 765–770.

Grogan, P. M., M. V. Alvarez, and L. Jones. 2013. "Headache direction and aura predict migraine responsiveness to rimabotulinumtoxin B". *Headache* vol. 53 (1), pp. 126–136.

Hadka, D. 2019. "MOEA framework". http://moeaframework.org/. Accessed on 06-06-2019.

IHS 2013. "The International Classification of Headache Disorders, 3rd edition (beta version)". *Cephalalgia* vol. 33 (9), pp. 629–808. PMID: 23771276.

Jakubowski, M., P. J. McAllister, Z. H. Bajwa, T. N. Ward, P. Smith, and R. Burstein. 2006. "Exploding vs. imploding headache in migraine prophylaxis with botulinum toxin A". *Pain* vol. 125 (3), pp. 286–295.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. "Optimization by simulated annealing". *science* vol. 220 (4598), pp. 671–680.

Kukkonen, S., and J. Lampinen. 2005. "GDE3: The third evolution step of generalized differential evolution". In *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, Volume 1, pp. 443–450. IEEE.

Lainez, M., R. Gil, A. Salvador, A. Piera, and B. Lopez. 2006. "Unilateralism as a predictor of response in treatment of chronic headache patients with botulinum toxin". *Headache* vol. 46 (5), pp. 846:F12.

Lipton, R., S. Varon, B. Grosberg, P. McAllister, F. Freitag, S. Aurora, D. W. Dodick, S. Silberstein, H. Diener, R. DeGryse et al. 2011. "OnabotulinumtoxinA improves quality of life and reduces impact of chronic migraine". *Neurology* vol. 77 (15), pp. 1465–1472.

Lovati, C., and L. Giani. 2017. "Action mechanisms of Onabotulinum toxin-A: hints for selection of eligible patients". *Neurological Sciences* vol. 38 (1), pp. 131–140.

Mathew, N. T., and S. F. A. Jaffri. 2009. "A double-blind comparison of OnabotulinumtoxinA (Botox) and Topiramate (Topamax) for the prophylactic treatment of chronic migraine: a pilot study". *Headache* vol. 49 (10), pp. 1466–1478.

Mathew, N. T., J. Kailasam, and L. Meadors. 2008. "Predictors of response to botulinum toxin type A (BoNTA) in chronic daily headache". *Headache* vol. 48 (2), pp. 194–200.

Mirjalili, S., S. M. Mirjalili, and A. Hatamlou. 2016. "Multi-verse optimizer: a nature-inspired algorithm for global optimization". *Neural Computing and Applications* vol. 27 (2), pp. 495–513.

Natoli, J., A. Manack, B. Dean, Q. Butler, C. Turkel, L. Stovner, and R. Lipton. 2010. "Global prevalence of chronic migraine: a systematic review". *Cephalalgia* vol. 30 (5), pp. 599–609.

Nebro, A. J., J. J. Durillo, J. Garcia-Nieto, C. C. Coello, F. Luna, and E. Alba. 2009. "Smpso: A new pso-based metaheuristic for multi-objective optimization". In *Computational intelligence in miulti-criteria decision-making, 2009. mcdm'09. ieee symposium on*, pp. 66–73. IEEE.

Oterino, A., C. Ramón, and J. Pascual. 2011. "Experience with onabotulinumtoxinA (Botox) in chronic refractory migraine: focus on severe attacks". *The Journal of Headache and Pain* vol. 12 (2), pp. 235–238.

Pagola, I., P. Esteve-Belloch, J. Palma, M. Luquin, M. Riverol, E. Martinez-Vila, and P. Irimia. 2014. "Predictive factors of the response to treatment with onabotulinumtoxinA in refractory migraine". *Revista de Neurologia* vol. 58 (6), pp. 241–246. PMID:9925225.

Parrales, F., A. A. Del Barrio García, M. Gallego de la Sacristana, Mercedes, J. Guerrero Peral, A, and J. L. Ayala Rodrigo. 2019. "Prediction of Patient's Response to OnabotulinumtoxinA Treatment for Migraine". *Heliyon* vol. 5 (3), pp. e01043.

Parrales Bravo, F., A. Del Barrio García, M. Gallego de la Sacristana, L. López Manzanares, J. Vivancos, and J. Ayala Rodrigo. 2017. "Support system to improve reading activity in parkinson's disease and essential tremor patients". *Sensors* vol. 17 (5), pp. 1006.

Ram, D. J., T. Sreenivas, and K. G. Subramaniam. 1996. "Parallel simulated annealing algorithms". *Journal of parallel and distributed computing* vol. 37 (2), pp. 207–212.

Sandrini, G., A. Perrotta, C. Tassorelli, P. Torelli, F. Brighina, G. Sances, and G. Nappi. 2011. "Botulinum toxin type-A in the prophylactic treatment of medication-overuse headache: a multicenter, double-blind, randomized, placebo-controlled, parallel group study". *The Journal of Headache and Pain* vol. 12 (4), pp. 427–433.

Schulman, E. A., A. E. Lake, P. J. Goadsby, B. L. Peterlin, S. E. Siegel, H. G. Markley, and R. B. Lipton. 2008. "Defining refractory migraine and refractory chronic migraine: proposed criteria from the Refractory Headache Special Interest Section of the American Headache Society". *Headache* vol. 48 (6), pp. 778–782.

Silberstein, S. D., D. W. Dodick, S. K. Aurora, H.-C. Diener, R. E. DeGryse, R. B. Lipton, and C. C. Turkel. 2015. "Per cent of patients with chronic migraine who responded per onabotulinumtoxinA treatment cycle: PREEMPT". *J Neurol Neurosurg Psychiatry* vol. 86 (9), pp. 996–1001.

Van Veldhuizen, D. A., and G. B. Lamont. 2000. "Multiobjective evolutionary algorithms: Analyzing the state-of-the-art". *Evolutionary computation* vol. 8 (2), pp. 125–147.

Witten, I. H., E. Frank, M. A. Hall, and C. J. Pal. 2016. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Wu, X., V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, S. Y. Philip et al. 2008. "Top 10 algorithms in data mining". *Knowledge and Information Systems* vol. 14 (1), pp. 1–37.

Yang, M., R. Rendas-Baum, S. F. Varon, and M. Kosinski. 2011. "Validation of the Headache Impact Test (HIT-6^{TM}) across episodic and chronic migraine". *Cephalalgia* vol. 31 (3), pp. 357–367.

Zitzler, E., K. Deb, and L. Thiele. 2000. "Comparison of multiobjective evolutionary algorithms: Empirical results". *Evolutionary computation* vol. 8 (2), pp. 173–195.

Zitzler, E., M. Laumanns, and L. Thiele. 2001. "SPEA2: Improving the strength Pareto evolutionary algorithm". *TIK-report* vol. 103.

Zitzler, E., and L. Thiele. 1999. "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach". *IEEE transactions on Evolutionary Computation* vol. 3 (4), pp. 257–271.

AUTHOR BIOGRAPHIES

FRANKLIN PARRALES BRAVO achieved the M.Sc. degree in computer science from Complutense University of Madrid (UCM), Madrid, Spain, in 2015. His current research interests include eHealth, data mining, and machine learning. He is PhD student at UCM and also he works as a Associate Professor of Computer Science at University of Guayaquil, Ecuador. He is currently focused on Data processing methodologies in the area of E-Health for categorizing therapeutic responses in patients with migraine. His email address is fparrale@ucm.es.

ALBERTO A. DEL BARRIO received the Ph.D. degree in Computer Science from the Complutense University of Madrid (UCM), Madrid, Spain, in 2011. Since 2017, he has been an Interim Associate Professor of Computer Science with the Department of Computer Architecture and System Engineering, UCM. His research interests include Design Automation, Arithmetic as well as Video Coding Optimizations. His email address is abarriog@ucm.es.

JOSE LUIS AYALA received the MSc and PhD degrees in Telecommunication Engineering from the Technical University of Madrid, Spain, in 2001 and 2005, respectively. He is currently an Associate Professor in the Department of Computer Architecture and Automation, UCM. His research interests include thermal and energy-aware design and management of processor-based systems, design of embedded processors, thermal estimation, 3D integration, health monitoring, and wireless sensor networks. His email address is jayala@ucm.es.