IMPLEMENTATION OF MACHINE LEARNING BASED PASS/FAIL CHECKS TO CHIP BEHAVIOUR

(ÇİP DAVRANIŞINLARINA MAKİNE ÖĞRENMESİ TEMELLİ BAŞARILI/BAŞARISIZ KONTROLÜNÜN UYGULANMASI)

by

Hüsnü Murat Koçak, B.S.

Directed Research

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

in

COMPUTER ENGINEERING

in the

GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

of

GALATASARAY UNIVERSITY

Supervisor: Assist. Prof. Dr. Ahmet Teoman Naskali

June 2021

TABLE OF CONTENTS

LIST OF SYMBOLS

ANN	**:** Artificial Neural Network
DL	**:** Deep Learning
EDA	**:** Electronic Design Automation
GA	**:** Genetic Algorithm
KUL	**:** Catholic University of Leuven
ML	**:** Machine Learning
NDA	**:** Non-Disclosure-Agreements
NN	**:** Neural Network
PDK	**:** Process Design Kit
R&D	**:** Research and Development

ABSTRACT

Compact models of transistors, which we call prototypes of development results of semi-conductor technology and circuit design, are challenging and time-consuming with current development tools. Therefore research and development (R&D) departments or chip manufacturers that develop these prototypes want to keep the model details confidential. Full access to details of compact models of transistors is restricted in commercial agreements (e.g., Non-Disclosure-Agreements). Machine Learning (ML) methods also exhibit black-box nature, which not possible to understand how it generates results, just like restricted compact models by confidentiality agreements. In this proposal, we aim to fill these knowledge gaps in compact models from the early development stage to the final product with Machine Learning methods. In our approach, we propose an Artificial Neural Network (ANN), a more specific Deep Learning (DL) model that imitates the output signals based on the transistor signal inputs without using semi-conductor physics. Then we will evolve our Machine Learning model to use it on compact model of transistor variants to get comparable data. To evaluate the results, we will compare the output signals over real compact models and evaluate the predictability performance of our machine learning model. As a result of these evaluations, we will discuss the improvements to increase efficiency and the potential optimization of our machine learning model for different compact models and semi-conductor systems.

Keywords: Machine learning, Neural networks, Compact model, Transistor model, Reliability test, Emerging technology

ÖZET

Kompakt transistör olarak adlandırdığımız modeller, yarı iletken teknolojisi ve devre tasarımı ile birlikte oluşan bir prototiptir ve geliştirilmesi oldukça zorlu ve zaman alan bir süreçtir. Bu prototipleri geliştiren ar-ge departmanları veya üreticiler model detaylarını çeşitli sebeplerden ötürü gizli tutmak isterler. Bu nedenle, ticari anlaşmalarda (örneğin Non-Disclosure-Agreements) kompakt transistör modellerine erişim kısıtlandıran maddeler bulunur. Makine öğrenmesi (ML) yöntemleri de aynı gizlilik anlaşması yapılmış kompakt modeller gibi kara kutu davranışı sergiler. Bu da onun içeride nasıl sonuç ürettiğini gizli tutar. Bu raporda, kompakt modellerin erken geliştirme safhasından son ürüne kadar ki gelişiminde oluşan bu bilgi boşluklarını Makine Öğrenmesi yöntemleriyle doldurmayı amaçlıyoruz. Yaklaşımımızda yarı iletken fiziği kullanmadan çip içerisindeki transistör sinyal girdilerine göre çıktıyı taklit eden bir Yapay Nöral Ağ (ANN), daha spesifik bir Derin Öğrenme (DL) modeli sunuyoruz. Daha sonra bu modeli çeşitli transistör varyantlarında denemek üzere evrimleştireceğiz. Sonucu değerlendirmek adına gerçek kompakt modeller üzerinden çeşitli sinyal girdilerinin çıktılarını karşılaştıracağız ve makine öğrenmesi modelimizin isabetlilik performansını değerlendireceğiz. Bu değerlendirmeler sonucu hem isabetliliğini arttırmaya yönelik gelişmeleri hem de makine öğrenmesi modelimizin farklı kompakt modeller ve yarı iletken sistemleri için potansiyel optimizasyonu üzerine konuşacağız.

Anahtar Kelimeler: Makine öğrenmesi, Nöral ağlar, Kompakt modeller, Transistör modelleri, Güvenirlilik testi, Gelişmekte olan teknoloji

1. INTRODUCTION

Recently, many electronic devices require containing chips with the kernel inside that will enable them to perform complex tasks. These chips are specially designed according to the device to be used as a result of cumulative experiences. The design phase of transistors, which do the main work inside the chips and whose structures are based on semiconductor technology, is a complex process. It is designed within multiple fundamental world factors (e.g., chip size, number of transistors, performance, heating). The compact model of transistors formed with various transistor types, and it is subjected to circuit analysis with SPICE-like software. Circuit analyses measure the response of the model to physical, electrical signals input. Such processes incur high monetary and time costs. Therefore, chip manufacturers or researchers want to keep this information a trade secret. Trade agreements (e.g. Non-Disclosure-Agreements) serve to keep the details confidential and hide the contents of the compact models for block reverse engineering. This is evident by the frequent use of random values in publications [1]-[3] and NDA blocking Process Design Kits (PDK) access.

With the end of Dennard's scaling and the current era of innovation in semiconductor technology, transistors are no longer just scalable, but their basic structure can also changeable. These innovations have allowed computer systems to improve continuously in terms of low power consumption and high performance; however, this innovation has brought a severe problem in the compact model because each innovation requires a new compact model instead of updated parameters on the model card. Therefore the compact model may take several years to develop. Development cannot begin until the basic technology has matured and its physical principles are understood. This delay in model release makes it difficult to use standard Electronic Design Automation (EDA) tools to evaluate the circuits in these emerging technologies, making it difficult to market and increase the time to market for new technologies.

Machine learning has created a revolution, not only in computer science but in almost all branches of science. Instead of making calculations with predetermined formulas, it provides learning and results by finding related information (features) directly from data. Machine learning algorithms have been known for decades. However, with the increase in computing power with the power of processors and the increase of memory size, it has allowed us to use it efficiently, as mentioned in [4]: "Machine learning is a new ingredient in this "virtuous cycle" of computing, with the potential of establishing more and stronger positive feedbacks. If using machine learning for designing computers will result in more computational power becoming available, machine learning tasks themselves will be the first ones to profit, enabling more advanced types of machine learning and even better effects on machine learning-enabled design automation and all the other machine-learning-enabled applications."

Designing electronic circuits is a complex and distributed process that can be divided into multiple steps. We focus ML on an essential part of this process: achieving circuit quality directly after manufacturing (test) and throughout its lifetime (reliability). Reliability is an essential feature of modern electronic products, which are usually designed for critical business applications ranging from industrial automation to self-driving cars. Considering these characteristics, the scale and architectural complexity of circuits designed to solve complex problems are growing exponentially. For example, traditional analytical models for reliability prediction begin to lose their accuracy as second and third-order effects become common and require replacing equations and trained ANN. Also, traditional algorithms, e.g., fault diagnosis or test point placement, can reach their limits for billion-gate designs, and alternative ML-based approaches may provide reasonable alternatives.

In this report, we propose Machine Learning, specifically a Neural Network based compact model of transistors that will act as an intermediary between early measurements from early prototypes and accurate yet complex compact models. Our goal is to provide neural network based transistor models to solve the above three problems. First, the neural network based transistor model is essentially a black-box modelling method, so the details of transistor manufacturing cannot be disclosed, thereby protecting the intellectual property rights of the manufactory. In addition, neural

network based transistor models are universal; thus, different transistor technologies can be handled uniformly, and there is no need to develop new models constantly. Finally, neural network based modelling can work without any domain knowledge such as physics, materials. That means that it enables to increase the development speed of this type of neural network based models with innovative transistors. Combined with a transistor modelling capability based on limited data from early prototype measurements, neural network-based transistor models are rapid prototyping for early design space exploration of technology. Naturally, these limited datasets result in faulty and constrained models at such an early stage. This, in turn, could drive or deter the development of more compact models for the technology.

2. LITERATURE REVIEW

Publications on this field can be categorized into two main classes; works using transistor knowledge and ones that do not.

2. 1 ML-based Transistor Models with Domain Knowledge

Early work on transistor modelling based on ML was resource-limited. The number of neurons and the number of layers in NN were limited due to computing power and instrument support. Therefore, many studies used domain knowledge to enhance the structure and weight of NN to improve its accuracy. Although this process brings performance improvements, it cannot be applied to new technologies (not yet fully explored) and optimize (and constrain) models of specific technologies.

[5] shows one of the early methods of modelling transistors using neural networks. NN-based ML model is used to predict current voltage, and it has done only 12 neurons. In 1996, ML methods were still in their infancy. Specific characteristics like monotonicity were defined by the training algorithms. Device domain knowledge used to reduce errors and reduce the size of required data for training.

A recent application of transistor modelling with neural network is presented in [6]. This time, a small neural network with less than 15 nodes and transistor physic information is used here. The nodes are connected manually, taking into account the physical dependences inside the transistor. The results in specific nodes in the model the sub-threshold area and nodes were modelling only the area above the voltage threshold. In order to keep the data they need to a minimum, they use sparse and non-uniform data sets. Another NN-based method introduced in [9] mainly focuses on the physical integration of the device and is aimed at the emerging ThinTFET transistor. Two

independent NNs, less than 10 per node, form the core structure is used for modelling. The control function set to follow the linear behaviour of the transistor in the sub-threshold range and the saturation behaviour above the threshold voltage. All these steps are taken to avoid learning physically impossible (against the law of physics) behaviours.

2. 2 ML-based Transistor Models without Domain Knowledge

It is possible to modelling without domain knowledge while it requires more computational power to achieve accurate results.

In [7], a combination of the Genetic Algorithm (GA) and NN is used. GA used for finding the best structure of NN in the training process, and NN used for transistor modelling. However, It can only apply to specific transistors. Therefore, this work could not be used in emerging technologies due to lacks of information and adequate validation.

[3] presents an NN-based imitate model as a compact model alternative for novel transistors. While they demonstrate their framework for fin effect transistor as an established technology and tunnel field-effect transistor as an emerging technology with no available compact model, the prediction accuracy is not discussed on some transistor curves. Therefore, their validation is lacking, and we cannot estimate their achieved modelling accuracy concerning the transistor parameters. In addition, traditional and emerging technologies are evaluated according to different schemes, and neither direct comparison nor the general validity of their methods is possible.

With the spread of ML techniques, commercial tools using ML techniques for compact model generation are recently available. [8] uses ML techniques to build models from input samples. Unfortunately, it is not possible to compare it to other works due to limited publicly available information.

3. CONCLUSION

In this proposal, we discussed the potential of NN-based transistor modelling, and by examining similar works, we have made a comparative analysis including pros and cons. Our work looks promising because the vast majority of the research on transistor modelling based on ML are current and, it is expected that it will produce more accurate results as the computation speed increases in processors.

Result of the black box characteristics of NNs, the extraction of technical details is challenging or even impossible. Compared with traditional compact models, machine learning solutions can be developed in a shorter time and be shared with others easily. With these advantages, ML provides a great way to accelerate technological progress, shorten time to market, and increase customer adoption through easier access.

REFERENCES

[1] S.-Y. Wu, C. Lin, M. Chiang, J. Liaw, J. Cheng, S. Yang, C. Tsai, P. Chen, T. Miyashita, C. Chang et al., *"A 7nm cmos platform technology featuring 4 th generation finfet transistors with a 0.027 um 2 high density 6-t sram cell for mobile soc applications,"* in 2016 IEEE International Electron Devices Meeting (IEDM). IEEE, 2016, pp. 2–6.

[2] Z. Zhang, R. Wang, C. Chen, Q. Huang, Y. Wang, C. Hu, D. Wu, J. Wang, and R. Huang, *"New-generation design-technology cooptimization (dtco): Machine-learning assisted modeling framework,"* in 2019 Silicon Nanoelectronics Workshop (SNW). IEEE, 2019, pp. 1–2.

[3] D. Ha, C. Yang, J. Lee, S. Lee, S. Lee, K.-I. Seo, H. Oh, E. Hwang, S. Do, S. Park et al., *"Highly manufacturable 7nm finfet technology featuring euv lithography for low power and high performance applications,"* in 2017 Symposium on VLSI Technology. IEEE, 2017, pp. T68–T69.

[4] Amrouch, H., Chowdhury, A. B., Jin, W., Karri, R., Khorrami, F., Krishnamurthy, P., ... & Tan, S. X. D. (2021, April). Special Session: *Machine Learning for Semiconductor Test and Reliability*. In 2021 IEEE 39th VLSI Test Symposium (VTS) (pp. 1-11). IEEE.

[5] P. B. L. Meijer, *Neural network applications in device and subcircuit modelling for circuit simulation.* Philips Electronics, 1996.

[6] L. Zhang and M. Chan, *"Artificial neural network design for compact modeling of generic transistors,"* Journal of Computational Electronics, vol. 16, no. 3, pp. 825–832, 2017.

[7] K. Lamamra and S. Berrah, "*Modeling of mosfet transistor by mlp neural networks,"* in International Conference on Electrical Engineering and Control Applications. Springer, 2016, pp. 407–415.

[8] T. Blaesi, *"Spice model generation by machine learning,"* 2019. [Online]. Available: https://silvaco-blog.com/circuit_simulation_modeling/spice-model-generation-by-machine-learning-html/

BIOGRAPHICAL SKETCH

Hüsnü Murat Koçak was born on August 10, 1996 in Ankara / Turkey. After graduating from Etimesgut Anatolian High School in 2014, he entered the Konya Technical University in Konya and received his B.S. degree from the Department of Computer Engineering in 2020. After that, in the same year, he enrolled in the M.Sc. program in the Computer Engineering department at Galatasaray University and is pursuing his studies here.

He attended to respectively University of Lorraine and KU Leuven as a research intern. Currently, he is writing up his thesis on machine learning techniques on chip behaviour at KU Leuven.