ISSN: 2395-1680 (ONLINE)
DOI: 10.21917/ijme.2023.0254

A DEEP LEARNING BASED POWER ESTIMATION MECHANISM FOR CMOS VLSI CIRCUITS

N. Sivakumar[1], N.S. Suresh[2] and G.K. Arpana[3]

[1]*Department of Electrical and Electronics Engineering, Varuvan Vadivelan Institute of Technology, India*
[2]*Department of Electrical and Electronics Engineering, Saveetha Institute of Medical and Technical Science, India*
[3]*Department of Electronics and Communication Engineering, East West College of Engineering, India*

Abstract

In general, power profile flattening solutions have a higher power consumption; nevertheless, when adopting increased levels of security, this is a sensible choice to make. There have been a variety of various issues raised in relation to each of these tactics, including the fact that they are difficult to scale, that they use a significant amount of energy, that they result in a performance decrease. In this paper, we develop a deep learning-based power estimation mechanism to evaluate the performance of the circuits using VLSI circuits. Various cells of VLSI circuits are evaluated to perform the evaluation of the power supply performance. The results show that the proposed method achieves higher degree of power consumption than the other methods.

Keywords:

Deep Learning, Power Estimation, VLSI Circuits

1. INTRODUCTION

In today very large-scale integrated circuits (VLSI), the power distribution network (PDN) could have as many as several million different nodes. The variations in voltage that occur at the different nodes of the on-chip power grid are the result of several different factors [1]. These factors include the effective impedance that exists between the various nodes of the grid, non-linear current loads, and distributed on-chip voltage sources. It is essential to keep in mind that on-chip PAAs are the only ones that are taken into consideration, and that any references to PDNs or power grids should be construed as referring to PDNs or power grids that are located on the same chip as the PAAs that are in question. This is because on-chip PAAs are the only ones that are taken into consideration.

To monitor and record the voltage changes that take place between a device power or ground (P/G) pin and an off-chip power supply, an extremely small resistor with a resistance of up to fifty ohms is inserted externally between the two components.

Since the first-time power analysis was used, back in 1998 [1] a variety of various PAA mitigation techniques have been suggested. Hiding and masking are two preventative measures that are commonly employed to improve the robustness of modern integrated systems when PAAs are present [2]. Existing techniques of concealment need regular adjustment of IC power consumption, which results in the same amount of power being squandered regardless of the operation that is being carried out. There is a need for a method that does not require constant adjustment of IC power consumption. The sense amplifier-based logic (SABL), the dual-spacer dual-rail logic (DSDR), and the three-phase dual-rail pre-charge logic are all examples of DRP techniques (TDPL).

The symmetric differential nature of these DRP techniques and their capacity to maintain a consistent power output over the course of time [3] are the distinguishing characteristics that set them apart from one another. Current-mode logic (CML) with fluctuations in the supply current, such as MOS current mode logic (MCML) and dynamic current-mode logic (DyCML), have also been investigated for the purpose of secret data storage. Examples of these include MOS current mode logic (MCML) and dynamic current-mode logic (DyCML). MOS current-mode logic (MCML) and dynamic current-mode logic are two examples of the current-mode logic found in electronic devices (DyCML).

Alternately, wave dynamic and differential (WDDL) logic were developed in [4], and it was shown to be effective in [5] for flattening power profiles. In reference [5], a further method for eliminating power utilisation that is reliant on the data being processed is given. This strategy calls for the utilisation of switching capacitors to achieve current equality throughout the circuit.

By introducing random characteristics into power profiles, masking techniques attempt to reduce how dependent overall power consumption is on the data that is processed. This is done to realise our objective of lowering our overall power consumption as much as possible. For instance, in [6], the power profile is randomised by encrypting sensitive information with Boolean and/or arithmetic operations. This is done so that the information cannot be read. These citations are in the bibliography that was provided. Instead, then relying on the real and pertinent data, this technique makes use of the masked data to calculate the total amount of power that is spent. The elimination of PAAs using a single mask, on the other hand, is often insufficient to fulfil the task. The performance of the integrated circuit suffers as a direct result of the increased computational cost of the cryptographic operation, which is caused by the presence of many masks that need to be processed simultaneously. The use of delay is one of the most common and effective ways that masking techniques provide unpredictability. It is also one of the most widespread ways that masking techniques provide unpredictability.

In the paper [7], the author describes a method for producing power profiles that are unpredictable at each cycle by arbitrarily introducing delays to data paths. You may learn more about this strategy by reading the article. It introduces the concept of data-dependent propagation delays and demonstrates how these delays can be exploited to scramble the order in which output bits are received. The idea of data-dependent propagation delays is introduced. The employment of ring oscillators has resulted in the development of additional random power traces because of the randomization of the power profiles of the AES core. This was accomplished by randomising the power profiles. The secure double rate register (SDRR) is a mechanism that was proposed in [28] for the purpose of randomising information at the register transfer level in combinational and sequential logic. This was

accomplished by doubling the rate at which information is transferred from one register to another. (RTL). The randomised multi-topology logic (RMTL) that was disclosed and it gives additional evidence in support of the concept of randomised power profiles. RMTL displays a reconfigurable logic format and dynamic structure. Random pre-charge logic (RPL), sometimes known as Random Precharge Logic, is a factor that adds noise to the actual power profiles, as detailed in [7], where it is also illustrated visually.

It has been proved that modern systems are less susceptible to power attacks because of the countermeasures that are now in place. The reason for this is since countermeasures are now in place. However, these methods do not provide protection that is against sophisticated PAAs, and they come with a price tag in terms of power consumption, performance, space requirements, complexity of design, and scalability of the underlying infrastructure [4]. Additionally, these methods are not scalable. In addition to this, these methods do not lend themselves well to scaling. Because the current countermeasures are intended to be preventative rather than detective, they are unable to discover ongoing hazards to the organization security. This is because the current countermeasures were designed to be preventative rather than detective. To successfully improve the resistance of integrated systems against attacks on their security, it is required to have a collection of power-efficient IC design approaches that can detect and mitigating PAAs in real time.

2. BACKGROUND

In recent years, it has become abundantly clear that algorithms based on machine learning hold the potential to significantly improve the efficacy of attack and prevention mechanisms in relation to the hardware of Internet of Things devices. This potential has been highlighted by recent developments in the field. In [8], we built a taxonomy of the hardware security challenges that can be overcome by deploying ML techniques. These problems are categorised according to the type of solution they provided. We did this so that we could better organise the data by having it all in one place.

In the article [9], we introduce the concept of utilising machine learning classifiers in conjunction with hardware-assisted malware detectors (HMD) to determine the possible dangers that could be caused by malware in low-resource embedded systems. Because real-time software-based malware detectors require frequent updates to the database, which must incorporate the known threats, HMD solutions are preferable. On low-resource devices, you ought to avoid doing things like this as much as possible. In addition, the authors proposed considering the operation of the hardware performance counters (HPCs), not to measure the performance of the system, but rather as a means of training the ML classifiers for the possible detection of real-time malware using data generated by a variety of actual malware threats. This was done to consider the operation of the HPCs to detect real-time malware using data generated by a variety of actual malware threats. The authors of this study anticipated that by doing so, they could enhance the accuracy with which real-time malware could be identified by the classifiers. The authors concluded that several 4 HPCs are sufficient to meet the requisite categorization of malware threats achieved by the used ML algorithm.

To establish which machine learning classifiers were the most successful for each category of malware, evaluation tests were carried out. These tests measured the level of accuracy achieved in malware detection in addition to the amount of hardware overheads that were incurred. Power profiling is an intrusive technique that requires direct measurements of current, voltage, and other power characteristics at the IoT device.

This method, which can be used to detect covert channel attacks as well as power depletion attacks that are carried out on resource constrained IoT devices, is based on a machine learning (ML) algorithm that categorises the power profiles generated by the operational behaviour calculated by the measurement of the attributes. Using this technology, these assaults are able to be uncovered and exposed. When performing an analysis of the process, having the capability to collect the necessary measurements for the power profile in a manner that does not call for any invasive procedures can be of great assistance. The research that suggests that it is possible to clone a PUF when ML approaches are applied calls attention to the significance of recognising the vulnerabilities that are associated with each PUF. The findings of this research underline the importance of becoming familiar with each PUF vulnerability.

Cloning big XOR arbitrator PUFs is becoming an increasingly difficult and time-consuming task. XOR arbitrator PUFs provide a high level of security. The proposed solution makes use of an artificial neural network. This network implements an optimization strategy that enables it to process training datasets that are larger than the amount of computer memory that is available to it. This is made possible because the network employs an optimization strategy. The performance of the authors' implementation of this neural network in a massive XOR arbitrator PUF is superior to the performance of the ML code from another study when the findings of both experiments are compared to one another. In a matter of hours, a neural network that has been optimised can produce the same responses as the PUFs, whereas the performance of the ML code could deliver the same results in a matter of days. The findings substantiate this assertion. This method can be used to generate critical knowledge regarding the vulnerabilities of PUFs, which can then help in the development of IoT devices that have a higher level of security.

In the article [10], the authors address the limitations of previously developed strategies for multi-party communications amongst IoT devices that have limited resources. In addition, the authors provide an innovative strategy for key sharing that takes use of PUFs. This is done to get around these restrictions as much as possible. In crossover PUFs, the authors describe a configuration of challenges and inter-stage crossing structures. This configuration is malleable to accommodate the user preferences and has the potential to result in the devices generating the same key (and so sharing it). When applied to gadgets that are part of the Internet of Things, this technique is viewed as a low-priced, hardware-based shared-key mechanism that offers increased security (IoT).

In conclusion, we present the low-cost, low-power radiofrequency (RF)-PUF, which is a deep neural network (DNN)-based detection method that enables real-time authentication of IoT devices by exploiting inherent features

embedded in the RF signals. This is accomplished by utilising the inherent features embedded in the RF signals. Utilizing the inbuilt properties of the RF waves allows for the successful completion of this task. It is possible to implement the RF-PUF without adding any new hardware to the transmitter end, and the utilisation of DNNs will only result in a marginal increase in the amount of power that is consumed by the receiver. However, the addition of new hardware to the transmitter end is not required.

3. PROPOSED MODEL

Estimating power consumption can be challenging because it is difficult to determine the typical amount of power used by an electronic system. This is a separate issue from the problem of the voltage drop, which requires estimating the instantaneous power in the most catastrophic event that may ever occur. The average quantity of power that is extracted from the chip is proportional to the amount of heating and temperature that is generated by the chip.

We have already touched on a way that is easy to understand for evaluating power, and that method is simulation. Carry out a simulation of the circuit using the design as the basis, all the while keeping a close eye on the current waveform being produced by the power supply. After that, the current is used to compute the waveform mean value to acquire the power. This is done so that the average power may be determined. The two key advantages that this technology possesses are the first being its accuracy and the second being its versatility.

It is possible to arrive at an accurate estimation of the amount of power that is consumed by a circuit without respect to the underlying technology, design philosophy, functionality, or architecture of the system. On the other hand, the relevance of the input signals that are given to the simulator is directly proportional to the outputs that it generates. These outputs can be thought of as the results of the simulation. In addition to this, you will require information on the voltage waveforms of the signals that are being input into the system.

The issue of relying on patterns is a very important one that needs to be addressed. It is standard practise to generate an estimate of the power consumption of a functional block before the remaining component of the chip has been fabricated or even fully specified. This can be done at any time during the design process. When this occurs, there is only a fundamental grasp of the inputs to this functional block, and it is plainly difficult to acquire a more in-depth comprehension of those inputs. However, there is a fundamental grasp of the inputs to this functional block.

Even if one is willing to make educated predictions about the waveforms in question, it is possible that it will be impossible to determine whether a specific collection of input waveforms represents a typical example. This is the case even if one is prepared to make educated predictions regarding the waveforms in question. This method cannot be used in the design of complicated circuits because the amount of computing resources needed to simulate a large number of input patterns could very rapidly become prohibitive. As a result, this method is not relevant to the design of complex circuits.

To begin, the issue is simplified in one of three distinct ways, which is the case for most of the many methods of power estimation that are more effective. To begin calculating the power, the first thing that must be done is to assume that the voltage levels across the chip from the power source and ground are always the same. This will allow the power to be calculated accurately. Because of this, it is much simpler to estimate the current that is being utilized by each sub-circuit at a given voltage that has been specified for the power supply. This is the mechanism that is ultimately used to determine who holds power. Second, it is hypothesized that a synchronous sequential design approach will be implemented, with logic gates and latches functioning as the primary constituents of the finished product architecture.

When the clock activates latches, those latches will transition between states, which will result in them expending energy because of the transition. Because of this, the amount of power that the latch consumes changes in tandem with the passage of time. The gates used in combinational logic do not have to adhere to the same restrictions as those used in other kinds of logic. When a latch operation is done (in synchrony with the clock), the inputs of a combinational logic block are updated; however, the block internal gates may go through several transitions before obtaining their steady state values for the period.

This exemplifies both the conventional method of utilising circuit modelling as well as the alternative method of applying probability. It has been demonstrated that both approaches are successful in resolving the issue. The technique of averaging needs to be finished first, and only then can one move on to the actual analysis. Therefore, a single iteration of a probabilistic analysis tool can replace several circuit simulations runs if the level of accuracy is prepared to accept a slight loss. This is granted, however, that the level of accuracy is willing to accept a minor reduction. The questions concern the kind of analysis that needs to be done, the process by which the necessary probabilities are to be produced, and the ultimate application of the results of the study.

When the algorithm is being used, it is not necessary to provide a comprehensive collection of extremely specific input patterns because the necessary input probabilities can be provided in a straightforward manner. This is because it is not necessary to provide a comprehensive collection of input patterns. The probabilities that were placed into the study had a significant impact on the findings of the research. Because of this, the process continues to rely on patterns that have been formed in the past, and the user is required to submit data describing the typical behaviour that occurs at the circuit inputs, stated in terms of probabilities. In addition, the procedure continues to rely on patterns that have been formed in the past. We refer to these methods as weakly pattern dependent because they do not require exhaustive information regarding the input signals to function correctly.

The idea of signal likelihood has been around for quite some time and was first utilised in the process of evaluating the testability of circuits. Remember that both distinct approaches to calculating probability result in the same value.

The circuit has some delays that are inherent to it. Even considering a timing model with no delay at all makes no obvious difference in the outcomes. On the other side, the toggle power will be ignored immediately if this action is executed. We demonstrated why this is such a serious problem for approaches that rely on the measures described in the previous paragraph. If a model with no delay is employed and transition probabilities for

that model are established, then it is possible to calculate power for that model.

$$P_{av} = \frac{1}{2T_c} V_{dd}^2 \sum_{i=1}^{n} C_i P_t(x_i)$$

where

T_c - clock period,

C_i - total capacitance, and

n - total circuit nodes.

As things are right now, we need to talk about the issue of signals not being dependent on one another, and the moment has come to do so. For example, two signals may be connected in such a way that it is physically impossible for to be high at the same time. This may be the case if the signals are linked. Due to the significant processing cost involved in computing these correlations, it is common practise in the real world to think of input and internal nodes of a circuit as being uncorrelated. This is because of the way these nodes are used. This is due to the significant amount of processing power that is required to compute these relationships. An assumption of independence from space in terms of location has been made here. Another example of independence is the question of whether the values of a signal in two successive clock cycles can be considered independent from one another. This question illustrates the concept of whether the values of a signal can be considered independent from one another. Assuming that the transitions are unrelated to one another, using the following formula to get the transition probability from the signal probability is a simple process:

$$P_t(x) = 2P_s(x)P_s(\bar{x})$$

3.1 DEEP LEARNING MODELLING

SqueezeDet is a fully convolutional neural network (CNN) that is both compact and rapid, and it was designed specifically for the purpose of object detection for autonomous vehicles. Deep CNN can be used for real-time object recognition, but only if the model can handle some critical problems, such as those relating to speed, accuracy, model size, and power efficiency. If the model can handle these concerns, then Deep CNN can be used for real-time object identification.

If the model is able to address these concerns, then Deep CNN may be utilised. The SqueezeDet model performs a respectable job of taking into consideration the limits that have been imposed on the system. To derive high-dimensional but low-resolution feature maps from the input photos, this object detector needs only a single forward pass to be applied. This is because to the layered nature of the convolution filters that it utilises, which makes this possible. In addition to this, it utilises a convolutional layer that is referred to as the ConvDet layer.

This layer accepts a feature map as its input and generates many bounding boxes as its output. These bounding boxes are subsequently use to solve the challenge of establishing the object category. It generates filtered item detections based on the bounding boxes that have been provided, which is the very last step but certainly not the least. The fundamental model that is utilised by SqueezeDet is considerably less than 8 megabytes, in contrast to the model that is utilised by AlexNet, which is greater than 8 megabytes.

4. PARAMETER MODELLING

We put our simulator through its paces by recreating a low-power CMOS library, which consisted of two-stage cells and sophisticated gates, as a means of evaluating its capabilities. To characterising each library cell, electrical simulations were carried out in HSPICE. The paper in question contains the model that was utilised for the simulations presented here.

Our team conducted an initial series of research to ensure that our energy model delivers an accurate depiction of a single cell or pattern. This was done so that we could guarantee its accuracy. A simulation was run on each library cell, and it contained a thorough variety of fan-in and fan-out circumstances, in addition to every viable test combination. This was done so that the most accurate results could be obtained. The worst-case scenario had an average absolute inaccuracy of 4% from HSPICE, with a standard deviation of 0.2%. This was determined to be the absolute worst-case situation. Applying a series of one hundred randomly generated test vectors, each of which contained fifty percent of its input transitions that were mismatched, allowed us to achieve the same level of accuracy as before.

The precharacterized test library had a vast collection of benchmark circuits superimposed on top of it so that the accuracy of the library as well as its performance could be evaluated. Combinatorial and sequential circuits were both put through their paces by being put through simulated random sequences that lasted for thirty nanoseconds and contained one hundred test vectors each.

Table.1. Power Consumption

Number of Cells	HSPICE	Proposed
1	0.052	0.023
2	0.057	0.031
3	0.062	0.035
4	0.085	0.035
5	0.092	0.052
6	0.099	0.068
7	0.125	0.075
8	0.152	0.082

Table.2. Delay

Number of Cells	HSPICE	Proposed
1	0.0531	0.0235
2	0.0583	0.0317
3	0.0634	0.0358
4	0.0869	0.0358
5	0.0940	0.0531
6	0.1012	0.0695
7	0.1278	0.0767
8	0.1553	0.0838

Table.3. Power Dissipation

Number of Cells	HSPICE	Proposed
1	0.0006	0.0003
2	0.0007	0.0004
3	0.0007	0.0004
4	0.0010	0.0004
5	0.0011	0.0006
6	0.0012	0.0008
7	0.0015	0.0009
8	0.0018	0.0010

The precision of the current waveforms in the time domain that can be generated through the application of our technique is typically within a margin of error of around 20%. The time resolution that was applied while presenting and analysing the data has a considerable bearing on this result; however. The average absolute error would drop to the accuracy of the single-pattern energy estimate (about 5%) for larger time steps; however, for smaller time steps, most of the error would be caused by modest time misalignments between the two waveforms.

This is because the larger the time step, the more accurate the single-pattern energy estimate will be. This would be the circumstance when dealing with time steps that were of a shorter duration. We made the decision to use a temporal resolution that was exceedingly exact so that we could record each peak in the current. HSPICE has produced estimates with an accuracy of 9 and 6%, respectively, for the peak value and length of the overall current drawn by the circuit for each transition at the principal inputs. These estimates pertain to the peak value of the current and the length of the transition. These numbers show the peak value of the current and the length of time that the changeover lasted. Calculations of the local current have been made with an unprecedented level of accuracy as a direct consequence of random circuit division.

5. CONCLUSION

This study has highlighted essential countermeasures and placed an emphasis on ML-based solutions that optimise the process of mitigating various threats. This is because hardware security represents a highly significant sector in the process of achieving the IoT goal. The study has highlighted essential countermeasures.

The completion of future work should have as its primary purpose the promotion of additional improvements in the level of security offered to the hardware of Internet of Things devices. This should be the case since the completion of this work will be necessary to achieve the primary objective. It is possible to improve the level of security provided by IoT devices by making modifications to the components used in their construction as well as the overall dimensions of these devices.

It is recommended that countermeasures that are co-designed be explored to make up for any potential bottlenecks that may be identified in separate software and hardware solutions. This is because separate software and hardware solutions are more likely to experience performance issues.

Despite the significant amount of research that has been put into HT threats at the chip-layer in the context of ML, it is anticipated that the information that has been gleaned will also be used to improve the device defences at the higher layers. This is because HT threats are becoming increasingly sophisticated. In a similar line, understanding how adversaries utilise the techniques of machine learning (ML) can be beneficial in the process of building solutions for mitigating risks.

REFERENCES

[1] P.J. Kumar, "Machine Learning based Workload Balancing Scheme for Minimizing Stress Migration Induced Aging in Multicore Processors", *International Journal of Information Technology*, Vol. 12, No. 1, pp. 1-12, 2022.

[2] G. Sapone and G. Palmisano, "A 3-10-GHz Low-Power CMOS Low- Noise Amplifier for Ultra- Wideband Communication", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 59, No. 3, pp. 678-686, 2011.

[3] Behzad Razavi, "*RF Microelectronics*", New York: Prentice Hall, 1998.

[4] K.G. Devi and N.T.D. Linh, "*Artificial Intelligence Trends for Data Analytics using Machine Learning and Deep Learning Approaches*", CRC Press, 2020.

[5] S. Zheng and S. Yin, "An Ultra-Low Power Binarized Convolutional Neural Network-Based Speech Recognition Processor with On-Chip Self-Learning", *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 66, No. 12, pp. 4648-4661, 2019.

[6] Y. Shen and V. Chen, "Class-E Power Amplifiers Incorporating Fingerprint Augmentation with Combinatorial Security Primitives for Machine-Learning-based Authentication in 65 nm CMOS", *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 69, No. 5, pp. 1896-1909, 2022.

[7] R.M. Weng, C.Y. Liu and P.C. Lin, "A Low-Power FullBand Low- Noise Amplifier for Ultra-Wideband Receivers", *IEEE Transactions Microwave Theory and Techniques*, Vol. 58, No. 8, pp. 2077-2083, 2010.

[8] P. Kumar and C.S. Thakur, "Hybrid Architecture based on Two-Dimensional Memristor Crossbar Array and CMOS Integrated Circuit for Edge Computing", *NPJ 2D Materials and Applications*, Vol. 6, No. 1, pp. 1-8, 2022.

[9] C. Xu, "SNS's not a Synthesizer: a Deep-Learning-based Synthesis Predictor", *Proceedings of Annual International Symposium on Computer Architecture*, pp. 847-859, 2022.

[10] T. Alnuayri, "A Support Vector Regression-Based Machine Learning Method for On-Chip Aging Estimation", *Proceedings of International Conference on Computing and Information Sciences*, pp. 1-6, 2021.