

Прогнозирование динамического энергопотребления за счет переключений на этапе планировки физического проектирования ИС с использованием машинного обучения

В.А. Джанполадов

Национальный исследовательский университет «МИЭТ», г. Москва,
vakhtang.janpoladov@gmail.com

Аннотация — В данной работе разбирается применение алгоритмов машинного обучения для прогнозирования мощности рассеивания энергии за счет переключений компонентов схем на начальном этапе физического проектирования интегральных схем (ИС) для конкретной архитектуры. Реалистичная оценка потребляемой мощности возможна на заключительных этапах маршрута проектирования ИС, что может создать дополнительную итеративность в маршруте для оптимизации энергопотребления. Предложенный метод позволяет довольно точно спрогнозировать конечное значение рассматриваемого вида энергопотребления с высокой точностью для различных типов стандартных ячеек при различных сценариях и конфигурациях планировки. Недостатком метода является необходимость прохождения полного маршрута проектирования выбранной схемы с выбранным диапазоном параметров для сбора данных, нужных для обучения моделей машинного обучения, что требует дополнительных машинных и временных ресурсов.

Ключевые слова — динамическая мощность, прогнозирование, машинное обучение, интегральные схемы, регрессия, нейронные сети, арифметико-логический блок.

I. ВВЕДЕНИЕ

Энергопотребление является одним из ключевых параметров, характеризующих качество проектируемой ИС, наряду с быстродействием и занимаемой площадью. Роль энергопотребления растет с развитием индустрии и масштабированием технологических процессов. Известно, что общее энергопотребление схем можно разделить на две составляющие: статическое и динамическое. Его можно представить в виде суммы этих компонентов, как показано в формуле (1).

$$P_{\text{общее}} = P_{\text{стат}} + P_{\text{дин}}, \quad \#(1)$$

где $P_{\text{общее}}$ — это общее энергопотребление, $P_{\text{стат}}$ — статическое энергопотребление или мощность утечки, а $P_{\text{дин}}$ — соответственно динамическая мощность. Статическая мощность рассеивается в статическом состоянии схемы и обусловлена в основном токами утечки. Динамическое энергопотребление же является следствием переходных процессов в ИС.

Динамическую мощность, в свою очередь, разделяют на еще два компонента: мощность, рассеиваемую за счет переключений сигналов на выходах элементов схемы и заряда или разряда емкостей нагрузки ($P_{\text{перекл}}$), и мощность, обусловленную сквозными токами при переключениях в самих узлах элементов ($P_{\text{сквоз}}$). [1] Таким образом, в развернутом виде формула 1 имеет вид:

$$P_{\text{общее}} = P_{\text{стат}} + P_{\text{дин}} = P_{\text{стат}} + P_{\text{перекл}} + P_{\text{сквоз}} \cdot \#(2)$$

Расчет точного значения энергопотребления проектируемой ИС требует значительных машинных и временных ресурсов. Кроме того, для реалистичной оценки потребляемой мощности необходимо выполнять расчеты на поздних этапах маршрута проектирования, где уже готова топология проектируемой ИС и может быть произведено моделирование на уровне вентилях с учетом паразитных эффектов. Из вышесказанного становится ясно, что при неудовлетворительных значениях энергопотребления проектируемой ИС, необходим возврат к предыдущим этапам и дополнительная оптимизация. Это вводит дополнительную итеративность в маршрут проектирования и, как следствие, приводит к увеличению общего времени проектирования и экономическим рискам.

В свою очередь, этап физического проектирования в маршруте является довольно трудоемким процессом, состоящим из нескольких крупных шагов: декомпозиция, планировка, размещение, синтез дерева синхросигнала, трассировка и проверка проектируемой топологии на соответствие заданным требованиям. Резюмируя вышесказанное можно сделать вывод, что разработка методов, позволяющих спрогнозировать энергопотребление ИС на ранних этапах проектирования является актуальной задачей. Для решения подобного рода задач, где на основе входных данных нужно спрогнозировать какое-то выходное значение, широко используется машинное обучение (МО).

На сегодняшний день существуют два основных подхода оценки энергопотребления ИС: статический и динамический анализ. Статические методы основаны на вычислении вероятностей переключений сигналов в

схеме, для чего не требуются входные векторы [2]. В основе динамических методов лежит моделирование различных сценариев оперирования проектируемой схемы. Благодаря динамическому анализу можно реалистично оценить потребляемую мощность схемы, но для этого нужны адекватные входные тестовые вектора, максимально покрывающие область возможных входных значений [3]-[4], а также значительные временные и машинные ресурсы. Несмотря на то, что статические методы оценки уступают в точности динамическим, они позволяют проводить оценку энергопотребления на ранних этапах проектирования.

В настоящее время активно исследуются пути интеграции методов МО для автоматизации проектирования ИС, в частности – в физическое проектирование [5]-[9]. В работе [10] алгоритмы МО используются для прогнозирования производительности схемы на этапе планировки ИС.

В данной статье рассматриваются возможности использования моделей МО для оценки динамического энергопотребления за счет переключений компонентов схем на этапе планировки ИС для спроектированного блока.

II. ГЕНЕРАЦИЯ БАЗЫ ДАННЫХ

Для обучения моделей МО необходима база данных, на основе которой и будет происходить обучение.

Целью данной работы является создание моделей МО, способных спрогнозировать конечное значение энергопотребления при различных сценариях, для чего необходимо иметь как данные с различных этапов проектирования ИС, так и конечное значение потребляемой энергии для каждого сценария. В этой главе рассматривается процесс создания набора данных, состоящий из 4 основных пунктов.

A. Начальный набор входных данных

На начальном этапе определены основные параметры, при которых должно быть спроектировано устройство для исследований. После определения необходимо выбрать разброс выбранных параметров. Определение вышеперечисленных параметров было основано на исходных данных, доступных для выбранной технологии. В начальный набор данных вошли следующие параметры: напряжение питания, температура, тактовая частота, а также пороговое напряжение транзисторов. Выбрано три типа библиотек на основе следующих пороговых напряжений: стандартное пороговое напряжение (SVT), низкое пороговое напряжение (LVT) и супернизкое пороговое напряжение (SLVT). Разброс значений выбранных параметров показан в табл. 1. Остальные входные параметры будут получены на более поздних этапах маршрута проектирования цифровых ИС. Для прохождения всего маршрута были использованы программные средства компании Synopsys Inc.

B. Проектирование устройства для проведения исследований

Для оценки эффективности предложенного метода разработан модуль с использованием языка описания аппаратуры Verilog, состоящий из Арифметико-логического устройства (ALU) и Регистров общего назначения (GPRS). Разработанное устройство поддерживает 16 логических операций с 64-битными операндами. Управление логическими операциями выполняется с помощью 4-разрядного входного сигнала ALU_Control. Модуль GPRS представляет собой массив регистров 32x64, откуда считываются операнды и куда записываются результаты из блока ALU. Этот процесс контролируется управляющими входными сигналами записи и чтения.

Для проверки корректности работы устройства необходимо выполнить функциональное моделирование, для чего требуется создание входных векторов. Корректность работы была проверена с использованием программы VCS [11]. Логический синтез был выполнен с помощью программы Design Compiler [12] для различных сценариев. Всего для разработки данной базы данных было выполнено 54 логических синтеза. Графическое представление логической схемы представлено на рис. 1а. Для проверки соответствия полученных описаний на уровне вентилей с исходным описанием была использована программа Formality [13]. Данный этап называется формальной верификацией. Для физического проектирования использовалась программа IC Compiler II [14].

На этапе планировки в число входных параметров для обучения моделей МО добавляется значение утилизации блока, определяющее плотность размещения, а также соотношение сторон проектируемого чипа, определяющее его форму. Всего было получено 486 топологий с учетом различных параметров. Пример топологии приведен на рис. 1б.

После топологического синтеза был произведен анализ энергопотребления с учетом паразитных элементов и информации о переключениях, полученной из моделирования описания на уровне вентилей после физического проектирования. Используемые в данной работе значения вышеуказанных параметров показаны в табл. 1.

Как уже было сказано выше, для адекватной оценки энергопотребления на основе моделирования схемы, необходимо создание последовательности тестовых векторов, по возможности покрывающих область возможных входных комбинаций. Для этого был использован подход генерации случайных входных значений для каждого входа схемы с периодом в один такт. Всего количество циклов для одного моделирования составляет 5000. Данное число выбрано опытным путем исходя из временных затрат и места на жестком диске. Пример полученных результатов моделирования показан на рис. 1в.

Таблица 1

Параметры, входящие в базу данных

Параметр	Значения	Источник
Напряжение питания	0.825В, 0.935В	Угол библиотеки
Температура	-40°C, 0°C, 125°C	Угол библиотеки
Частота	2 ГГц, 2.67 ГГц, 4 ГГц	Ограничения
Тип библиотеки	SVT, LVT, SLVT	Библиотека
Утилизация	0.15, 0.3, 0.45	Планировка
Соотношение сторон	1:1, 1:2, 1:3	Планировка
Энергопотребление последовательных ячеек	Получены после расчета	Отчет в PrimeTime
Энергопотребление комбинационных ячеек	Получены после расчета	Отчет в PrimeTime
Энергопотребление ячеек в сети синхросигнала	Получены после расчета	Отчет в PrimeTime
Энергопотребление регистров	Получены после расчета	Отчет в PrimeTime

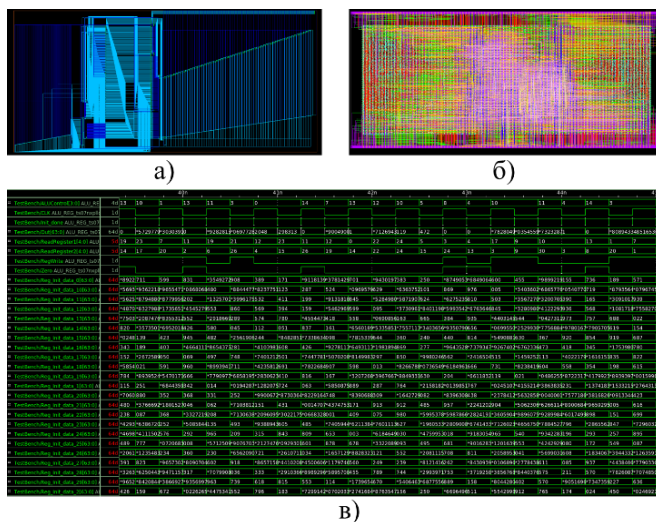


Рис. 1. Результаты проектирования тестового устройства: а) логическая схема, б) топология, в) моделирование описания на уровне вентилях

Для экстракции паразитики использовалась программа StarRC [15], а для конечного расчета энергопотребления – PrimeTime [16].

В качестве входных параметров для моделей МО также предполагалось использовать количество ячеек и количество межсоединений. Поскольку архитектура устройства фиксирована, значительного разброса значений упомянутых параметров не было обнаружено после логического синтеза. Эти числа можно считать постоянными вследствие малого отклонения от среднего значения, поэтому их влияние на процесс обучения будет незначительным и нет нужды добавлять их в набор данных.

Стоит отметить, что в число входных параметров не входят данные, зависящие от используемого технологического узла, что делает предлагаемый маршрут универсальным для различных технологий.

С. Набор выходных параметров

В качестве выходных данных для обучения моделей МО нужно собрать отчеты о потреблении мощности, рассчитанной на основе переключательной активности, полученной из моделирования, а также с учетом паразитных элементов в топологии.

PrimeTime позволяет собрать отчеты для следующих видов ячеек в блоке: регистры, комбинационные ячейки, последовательные ячейки и ячейки, входящие в состав дерева синхросигнала. Значения энергопотребления были собраны для всех вышеперечисленных групп соответственно. Данный подход позволит повысить точность прогнозирования, но требует дополнительных ресурсов, так как для каждой группы ячеек должна быть обучена отдельная модель.

Итоговые входные и выходные данные, входящие в созданную базу данных, со значениями, использованными для данной работы, продемонстрированы в табл. 1.

Д. Обработка данных

Предварительная обработка данных является важным начальным этапом в обучении моделей. Данный этап позволяет улучшить производительность моделей МО. Вид обработки зависит от конкретного типа данных. Данные могут быть разных типов: отсутствующие данные, шумовые данные, категориальные данные и т.д. Простейший способ обработки категориальных данных, таких как тип библиотеки со значениями SVT, LVT и SLVT – это присвоение каждой категории индивидуального числового значения. В данной работе было выполнено следующее присвоение: SVT – 1, LVT – 2, SLVT – 3.

Полученная база данных разделена на случайные обучающие и тестовые наборы в процентном соотношении 60:40. В итоге разделения получилось 291 обучающих наборов и 195 тестовых. Входные параметры в базе данных имеют широкий спектр значений, поэтому необходимо провести их обработку в виде нормализации значений или же стандартизации. Если обработка входных параметров не будет выполнена, то параметры с более крупными значениями могут доминировать над другими в процессе обучения. В данной работе стандартизация Z-score используется в качестве метода масштабирования для набора входных параметров, которая рассчитывается по приведенной формуле ниже.

$$z = (x - u) / s, \#(3)$$

где u - среднее значение обучающего набора, а s - его стандартное отклонение. Масштабирование выполняется независимо для каждого входного

параметра. Среднее значение и стандартное отклонение сохраняются для последующего масштабирования тестовых и новых входных данных. Масштабирование возможно применить и для выходных данных, используя свои собственные среднее значение и стандартное отклонение.

III. ОБУЧЕНИЕ МОДЕЛЕЙ МО

МО в основном делится на три направления: обучение без учителя, обучение с подкреплением и обучение с учителем. Обучение без учителя происходит на основе входных данных, где алгоритм вычисляет особенности данных. Примером таких алгоритмов является кластеризация. Обучение с подкреплением основывается на логике вознаграждений при взаимодействии со средой. В обучении с учителем же присутствуют конечные значения, которые должны быть спрогнозированы. В данной работе используются алгоритмы обучения с учителем, так как имеются выходные данные, на основе которых должно проходить обучение. [17]. Алгоритмы обучения с учителем, где имеются векторы входных параметров и вектор выходных значений, разделяются на два вида: алгоритмы регрессии и классификации. Алгоритмы регрессии используются для прогнозирования непрерывных значений, когда как алгоритмы классификации прогнозируют дискретные значения. Алгоритмы регрессии лучше подходят для решения поставленной задачи, так как целью данной работы является прогнозирование значения энергопотребления за счет переключений. Для проведения исследований были выбраны следующие 5 алгоритмов: полиномиальная регрессия (англ. Polynomial Regression), регрессия дерева решений (англ. Decision Tree Regression), регрессия случайного леса (англ. Random Forest Regression), регрессия опорных векторов (англ. SVR, Supported Vector Regression) и многослойный перцептрон (англ. MLP, Multi-layer perceptron). Исследование эффективности вышеперечисленных алгоритмов кроме MLP для прогнозирования непрерывных значений на примере базы данных из открытого доступа представлено в работе [18]. В данной работе также был использован многослойный перцептрон [19], который является структурой нейронной сети с обратным распространением в процессе обучения. Так как база данных содержит не один входной параметр, а 6, базовый алгоритм линейной регрессии не был рассмотрен. Также существуют иные популярные алгоритмы регрессии, такие как логистическая регрессия (англ. Logistic Regression), гребневая регрессия (англ. Ridge regression) и т. д., которые не были использованы в данной работе.

Обучение и выбор удовлетворительной модели осуществляется с помощью следующих шагов: создаются обучающие и тренировочные наборы данных для регистров, комбинационных и последовательных ячеек, а также для ячеек, входящих в дерево синхросигнала. На этапе обработки данных производится стандартизация входных данных, которая была описана в прошлой главе. Далее строится

модель, основанная на вышеперечисленных алгоритмах. Настраиваются параметры моделей для наилучшей производительности. Это делается с помощью этапа валидации, когда производительность модели с выбранными параметрами оценивается на определенной выборке данных. Настройка параметров делалась в особенности для моделей полиномиальной регрессии и многослойного перцептрона. Далее производится обучение модели на тренировочном наборе и производится оценка ее производительности на тестовом наборе данных. В тестовых наборе данных не входят данные, используемые во время обучения.

Для оценки производительности были выбраны следующие метрики: корень из средней квадратичной ошибки (англ. RMSE, Root Mean Squared Error), коэффициент детерминации или оценка R^2 (англ. R-Squared) и средняя абсолютная процентная ошибка (англ. MAPE, Mean Absolute Percentage Error). Расчет метрик RMSE, R^2 и MAPE продемонстрирован в формулах 4, 5 и 6 соответственно.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - y_i)^2}, \#(4)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (a_i - y_i)^2}{\sum_{i=1}^n (a_i - \bar{a})^2}, \#(5)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{a_i - y_i}{a_i} \right| * 100\%, \#(6)$$

где n – количество наборов, y_i – i -тое спрогнозированное значение, a_i – i -тое реальное значение, $\bar{a}$ – среднее реальных значений. Чем меньше значения RMSE и MAPE, тем лучше созданная модель. Значение R^2 может принимать значения от -1 до +1. Если значение близко к +1, то модель хорошо подходит для выбранных наборов данных и прослеживается строгая связь между параметрами. Значения, близкие к -1, говорят о том, что наблюдается строгая обратная связь. В случае значений, близких к 0, можно сказать, что особой связи между параметрами не наблюдается. В качестве метрики, по которой будут фильтроваться полученные модели МО, была выбрана MAPE. Блок схема алгоритма обучения и оценки моделей МО показан на рис 2.

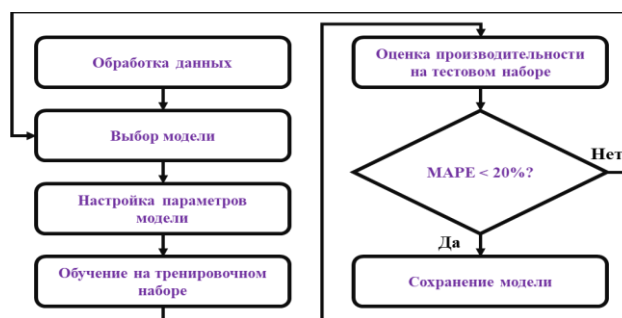


Рис. 2. Блок-схема алгоритма обучения и оценки моделей МО

IV. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Прогнозирование конечного энергопотребления на этапе планировки ИС является сложной задачей в виду отсутствия важных входных данных, таких как размещение стандартных ячеек, оптимизированное дерево синхросигнала, длина межсоединений, паразитные элементы и т. д. Машинное обучение как раз подходящий инструмент для решения задач такого рода.

Данная работа нацелена на прогнозирование динамического энергопотребления за счет переключений для 4 групп ячеек: регистров, последовательных и комбинационных ячеек, а также ячеек, входящих в состав дерева синхронизации. Разделение общего энергопотребления на эти 4 группы с одной стороны упростит анализ проектируемой схемы, позволит идентифицировать проблемную группу ячеек, нуждающуюся в оптимизации, а с другой – требует создания, оптимизации и анализа различных моделей МО.

Как было сказано выше, имеется 6 входных и один выходной параметр для каждой отдельной модели МО, зависящий от выбранной группы ячеек. Все вышеописанные модели и метрики для оценки

производительности были использованы из библиотеки Scikit-Learn [20]. Результаты метрик для вышеописанных алгоритмов показаны в таблицах 2, 3, 4, 5 и 6. Для краткости, в таблицах приведены результаты, полученные на тестовом наборе данных, который не был использован в процессе обучения. Для алгоритма полиномиальной регрессии приведены результаты, полученные при степени полиномов равной 4, так как при этом значении были получены лучшие результаты этой модели. Всего были протестированы полиномы 2, 3, 4, и 5 порядка. При высокой степени полиномов наблюдалось переобучение модели, при котором производительность на обучающем наборе данных хороша, но на тестовых наборах неудовлетворительна.

Для всех алгоритмов прослеживаются значения метрики R^2 близкие к 1, что говорит о положительной тенденции зависимости выходного значения от входных параметров. Метрика MAPE, которая используется для фильтрации пригодных моделей, имеет свои недостатки: чувствительность к значениям, сильно отклонившимся от остального набора данных.

Таблица 2

Результаты алгоритма Polynomial Regression

Группа ячеек	RMSE (нВт)	R^2	MAPE (%)
Регистры	434841.4015	0.8487	39.7733
Комбинационные	548669.0765	0.9805	54.1899
Последовательные	146245.5608	0.8172	73.6227
В составе дерева синхросигнала	81047.5340	0.9899	45.9959

Таблица 3

Результаты алгоритма Decision Tree Regression

Группа ячеек	RMSE (нВт)	R^2	MAPE (%)
Регистры	406499.9979	0.8673	39.0110
Комбинационные	594017.7650	0.9804	57.0184
Последовательные	176141.5631	0.7763	79.4085
В составе дерева синхросигнала	88149.8496	0.9874	44.2601

Таблица 4

Результаты алгоритма Random Forest Regression

Группа ячеек	RMSE (нВт)	R^2	MAPE (%)
Регистры	280948.5542	0.9382	39.3952
Комбинационные	400676.1174	0.9904	54.9591
Последовательные	129855.4714	0.8531	71.7232
В составе дерева синхросигнала	69453.3060	0.9915	44.0952

Таблица 5

Результаты алгоритма SVR

Группа ячеек	RMSE (нВт)	R^2	MAPE (%)
Регистры	338122.2702	0.9164	40.8842
Комбинационные	771037.7082	0.9661	55.0559
Последовательные	142679.7355	0.8264	71.6161
В составе дерева синхросигнала	114812.9642	0.9799	46.3582

Таблица 6

Результаты структуры нейронной сети MLP

Группа ячеек	RMSE (нВт)	R^2	MAPE (%)
Регистры	366755.5713	0.9044	8.3563
Комбинационные	492476.0877	0.9876	3.7758
Последовательные	122032.4353	0.8682	18.1301
В составе дерева синхросигнала	79779.1329	0.9891	2.3604

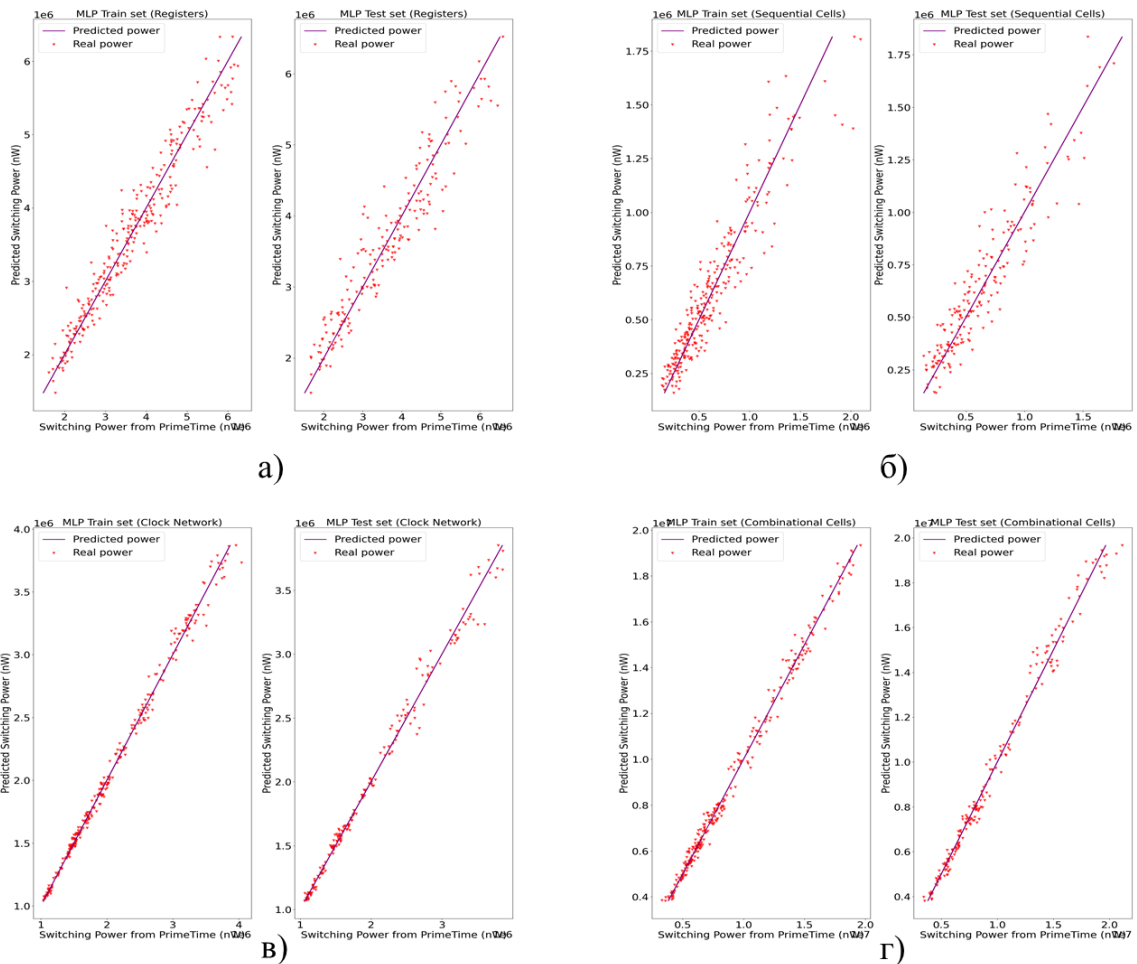


Рис. 3. Результаты работы MLP на обучающем и тестовом наборах данных: а) для регистров, б) для последовательных ячеек, в) для ячеек в составе дерева синхросигнала, г) для комбинационных ячеек

Из результатов можно сделать вывод, что поставленному критерию фильтрации удовлетворительных моделей (MAPE < 20%) соответствуют модели MLP, основанные на структуре нейронной сети. На рис. 3 продемонстрированы результаты для тренировочного и тестового наборов.

ЗАКЛЮЧЕНИЕ

В данной статье продемонстрировано использование алгоритмов МО для прогнозирования энергопотребления ИС за счет переключений в узлах на этапе планировки в физическом проектировании. Прогнозирование делается отдельно для 4 различных групп ячеек, что позволит идентифицировать критические группы с точки зрения энергопотребления, проанализировать результаты и оптимизировать схему.

Было протестировано 5 алгоритмов МО для решения поставленной задачи. Лучшую производительность удалось получить с использованием структуры нейронной сети MLP. Метрика MAPE для регистров равна ~8.35%, для последовательных ячеек ~18.13%, для сети тактового синхросигнала ~2.36% и для комбинационных ячеек ~3.77%.

ЛИТЕРАТУРА

- [1] Черемисинова Л.Д. Методы оценки энергопотребления при проектировании кмп-микросхем // Образовательные ресурсы и технологии. 2018. №4 (25). URL: <https://cyberleninka.ru/article/n/metody-otsenki-energopotrebleniya-pri-proektirovanii-kmp-mikroshem> (дата обращения: 01.08.2022).
- [2] R. Burch, F. Najm, P. Yang and D. Hoeser, "Pattern-independent current estimation for reliability analysis of CMOS circuits," 25th ACM/IEEE, Design Automation Conference.Proceedings 1988., 1988, pp. 294-299, doi: 10.1109/DAC.1988.14773.
- [3] S. M. Kang, "Accurate simulation of power dissipation in VLSI circuits," in IEEE Journal of Solid-State Circuits, vol. 21, no. 5, pp. 889-891, Oct. 1986, doi: 10.1109/JSSC.1986.1052622.
- [4] G. Y. Yacoub and W. H. Ku, "An accurate simulation technique for short-circuit power dissipation based on current component isolation," IEEE International Symposium on Circuits and Systems., 1989, pp. 1157-1161 vol.2, doi: 10.1109/ISCAS.1989.100558.
- [5] Kahng, A.B. (2018). Machine Learning Applications in Physical Design: Recent Results and Directions. Proceedings of the 2018 International Symposium on Physical Design.
- [6] Lin, Y., & Pan, D.Z. (2019). Machine Learning in Physical Verification, Mask Synthesis, and Physical Design. Machine Learning in VLSI Computer-Aided Design.
- [7] Yibo, L., Xiaohan, G., Tinghuan, C., & Bei, Y. (2022). Machine learning for digital circuit backend design.
- [8] Yan, J., Lyu, X., Cheng, R., & Lin, Y. (2022). Towards Machine Learning for Placement and Routing in Chip Design: a Methodological Overview. ArXiv, abs/2202.13564.
- [9] L. Wang and M. Luo, "Machine Learning Applications and Opportunities in IC Design Flow," 2019 International Symposium on VLSI Design, Automation and Test (VLSI-DAT), 2019, pp. 1-3, doi: 10.1109/VLSI-DAT.2019.8742073.
- [10] L. Bai and L. Chen, "Machine-Learning-Based Early-Stage Timing Prediction in SoC Physical Design," 2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), 2018, pp. 1-3, doi: 10.1109/ICSICT.2018.8565778.
- [17] M. A. El Mrabet, K. El Makkaoui and A. Faize, "Supervised Machine Learning: A Survey," 2021 4th International Conference on Advanced Communication Technologies and Networking (CommNet), 2021, pp. 1-10, doi: 10.1109/CommNet52204.2021.9641998
- [18] D. Kinaneva, G. Hristov, P. Kyuchukov, G. Georgiev, P. Zahariev and R. Daskalov, "Machine Learning Algorithms for Regression Analysis and Predictions of Numerical Data," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-6, doi: 10.1109/HORA52670.2021.9461298.
- [19] Popescu, M., Balas, V.E., Perescu-Popescu, L., & Mastorakis, N.E. (2009). Multilayer perceptron and neural networks. WSEAS Transactions on Circuits and Systems archive, 8, 579-588.
- [20] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Machine Learning-Based Switching Power Prediction at Floorplan Stage of IC Physical Design

V.A. Janpoladov

National Research University of Electronic Technology, Moscow,

vakhtang.janpoladov@gmail.com

Abstract — This paper demonstrates the effectiveness of using machine learning algorithms to predict the switching power at the floorplan stage of the physical design for a specific block architecture. Since the design process of integrated circuits (ICs) includes iterative stages for circuit optimization and time-to-market is critical for the industry,

it is crucial to develop effective methods to evaluate the parameters of the IC at the early stages of the design process. The effectiveness of the proposed method was demonstrated for a block containing a 64-bit Arithmetic Logic Unit (ALU) coupled with General-Purpose Registers (GPRs). This block has been designed for 486 different configurations and

scenarios. The study was carried out for the following cell groups: registers, sequential cells, combinational cells, and cells in clock network. To solve the regression problem of switching power prediction five algorithms were selected and studied: Polynomial Regression, Decision Tree Regression, Random Forest Regression, Supported Vector Regression and Multi-layer Perceptron (MLP). To evaluate prediction accuracy Root Mean Squared Error, R2 and Mean Absolute Percentage Error (MAPE) metrics were used. The main goal was to identify the algorithm with a minimum MAPE value for each cell group. The best performance with lowest MAPE values was obtained with MLP algorithms. The MAPE value for registers is ~8.35%, for serial cells ~18.13%, for the clock network ~2.36%, and for combinational cells ~3.77%. The proposed method does not use technology-dependent data, which makes it universal for any technology nodes used to design different blocks. The disadvantage of the method is the need for implementation of the whole design flow for a selected design with the selected range of parameters for collecting necessary training data, which requires additional time and machine resources.

Keywords — dynamic power, prediction, machine learning, integrated circuits, regression, neural networks, arithmetic logic unit

REFERENCES

- [1] Cheremisinova L.D. Metody otsenki energopotrebleniya pri proektirovanii kmop mikroskhem (Methods for Estimating Power Consumption in CMOS Circuits Design) // Obrazovatel'niye resursi I tekhnologii. 2018. №4 (25). URL: <https://cyberleninka.ru/article/n/metody-otsenki-energopotrebleniya-pri-proektirovanii-kmop-mikroskhem> (access date: 01.08.2022).
- [2] R. Burch, F. Najm, P. Yang and D. Hoeser, "Pattern-independent current estimation for reliability analysis of CMOS circuits," 25th ACM/IEEE, Design Automation Conference.Proceedings 1988., 1988, pp. 294-299, doi: 10.1109/DAC.1988.14773.
- [3] S. M. Kang, "Accurate simulation of power dissipation in VLSI circuits," in IEEE Journal of Solid-State Circuits, vol. 21, no. 5, pp. 889-891, Oct. 1986, doi: 10.1109/JSSC.1986.1052622.
- [4] G. Y. Yacoub and W. H. Ku, "An accurate simulation technique for short-circuit power dissipation based on current component isolation," IEEE International Symposium on Circuits and Systems., 1989, pp. 1157-1161 vol.2, doi: 10.1109/ISCAS.1989.100558.
- [5] Kahng, A.B. (2018). Machine Learning Applications in Physical Design: Recent Results and Directions. Proceedings of the 2018 International Symposium on Physical Design.
- [6] Lin, Y., & Pan, D.Z. (2019). Machine Learning in Physical Verification, Mask Synthesis, and Physical Design. Machine Learning in VLSI Computer-Aided Design.
- [7] Yibo, L., Xiaohan, G., Tinghuan, C., & Bei, Y. (2022). Machine learning for digital circuit backend design.
- [8] Yan, J., Lyu, X., Cheng, R., & Lin, Y. (2022). Towards Machine Learning for Placement and Routing in Chip Design: a Methodological Overview. ArXiv, abs/2202.13564.
- [9] L. Wang and M. Luo, "Machine Learning Applications and Opportunities in IC Design Flow," 2019 International Symposium on VLSI Design, Automation and Test (VLSI-DAT), 2019, pp. 1-3, doi: 10.1109/VLSI-DAT.2019.8742073.
- [10] L. Bai and L. Chen, "Machine-Learning-Based Early-Stage Timing Prediction in SoC Physical Design," 2018 14th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), 2018, pp. 1-3, doi: 10.1109/ICSICT.2018.8565778.
- [17] M. A. El Mrabet, K. El Makkaoui and A. Faize, "Supervised Machine Learning: A Survey," 2021 4th International Conference on Advanced Communication Technologies and Networking (CommNet), 2021, pp. 1-10, doi: 10.1109/CommNet52204.2021.9641998
- [18] D. Kinaneva, G. Hristov, P. Kyuchukov, G. Georgiev, P. Zahariev and R. Daskalov, "Machine Learning Algorithms for Regression Analysis and Predictions of Numerical Data," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-6, doi: 10.1109/HORA52670.2021.9461298.
- [19] Popescu, M., Balas, V.E., Perescu-Popescu, L., & Mastorakis, N.E. (2009). Multilayer perceptron and neural networks. WSEAS Transactions on Circuits and Systems archive, 8, 579-588.
- [20] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.