# An Optimal Approach for Low-Power Migraine Prediction Models in the State-of-the-Art Wireless Monitoring Devices

Josué Pagán*†, Ramin Fallahzadeh‡, Hassan Ghasemzadeh‡, José M. Moya§†, José L. Risco-Martín*
and José L. Ayala*
*DACYA, Complutense University of Madrid, Madrid, Spain.
E-mail: {jpagan,jlrisco,jayala}@ucm.es
†CCS-Center for Computational Simulation, Madrid, Spain
‡EPSL, Washington State University, Pullman, Washington, USA.
E-mail: {rfallahz,hassan}@eecs.wsu.edu
§LSI, Technical University of Madrid.
E-mail: {josem}@die.upm.es

***Abstract*—Wearable monitoring devices for ubiquitous health care are becoming a reality that has to deal with limited battery autonomy. Several researchers focus their efforts in reducing the energy consumption of these motes: from efficient micro-architectures, to on-node data processing techniques. In this paper we focus in the optimization of the energy consumption of monitoring devices for the prediction of symptomatic events in chronic diseases in real time. To do this, we have developed an optimization methodology that incorporates information of several sources of energy consumption: the running code for prediction, and the sensors for data acquisition. As a result of our methodology, we are able to improve the energy consumption of the computing process up to $90\%$ with a minimal impact on accuracy. The proposed optimization methodology can be applied to any prediction modeling scheme to introduce the concept of energy efficiency. In this work we test the framework using Grammatical Evolutionary algorithms in the prediction of chronic migraines.**

## I. Introduction

Health care in unobtrusive monitoring scenarios has become one of the most challenging areas of the Internet of Things (IoT). Health care applications are being exploited in both commercial and research environments. We can find examples in proactive personal eHealth control (1), health-related timely interventions (2), monitoring of illnesses related to functional impairment (3) or Parkinson disease (4), and it also plays a crucial role in Ambient Assisted Living tools designed for elders (5). Among all these applications, the core functionality is detection and classification of activities or events such as heart failure. In chronic diseases with symptomatic events the prediction of incoming events is crucial. Per-patient modeling of the behavior of chronic diseases has proved that the implementation of such kind of methodologies is applicable in a real ubiquitous health care system. Examples of prediction systems are the modeling of glycemia in humans (6), or the prediction of migraines (7).

The migraine is one of the most disabling neurological diseases. It is a chronic headache that presents symptomatic crisis. The prevalence of the migraine is around the 10% worldwide (8; 9), and leads to high economic costs for national and private health systems (10). A cascade of neurological processes precedes a migraine, but not always these processes are perceived by the patient. The most efficient way to avoid the pain is the intake in advance of specific drugs. Thus, a prediction system becomes necessary. The problem of the prediction of the migraine has already been tackled by the authors. First approaches using classical state-space algorithms (7) and heuristic models using GE—Grammatical Evolution (11)—have shown that the predictive modeling of the migraine is possible. In spite of this paper shows our work in migraine prediction using a modeling-box that applies GE models (see Figure 1), the main goal of this paper is focused on a methodology framework to optimize energy for wearable devices used in ambulatory monitorization for chronic diseases presenting symptomatic crises.

Longer battery operation means less disruptions and data losses. Several research teams work on reducing the energy consumption in order to enlarge the battery life, from architecture (12) to system level (13; 14). Signal management and processing studies also focus in reducing the energy consumption of these monitoring devices, as well as the study of dependence of battery life with the sampling rate of the biomedical signals (15) or pre-processing techniques such as compressed sensing (16).

The form factor of the battery limits the size of the monitoring device such that high density integration is still a challenge to deal with. Current commercial approaches use low power microcontroller architectures and high

978-3-9815370-8-6/17/$31.00  

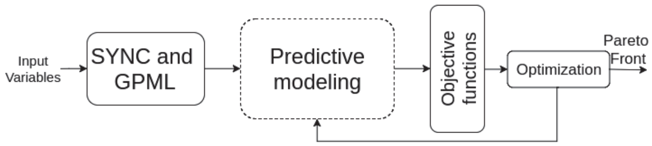
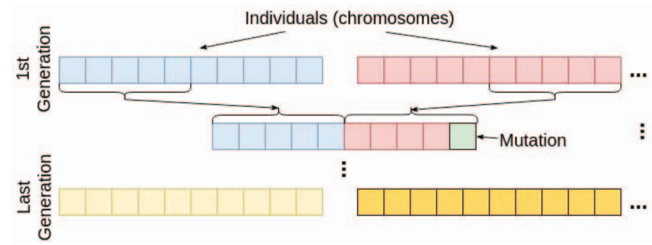



Fig. 1. Overview of the workflow of the proposed methodology. For further details of the SYNC and GPML block, see (7).




(a)

Fig. 2. Example of the population, individuals, crossover and mutation of the GE predictive modeling.

performance microcontrollers[1], as well as more efficient wireless communication interfaces.

This work is based on a previous study (11), and targets the goal of increasing the operation time of the battery by the reduction of the energy consumption. This is achieved through optimization of two sources of energy consumption: i) reducing the complexity of the processing in the embedded microprocessor of a monitoring device, that leads to reduce the number of clock cycles required to execute the code and, ii) reducing the consumption of peripherals using the minimum number of sensors. In this work, we focus on low performance microcontrollers to deal with the migraine prediction problem.

## II. Methodology overview

This section describes the workflow followed in the optimization process depicted in Figure 1. After the data acquisition and the offline data processing—see details for Synchronization and Gaussian Process Machine Learning blocks in (7)—the predictive modeling box generates an output that represents the prediction of the system. This output depends on past values of the input variables. To illustrate the process along this paper, we use previous knowledge of our research group and implement Grammatical Evolutionary algorithms in the prediction modeling box. In the same way, we are going to use biometric variables as system inputs to model the output of a migraine event. The energy optimization framework proposed in this paper can be applied to other prediction modeling systems.

The GE algorithms create a set of mathematical expressions from which we extract three optimization objectives. These objectives are: i) the accuracy or *fit* of the predicted values, ii) the number of clock cycles #clk that the expression takes to be computed in the embedded microcontroller, and iii) the energy consumption $\varepsilon_{sensing}$ of the sensors used.

Some constrains have been set: $0 \leq fit \leq 100$, and only solutions that use data provided by any of the sensors are considered. The multi-objective problem can be formulated as follows:

$$\min\left(-fit, \#clk, \varepsilon_{sensing}\right) \quad (1)$$

[1]Snapdragon: https://www.qualcomm.com/products/snapdragon/

### A. Energy optimization in the loop

In previous works, the authors have studied different prediction modeling frameworks that can be introduced in the modeling-box in Figure 1. Classic state-space algorithms were used in the work presented in (7). In (11), the authors found that the predictive modeling of the symptomatic events of a migraine can be detected through mathematical expressions generated by GE algorithms. Because the goal of this paper is the energy optimization methodology and not the models themselves, in this paper we introduce the energy optimization loop in the already developed GE predictive model where minor changes have been introduced. For the sake of clarity, in the following we show the basics of our multi-objective optimization using GE.

GE (17) is a grammar-based form of Genetic Programming (GP), used to generate programs in any language. GE algorithms are based on Genetic Algorithms (GAs). A GA selects a group of production rules—encoded as integers—expressed in a Backus Naur Form (BNF). Based on biology, the GA evolves a population formed by a set of individuals. A chromosome is an array of integer numbers (genes) that represents an individual, and each one of these define the rule of the BNF to be applied. As shown in Figure 2a, individuals mutate and mix each other to create new ones in every generation. Each individual is evaluated, and those of each generation that better fit with the objective, evolve and will survive with a higher probability in future generations.

The optimization is based on the Non-dominated Sorting Genetic Algorithm II (NSGA-II) (18). NSGA-II is an elitist approach, i.e., a small part of the best candidates remain unchanged into the next generation—they remain as parents of the next generation. Instead of a single best solution, the NSGA-II selects a set of non-dominated solutions (usually called Pareto front). An individual belongs to the Pareto front and is called non-dominated if any of its objectives can be improved (minimized in our case) without degrading the others. To know more about multi-objective GE implementation see (19).

To train the models, 500000 generations with a population of 250 individuals each have been used. The probability of crossover was set to 0.9 and the probability of mutation was set to 0.083—the inverse of number of rules in the BNF. The length of the chromosomes is 100, and wrapping is not allowed during the decoding process.

To understand the basics of BNF see (11). The BNF used in (11) is the first record in literature of the approach used in this work. However, a minor modification has been introduced to seek for more realistic solutions: autoregressive expressions using information of the real pain reported by the migraine sufferer are not allowed.

### B. The objective functions

In the following lines we define the objective functions in Eq. 1, and we give a detailed explanation of them.

*1) The fit:* measures the likeness of the predicted value by the system $\hat{y}$—the migraine pain level in our case—and the known real output $y$—in our case, the symptomatic curve reported by the patient. The *fit* definition is based on the Normalized Root Mean Square Error (NRMSE) as defined in (20). As we want to maximize this value, the optimization process tries to minimize its opposite value.

$$fit = 100 \times \left(1 - \frac{\|y - \hat{y}\|}{\|y - mean(y)\|}\right) \tag{2}$$

*2) The number of clock cycles:* The main goal of this paper is the reduction of the energy consumption of an ambulatory monitorization device. As aforementioned, one of the ways to achieve it is through the reduction of $\varepsilon_{\mu C}$, the energy consumption of the microcontroller that eventually will compute the migraine prediction in real time. Given the impossibility of measuring the real consumption of the monitoring device to introduce this value into the optimization loop, several assumptions were taken. The authors did not find neither a simulator or instruction level energy consumption table for low performance microcontrollers.

The main assumption was that the number of clock cycles that the microcontroller takes in the execution of the prediction is proportional to its energy consumption, $\varepsilon_{\mu C} \propto \#clk_S$. The predictive modeling box is, in most of the cases, data dependant. In our case using GE, this dependence appears in the output range of the mathematical expressions generated. Data from sensors might be bounded, but not the result when they are introduced in a mathematical function. Thus, trying to make a profile via a static code analysis is not possible. For the same reason, any strategy of branch prediction could not be carried out.

We also assumed that the total number of clock cycles for a GE solution $S$ is computed as the sum shown in Eq. 3. Where the number of clock cycles for a given function $f_i \in F$ is the average value of the clock cycles of all branches $B$ in the compiled code for a function (Eq. 4).

$$\#clk_S = \sum_{i=1}^{\forall f \in F in S} \#clk_{f_i} \tag{3}$$

$$\#clk_{f_i} = \frac{\sum_{b=1}^{B} \#clk_b}{B} \tag{4}$$

The set of functions $F$ is defined in the BNF. The number of cycles for the execution of a code depends strongly on the microcontroller architecture: hardware multipliers, float point unit, vector unit, *etc*. We have based our study in the well-known microcontroller MSP43016F1101, a low-performance 16-bit microcontroller suitable for monitoring devices. Its maximum frequency is 8 MHz, and it has 10 kB of RAM and 48 kB of flash memory.

The assembler code for each instruction in $F$ was obtained using the `msp430-gcc` cross-compiler. All clock cycles and branches were calculated manually, according to the technical specifications of the microcontroller. The number of clock cycles for unresolved calls of internal functions of the tool-chain were calculated as well. In those cases where calls to kernel functions were not able to be resolved, the mathematical functions where approximated by 10-degree Taylor series—only $sine$ and $cosine$ functions were approximated. The standard `math` library was used for the others.

*3) The sensing energy consumption:* $\varepsilon_{sensing}$ draws the amount of energy consumption due to the use of every sensor. As the sampling frequency and hardware of each sensor is different, we would prefer to use the minimum number of the least consuming sensors that allow the maximum *fit*.

## III. Experimental setup

In this paper we use real data from two migraine patients acquired with the commercial monitoring device PLUX-Wireless Biosignals[2] and a smartphone application. For more information about these data see (7). We have focused our optimization in the architecture of the microcontroller available in the open source commercial Shimmer device[3] that integrates the microprocessor MSP43016F1101. With this choice, we will be able to implement the obtained models in the device to compare their energy performance against the baseline scenario: monitoring devices without predictive intelligence that sense using all sensors and transmit the data wirelessly.

The biometric variables and their associated sensors are: i) the skin surface temperature (TEMP) using a thermistor, ii) the electrodermal activity (EDA) using two differential electrodes, iii) the electrocardiogram (ECG) to extract the Heart Rate (HR) using two leads with one reference, and iv) the blood oxygen saturation (SpO2) using the 8000R SpO2 sensor and the OEM-III board (both from NONIN[4]). Data processing is performed offline as shown in (7).

For the optimization process, the number of clock cycles for each function, and the energy consumption of each sensor are needed. According to the technical reports of the microcontroller and the criteria established in Section II, the clock cycles for the base functions of the microcontroller are: 184 for `addition`, 177 for `subtraction`, 153 for

[2]BioPlux: http://www.biosignalsplux.com
[3]Shimmer: http://www.shimmersensing.com/
[4]NONIN: http://www.nonin.com

`HW multiplication`, 37 for `comparison`, and 405 for `division`. For each mathematical function individually compiled from the $F$ set, the average number of clock cycles of all possible paths are: 4443 for $e$, 6416 for `cos`, 6612 for `sin`, 1079 for `sqrt`, 12344 for `pow`, 4890 for `log` and, `max` and `min` values are proportional to the number of `comparisons`. Clock cycles for the computation of the HR were introduced too. To this end we based our work on the implementation by Boichat *et al.*in (21); it takes 4672432 cycles (almost 0.5 seconds in a microcontroller running at 8 MHz).

For the third objective of the optimization problem we have used a HAMEG HM8012 digital multimeter to measure the consumption of the sensors. TEMP and EDA are analog sensors, sampled every 5 seconds, that consume $0.32mJ(-4.9dBm)$ each in the Shimmer mote. The ECG is gathered at 250 Hz leading to an energy consumption of $396mJ(26dBm)$. The consumption of the OEM-III and the SpO2 sensor is far from these values, and it is $3665mJ(35.6dBm)$. The digital module OEM-III samples the SpO2 at 75 Hz, and sends through the microcontroller 3 Kbit of data per second. The presence or absence of this sensor will have a high impact in the optimization process of the GE expressions. The HR computation leads to $34mJ(15.3dBm)$ extra, approximately.

In this work we have used the HERO Java library[5] published under GPL license. Based on the results provided in (11), only 10 and 20 minutes prediction horizon models have been explored. The experiments lasted several days and were executed in a heterogeneous distributed cluster of PCs with Intel Core i7-4710HQ and AMD A8-6600K CPU, with 16 GB and 8 GB of memory and, Ubuntu and CentOS respectively.

## IV. Evaluation and results

In this section we evaluate the proposed energy optimization methodology. The results of the optimization problem are shown. As mentioned in Section II, the main goal of this paper is to optimize the energy consumption of the target system as an additional objective of thecost function. Therefore, we will not analyze in detail the obtained GE models for migraine prediction.

To create the models, we have used 10 and 5 random migraines from data of two female migraineurs, labeled as Patient A and Patient B respectively. All models have been trained for predictions of 10 and 20 minutes ahead. As a result of each training experiment we obtain a tri-dimensional Pareto front. The Pareto front ideally contains 250 different solutions (the size of the population). In most of the cases, repeated solutions appear; thus, a Pareto front reduction is carried out to remove them.

The goodness of all the solutions in the Pareto front are, objectively, the same, because they are non-dominated solutions. We must choose subjectively one criterion to select the models to be embedded in the monitoring device. To maintain the maximum benefit for the patients, we have selected those models with maximum fit. Figures 3a through 3c show the 3D Pareto front and one of its 2D projections for models trained with 10 minutes of prediction horizon for both of the patients. The solutions chosen are rounded by a red circle. All Pareto fronts obtained are globally convex: the solutions tend to minimize the error, the number of clock cycles and the energy due to the sampling of sensors. To improve the visualization and reading, the *#clk* is shown as thousands of cycles and the energy of sampling is shown in *dBm*.

TABLE I
BEST FIT SOLUTIONS IN THE PARETO FRONTS FOR PATIENT A.

| idx | fit (%) | | #clk (k-cycles) | | $\varepsilon_{sensing}(dBm)$ | | $\overline{U}$ | |
|---|---|---|---|---|---|---|---|---|
| | 10' | 20' | 10' | 20' | 10' | 20' | 10' | 20' |
| 1 | 79.2 | 77.4 | 74.3 | 24.8 | 35.6 | -1.9 | 0.56 | 0.51 |
| 2 | 94.8 | 73.9 | 61.9 | 17.4 | 35.6 | -1.9 | 0.52 | 0.00 |
| 3 | 91.6 | 92.4 | 11.6 | 106.1 | 35.6 | -4.9 | 0.32 | 0.55 |
| 5 | 79.3 | 64.6 | 1.9 | 28.6 | -4.9 | 35.6 | 0.64 | 0.52 |
| 6 | 75.2 | 76.7 | 21.8 | 21.7 | -4.9 | 35.6 | 0.00 | 0.00 |
| 7 | 75.7 | 75.7 | 23.0 | 16.6 | -1.9 | 35.6 | 0.34 | 0.41 |
| 10 | 77.1 | 76.3 | 21.7 | 75.7 | 35.6 | -1.9 | 0.00 | 0.51 |
| 11 | 81.4 | 75.8 | 21.7 | 51.7 | 35.6 | 35.6 | 0.3 | 0.51 |
| 12 | 72.9 | 69.5 | 129.6 | 24.3 | -1.9 | 35.6 | 0.36 | 0.53 |
| 15 | 76.2 | 84.2 | 38.2 | 86.3 | -4.9 | 35.6 | 0.61 | 0.44 |

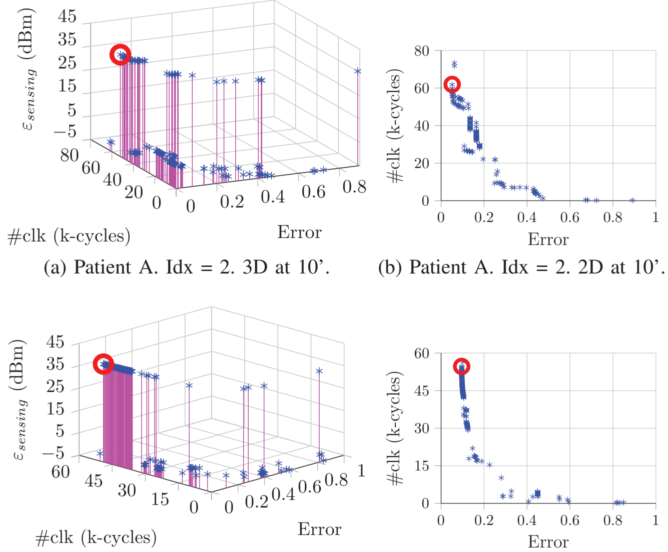
TABLE II
BEST FIT SOLUTIONS IN THE PARETO FRONTS FOR PATIENT B.

| idx | fit (%) | | #clk (k-cycles) | | $\varepsilon_{sensing}(dBm)$ | | $\overline{U}$ | |
|---|---|---|---|---|---|---|---|---|
| | 10' | 20' | 10' | 20' | 10' | 20' | 10' | 20' |
| 2 | 81.6 | 83.6 | 16.6 | 52.1 | -1.9 | 35.6 | 0.00 | 0.22 |
| 3 | 82.0 | 77.5 | 4898.7 | 4697.3 | 26 | 36.1 | 0.49 | 0.48 |
| 4 | 80.3 | 81.1 | 31.6 | 26.4 | 35.6 | -4.9 | 0.61 | 0.62 |
| 5 | 90.4 | 81.4 | 54.7 | 36.2 | 35.6 | -1.9 | 0.57 | 0.73 |
| 7 | 89.4 | 81.4 | 4773.1 | 48.1 | 26 | -4.9 | 0.63 | 0.54 |

In Table I and Table II we show all the models for Patient A and Patient B at 10 and 20 minutes of prediction horizon. In general, the average *fit* is higher at 10 minutes. Despite the average fit is still high at 20 minutes, the average number of clock cycles is higher. This leads to a longer computation time to get the prediction and thus a higher energy consumption. Regarding the sampling, the presence or not of the SpO2 sensor—consuming $35.6dBm$—makes the energy results vary quite much. For Patient A, it seems that the SpO2 sensor is very important for the prediction of her migraine. For Patient B it seems that the HR is an important variable, that makes increase rapidly the number of clock cycles.

From these solutions with the best fit, we can compute the average symmetric uncertainty $\overline{U}$. With this, following the ideas that Ghasemzadeh *et al.* pose in (14), we can split our GE expressions into a combination of features, if possible, and we can estimate wether the features that compose the GE expressions are representative, or not. The symmetric uncertainty $U(X,Y)$ is a normalization of the mutual information $I(X,Y)$ between two discrete random variables $X$ and $Y$, defined as in Eq. 5. $U$ is bounded such that

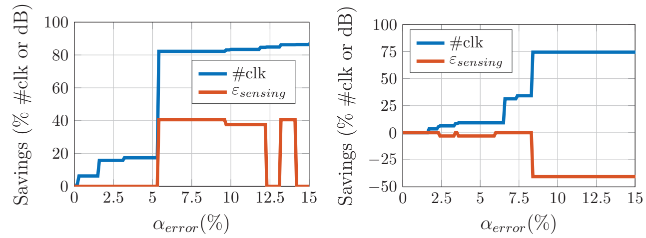
[5]HERO: https://github.com/jlrisco/hero




(a) Patient A. Idx = 2. 3D at 10'. (b) Patient A. Idx = 2. 2D at 10'.

(c) Patient B. Idx = 5. 3D at 10'. (d) Patient B. Idx = 5. 2D at 10'.

Fig. 3. 3D and 2D views of the Pareto Front result of the optimization process.




(a) Patient A. Idx = 3 at 10'. (b) Patient B. Idx = 4 at 20'.

Fig. 4. Savings in energy consumption from the perspective of number of clock cycles in the model execution and the energy consumed by the sensors in the sensing process.

$U \in [0, 1]$, and the lower the $U$, the better. The idea can be introduced for efficient feature selection and can be adapted to any problem that the reader desires. For the sake of simplicity, we compute the average symmetric uncertainty between all pairs of features—as a worst case of overlapping of the individual entropy—instead of the complex computation of the multivariate symmetric uncertainty.

$$U(X, Y) = \frac{2I(X, Y)}{H(X) + H(Y)} \tag{5}$$

In Eq. 5 $H(X)$ and $H(Y)$ represent the entropy of the random variables $X$ and $Y$, and $I(X, Y)$ is the mutual information defined as:

$$I(X, Y) = H(X) - H(X|Y) \tag{6}$$

As we can see in Table I and Table II, $\overline{U}$ has medium values, that in most of the cases is lower than $0.5$. With this, we can say that the features that compose the GE expressions are relevant enough and they do not overlap each other. This means that the GE expressions do not compute unnecessary features and the energy consumption of the microcontroller is thus optimized.

### A. The concern of the energy consumption

Peculiarities of a GE predictive modeling box in the optimization framework is that the selected models are a result of the casuistic of the metaheuristic process that it is not deterministic. Nevertheless, with any other modeling solution the reader can also follow the underneath study.

Depicted Figures 4a through 4b show an example of all the solutions of a Pareto front that satisfy a maximum error difference of $\alpha_{error}$ with the best-fit solution. These figures represent the savings of clock cycles %—proportional to the energy consumption of the microcontroller—and the energy savings (in dB) due to the selection of another set of sensors. The abscissa represents $\alpha_{error}$, or the degree of error when we select a different solution of the Pareto front.

In Figure 4a we can see solutions for migraine *idx = 3* of Patient A at 10 minutes. We can see that we can save clock cycles if we tolerate more error, and in addition, the selected sensors consume as much as the ones selected in the best-fit solution. At a certain point, the savings in clock cycles get over the $90\%$ with the good news that the set of selected sensors does not require the SpO2 sensor anymore, and the savings due to sensing reaches 40 dB. The opposite occurs in Figure 4b—$4th$ migraine from Patient B at 20 minutes. Any of the remaining solutions in the Pareto front saves energy from sampling; furthermore, when the saving in the number of clock cycles gets almost the $80\%$, to maintain that needs from the SpO2 sensor and the consumption due to sampling increases (negative savings in dB). The sacrifice of error tolerance at 20 minute does not have such reward but a penalty of higher consumption. With this, we can apply a methodology for model selection as the one presented in (22).

## V. Conclusions

In this work we take an existing predictive modeling methodology for unobtrusive ambulatory monitorization in the health care in the IoT, and we integrate the concept of energy efficiency as a target objective. To tackle this problem we have considered two main sources of energy consumption: i) the consumption due to the code execution to perform the prediction in real time (measured through clock cycles execution) and, ii) the set of sensors that perform the input data sensing process. This leads to a multi-objective optimization framework that can be applied to any existing predictive modeling technique and application. As a case study, the authors have used an already validated migraine predictive model using Grammatical Evolutionary algorithms. The results show that for 10 and 20 minutes of prediction horizon we can reach savings up to $90\%$ in the execution time barely degrading the accuracy of the models, what it directly impacts on the energy consumption of the system.

ACKNOWLEDGEMENTS

Research by Josué Pagán has been funded by the EU (FEDER) under Research grant PI15/01976 and the Spanish Ministry of Economy and Competitiveness under Research Grants TIN 2015-65277-R and TEC2012-33892.

REFERENCES

[1] Y.-L. Zheng, X.-R. Ding, and et al., "Unobtrusive sensing and wearable devices for health informatics," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 5, pp. 1538–1554, 2014.

[2] H. Alemdar and C. Ersoy, "Wireless sensor networks for healthcare: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[3] B. G. Steele, L. Holt, B. Belza, S. Ferris, S. Lakshminaryan, and D. M. Buchner, "Quantitating physical activity in copd using a triaxial accelerometer," *CHEST Journal*, vol. 117, no. 5, pp. 1359–1367, 2000.

[4] J. Klucken, J. Barth, and et al., "Unbiased and mobile gait analysis detects motor impairment in parkinson's disease," *PloS one*, vol. 8, no. 2, p. e56956, 2013.

[5] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 579–590, 2013.

[6] J. I. Hidalgo, J. M. Colmenar, J. L. Risco-Martin, A. Cuesta-Infante, E. Maqueda, M. Botella, and J. A. Rubio, "Modeling glycemia in humans by means of grammatical evolution," *Applied Soft Computing*, vol. 20, pp. 40–53, 2014.

[7] J. Pagán, M. I. De Orbe, A. Gago, M. Sobrado, J. L. Risco-Martín, J. V. Mora, J. M. Moya, and J. L. Ayala, "Robust and accurate modeling approaches for migraine per-patient prediction from ambulatory data," *Sensors*, vol. 15, no. 7, p. 15419, 2015.

[8] R. B. Lipton and A. I. Scher, "Epidemiology and economic impact of migraine," *Current Medical Research and Opinion®*, vol. 17, no. S1, pp. s4–12, 2001.

[9] WHO, "Headache disorders, fact sheet n°277," http://www.who.int/mediacentre/factsheets/fs277/en/, 2012, accessed: 2016-01-01.

[10] M. Linde, A. Gustavsson, L. Stovner, T. Steiner, J. Barré, Z. Katsarava, J. Lainez, C. Lampl, M. Lantéri-Minet, D. Rastenyte *et al.*, "The cost of headache disorders in europe: the eurolight project," *European journal of neurology*, vol. 19, no. 5, pp. 703–711, 2012.

[11] J. Pagán, , J. L. Risco-Martín, J. M. Moya, and J. L. Ayala, "Grammatical evolutionary techniques for prompt migraine prediction," in *Genetic and Evolutionary Computation Conference (GECCO), 2016*.

[12] R. Braojos Lopez and D. Atienza, "An ultra-low power nvm-based multi-core architecture for embedded bio signal processing," in *ICT-Energy conference 2016*, no. EPFL-CONF-222859, 2016.

[13] A. Tobola, O. Korpok, H. Leutheuser, B. Schmitz, C. Hofmann, M. Struck, C. Weigand, B. Eskofier, A. Heuberger, and G. Fischer, "System Design Impacts on Battery Runtime of Wearable Medical Sensors," in *Proceedings of International Conference on Mobile and Information Technologies in Medicine and Health*, D. Novák, Ed., 2014, pp. 1–4.

[14] H. Ghasemzadeh, N. Amini, R. Saeedi, and M. Sarrafzadeh, "Power-aware computing in wearable sensor networks: An optimal feature selection," *IEEE Transactions on Mobile Computing*, vol. 14, no. 4, pp. 800–812, 2015.

[15] A. Tobola, F. J. Streit, C. Espig, O. Korpok, C. Sauter, N. Lang, B. Schmitz, C. Hofmann, M. Struck, C. Weigand *et al.*, "Sampling rate impact on energy consumption of biomedical signal processing systems," in *Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on*. IEEE, 2015, pp. 1–6.

[16] R. Braojos Lopez, H. Mamaghanian, R. Vallejos, F. Javier, A. Dias Junior, G. Ansaloni, S. Murali, and D. Atienza Alonso, "Ultra-low power design of wearable cardiac monitoring systems," in *Proceedings of the IEEE/ACM 2014 Design Automation Conference (DAC)*, vol. 1. IEEE/ACM Press, 2014, pp. 140–145.

[17] C. Ryan and M. O'Neill, "Grammatical evolution: A steady state approach," *Late Breaking Papers, Genetic Programming*, vol. 1998, pp. 180–185, 1998.

[18] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

[19] J. M. Colmenar, J. L. Risco-Martin, D. Atienza, and J. I. Hidalgo, "Multi-objective optimization of dynamic memory managers using grammatical evolution," in *Proceedings of the 13th GECCO*. ACM, 2011, pp. 1819–1826.

[20] MATLAB, *version 8.5.0.197613 (R2015a)*. Natick, Massachusetts: The MathWorks Inc., 2015.

[21] N. Boichat, N. Khaled, F. Rincon, and D. Atienza, "Wavelet-based ecg delineation on a wearable embedded sensor platform," in *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*. IEEE, 2009, pp. 256–261.

[22] J. Pagán, J. L. Risco-Martín, J. M. Moya, and J. L. Ayala, "Modeling methodology for the accurate and prompt prediction of symptomatic events in chronic diseases," *Journal of biomedical informatics*, 2016.