Gaussian Process Regression for Improving the Performance of Self-powered Time-of-Occurrence Sensors

Liang Zhou*, Kenji Aono* and Shantanu Chakrabartty†
Department of Computer Science and Engineering*
Department of Electrical and Systems Engineering†
Washington University in St. Louis
St. Louis, MO, U.S.A.
{liang.zhou, aono, shantanu}@wustl.edu

***Abstract*—In our previous work, we had demonstrated a CMOS timer-injector integrated circuit for self-powered sensing of time-of-occurrence of mechanical events. While the sensor could achieve an improved time-stamping accuracy by averaging the output across over multiple channels, the mismatch between the channels made the calibration process cumbersome and time-consuming. In this paper, we propose the use of non-parametric machine learning techniques to achieve more robust and accurate event reconstruction. This is demonstrated using training and testing data that were obtained from fabricated prototypes on a 0.5-μm CMOS process; the model trained using Gaussian process regression can achieve an average recovery accuracy of 3.3% on testing data, which is comparable to the performance of using an averaging technique on calibrated injection results. The experimental results also validate that scalable performance can be achieved by employing more injection channels.**

I. Introduction

Self-powered sensors are attractive for applications where access to batteries is considered impractical, such as embedded mechanical variation monitoring or implanted health sensors. A self-powered sensor operates by harvesting energy from ambient environment instead of extracting energy from a battery. For instance, the self-powered sensor-data-logger proposed in [1] harvests the energy from mechanical events to compute, store and update the event statistics on a non-volatile memory; however, it remains a challenge to sense and record the events' time-of-occurrence because it requires a continuous system reference clock — and self-powered systems cannot guarantee continuous powering of such references. Watch-dog timers [2] have been proposed in literature for ultra-low-power applications such as wireless Internet-of-Things (IoT), yet they operate in a synchronous manner that is only functional when external power is accessible, which obviates the asynchronous self-powering paradigm, thus rendering it unable to sense time-of-occurrence.

In [3], we proposed a fully integrated CMOS timer-injector system which can sense the time-of-occurrence of mechanical events. The system combines a robust self-powered timer system-op-chip (SoC) based on Fowler-Nordheim (FN) tunneling [4] (as illustrated in Fig. 1(b)) which can continuously track time without external powering, with a linear injector SoC which employs the physics of piezoelectricity-driven impact-ionized hot-electron injection (p-IHEI) [5], [6]

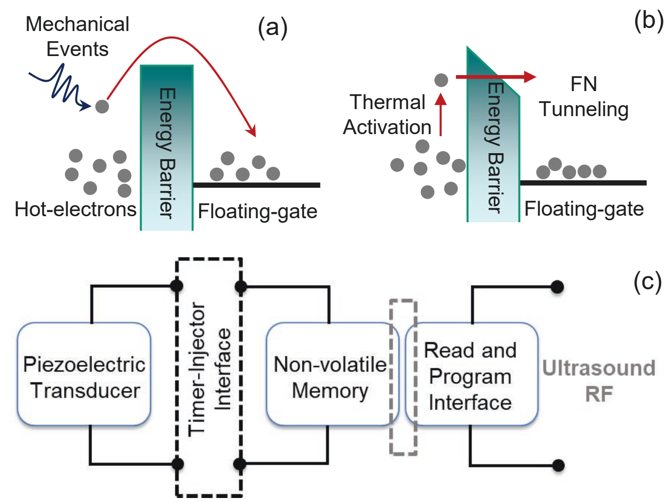

Fig. 1. Illustration of using PFG sensors for time-stamped sensing of mechanical events: (a) hot-electron injection, (b) FN tunneling for self-powered timing and (c) system architecture of a time-stamped sensor.

(as shown in Fig. 1(a)) to sense the time-of-occurrence and store in non-volatile memories, as illustrated in Fig. 1(c). Based on the mathematical model derived from the self-powered timer model and the linear injection model, the proposed self-powered sensing modality can achieve an average timing recovery accuracy of 6.9% for time stamping of mechanical events when using single injection channel, and an accuracy of 3.2% when averaging over five injection channels.

While the model proposed in [3] captures the dynamic time-stamping behavior, it neglects the high-order effects in the model of the timer-injector circuits, thereby introducing systematic error to the data recovery process. In addition, the random variations in the system from sensing to measurement can degrade the time-recovering performance. Random error sources can be minimized by enlarging the event duration, which are unknown a priori in most cases, or using an averaging technique across multiple sensing channels, yet the mismatch across different channels requires careful calibration to achieve improved performance. In this paper, we propose to use machine learning techniques which can eliminate the calibration process yet maintain the data recovery accuracy performance. The proposed technique is built based on a nonparametric learning process, therefore the issue of high-order nonideal factors will not affect the performance. The operation principle of the CMOS timer-injector system and the data reconstruction will be introduced briefly in Section II. In Section III, the experimental results based on the timer-injector and the proposed technique will be presented to validate the performance. The paper is concluded in Section IV.

II. Self-powered Time-stamped Injectors

A. Operation of the Timer-Injector

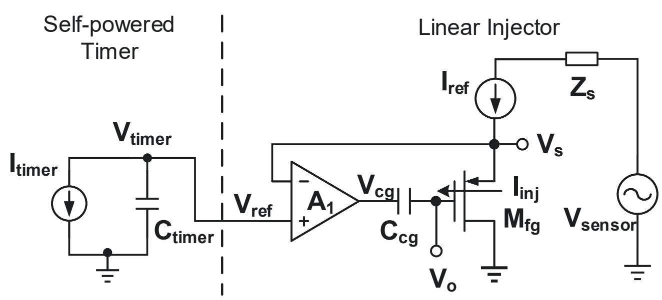

Fig. 2. Simplified schematic of the timer-injector system.

The proposed time-stamped sensor consists of a robust self-powered timer and a self-powered linear injector, as illustrated in Fig. 2. As described in [4], the FN tunneling based self-powered timer can be modeled as a charge storage capacitor C_{timer} and an FN tunneling current source I_{timer}. By programming the initial value of V_{timer} to the FN tunneling region, V_{timer} will demonstrate a dynamic response which can continuously keep track of time, which has been shown to be reliable and durable for long-term operation (as long as three years) [4]. At the core of the timing device is a PMOS floating-gate transistor, where the charge is stored at the floating-gate node and a high-quality thin gate oxide provides the tunneling junction.

The output of the timer is fed into a linear variant of p-IHEI [6], as shown in the right side of Fig. 2. The injection current I_{inj} from the channel to the floating-gate node is a function of the source-to-drain voltage of M_{fg} and channel current I_{ref}. By employing an operational amplifier A_1 to form a negative feedback loop, the output of the timer V_{timer} is isolated from the injection current and I_{inj} is only modulated by V_{timer}. Once a mechanical event activates the transducer and generates an electrical signal V_{sensor}, the injector is activated and generates an injection current corresponding to V_{timer}. As a result, the change in V_{o} noted as ΔV_{o} is a function of both the duration of the event and the instance of the time-of-occurrence. Omitting some details here for the sake of brevity, the mathematical model was derived in [3] as:

$$t \approx \frac{1}{k_1} \exp\left(\frac{\gamma}{\log\left[-\frac{1}{\lambda}\frac{\Delta V_{\mathrm{o}}}{\Delta t}\right]}\right) - \frac{k_0}{k_1}. \quad (1)$$

where t is the time-of-occurrence of the event, Δt is the duration of the event, γ, λ, k_0 and k_1 are model coefficients determined by the timing and injector characteristics and physics constants. The duration of the event can be easily measured and recorded using a linear injector with a constant V_{ref}. ΔV_{o} is a function of the charge stored on the floating-gate node in a nonvolatile manner and can be retrieved at a later stage for data reconstruction based on the model (1).

B. Data Reconstruction

One may recover an event's information based on (1), the coefficients of which can be obtained by training the model over a training set. Once we have the model coefficients, it will be easy to calculate t from ΔV_{o}. The performance of this technique has been extensively characterized in [3], with the average accuracy of recovering a time-of-occurrence for events with a one second duration has been characterized to be between 6%–8% using fabricated prototypes on 0.5-μm CMOS process.

Considering the fact that random noise in the system can not be neglected and has significant impact on the performance, [3] also analyzed the noise sources and proposed techniques to compensate for it. By averaging over multiple channels, the accuracy can be improved significantly at the cost of power and area overhead. An implementation of five injection channels can improve the average time recovery accuracy from 6.9% to 3.2% which approximately follows the statistical rule:

$$p_N \approx \frac{p_1}{\sqrt{N}} \quad (2)$$

where p_1 and p_N are the accuracy performance of single channel and N channels, respectively. While the averaging technique is proven effective, the implementation process is time-consuming because the mismatch across different injection channels need to be calibrated beforehand.

In this work, we propose to use machine learning techniques to train a non-parametric model which is robust to mismatch. Gaussian process has been shown to be an efficient and effective technique for nonlinear curve fitting in the area of supervised learning when the training set is not too large. The model starts from a nonlinear transformation function $\phi(\mathbf{x})$ and assumes the label y and feature vector x follows a linear model with random noise as [7]

$$y = \phi(x)^T \mathbf{w} + \epsilon \quad (3)$$

where $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$ represents Gaussian noise and it is identical for different x. Assuming we have a training data set $\mathcal{D} = (\mathbf{X}, \mathbf{y})$ and the prior distribution of the weight vector $\mathbf{w}$ follows a zero mean Gaussian distribution as

$$\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma_p), \quad (4)$$

then the prediction $\mathbf{y}_*$ of a test vector $\mathbf{x}_*$ also follows a Gaussian distribution as [7]:

$$y_* | x_*, \mathbf{X}, \mathbf{y} \sim \mathcal{N}(\mu_*, \sigma_*^2), \quad (5)$$

$$\mu_* = \phi_*^T \Sigma_p \Phi (K + \sigma_n^2 I)^{-1} \mathbf{y}, \quad (6)$$

$$\sigma_*^2 = \phi_*^T \Sigma_p \phi_* - \phi_*^T \Sigma_p \Phi (K + \sigma_n^2 I)^{-1} \Phi^T \Sigma_p \phi_*, \quad (7)$$

where $K = \Phi^T \Sigma_p \Phi$ and Φ is the matrix version of the transformed $\phi(\mathbf{x})$. If we define a kernel function $k(\mathbf{x}, \mathbf{x}')$ as

$$k(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x})^T \Sigma_p \phi(\mathbf{x}') \quad (8)$$

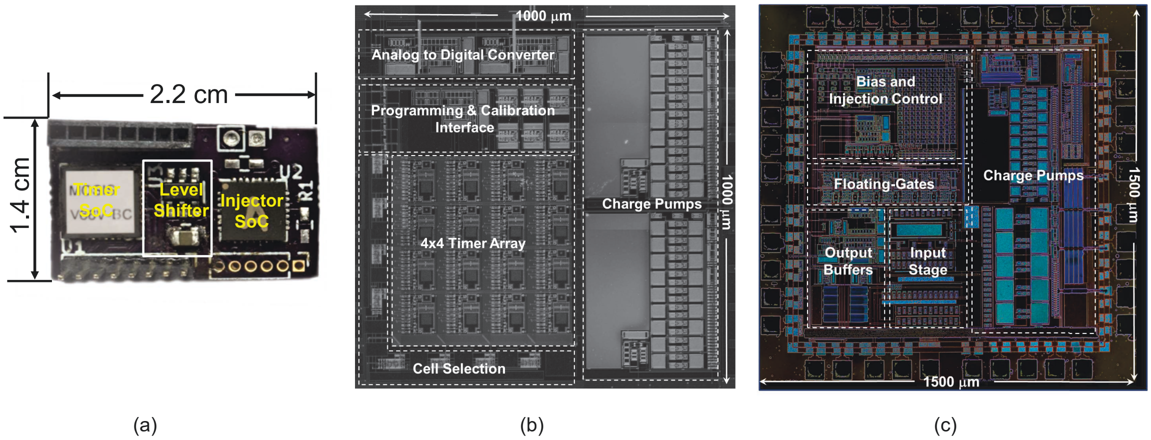

Fig. 3. Implementation of the self-powered timer-injector system: (a) PCB, (b) die photograph of the timer SoC and (c) die photograph of the injector SoC.

then the distribution in (5) can be calculated using the kernel function. A reasonable choice for the kernel is the squared exponential kernel of the form

$$k_{\mathrm{SE}}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{|\mathbf{x}-\mathbf{x}'|^2}{2l^2}\right). \tag{9}$$

In this work, the problem is defined as a nonlinear regression problem with training data set $\mathcal{D} = (\mathbf{\Delta V}, \mathbf{T})$, where $\mathbf{\Delta V}$ is a matrix formed by each feature vector $\Delta V_i = (\Delta V_{i1}, \Delta V_{i2}, \ldots\ \Delta V_{in})$ obtained from the measured output. $(\Delta V_{i1}, \Delta V_{i2}, \ldots\ \Delta V_{in})$ are the corresponding output from channel 1, … channel n at time instant T_i. $\mathbf{T}$ is the vector corresponding to true time-of-occurrence. The squared exponential kernel takes the distance across two vectors, therefore canceling out the offset across different channels. In addition, the random noises existing in each dimension of the feature vector will average out across multiple channels, making it more robust as the number of injection channels is increased. To predict the time-of-occurrence, a group of training data will be required to train the parameters in the Gaussian process regression model, such as the distance parameter l in the kernel function, and the noise parameter σ_n.

III. Experimental Verification

The timer and injector were fabricated separately on two different silicon dies using a 0.5-µm CMOS standard process. A level shifter was employed to merge the gap between the timer output and injection reference voltage. The microphotograph of the dies and the PCB are shown in Fig. 3. An array of timers were implemented for robust time extraction. Programming of the floating-gate transistor in the timer-injector circuits can be achieved using a combination of FN tunneling and hot-electron injection. FN tunneling removes the electrons from the floating-gate node by applying a high voltage (≥15 V in 0.5-µm CMOS process) across a parasitic nMOS capacitor. Although FN tunneling can be used to program FG memories individually [8], it is typically used as a global programming, because the isolation of high voltages in a standard CMOS process is arduous. Hot-electron injection, on the other hand, requires lower voltages (4.2 V in 0.5-µm CMOS process) than tunneling and, hence, is the preferred mechanism for precise programming of floating gates. Because hot-electron injection

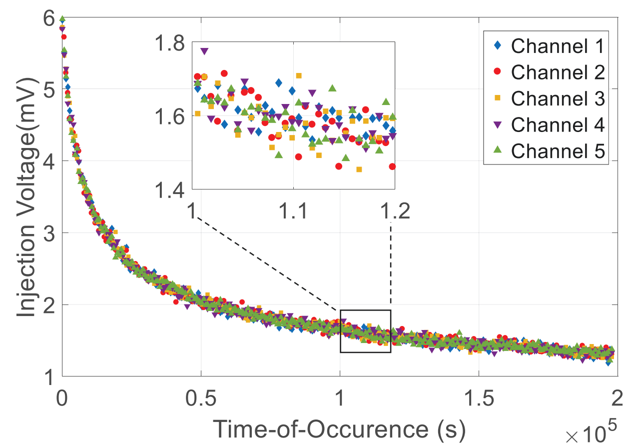

Fig. 4. Measured response of the timer-injector sensors from five channels.

in a pMOS transistor is a positive feedback process and can only be used to add electrons to the floating gate, the process needs to be carefully controlled and periodically monitored to ensure that the FG voltage is programmed to the desired value. The methods proposed in the literature [5] achieve a desired value either by adjusting the duration for which the FG transistor is injected or by adjusting the magnitude of the injection pulses.

To validate the data recovery technique based on Gaussian process regression, we first characterize the response of the time-injector system. The timer was programmed to the FN tunneling range and the output is left to dynamically evolve with respect to time. To emulate mechanical events, the system was periodically activated every 10 minutes for a duration of one second. Five injection channels on the linear injector SoC were activated, using the time signal as a reference to modulate the injection process. Fig. 4 shows the measured response for a total monitoring duration of 3,000 minutes with 300 measured points. The voltage change in the floating-gate nodes caused by injection shows monotonic dependence on the time-of-occurrence. The measured results also validate the assumption that random noise degrades the system performance and introduces error in the data reconstruction process.

To verify the performance of using Gaussian process regression for data reconstruction, the measured data were

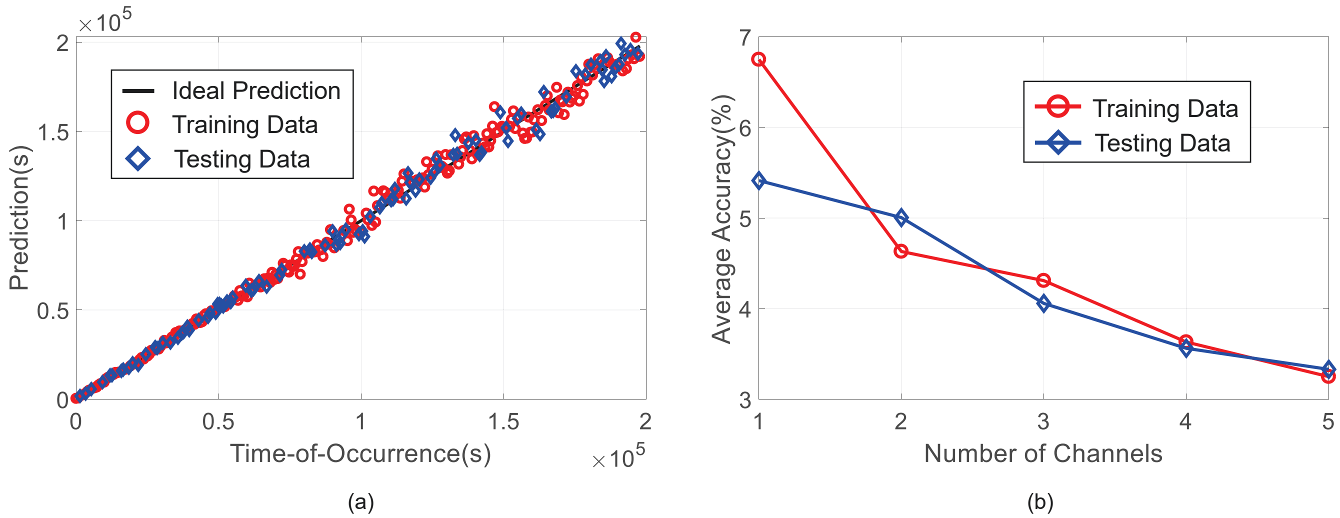

Fig. 5. Recovering time-of-occurrence using Gaussian process regression: (a) predicted time using the trained Gaussian model for training and testing data set with measured data from 5 channels; (b) dependence of the data reconstruction accuracy on number of injection channels.

randomly split into two groups, with 200 points as training set, and the remaining 100 points set aside as testing data. A squared exponential kernel was employed for training the model. Fig. 5(a) shows the predicted results of the trained model using the measured results from five channels, where the ideal prediction is marked as solid, diagonal black line. The predicted time-of-occurrence of training data and testing data are marked with red circles and blue rhombuses respectively. The average recovery accuracy for the training data and testing data are 3.25% and 3.33%, showing good learning transfer without overfitting. The performance is comparable to that of [3] using averaged data for parametric learning. Therefore, the Gaussian process regression is verified to be valid and robust in reconstructing the time-of-occurrence.

The next experiment was conducted to verify the impact of number of injection channels on the reconstruction performance. As discussed in Section II, the dimension of the feature vector, namely the number of injection channels will help improve the prediction accuracy. Fig. 5(b) shows the dependence of the average prediction accuracy on the number of channels used for model training. The monotonically decreasing relationship validates the conclusion. It also implies that the accuracy can be further improved by implementing more channels, however at the cost of power and area overhead. Notice that by using a squared exponential kernel function, the offset between different channels caused by mismatch in the injection rate can be compensated and does not affect the performance, therefore eliminating the calibration process and making the data processing easy to implement.

IV. Conclusions

In this paper, we proposed to use a machine learning technique to learn nonparametric models to predict time-of-occurrence from the timer-injector sensor proposed in [3]. The nonparametric learning process can eliminate the process of calibrating the mismatch when employing multiple injection channels to improve accuracy, which can significantly relieve the data processing task. When using Gaussian process regression to train the model with a squared exponential kernel function, the performance of the model prediction is comparable to the benchmark previously reported in literature. Experimental results validate that with more injection channels, the accuracy can be further improved. Note that other machine learning techniques such as k-NN and radial basis function should also work for the time-of-occurrence reconstruction if provided with a sufficient training set.

V. Acknowledgment

This work was supported in part by the National Science Foundation under Grant ECCS-1550096,CNS-1646380, DGE-0802267 and DGE-1143954.

References

[1] C. Huang and S. Chakrabartty, "An asynchronous analog self-powered cmos sensor-data-logger with a 13.56 mhz rf programming interface," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 2, pp. 476–489, 2012.

[2] Y. Lee, B. Giridhar, Z. Foo, D. Sylvester, and D. B. Blaauw, "A sub-nW multi-stage temperature compensated timer for ultra-low-power sensor nodes," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 10, pp. 2511–2521, 2013.

[3] L. Zhou, K. Aono, and S. Chakrabartty, "A cmos timer-injector integrated circuit for self-powered sensing of time-of-occurrence," *IEEE Journal of Solid-State Circuits*, 2018.

[4] L. Zhou and S. Chakrabartty, "Self-powered timekeeping and synchronization using fowler–nordheim tunneling-based floating-gate integrators," *IEEE Transactions on Electron Devices*, vol. 64, no. 3, pp. 1254–1260, 2017.

[5] C. Huang, P. Sarkar, and S. Chakrabartty, "Rail-to-rail, linear hot-electron injection programming of floating-gate voltage bias generators at 13-bit resolution," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 11, pp. 2685–2692, 2011.

[6] T. Feng, K. Aono, T. Covassin, and S. Chakrabartty, "Self-powered monitoring of repeated head impacts using time-dilation energy measurement circuit," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 9, no. 2, pp. 217–226, 2015.

[7] C. K. Williams and C. E. Rasmussen, "Gaussian processes for machine learning," *the MIT Press*, vol. 2, no. 3, p. 4, 2006.

[8] L. Zhou and S. Chakrabartty, "A 7-transistor-per-cell, high-density analog storage array with 500μv update accuracy and greater than 60db linearity," in *Circuits and Systems (ISCAS), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 1572–1575.