

И.А. БОРДАНОВ, С.А. ЩАНИКОВ

**Перспектива применения  
имитационного моделирования при  
проектировании искусственных  
нейронных сетей на базе  
мемристоров**

УДК 004.383.8.032.26

Муромский институт  
(филиал) ФГБОУ ВО  
«Владимирский  
государственный  
университет имени  
А.Г. и Н.Г. Столетовых»,  
г. Муром

*В статье рассмотрены основные преимущества аналоговой реализации искусственных нейронных сетей (ИНС) по сравнению с программной. Выделены некоторые недостатки реализации синапсов на транзисторах, резисторах и конденсаторах. Отмечены основные преимущества аналоговой реализации ИНС на базе мемристоров (ИНСМ). Выделены основные недостатки ИНСМ и рассмотрены некоторые подходы по их решению. Предложен собственный метод решения данных проблем, основанный на применении имитационного моделирования.*

*Ключевые слова: мемристор, нейроморфные системы, искусственные нейронные сети, имитационное моделирование*

Искусственные нейронные сети (ИНС) созданы по подобию их биологического аналога – сетей нейронов нервной системы живого организма. Биологическая нейронная сеть (БНС) представляет собой очень сложную, нелинейную и параллельную систему обработки информации, которая может обучаться. Эти особенности БНС стали мотивирующим фактором и привели к созданию ИНС эмулируемых на ЭВМ [1,2].

ИНС состоят из нейронов, соединённых направленными дугами – синапсами. Совокупность нейронов и связей между ним

составляют топологию сети. Каждый синапс имеет числовой вес. Нейрон ИНС представляет собой взвешенный сумматор, результат работы которого пропускается через нелинейную функцию активации [1, 3].

Первая ИНС, реализованная на ЭВМ, была разработана Розенблаттом в 1958 [3], и по архитектуре представляла собой однослойный перцептрон. Несмотря на её простую топологию, она могла решать сложные задачи, такие как распознавание букв.

В настоящее время существует множество архитектур ИНС. Они используются для решения большого количества задач таких как классификация, кластеризация, распознавание образов, прогнозирование и т. д. Их преимущества перед другими методами становятся особенно явными при обработке сложных неструктурированных данных, которые могут представляться не только в числовой форме, но и в виде изображений, текста и т.д. [4, 5].

Объемы информации, растущие с каждым годом, требуют всё большей скорости обработки от различных устройств и нейронные сети в этом плане не являются исключением. Программная эмуляция ИНС на ЭВМ с архитектурой фон Неймана является наиболее распространённой на сегодняшний день [4]. Такие ИНС потребляют много энергии и для их реализации необходимы значительные вычислительные мощности, что приводит к снижению временных показателей при работе [6]. Поэтому в настоящее время все большее количество научных коллективов исследуют варианты аналоговой реализации ИНС [3-6].

Переход от цифровых сигналов к аналоговым позволит увеличить скорость работы ИНС и снизить их энергопотребление. Одной из основных проблем в данном случае является учет погрешностей элементной базы, которые неизбежно возникают в процессе работы аналогового оборудования [7]. Эти погрешности не могут быть проигнорированы, так как они непосредственно влияют на значения весов синапсов ИНС, которые отвечают за вычисления в ИНС. Погрешности в весовых коэффициентах ИНС приводят к ухудшению точности её работы.

Другой не менее важной проблемой является выбор элементной базы, на которой будут реализованы синапсы ИНС. В аналоговых аппаратных реализациях, веса обычно хранятся в

резисторах, конденсаторах, и транзисторах с плавающим затвором [5]. Резисторы имеют статическое сопротивление, которое не может быть изменено в процессе работы. При реализации синапсов на базе конденсаторов, значения их весовых коэффициентов будут храниться лишь короткий промежуток времени из-за утечек заряда и требовать динамического обновления веса через определённые промежутки времени. Транзисторы же, в свою очередь, страдают от высокой нелинейности в синаптических взвешиваниях [5]. Поэтому ни одно из этих устройств не является достойным кандидатом на реализацию синапсов [6].

Появление же мемристора – энергонезависимого переменного резистора – открыло новые возможности в области аппаратной реализации ИНС. Он может изменять своё сопротивление и хранить его продолжительное время без затрат энергии. Программирование мемристора осуществляется напряжением или током. Запрограммированная информация сохраняется до тех пор, пока не будет применено следующее входное напряжение или ток. Другой особенностью мемристора является аналоговое умножение токов по закону Ома.

Совокупность вышеописанных преимуществ, делает его лучшим кандидатом на роль аппаратной альтернативы синапсов ИНС, по сравнению с имеющимися [8].

Мемристоры позволяют реализовать различные типы ИНС на их основе (ИНСМ) от самых простых (многослойный персептрон [9]) до сетей третьего поколения (спайковая нейронная сеть [8]), аналоговая реализация которой возможна, только благодаря тому, что с их помощью можно эмулировать синаптическую пластичность.

Но несмотря на все вышеназванные преимущества, мемристоры имеют и недостатки, которые связаны с изменчивостью их сопротивления в процессе работы, ограниченным числом рабочих состояний, небольшим рабочим диапазоном модуляции сопротивления, нелинейностью вольтамперных характеристик, нелинейностью модуляции сопротивления с импульсами напряжения [4]. Все эти недостатки не могут быть проигнорированы из-за того, что они приведут к снижению точности вычислений [3].

Из-за этих недостатков мемристоров, невозможно обучить нейронную сеть на компьютере, а затем перенести веса на данную элементную базу без последствий для точности её работы [6]. Для

некоторых задач это может быть не критично, но там, где важна высокая точность, погрешности могут сделать создаваемые устройства бесполезными.

Поэтому при разработке ИНСМ, необходимо учитывать погрешности элементной базы, чтобы нивелировать негативное влияние на итоговую точность работы нейроморфного устройства.

Одним из возможных решений данной проблемы является обучение ИНС с учётом различных погрешностей элементной базы. В свою очередь оно делится на три вида: программное, аппаратное и гибридное.

Первое из них предполагает программное обучение ИНС с дальнейшим переносом весов на элементную базу. Применяя данный метод, можно добиться уменьшения негативного влияния погрешностей элементной базы, тем самым приблизив точность работы к программной модели, но не достигнув её [10]. Причём наилучший результат обеспечивается при условии, что все погрешности аппаратной базы были учтены в модели ИНС [11].

Применение аппаратного обучения позволяет учитывать не идеальности синапсов из мемристоров в процессе обучения без моделирования. Это позволяет создавать ИНСМ сопоставимые по точности работы с программными моделями [5]. Процесс обучения искусственной нейронной сети на базе мемристоров является более долгим, чем программным методом [11], а аппаратная реализация наиболее эффективных алгоритмов является достаточно сложной [8].

При гибридном обучении модель ИНСМ обучается с начала на компьютере. Затем полученные значения весовых коэффициентов переносятся на элементную базу и там происходит дообучение аппаратно ориентированным методом. При данном подходе необходимо учитывать некоторые особенности элементной базы, такие как допустимый диапазон весов, вид функции активации нейронов и т. д [6]. Данный метод позволяет достичь лучшей точности работы ИНСМ [11].

Однако несмотря на то, что обучение с учётом погрешностей элементной базы позволяет обучить ИНСМ до нужной точности, оно не делает её весовые коэффициенты устойчивыми к погрешностям, которые возникают в процессе работы самого нейроморфного устройства и будут неизбежно влиять на точность его.

В качестве одного из методов решения данной проблемы, может выступать имитационное моделирование [3]. Если создать программную модель ИНС с учетом особенностей элементной базы [12] (допустимый диапазон значений весовых коэффициентов, вид функции активации и т.д.) и затем проэмулировать её работу с учётом погрешностей мемристоров, то можно определить как изменится точность работы ИНСМ и будет ли она, в контексте текущей задачи, удовлетворять техническому заданию.

Таким образом, необходимо реализовать программу, которая бы позволила подключать различные модели ИНСМ, и оценивать их точность работы при дестабилизирующих воздействиях, характерных для мемристоров. Для этого необходимо решить следующие задачи:

1. Разработать методы определения основных показателей качества функционирования подлежащих оценке при инженерном проектировании ИНСМ.

2. Разработать модели дестабилизирующих факторов, возникающих в процессе функционирования искусственных нейронных сетей на базе мемристоров.

3. Разработать имитационные модели ИНСМ для определения основных показателей качества их функционирования в условиях дестабилизирующих факторов.

4. Разработать планы экспериментов и алгоритмы визуализации результатов исследований.

5. Программно реализовать разработанные методы и алгоритмы в виде системы имитационного моделирования ИНСМ, со следующим функционалом:

- подключения к системе различных моделей ИНСМ, с целью дальнейшего исследования;

- оценка точности функционирования ИНСМ по различным критериям точности в номинальном режиме работы;

- оценка точности функционирования ИНСМ при воздействии дестабилизирующих факторов;

- определение допусков на значения параметров ИНСМ при которых она работает с требуемой точностью;

- сохранение результатов полученных в ходе экспериментов;

- загрузка и просмотр результатов экспериментов;

- реализация графического интерфейса системы, для удобства работы пользователя;
- визуализация результатов экспериментов;
- вывод результатов эксперимента в лог;
- возможность сохранения лога, полученного в ходе проведения экспериментов, в виде текстового файла;
- настройка параметров экспериментов.

6. Провести исследование разработанных методов и алгоритмов оценки основных показателей качества функционирования на примере моделей ИНСМ, решающих разные задачи.

Таким образом, мемристоры являются многообещающей технологией, на базе которой могут быть реализованы ИНС. Многие научные коллективы в настоящее время занимаются разработкой мемристоров, архитектур синапсов и ИНС на их основе, а также алгоритмов обучения. Однако недостатки, описанные в данной статье, препятствуют внедрению ИНСМ в состав сложных систем с гарантированной надежностью. Необходимо разработать методы, позволяющие учитывать данные недостатки на этапе проектирования ИНСМ. Проведенные авторами исследования [13,14] показывают, что наиболее подходящими технологиями для этих целей являются имитационное моделирование и планирование эксперимента.

Работа выполнена при поддержке гранта РФФИ №19-07-01215 (в части определения отказоустойчивости ИНСМ методами имитационного моделирования) и гранта Президента Российской Федерации МК-3927.2019.9 (в части применения методов имитационного моделирования для отказоустойчивой настройки ИНСМ).

### Литература

1. Галушкин А. И. Нейронные сети: основы теории. М.: Горячая линия - Телеком, 2015. 496 с.
2. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика / Перевод с англ. Ю. А. Зуева, В. А. Точенова. Под ред. А. И. Галушкина. М.: Мир, 1992. 236 с.
3. Yeung D et al. Sensitivity Analysis for Neural Networks. Natural Computing Series. Berlin: Springer-Verlag, 2010. 89 с.
4. Mehonic A. et al. Simulation of Inference Accuracy Using Realistic RRAM Devices. // Front. Neurosci. 2019. Vol. 13.

5. Adhikari et al. Memristor Bridge Synapse-Based Neural Network and Its Learning // IEEE Transactions on neural networks and learning systems. 2012. Vol. 23(9). P. 1426 – 1435.
6. Yang C. et al. A circuit-based neural network with hybrid learning of backpropagation and random weight change algorithms // Sensors. 2017. Vol. 17(1). №. 16.
7. Zeng X., Yeung D. Sensitivity analysis of multilayer perceptron to input and weight perturbations // IEEE Transactions on Neural Networks. 2001. Vol. 12(6). P. 1358 – 1366.
8. Guo Y. Unsupervised Learning on Resistive Memory Array Based Spiking Neural Networks // Front. Neurosci. 2019. Vol. 13.
9. Hasan R., Taha T., Yakopcic C. On-chip training of memristor crossbar based multi-layer neural networks // Microelectronics Journal. 2017. Vol. 66 P. 31-40.
10. Merrikh-Bayat F. Implementation of Multilayer Perceptron Network with Highly Uniform Passive Memristive Crossbar Circuits // Nature Communications. 2017. Vol. 9(1).
11. Merrikh-Bayat F. Memristor-Based Perceptron Classifier: Increasing Complexity and Coping with Imperfect Hardware // 2017 IEEE/ACM International Conference on Computer-Aided Design (ICCAD). 2017.
12. Галушкин А. И. Нейрочипы и нейроморфные ЭВМ: проблемы моделирования // Информационные технологии. 2015. Т. 21. № 12. С. 942-949.
13. Данилин С. Н. и др. Состояние исследований в области инженерного проектирования и производства нейрокомпьютеров // Алгоритмы, методы и системы обработки данных. 2019. № 1 (39). С. 14-45.
14. Данилин С. Н. и др. Разработка имитационных моделей искусственных нейронных сетей на базе мемристоров // Нейрокомпьютеры и их применение XVII Всероссийская научная конференция: тезисы докладов. 2019. С. 203-204.

ЩАНИКОВ С.А. SEACH@INBOX.RU

БОРДАНОВ И. А.  
BORDANOV2011@YANDEX.RU

Рецензент: Данилин Сергей Николаевич, к.т.н., доцент, доцент кафедры «Программная инженерия» Муромского института (филиала) ФГБОУ ВО «Владимирский государственный университет имени А.Г. и Н.Г. Столетовых», г.Муром.