

ОПТИМИЗАЦИЯ КОНФИГУРАЦИИ НЕЙРОННОЙ СЕТИ ТИПА МНОГОСЛОЙНЫЙ ПЕРСЕПТРОН В ЗАДАЧЕ МОДЕЛИРОВАНИЯ СТАТИЧЕСКИХ ХАРАКТЕРИСТИК LDMOS ТРАНЗИСТОРА

М. И. Черных¹, С. А. Победа¹, Д. А. Жукалин²

¹ – АО «НИИЭТ»;

² – Воронежский Государственный университет

Поступила в редакцию 25.06.2021 г.

Аннотация. В статье представлен точный метод создания поведенческой модели латеральных МОП транзисторов (LDMOS), основанный на искусственных нейронных сетях (ИНС). Показан процесс оптимизации ИНС типа многослойный персептрон. Освещены вопросы обработки данных, необходимых при обучении ИНС и выбора гиперпараметров. Полученные результаты показывают перспективность использования ИНС для моделирования нелинейных полупроводниковых приборов.

Ключевые слова: LDMOS, искусственная нейронная сеть, оптимизация, вольт-амперные характеристики.

OPTIMIZATION OF A NEURAL NETWORK CONFIGURATION OF MULTILAYER PERSEPTRON TYPE IN THE PROBLEM OF MODELING THE STATIC CHARACTERISTICS OF A LDMOS TRANSISTOR

M. I. Chernykh, S. A. Pobeda, D. A. Zhukalin

Abstract. The article presents an accurate method for creating a behavioral model of lateral MOS transistors (LDMOS), based on artificial neural networks (ANN). The process of optimization of an ANN of the multilayer perceptron type is shown. The issues of data processing necessary for training the ANN and the choice of hyperparameters are highlighted. The results obtained show the promise of using ANNs for modeling nonlinear semiconductor devices.

Keywords: LDMOS, artificial neural network, optimization, current-voltage characteristics.

ВВЕДЕНИЕ

Важность транзисторов в современном мире сложно переоценить. На данный момент практически не осталось областей, в которых бы не применялись устройства, основанные на достижениях полупроводниковой промышленности. Из-за постоянной потребности к развитию и все большему усложнению устройств, повышению их надежности, мощности, удобства. Растет и число разрабатываемых транзисторов.

Однако недостаточно просто разработать хороший прибор, его еще нужно правильно применить, спроектировать устройство, которое будет реализовывать весь потенциал, заложенный в транзисторе.

Для решения этой задачи необходимо создавать реалистичные модели полупроводниковых изделий. Которые будут учитывать нелинейное поведение приборов, оказывающее пагубное воздействие на всю систему в целом. Модели подобного рода должны соответствовать нескольким основным критериям: во-первых модели необходимо создавать так, чтобы они могли предельно точно описывать поведение полупроводниковых устройств во всем диапазоне напряжений и частот, отражая все особенности и нелинейности в их поведении, во-вторых процесс создания моделей должен быть максимально упрощен, существующие на данный момент модели, основанные на системе уравнений аппроксимирующих все основные зависимости транзистора сложны в своей экстракции, требуют значительных временных ресурсов, и дорогостоящего оборудования. Также стоит учесть, что из-за огромного числа полупроводниковых устройств, традиционная разработка, основывающаяся на физике модели [1, 2] не поспевает за темпами развития технологий. Более того, модель должна хорошо предсказывать поведение транзистора за пределами тех данных, на которых она была создана.

Решить эту сложную и комплексную проблему, могут искусственные нейронные сети. Целью данной работы является разработка и оптимизация ИНС способной быстро обучаться и точно аппроксимировать ВАХ LDMOS транзистора.

ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

Способность к обучению и обобщению зависимостей, поиску нетривиальных закономерностей, а также нелинейный характер обработки данных делают искусственные нейронные сети (ИНС) хорошим инструментом в построении моделей.

Согласно универсальной теореме аппроксимации, точность аппроксимации ИНС может быть сколько угодно большой, при условии существования детерминированной связи между входными и выходными сигналами [3]. Однако эта же теорема не говорит нам о том, что наиболее эффективным способом аппроксимации является применение ИНС всего лишь с одним скрытым слоем. Наиболее распространённая топология ИНС в задачах моделирования нелинейных СВЧ компонентов — это многослойный персептрон (MLP), который обучается по средствам алгоритма обратного распространения ошибок.

Модели, созданные на основе ИНС, имеют преимущество в скорости расчетов и могут быть значительно точнее моделей основанных на стандартных методиках. В работах [4] и [5], было продемонстрировано, что ИНС прекрасно показывают себя в таких задачах как оптимизация СВЧ цепей, и моделирование статических характеристик приборов. ИНС хорошо зарекомендовали себя в вопросах моделирования, оптимизации и проектирования СВЧ фильтров [6], пассивных микроволновых компонентов [7], моделирования усилителей мощности [8]. В [9] дан обзор структур нейронных сетей и алгоритмов обучения в приложении к моделированию ВЧ и СВЧ устройств.

Область ИНС бурно развивается на протяжении последних двух десятилетий, предлагаются новые архитектуры и методы обработки данных, новые области применения. В этой статье рассмотрено использование современных инструментов создания, обучения и оптимизации ИНС, позволяющих в кратчайшие сроки разработать поведенческую модель LDMOS транзистора.

ОБЪЕКТ ИССЛЕДОВАНИЙ

В качестве объекта исследований выступает линейный LDMOS транзистор — 2П9123Б, разработанный АО «НИИЭТ». Транзистор предназначен для работы в аппаратуре радиосвязи с высокими требованиями по линейности и передающих станциях эфирного телевидения. ВАХ были получены на установке для автоматизированного измерения статических параметров мощных полевых и биполярных транзисторов Agilent — B1505A.

МЕТОД ИССЛЕДОВАНИЯ

При аппроксимации вольт-амперных характеристик (ВАХ) использовалась ИНС типа многослойный перцептрон. ИНС была написана на языке программирования Python с подключенным фреймворком Pytorch. Но перед тем как обсуждать топологию ИНС и ее оптимизацию, стоит сказать несколько слов о подготовке обучающих данных, так как от них во многом будет зависеть точность предсказаний ИНС.

В качестве входных сигналов использовались значения напряжений затвор-исток и сток-исток. Выходным же сигналом являлся ток стока. Соответственно все данные можно записать в таблицу с тремя столбцами. Однако в таком виде использовать полученные значения для обучения ИНС нельзя. Во-первых, необходимо разделить весь объем данных на тренировочную и тестовую выборки. На первой выборке проводилось обучение ИНС, вторая выборку ИНС использовалась для проверки полученных данных, при помощи тестовых значений проверялась возможность ИНС к предсказанию и контролировалась ее склонность к переобучению. Во-вторых, данные необходимо нормализовать, преобразовать все значения напряжений и токов таким образом, чтобы все они укладывались в диапазон от 0 до 1, что ускорит процесс обучения ИНС. Еще одним крайне важным шагом является рандомизация. Нужно избавиться от любых возможных корреляций внутри тренировочной выборки, для этого необходимо в случайном порядке перемешать строки в таблице с данными, иначе сеть не сможет правильно обучиться и вместо ВАХ, будет возвращать прямую линию.

Так же не маловажным является подбор оптимизатора способного эффективно выполнять задачи схожие с аппроксимацией ВАХ транзистора. Для создания больших моделей и работы с большими объемами данных, в нашем случае объем тренировочных данных составляет порядка 18 тысяч точек, лучше всего подходит оптимизатор Adam [10]. Далее стоит определиться с методом вычисления ошибки, нами был использован метод определения среднеквадратичной ошибки (MSE)

$$MSE = \frac{1}{m} \sum_{i=1}^m |x_i - y_i|^2, \quad (1)$$

где x_i ожидаемый результат, а y_i прогноз модели.

Как наиболее подходящий в тех задачах, где ИНС должна прогнозировать на выходе число, а не определять класс объекта, как это делается в задачах классификации или при обучении ИНС без учителя.

Далее стоит определиться с функцией активации нейронов, в нашем случае применялся классический сигмоид

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

Так как при аппроксимации ВАХ мы работали только в диапазоне положительных напряжений. При расширении моделирования в область отрицательных значений напряжений можно использовать вместо сигмоида тангенциальную функцию активации

$$\text{tanh}(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (3)$$

Помимо прочего, к каждому нейрону применялось смещение, величина которого определялась ИНС в процессе тренировки, что положительно сказывалось на скорости ее обучения.

ЭКСПЕРИМЕНТ

Начальной точкой оптимизации сети послужила конфигурация с одним скрытым слоем и 20 нейронами в нем. Для первых тренировок использовался объем обучающей выборки содержащий, 630 точек включающих в себя пять веток ВАХ и такой же объем тестовой выборки. Так же варьировалась скорость обучения.

Последовательное увеличение числа слоев и количества нейронов в каждом слое, не приводит к значимому увеличению точности предсказаний ИНС. Каждая ИНС приведенная в таблице 1 обучалась в течении 5000 эпох.

Таблица 1. Значения ошибки ИНС в зависимости от ее топологии в первой итерации.

№ ИНС	Конфигурация ИНС	Средняя ошибка на тренировочной выборке, А	Средняя ошибка на тестовой выборке, А
1	2:20:1	0,059591	0,199996
2	2:10:10:1	0,026387	0,079575
3	2:20:20:20:1	0,020847	0,042373
4	2:10:10:10:10:1	0,022973	0,040318

Расширение набора обучающих данных до 13672 точек, 20 веток ВАХ в тренировочной и столько же в тестовой выборках. Позволило ИНС с 5 скрытыми слоями и 30 нейронами в каждом, за 6 часов обучения достигнуть хорошей точности. ИНС с одним скрытым слоем и 150 нейронами, после обучения в течении 18 часов достигла точности предсказаний.

Таблица 2. Значения ошибки ИНС в зависимости от ее топологии во второй итерации.

№ ИНС	Конфигурация ИНС	Средняя ошибка на тренировочной выборке, А	Средняя ошибка на тестовой выборке, А
5	2:30:30:30:30:1	0,001086	0,000981
6	2:150:1	0,009642	0,008891

Далее размер обучающей выборки был увеличен до 17872 точек. Что позволило охватить область пробоя транзистора. Из-за увеличения числа нелинейностей в тренировочных данных точность предсказаний ИНС снизилась до значений указанных в таблице 3.

Таблица 3. Значения ошибки ИНС в зависимости от ее топологии в третьей итерации.

№ ИНС	Конфигурация ИНС	Средняя ошибка на тренировочной выборке, А	Средняя ошибка на тестовой выборке, А
7	2:30:30:30:30:1	0,009742	0,009914

На графике отображена зависимость средней ошибки ИНС от ее номера, Двумя линиями обозначаются результаты работы ИНС, линия “Обучение” показывает, как менялась ошибка в предсказании тех данных, на которых происходил процесс обучения. Линия “Предсказание” отображает изменение ошибки в прогнозировании данных, на которых ИНС не обучалась.

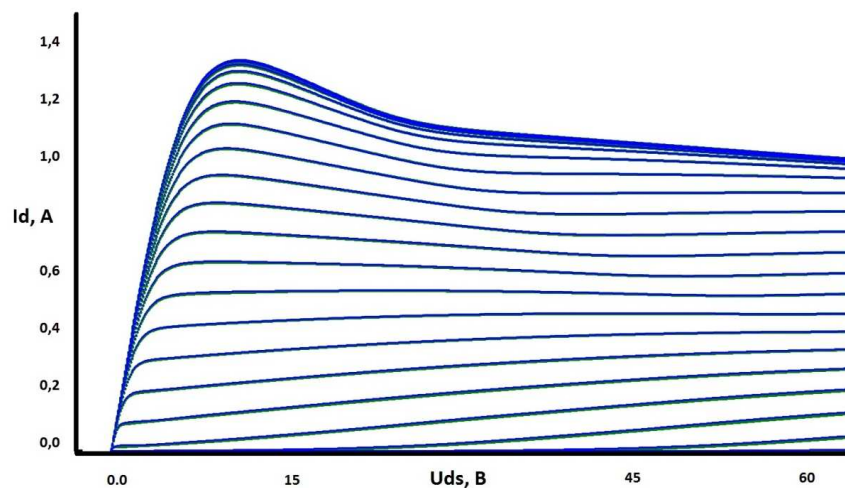


Рис. 1. Предсказания ИНС под номером 5.

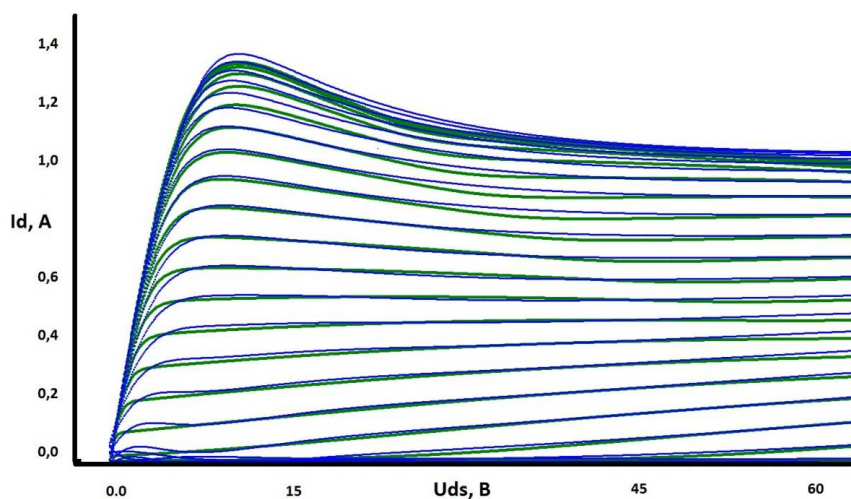


Рис. 2. Предсказания ИНС под номером 6.

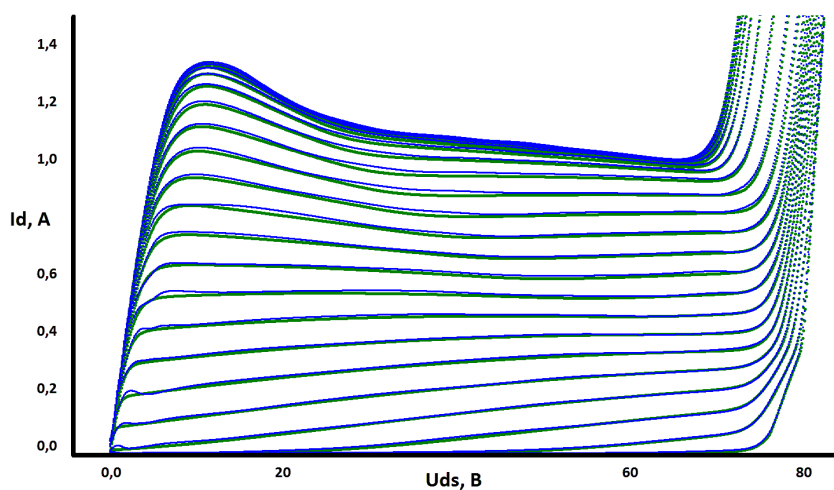


Рис. 3. Предсказания ИНС под номером 7

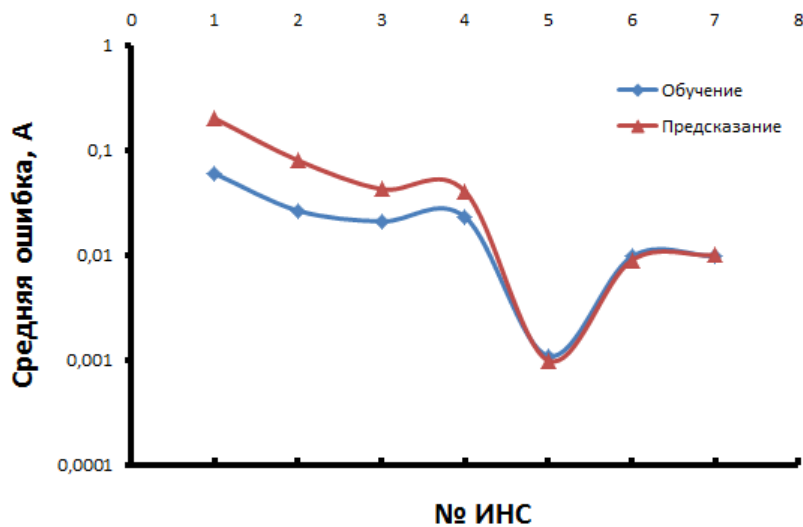


Рис. 4. График зависимости средней ошибки от номера ИНС.

ЗАКЛЮЧЕНИЕ

Путем увеличения числа скрытых слоев и количества нейронов в каждом слое, были получены несколько различных конфигураций ИНС. Было замечено, что проблемы с точностью у используемой ИНС возникают не из-за неправильной конфигурации сети, а из-за слишком маленького набора обучающих данных. Мы столкнулись с ситуацией, когда сеть достаточно точно описывала ветки ВАХ из тренировочной выборки, но с тестовой выборкой справлялась куда хуже. Так как в текущей версии ИНС число скрытых слоев было мало, это не могло быть переобучение, ситуация когда ИНС прекрасно запоминает обучающую выборку но теряет свою предсказательную силу. Решением этой проблемы было увеличение обучающих данных, так как в области малого напряжения сток-исток большая нелинейность не может быть корректно описана с имеющейся точностью.

Новый набор данных содержал уже порядка 14 тысяч точек. Сеть с 5 скрытыми слоями и 30 нейронами в каждом обучилась за 6 часов идеально моделировать не только тренировочные данные, но и давать точные предсказания на тестовой выборке. Сеть с одним слоем и 150 нейронами обучаясь в течении 18 часов, так и не смогла прийти к удовлетворительному результату.

Далее тренировочная выборка была расширена еще сильнее, до 18 тысяч точек, так, чтобы на ВАХ отображалась область пробоя транзистора. Здесь мы столкнулись с проблемой малого числа данных в области пробоя. Из-за этого не удается достичь той же точности моделирования, как в предыдущем случае. Причем неточности моделирования не локализуются только в области пробоя, а распространяются на весь диапазон напряжений. Однако ИНС все равно достаточно точно предсказывает поведение прибора. Для еще большего увеличения точности предсказаний ИНС, можно либо с максимальной плотностью точек исследовать область пробоя или же вручную аппроксимировать область с недостаточной плотностью точек, а затем на этих данных до обучить ИНС, таким образом можно добиться любой желаемой точности.

В работе рассмотрены основные вопросы обработки исходных данных необходимых для

обучения ИНС, а также уделено внимание процессу оптимизации архитектуры ИНС, а именно: выбору метода оптимизации, вычислению ошибок, функции активации, числу скрытых слоев и количеству нейронов. Предложенная архитектура ИНС способна за относительно небольшое часов обучения давать прогнозы с высокой точностью.

REFERENCES

1. Tsividis, Y. Operation and Modeling of the MOS Transistor / Y. Tsividis. — New York, 1999.
2. Khakifirooz, A. A simple semiempirical short-channel MOSFET current-voltage model continuous across all regions of operation and employing only physical parameters / A. Khakifirooz, O. M. Nayfeh, D. A. Antoniadis // IEEE Trans. Electron Devices. — 2009. — V. 56(8). — P. 1674–1680.
3. Artificial Neural Networks: Approximation and Learning Theory / H. White et. al. — Oxford, U.K.: Blackwell, 1992.
4. Zaabab, A. H. A neural network modeling approach to circuit optimization and statistical design / A. H. Zaabab, Q. J. Zhang, and M. S. Nakhla // IEEE Trans. Microwave Theory Tech. — 1995. — V. 43. — P. 1349–1358.
5. A neural network model for CAD and optimization of microwave filters / P. Burrascano, M. Dionigi, C. Fancelli, M. Mongiardo // IEEE MTT-S Int. Microwave Symp. Dig., Baltimore, MD, 1998. — P. 13–16.
6. Applications of artificial neural network techniques in microwave filter modeling, optimization and design / H. Kabir et. al. // PIERS Online. — 2007. — Т. 3, №. 7. — С. 1131–1135.
7. A novel deep neural network topology for parametric modeling of passive microwave components / J. Jin et. al. // IEEE Access. — 2020. — Т. 8. — С. 82273–82285.
8. Luongvinh, D. Behavioral modeling of power amplifiers using fully recurrent neural networks / D. Luongvinh, Y. Kwon // IEEE MTT-S International Microwave Symposium Digest, 2005. — IEEE, 2005. — P. 4.
9. Neural network structures and training algorithms for RF and microwave applications / F. Wang et. al. // International Journal of RF and Microwave Computer-Aided Engineering. — 1999. — Т. 9, №. 3. — С. 216–240.
10. Ruder S. An overview of gradient descent optimization algorithms / S. Ruder // arXiv preprint arXiv:1609.04747. — 2016.

Черных Максим Игоревич, канд. физ.-мат. наук, начальник лаборатории. АО “НИИ-ЭТ”, Воронеж, Россия
E-mail: m.i.chernykh@niiet.ru

Chernykh Maxim Igorevich, Cand. physical-mat. Sci., Head of the Laboratory. JSC “NIET”, Voronezh, Russia
E-mail: m.i.chernykh@niiet.ru

Победа Сергей Александрович, инженер-технолог 3 категории. АО “НИИЭТ”, Воронеж, Россия
E-mail: s.pobeda@niiet.ru

Pobeda Sergey Alexandrovich, process engineer of the 3rd category. JSC “NIET”, Voronezh, Russia
E-mail: s.pobeda@niiet.ru

*Жукалин Дмитрий Алексеевич, канд. физ.-мат. наук, доцент кафедры физики полупроводников и микроэлектроники, Воронежский государственный университет, Воронеж, Россия
E-mail: d.zhukalin@mail.ru*

*Zhukalin Dmitry Alekseevich, Cand. physical-mat. Sci., associate professor of the Department of physics of semiconductors and microelectronics. Voronezh State University, Voronezh, Russia
E-mail: d.zhukalin@mail.ru*