A Machine Learning Approach to the Estimation of Near-Optimal Electrostatic Force in Micro Energy-Harvesters

Masoud Roudneshin[1], Kamran Sayrafian[2], Amir G. Aghdam[1]

[1]Department of Electrical Engineering, Concordia University, Montreal, Canada

[2] Information Technology Laboratory, National Institute of Standards & Technology, Gaithersburg, MD, USA

***Abstract*—Wearable medical sensors are one of the key components of remote health monitoring systems which allow patients to stay under continuous medical supervision away from the hospital environment. These sensors are typically powered by small batteries which allow the device to operate for a limited time. Any disruption in the battery power could lead to temporary loss of vital data. Kinetic-based micro-energy-harvesting is a technology that could prolong the battery lifetime or, equivalently, reduce the frequency of recharge or battery replacement. Focusing on a Coulomb-Force Parametric Generator (CFPG) micro harvesting architecture, several machine learning approaches are presented in this paper to optimally tune the electrostatic force parameter, and therefore, maximize the harvested power.**

***Index Terms*—energy harvesting, wearable sensors, microgenerator, CFPG**

I. INTRODUCTION

Wearable and implantable medical sensors are considered a key component of future telemedicine systems, allowing clinicians to have remote access to real-time patient's data [1], [2]. These devices typically operate by using small batteries; therefore, frequent recharge or battery replacement might be necessary to keep the device functioning properly. Prolonging the lifetime of these batteries and reducing their frequency of recharge could have a paramount impact on their everyday use. This is especially important for implanted devices, where battery replacement is not easily possible.

Energy harvesting refers to the process of scavenging energy from external sources (ambient environment such as solar power, wind and kinetic energy) [3]. For wearable devices, kinetic energy can be a reliable solution for power generation in medical sensors. For cases of nonstationary vibrations (for example, as a result of the human body motion), Coulomb-force parametric generator (CFPG) architecture has been proposed as a promising solution to extract power form human movements. [4], [5]. In this type of system, a proof mass can move between upper and lower bounds $\pm Z_l$ as shown in Fig. 1. The summation of the device motion to the inertial frame $\xi(t)$ and the relative motion of the proof mass with respect to the device $z(t)$ make the absolute motion of the proof mass equal to $y(t) = \xi(t) + z(t)$.

To model the dynamics of the proof mass motion in CPFG, the following nonlinear differential equation was proposed in [6]:

$$m\ddot{y}(t) = -m\ddot{z}(t) - F \times \text{Relay}(z(t)) \tag{1}$$

where m is the proof mass, $\ddot{y}(t)$ is the acceleration of the frame of CPFG with respect to the inertial frame, z is the

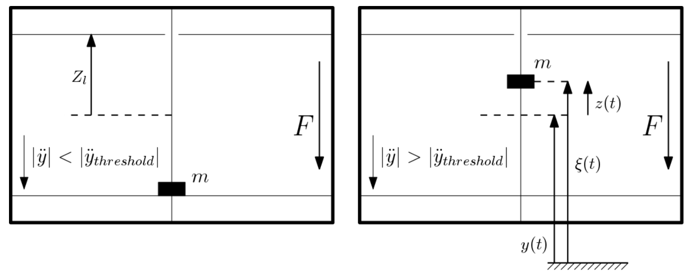

Fig. 1. Generic Model of a CPFG: (a) proof mass attached to one end, and (b) proof mass in flight

relative acceleration of the proof mass with respect to the frame of CFPG, and F is the electrostatic holding force which acts against the motion of the proof mass. The generated mechanical power of the system is equal to the product of the electrostatic holding force and the relative velocity of the proof mass with respect to frame and is calculated as:

$$P(t) = F \times \dot{z}(t) \tag{2}$$

Fig. 1 displays the process of energy generation in CPFG. The proof mass is initially located at either upper or lower plates. The mass does not move until the external acceleration exceeds a certain limit. To harvest power, the external acceleration must be strong enough to create a full displacement of the mass from one plate to the other. If the mass cannot complete a full flight to the other plate, all generated power is consumed in the electrical field of CFPG by the electrostatic force F.

A further investigation of the power equation of CFPG coupled with its dynamics reveals how adaptation of F affects the output power of the generator [7]. The power is a function of both relative velocity $\dot{z}(t)$ and force F. Meanwhile, $\dot{z}(t)$, as evident from equation (1), is a function of both external acceleration $\ddot{y}(t)$ and force F. Therefore, an optimal strategy that ensures proper adaptation of the electrostatic force F leads to the maximization of the average output power exists. As such, the following optimization problem is considered:

$$\underset{F_i}{\operatorname{argmax}} \left[\frac{1}{\Delta} \times \sum_{t=t_i}^{\Delta + t_i} P(t) \right], \tag{3}$$

with the constraints given by equations (1) and (2). In other words, it is desired to maximize the average harvested power during the time interval $[t_0 + (i-1)\Delta, t_0 + i\Delta]$ by selecting the optimal value of the electrostatic force F_i.

The authors in [6] demonstrate that the output power of a CFPG micro-harvester can be maximized by proper adjustment or adaptation of the electrostatic force F. A methodology for optimizing F by observing the input acceleration in the previous time interval is also proposed in [6]. The average output power for different values of the holding force F and various locations of the wearable sensor is evaluated in [7].

In this paper, we propose a novel method for the estimation of the suboptimal value of the electorstatic force in a micro energy-harvester, according to the current absolute acceleration of the CFPG frame. We use the frequency spectrum of the human body acceleration data in our analysis. Eight different machine learning classification schemes are then used and their performances are compared in terms of accuracy in estimating the suboptimal value of the holding force in the next time step for power maximization. To the best of the authors' knowledge, this is the first time the frequency spectrum of the human body acceleration is used for power maximization in CFPG microgenerators.

The rest of the paper is organized as follows. In Section II, we illustrate our proposed method, and the procedure for generating artificial data is explained. Then in Section III, we discuss various methods of classification of labeled data. In Section IV, results for eight classification structures is obtained and their accuracy is discussed. Finally, conclusions are drawn in Section V.

II. Problem Definition

A. Acceleration in Human Body

The authors in [8] demonstrate that during normal daily activities, bulk of the frequency content of the human motion acceleration in the upper extremity is within the range 0.8-5Hz. Also, it is shown in [9] that 99% of the acceleration power spectral density, when walking barefoot, is concentrated below 15Hz. Based on these findings, we make the following assumption.

Assumption 1 *The acceleration signal for time intervals of length Δ can be approximated by the following cosine series:*

$$\ddot{y}(t) \approx \sum_{n=0}^{20} A_n(i)\cos(2\pi f_n t) = \sum_{n=0}^{20} A_n(i)\cos(2\pi n t), \quad t \in [t_0 + (i-1)\Delta, t_0 + i\Delta], \tag{4}$$

where $A_n(i)$ is the amplitude of the frequency component corresponding to f_n in the ith time interval. Having the amplitudes of the frequency components in every time interval, the problem reduces to identifying the mapping ϕ such that:

$$\phi : [A_0(i), ..., A_{20}(i)] \to \tilde{F}_\Delta(i), \tag{5}$$

The ultimate goal is to maximize the harvested mechanical power in equation (2) for time interval i by finding a pseudo-optimal holding force $\tilde{F}_\Delta(i)$. Given the time dependency of

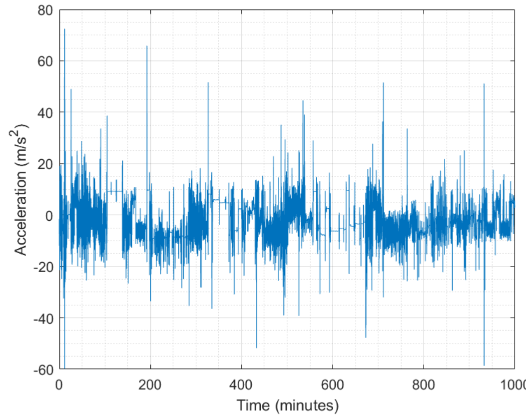

Fig. 2. A 1000-min acceleration sample of human arm motion

the parameters involved, we have divided this problem into two steps:

1) **Estimation of the pseudo-optimal holding force during time interval i:**
Assuming that we know the $A_n(i)$ coefficients for the acceleration signal in the time interval $[t_0+(i-1)\Delta, t_0+i\Delta]$, what is the pseudo-optimal holding force $\tilde{F}_\Delta(i)$ that maximizes average harvested power for the same time interval i.e. $[t_0 + (i-1)\Delta, t_0 + i\Delta]$?
2) **Estimation of the optimal interval size to maximize the average harvested power for the next time interval i.e. $[t_0 + i\Delta, t_0 + (i+1)\Delta]$:**
In practice, the information about $\tilde{F}_\Delta(i)$ can not be available for the current time interval $[t_0+(i-1)\Delta, t_0+i\Delta]$; and, at best, $\tilde{F}_\Delta(i)$ can be applied to the next time interval i.e. $[t_0+i\Delta, t_0+(i+1)\Delta]$. Therefore, depending on the temporal correlation of the estimated pseudo-optimal holding force, the achieved harvested power will be less than the value obtained in step 1. This step requires further study on optimizing the interval size Δ.

The focus of this paper is on solving the first step through machine learning algorithms. In the next subsection, we will present several methodologies for data classification in order to find a suitable mapping from the frequency spectrum of the acceleration to the electrostatic force F.

B. Acceleration Data Processing

The acceleration data of the human arm motion, as obtained in [10], is used in our analysis. Fig. 2 demonstrates a 1000-min data obtained by attaching an accelerometer to a male subject arm. The data is interpolated with 1ms sampling steps. Then, it is divided into intervals of length equal to one second for approximation with the cosine functions as stated in equation (4). To reduce the size of the action space, the values of the holding force F are divided into 1mN steps. Then, similar to the process outlined in [6] and [7], we choose a value for $\tilde{F}_\Delta(i)$ from the set $\{2, 3, ..., 10\}$.

To label each of the 1-sec intervals with their corresponding holding force which maximizes the harvested power in that

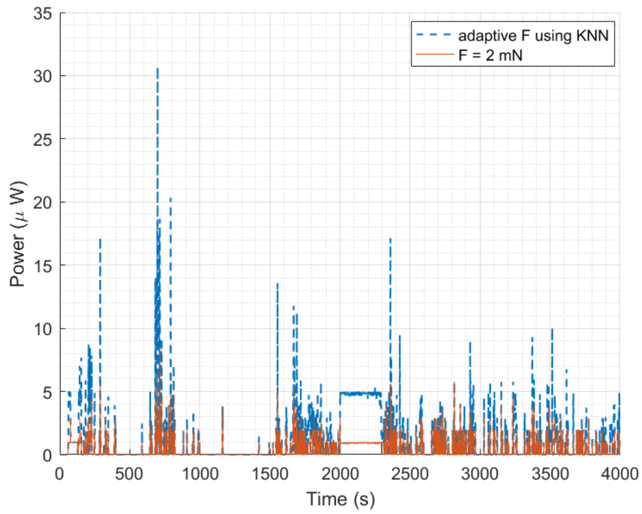

Fig. 3. Comparison of the harvested power for a 4000-sec data for the adaptive holding force with KNN algorithm and a constant holding force

time interval, an algorithm based on the work in [6] is implemented for each of the holding forces and the pseudo-optimal force is obtained accordingly. A set of 1950 labeled data for each of the nine classes (making a total of 17550 vectors) is obtained subsequently.

III. Data Classification

As discussed earlier, the estimation of the electrostatic holding force which maximizes the output power based on the spectral content of the acceleration signal as described in equation (4) can be studied in the context of a classification problem. In what follows, eight different classification schemes are briefly described.

- **Decision Tree Classifier [11]**
  A decision tree is a flowchart-like tree structure. An internal node represents a feature (or attribute), each branch represents a decision rule, and each leaf node represents an outcome. The best attribute is selected using an appropriate attribute selection measure (ASM) such as information gain or Gini index. Then, the selected attribute is used as a decision node, the dataset is broken into smaller subsets, and these steps are repeated recursively until a matching condition is satisfied.

- **Random Forest Classifier [12]**
  Random forest consists of a large number of individual decision trees that operate as an ensemble. A set of decision trees are created from a randomly selected subset of the training set. Then, the classifier aggregates the votes from different decision trees to decide the final class of the test object. Therefore, a classification is made based on the majority of votes received from each of the decision trees. It is to be noted that a single decision tree may be prone to noise, but aggregating many decision trees reduces the effect of noise, leading to more accurate results.
- **K-Neighbors Classifier [11]**
  In this method, an object is classified by a majority vote of its neighbors, with the object being assigned to the most common class among its k nearest neighbors.
- **Support Vector Machine (SVM) [13]**
  This is a discriminative classifier which is formally defined by a separating hyperplane. Given labeled training data, the algorithm outputs an optimal hyperplane which categorizes new entries. The hyperplane can be a linear, polynomial or exponential function of the weighted sum of inputs, where the weighting coefficients are optimized in the training process.
- **Multi Layer Perceptron (MLP) [14]**
  In this approach, each computational block consists of a weighting matrix that is multiplied by the inputs. The outputs are then passed to a function called *activation function*. These blocks are often repeated several times, and the output of each block is used as the input of the next one. Finally, the output, which is in the form of a vector, classifies the input in the form of an output vector. The elements of the weighting matrices are updated to minimize the error between the estimated vector and a label vector which is associated with the correct class.
- **Stochastic Gradient Descent (SGD) [15]**
  This is a linear classifier just like a linear hyperplane in SVM. The only difference is that the optimization technique applied to find the weighting parameters of the hyperplane is stochastic gradient descent. To update the weights, the gradient of loss function is needed. For the computation of the gradient of the cost function, the sum of the cost of each sample is needed, which could be inefficient for large training datasets. On the other hand, when applying SGD, the cost gradient of only one sample is used at each iteration (instead of the sum of the cost gradients of all training datasets) which can significantly reduce the computational complexity of the algorithm.
- **Passive Aggressive Classifier [16]**
  In this method, a linear function of the input multiplied by weighting elements is passed to an activation function (often a sign function for binary classifications). Then, a Hinge loss function is used to measure the error between the estimated and true values of labels. The update rule for weighting elements works in such a way that the algorithm is passive when a correct classification occurs (no weight change). For false classification cases, on the other hand, the algorithm becomes aggressive and updates the weights so as to minimize the loss for similar inputs that may occur in other instances.

- **Ridge Classifier [17]**
  This method often uses a linear function of the input vector and weighting parameters. The function used for updating the weights is the squared error of the estimated

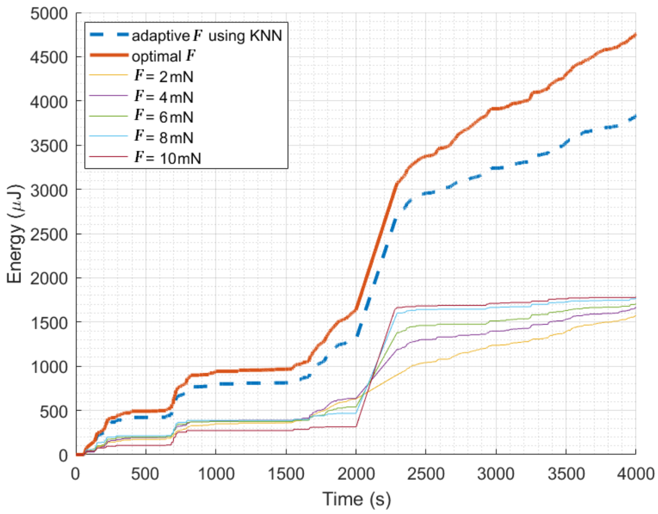

Fig. 4. Comparison of the harvested energy for a 4000-sec time sample for the ideal case with optimal value of F, adaptive holding force with KNN algorithm, and a constant holding force

outputs plus a regularization term which is a function of the weights. The regularization technique tends to reduce overfit in estimation.

IV. Simulation Results

The vectors containing the amplitudes of the cosine approximation for the nine classes of pseudo-optimal force $\tilde{F}_{\Delta}(i)$ are first divided by their maximum magnitude to normalize the vector elements between -1 and 1. Then, 90% of these 17550 vectors are randomly selected from each class in order to be used as the training data set, and the remaining 10% are used as the test data. All of the eight classification techniques described in Section III are compared by simulations, which are performed on a computer with an Intel R processor (Core i5) running at 2.5GHz using 16GB of RAM with Windows 10[1].

For training the classification algorithms and deploying machine learning models, we use the Google TensorFlow[1]platform. The results are given in Table I, indicating the running times for each input vector and accuracy of the learning methods in estimating the pseudo-optimal F. Fig. 3 shows the improvement in the generated power when the electrostatic force is changed using our approach, compared to the case of applying a constant force F=2mN. The amount of gain depends on many factors such as the dimension of the mass-spring-damper inside the CFPG micro-harvester. For these results, the size of the MSD and the distance between the two plates ($2Z_l$) were chosen as $15\times15\times1.5$mm^3 and 0.5mm, respectively.

[1]Intel R Processor, Windows 10 and Tensorflow are products of Intel Corp., Microsoft and Google, respectively. These products have been used in this research to foster research and understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that this product is necessarily the best available for the purpose.

TABLE I

Running time and accuracy of eight classification techniques for test data of the acceleration of human arm

	Running Time (s)	Accuracy (%)
Decision Tree	0.0009	45
Random Forest	0.3198	86
KNN	0.032	94
SVM	0.0866	73
MLP	0.0009	86
SGD	0.0010	43
Passive Aggressive	0.1149	50
Ridge	0.0001	43

Fig. 4 displays the gain in the harvested energy for a 4000-sec time interval. The harvested energy for different constant holding forces is also plotted. In addition, for this example, we have included the amount of harvested energy using the optimal value of the holding force as a reference. The curve corresponding to the optimal F serves as an upper bound for the amount of energy that can be generated. Although it is desirable for the adaptive holding force methodology to be as close as possible to the optimal curve, one should also take into consideration the computational complexity of the algorithms used. The higher this complexity is, the more energy it requires to perform, leading to less (or even no) gain in the harvested power. Therefore, a thorough analysis of the implementation complexity and the resulting power consumption of the added hardware is required to justify the addition of the optimization methodology to the micro-harvester circuitry. Our initial investigation shows that algorithms such as MLP can be implemented with reasonable runtime power consumption. Further details of our analysis will be provided in future publications.

V. Conclusion

Efficiency of eight machine learning classification techniques for adaptive estimation of the electrostatic force in a CFPG architecture according to the frequency spectrum of acceleration in human arm was investigated. It was observed that the net energy harvested using a suitable machine learning technique could lead to significant gain in the harvested power. The exact amount of the gain depends on many parameters such as the dimension of the MSD inside the harvester, adaptation time interval Δ, and the classification algorithm. The impact of these parameters will be studied in more details in future research.

Acknowledgement

This work was supported by the National Institute of Standards and Technology (NIST) [grant number 60NANB18D261].

References

[1] H. C. Koydemir and A. Ozcan, “Wearable and implantable sensors for biomedical applications,” *Annual Review of Analytical Chemistry*,vol. 11, pp. 127–146, 2018.

[2] Y. Khan, A. E. Ostfeld, C. M. Lochner, A. Pierre, and A. C. Arias,“Monitoring of vital signs with flexible and wearable medical devices,” *Advanced Materials*, vol. 28, no. 22, pp. 4373–4395, 2016.

[3] F. K. Shaikh and S. Zeadally, “Energy harvesting in wireless sensor networks: A comprehensive review,” *Renewable and Sustainable Energy Reviews*, vol. 55, pp. 1041–1054, 2016.

[4] P. D. Mitcheson, T. Sterken, C. He, M. Kiziroglou, E. Yeatman, and R. Puers, “Electrostatic microgenerators,” *Measurement and Control*, vol. 41, no. 4, pp. 114–119, 2008.

[5] T. Von Buren, P. D. Mitcheson, T. C. Green, E. M. Yeatman, A. S. Holmes, and G. Troster, “Optimization of inertial micropower generators for human walking motion,” *IEEE Sensors Journal*, vol. 6, no. 1, pp. 28–38, 2006.

[6] M. Dadfarnia, K. Sayrafian, P. Mitcheson, and J. S. Baras, “Maximizing output power of a CFPG micro energy-harvester for wearable medical sensors,” in *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies*, 2014, pp. 218–221.

[7] D. Budić, D. Šimunić, and K. Sayrafian, “Kinetic-based micro energy-harvesting for wearable sensors,” in *Proceedings of the 6th IEEE International Conference on Cognitive Infocommunications*, 2015, pp. 505–509.

[8] C. V. Bouten, K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen, “A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity,” *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 3, pp. 136–147, 1997.

[9] E. K. Antonsson and R. W. Mann, “The frequency content of gait,” *Journal of Biomechanics*, vol. 18, no. 1, pp. 39–47, 1985.

[10] N. Yarkony, K. Sayrafian, and A. Possolo, “Energy harvesting from the human leg motion,” in *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, 2014, pp. 88–92

[11] S. D. Jadhav and H. Channe, “Comparative study of K-NN, naive bayes and decision tree classification techniques,” *International Journal of Science and Research*, vol. 5, no. 1, pp. 1842–1845, 2016.

[12] L. Rokach, “Decision forest: Twenty years of research,” *Information Fusion*, vol. 27, pp. 111–125, 2016.

[13] D. Tomar and S. Agarwal, “A comparison on multi-class classification methods based on least squares twin support vector machine,” *Knowledge-Based Systems*, vol. 81, pp. 131–147, 2015.

[14] S. B. Wankhede, “Analytical study of neural network techniques: SOM, MLP and classifier-a survey,” IOSR *Journal of Computer Engineering*, ver. VII, vol. 16, no. 3, 2014.

[15] L. Bottou, “Large-scale machine learning with stochastic gradient descent,” in *Proceedings of COMPSTAT’2010*, 2010, pp. 177–186.

[16] J. Jorge and R. Paredes, “Passive-aggressive online learning with non-linear embeddings,” *Pattern Recognition*, vol. 79, pp. 162–171, 2018.

[17] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2011.