

DOI: 10.15514/ISPRAS-2023-35(5)-9



Быстрый анализ статического IR drop эффекта на базе методов машинного обучения

¹ Р.А. Соловьёв, ORCID: 0000-0003-0312-452X <roman.solovyev.zf@gmail.com >

¹ Д.В. Тельпухов, ORCID: 0000-0001-9551-0748 <telpukhov@ippm.ru >

² Е.Д. Демидов, ORCID: 0009-0009-0253-0228 <yevgeniy_demidov_1999@mail.ru>

¹ И.И. Шафеев, ORCID: 0009-0005-2323-3725 <ilya.tanc@gmail.com>

¹ Институт проблем проектирования в микроэлектронике РАН (ИППМ РАН),
124365, Россия, г. Москва, г. Зеленоград, ул. Советская, д.3.

² Национальный исследовательский университет «МИЭТ»,
124498, Россия, г. Москва, г. Зеленоград, пл. Шокина, д. 1.

Аннотация. В статье рассматривается решение проблемы быстрого проведения статического анализа падения напряжения с использованием нейронной сети. Рассматривается методика генерации базы данных необходимой для обучения ML-модели. Описывается методика обучения ML-модели для анализа статического IR drop эффекта. Описывается алгоритм получения входных данных для обучения нейронной сети из SPICE представления. Предложенное решение задачи попало в ТОП3 конкурса ICCAD Contest 2023.

Ключевые слова: IR drop эффект; падение напряжения; ML-модель; нейронная сеть; машинное обучение; база данных.

Для цитирования: Соловьёв Р.А., Тельпухов Д.В., Демидов Е.Д., Шафеев И.И. Быстрый анализ статического IR drop эффекта на базе методов машинного обучения. Труды ИСП РАН, том 35, вып. 5, 2023 г., стр. 127–144. DOI: 10.15514/ISPRAS–2023–35(5)–9.

Fast Analysis of Static IR Drop Effect Based on Machine Learning Methods

¹ R.A. Solovyev, ORCID: 0000-0003-0312-452X <roman.solovyev.zf@gmail.com>

¹ D.V. Telpukhov, ORCID: 0000-0001-9551-0748 <telpukhov@ippm.ru >

² E.D. Demidov, ORCID: 0009-0009-0253-0228 <yevgeniy_demidov_1999@mail.ru>

¹ I.I. Shafeev, ORCID: 0009-0005-2323-3725 <ilya.tanc@gmail.com>

¹ Institute of Design Problems in Microelectronics, RAS (IPPM RAS),
3, Sovetskaya st., Moscow, Zelenograd, 124365, Russia.

² National Research University of Electronic Technology (MIET),
1, Shokina Square, Moscow, Zelenograd, 124498, Russia.

Abstract. As part of the ICCAD Contest 2023 (Problem C) competition, the paper describes a methodology for applying ML models to perform static IR drop analysis. Methods for obtaining a database for training a neural network to solve this problem are given. We consider a technique for training an ML model to analyze the static IR-drop effect. The generation of input data for training a neural network from SPICE netlists is also discussed in this paper. This solution is ranked in the TOP 3 at the ICCAD Contest 2023 competition.

Keywords: IR drop analysis; ML-model; neural network; machine learning; database.

For citation: Solovyev R.A., Telpukhov D.V., Demidov E.D., Shafeev I.I. Fast analysis of static IR drop effect based on machine learning methods. *Trudy ISP RAN/Proc. ISP RAS*, vol. 35, issue 5, 2023. pp. 127-144 (in Russian). DOI: 10.15514/ISPRAS-2023-35(5)-9.

1. Введение

С уменьшением технологических норм проектирования увеличивается влияние различных негативных эффектов на разработанные микроэлектронные схемы. Одним из таких эффектов является IR drop (падение напряжения на шинах питания). IR drop эффект – это эффект, характеризующийся возникновением разности электрических потенциалов между двумя контактами проводника во время прохождения тока [1]. Эта разница потенциалов возникает за счёт падения напряжения при прохождении тока через участок металлической шины, имеющей сопротивление. Влияние на схему, которое оказывается IR drop эффектом, может привести к снижению производительности схемы из-за увеличения времени задержки переключения стандартных ячеек. Эффект падения напряжения может привести к функциональному сбою из-за нарушений синхронизации схемы, а именно времени установки и удержания [2].

В современном маршруте проектирования выделяют два типа IR drop эффекта: статический и динамический [3]. Статический IR drop представляет собой среднее падение напряжения в схеме, характеризующееся средним значением протекающего тока, который зависит от величины периода работы схемы. Основной вклад в статический IR drop вносит ток утечки в статическом состоянии ячеек. Динамическое падение напряжения по большей мере характеризуется высокой активностью переключения транзисторов и в меньшей степени зависит от величины периода.

Для упрощения анализа схемы используется метод статического анализа IR drop эффекта, который учитывает среднее количество переключений транзисторов за период тактового сигнала. Такой метод является менее точным, чем динамический, но при этом значительно сокращается время проведения анализа. Для проведения статического анализа падения напряжения топология схемы представляется в виде упрощённой эквивалентной модели схемы (рис. 1), которая включает в себя представление шин питания и земли в виде сетки резисторов, а также представление стандартных ячеек в виде источников постоянного тока.

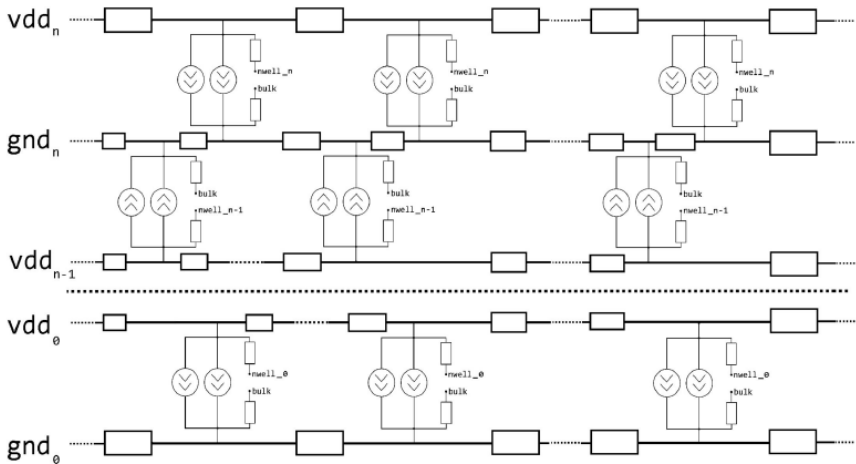


Рис. 1. Эквивалентная модель схемы.
Fig. 1. Equivalent circuit model.

Основные этапы построения эквивалентной модели и выполнения статического IR drop анализа:

- получение сетки резисторов, эквивалентной металлическим соединениям в схеме;
- вычисление средних значений потребляемой мощности каждой стандартной ячейки для получения источников постоянного тока и составления эквивалентной модели схемы;
- вычисление значений тока на каждом участке цепи путем решения системы линейных уравнений;
- вычисление значений напряжения на каждом участке цепи:

$$V = I * R$$

Целью моделирования эквивалентной модели схемы является вычисление напряжений во всех узлах сетки шин питания. Традиционно к решению этой задачи подходят с использованием метода решения системы линейных уравнений вида:

$$G * V = J, \text{ где}$$

G – матрица проводимостей;

V – вычисляемое значение напряжений в каждом узле;

J – матрица токов.

Проблема подобного подхода заключается во времени проведения расчётов, которое стремительно возрастает с увеличением узлов сетки питания в схеме. Одним из способов обойти вычислительную проблему при анализе статического IR drop эффекта является использование методов машинного обучения.

Ранее модели машинного обучения уже пытались применить в качестве средства для статического анализа падения напряжения [4, 5, 6, 7, 8]. В процессе исследования ранее предлагалось решение, определяющее статическое падение напряжения на всём кристалле на основе карты распределения мощности по сетке межсоединений и сетке шин питания [4]. Была разработана ML-модель, позволяющая быстро исследовать анализ падения напряжения, в несколько раз уменьшая время анализа по сравнению с коммерческими САПР [6]. Серьезной проблемой применения моделей машинного обучения является генерация базы данных для её обучения. Для полноценного обучения ML-модели необходимо анализировать все возможные варианты исследуемых схем. Отличия в исследуемых схемах могут начинаться от технологии, на которой они проектируются, и заканчиваться размерами и вариантами построения резистивной сети питания (power delivery network (PDN)).

В данной работе рассматривается:

- 1) архитектура и методика обучения нейронной сети;
- 2) методика сбора обучающих данных для увеличения датасета;
- 3) способ генерации двумерных данных из начального SPICE (Simulation Program with Integrated Circuit Emphasis) описания для шин питания;
- 4) сравнение скорости и точности работы предложенного метода в сравнении со SPICE моделированием.

Предложенное решение попало в ТОП3 конкурса ICCAD Contest 2023 (Problem C).

2. Конкурс ICCAD Contest 2023 (Problem C)

В 2023 году в рамках конференции ICCAD [9] проводился конкурс ICCAD Contest 2023 [10]. Конкурс проводится ежегодно на протяжении многих лет и на нем предлагаются для решения современные проблемы микроэлектроники и анализируются предлагаемые методы их решения. В качестве одной из трех задач этого года организаторы предложили для решения задачу по анализу статического эффекта падения напряжения на базе методов машинного обучения.

Глобальной задачей была разработка модели машинного обучения, которая позволяет за небольшое время определять статический IR drop эффект по всей площади схемы. Входными данными для обучения ML-модели выступают четыре матрицы для каждой схемы: матрица

источников тока, матрица эффективных расстояний источников напряжения, матрица плотности резистивной сетки и матрица значений IR drop. Подробное описание каждой из матриц будет рассмотрено далее.

В качестве обучающих данных организаторами конкурса были даны 10 реальных схем и 100 сгенерированных схем, информация о каждой из которых была представлена в виде четырех файлов: `current_map.csv`, `eff_dist_map.csv`, `pdn_density_map.csv`, `ir_drop_map.csv`. Также для каждой схемы было дано описание сети питания в SPICE формате.

В ходе предварительного исследования было обнаружено, что такого количества схем недостаточно для полноценного обучения ML-модели, поэтому в рамках работы был реализован маршрут для генерации дополнительных данных для обучения, как искусственно сгенерированных из представленных десяти схем, так и синтезированных из описаний RTL (Register Transfer Level) с помощью автоматизированного маршрута OpenROAD.

3. Предсказание результатов IR drop анализа с помощью методов машинного обучения

Распределение значений статического падения напряжения на цепях питания в микросхеме зависит от следующих трех факторов:

- расположение/распределение всех источников напряжения (силовых площадок) в схеме;
- топология сети земли и питания, значения сопротивлений каждого резистора (перехода/металлического слоя) в сети;
- распределение источников тока в схеме.

Каждый из обозначенных факторов связан с топологией и, следовательно, может быть представлен в виде изображения. Финальный результат IR drop анализа также является изображением, что позволяет использовать богатый опыт разработки различных моделей машинного обучения и нейронных сетей, работающих с изображениями. В работе был предложен маршрут для обучения и применения (инференса) моделей для анализа падения напряжения на цепях земли и питания (рис. 2).

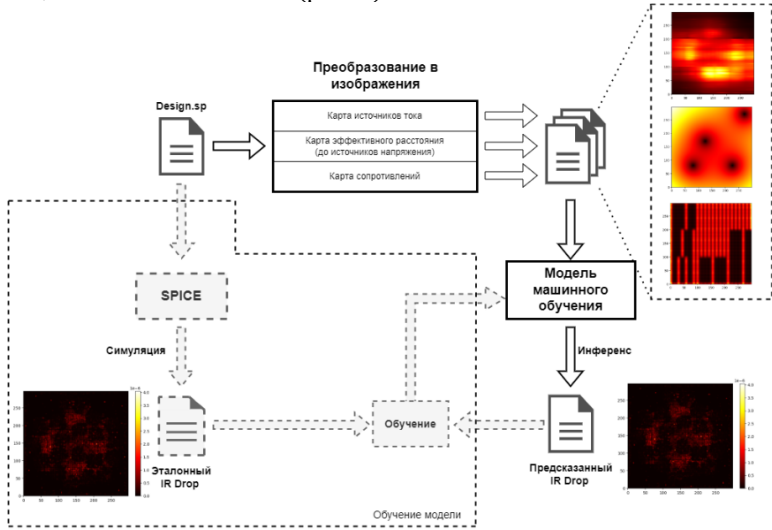


Рис. 2. Предложенный маршрут обучения и применения (инференса) моделей для предсказания результатов IR Drop анализа.

Fig. 2. A proposed route for training and applying (inferential) models for predicting IR Drop analysis results.

На вход разработанного маршрута поступает файл с описанием сети питания, который представляет собой SPICE файл с источниками внешнего напряжения, источниками тока, имитирующими стандартные ячейки и резисторами, которые определяют значения сопротивлений металлических проводников. В названиях цепей закодированы их координаты в топологии схемы, что позволяет полноценно воссоздать топологию сети питания (рис. 3).

```
...
R0 n1_m1_0_0 n1_m1_4000_0 4.463529
R1 n1_m1_4800_0 n1_m1_4000_0 0.892706
R2 n1_m1_4800_0 n1_m1_9600_0 5.356235
...
V0 n1_m9_160800_160800 0 1.100000
V1 n1_m9_452000_160800 0 1.100000
V2 n1_m9_250400_340000 0 1.100000
...
I0 n1_m1_9600_9600 0 2.140000e-08
I1 n1_m1_9600_28800 0 4.397143e-08
I2 n1_m1_9600_33600 0 2.539780e-08
...
```

Рис. 3. Пример SPICE файла с описанием сети питания.
Fig. 3. Example SPICE file with power delivery network description.

Подаваемый на вход файл поступает на конвертер, который должен сформировать изображения, отражающие все три фактора, обозначенные в начале данного раздела. После этого, изображения поступают на вход предсказательной модели, которая выдаёт карту падений напряжения на цепях питания. Обучение модели сводится к тому, чтобы сгенерировать датасет, в котором в качестве эталонных значений карт IR drop будут выступать результаты SPICE симуляции большого набора схем, после чего модель будет учиться предсказывать значения IR drop во всех узлах схемы (рис. 2).

Подобный подход, связанный с конвертацией сети питания в изображения, уже был предложен в научных работах [4-8]. Например, в [4] три вышеуказанные характеристики представлены в виде трех распределений, где распределение источников тока моделируется как карта токов, топология металлических проводников моделируется как функция плотности (расстояние между полосами питания на единицу площади) в различных регионах, а распределение источников напряжения моделируется как карта эффективных расстояний, которая представляет собой расстояние от каждого узла схемы до узла с источником напряжения. В статье предлагалось использовать свёрточные нейронные сети и UNet-сети [11], позволяющие выполнять предсказание IR drop с помощью модели, обученной на разнообразном наборе таких данных (три входа и один выход).

В работе предлагается схожий подход, отличающийся архитектурой нейронной сети, преобразователями SPICE файла в изображения, и некоторыми аспектами постобработки результатов нейронной сети. В следующем разделе описаны принципы генерации изображений для обучения и инференса нейронной сети из SPICE файла, описывающего сеть питания чипа.

4. Конвертация SPICE описания сети питания в набор изображений

Все три предоставленных входных изображения для каждой схемы, как и выходной IR drop, имеют одинаковое разрешение, которое пропорционально реальному размеру кристалла. В конкурсе, размер изображения в пикселях равен размеру схемы в микрометрах. Один пиксель

отражает общую площадь кристалла в один квадратный микрометр. Однако, в предоставленном SPICE файле «рис. 3», содержащем информацию о сопротивлениях, источниках напряжения и источниках тока, координаты цепей масштабированы с использованием коэффициента 2000. Этот коэффициент необходимо учитывать при конвертации. Иными словами, один пиксель итоговых изображений содержит 2000 условных единиц по оси x и 2000 условных единиц по оси y .

4.1. Получение карты эффективного расстояния

Способ получения карты эффективного расстояния заключается в отображении реальной матрицы с источниками напряжения в матрицу, которая не только будет сжата по размерам, но также будет показывать, насколько далеко или близко находится каждая ячейка матрицы, это позволит улучшить понимание влияния источников на все области PDN. Сделать это можно, используя следующий алгоритм:

1. Изначально выбирается шаг отображения по осям X, Y – $stepX$ и $stepY$ соответственно. Зная размер исходной реальной матрицы, можно вычислить размеры сжатой матрицы:

$$w' = \left(\frac{w}{stepX} \right) + 1, \quad h' = \left(\frac{h}{stepY} \right) + 1$$

2. Выполняется проход по всем ячейкам матрицы размера (w', h') . Для каждой ячейки вычисляется обратная сумма евклидова расстояния до всех источников напряжения:

$$S = \sum_0^n \frac{1}{\sqrt{(x_n - x)^2 + (y_n - y)^2}}, \text{ где}$$

n – количество источников напряжения;

x, y – координаты рассматриваемой ячейки;

x_n, y_n – координаты каждого источника напряжения в схеме.

3. После вычисления суммы, каждой ячейке присваивается значение эффективного расстояния:

$$cell = S * \frac{(stepX + stepY)}{2},$$

следовательно, если точка близко к источнику напряжения, ее значение в матрице будет высоким, и наоборот.

4.2. Получение карты сопротивлений сети питания

В рамках конкурса для представления данных о сети земли и питания было предложено использовать карту плотности PDN, которая отражает топологию сети питания и значения сопротивления каждого резистора в этой сети. Она моделируется как функция плотности (расстояние между полосами питания на единицу площади) сетки питания в разных областях чипа.

Однако в нашей работе было принято решение использовать саму топологию сети питания. Получаемая карта непосредственно отражает распределение сопротивлений металлических проводников цепи питания по площади чипа. Это позволяет более точно моделировать IR drop и использовать модель на очень нерегулярных сетях питания. Отличие данной карты от карты плотности PDN изображено на рис. 4.

Основные шаги получения карты сопротивлений PDN:

1. Извлечение координат $x1, y1, x2, y2$ узлов и значений сопротивления для всех резисторов из SPICE файла с описанием схемы.

2. Создание пустой матрицы размером $x_width \times y_width$, где x_width и y_width вычисляются исходя из размеров ядра и шага сетки $stepX$ и $stepY$, равного 2000.
3. Для каждого резистора добавляется его значение в ячейки матрицы, соответствующие координатам концов резистора, деленным на шаг сетки.
4. Применение размывки по Гауссу к матрице для сглаживания изображения.

Таким образом, в результате получается карта топологии PDN, отражающая распределение резисторов по площади чипа. Эта карта используется как один из входных признаков для обучения ML-модели.

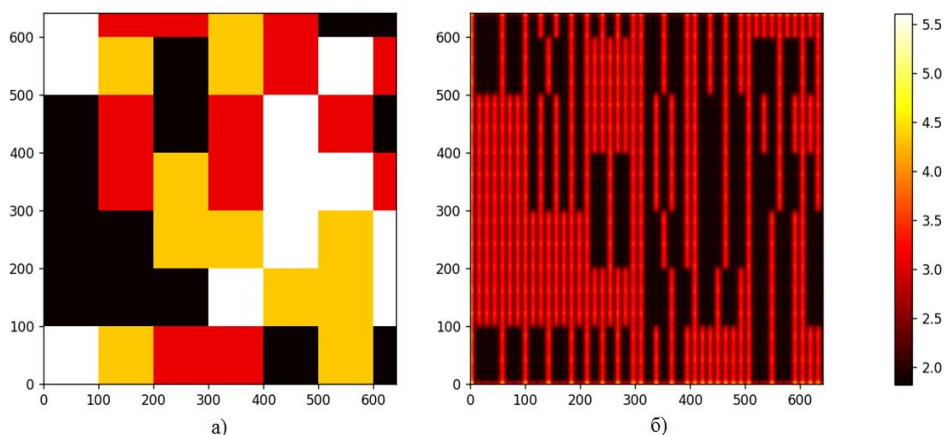


Рис. 4. Карта сети питания (карта плотности (а), карта сопротивлений (б)).

Fig. 4. Power network map (density map (a), resistance map (b)).

4.3. Получение карты источников тока

Основная идея получения карты источников тока заключается в отображении реальной матрицы с источниками тока в некоторую сжатую версию, при этом требуется минимизировать количество возможной потерянной информации. Сделать это можно, используя следующий алгоритм:

1. Изначально выбирается шаг отображения по осям x, y — $stepX$ и $stepY$ соответственно. Зная размер исходной реальной матрицы, можно вычислить размеры сжатой матрицы:

$$w' = \left(\frac{w}{stepX} \right) + 1, \quad h' = \left(\frac{h}{stepY} \right) + 1$$

2. Следующим шагом является обход реальной матрицы по ячейкам, которые хранят в себе информацию об источниках тока. Сначала вычисляется положение каждого источника тока в сжатой матрице:

$$x' = \left(\frac{x}{stepX} \right), \quad y' = \left(\frac{y}{stepY} \right)$$

В случае, когда координаты источника тока в сжатом матричном пространстве соответствуют координатам ячейки, расположенной на границе этой матрицы, происходит процедура аддитивной интеграции: значение тока, генерируемое источником, суммируется с существующим значением в указанной ячейке. В другом случае, когда позиция источника тока находится внутри матрицы и не прилегает к её границам, возникает необходимость в анализе диффузии токового значения от

ячейки (x', y') к соседним ячейкам в пределах заданного диапазона в пределах одной ячейки.

3. Для определения влияния текущего источника тока на окружающие его ячейки, нужно рассмотреть области пересечения ячейки с центром в точке $(xstepX - 0.5, ystepY - 0.5)$ и размерами $(1, 1)$ с ячейками, окружающими ячейку (x', y') в радиусе одной ячейки. Значение площади пересечения ячеек затем умножается на числовое значение источника тока и складывается с пересекающейся ячейкой.
4. Полученная сжатая матрица затем проходит через фильтр Гаусса для достижения более однородного распределения значений в пустых ячейках.

5. Метрики качества

В рамках конкурса было предложено 2 метрики качества решения: MAE (Mean Absolute Error) и F1-score. Наиболее важной при оценке качества прогнозирования IR drop считалась метрика MAE, её вклад в общую метрику качества составлял 60%. Метрика F1 давала вклад 30%, а оставшиеся 10% вносила скорость вычислений. Метрика MAE оценивает общее качество решения на всей схеме. Метрика F1 – оценивает качество только на критических участках схемы с максимальным значением IR drop. Проектировщикам микросхем важно знать координаты на топологии, где падение напряжения максимально. Именно в критичных местах будет необходимо вносить изменения в структуру схемы, чтобы решить проблему большого падения напряжения. Все остальные места схемы с минимальным значением IR drop и качество решения там играют гораздо меньшую роль.

5.1. Метрика MAE

MAE или, по-другому, средняя абсолютная ошибка [12] – это среднее значение абсолютной разницы между прогнозом и фактическим значением, рассчитанное для каждого примера в наборе данных. Цель состоит в том, чтобы минимизировать метрику MAE.

Пример: возьмем матрицу 2x2 для наглядности. Предположим, что мы получили значения IR drop:

2.1	6.4
3.8	0.1

При этом реальные значения IR drop:

2.5	5.3
4.9	0.0

Тогда матрица MAE будет выглядеть следующим образом:

0.4	1.1
1.1	0.1

Метрика MAE после усреднения будет равна:

$$MAE = 0.675$$

Это значение средней ошибки по всей матрице IR drop.

5.2. Метрика F1

F1 – это метрика бинарной классификации [13]. В этой задаче она использует 10 % максимального IR drop каждой схемы в качестве порога для выполнения классификации. Формулы для вычисления метрики:

$$F1 = 2 * Precision * \frac{Recall}{Precision + Recall}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}, \text{ где}$$

TP (True positive) – количество совпадающих значений «1» между предсказанной матрицей и реальной;

FP (False positive) – количество несовпадающих значений «1» между предсказанной матрицей и реальной;

TN (True negative) – количество совпадающих значений «0» между предсказанной матрицей и реальной;

FN (False negative) – количество несовпадающих значений «0» между предсказанной матрицей и реальной.

На карте IR drop значением 1 (True) отмечаются участки, на которых IR drop выше, чем 90% от максимального для заданной схемы. Ставится задача найти такие участки как можно точнее. Чем точнее результат, тем выше будет метрика $F1$.

Рассмотрим пример из предыдущего раздела. Максимальное значение IR drop для матрицы равно 5.3. Если мы возьмем 90% от него, получим 4.77. Таким образом, все значения больше, чем 4.77 и в предсказанной матрице, и в реальной матрице станут равными 1, а остальные значения станут равными 0.

Таким образом, матрица горячих точек для предсказанной матрицы будет равна:

$$\begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}$$

А для реальной матрицы:

$$\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}$$

Таким образом:

$$TP = 1, TN = 2, FP = 0, FN = 1$$

По формулам можно рассчитать:

$$Precision = 1 / (1 + 0) = 1,$$

$$Recall = 1 / (1 + 1) = 0.5, F1 = 2 * 1 * 0.5 / (1 + 0.5) = 0.67$$

6. Генерация датасета для обучения модели

Основная проблема, с которой сталкиваются разработчики нейросетевых моделей для решения задач в области физического синтеза – это отсутствие достаточного количества данных для обучения. Такая же проблема наблюдается и в данной задаче. Авторы конкурса предоставили в открытом доступе 100 сгенерированных схем для обучения и 10 реальных схем для отладки предлагаемого решения, а также ещё 10 реальных схем организаторы конкурса оставили закрытыми для внутреннего тестирования решений. Закрытые схемы нужны, чтобы быть уверенным, что разработанные модели не обучались на реальных схемах. Эти закрытые схемы использовались, чтобы сформировать финальную таблицу лидеров конкурса. Десяти реальных схем и ста сгенерированных схем явно недостаточно для обучения хорошей модели. Поэтому потребовалось искать и готовить дополнительные данные для обучения. Было обнаружено, что тестовые схемы были синтезированы с помощью технологии Nangate45 [14], поскольку сгенерированная сетка питания на данной технологии была схожа с предоставленными данными от организаторов. Организаторы рекомендовали использовать для обучения схемы, сгенерированные с помощью генеративной нейронной сети BeGAN [15-16]. В работе были сгенерированы новые схемы из уже существующих, а также синтезированы топологии нескольких реальных схем на

технологии Nangate45 с помощью автоматизированного маршрута физического синтеза OpenROAD [17-18].

6.1. Алгоритм получения карты падения напряжения

Результатом проведения этапа физического синтеза является топологическое представление схемы, представленное в общепринятом формате DEF (Design Exchange Format). Далее с помощью программы PDNSim, входящей в состав OpenROAD, генерировался SPICE нетлист. Моделируя нетлист с помощью SPICE симулятора, были получены значения напряжений в каждом узле резистивной сетки, которые в конечном итоге конвертировались в эталонные значения падения напряжения в каждом узле.

6.2. Алгоритм генерации новых схем из заданных

Для расширения набора данных для обучения на основе уже имеющихся схем, было использовано несколько подходов:

1. Перемещение источников тока на новые позиции.
2. Увеличение количества источников тока.
3. Изменение числового значения источников тока.

Для генерации новых схем, описанные выше подходы, применялись последовательно. Если все возможные перемещения источников тока уже были выполнены, начинается добавление новых источников. Как только все узлы на нижнем слое металлизации подключены к источникам тока, начинается изменение их значений до достижения сопоставимого с другими схемами по порядку значения падения напряжения.

Все модификации схемы происходят на основе входного SPICE описания, выходным форматом также является SPICE. В процессе генерации для расчета значений падения напряжения используются стандартные методы решения в схемотехнике – метод узловых потенциалов (МУП) [19].

Главное преимущество этого метода заключается в том, что он позволяет генерировать множество новых вариантов схем, которые также можно решить по методу узловых потенциалов, при этом, не изменяя резистивную сетку.

6.3. Получение новых тестов на реальных схемах

Для увеличения набора данных, необходимых для обучения ML-модели, также было принято решение о генерации дополнительного набора реальных схем. Разработка схем для обучения происходила на технологии Nangate45 [14].

Основными этапами в получении новых данных для обучения нейронной сети были:

1. Проведение физического синтеза схем, содержащихся в тестовом наборе для технологии Nangate45, а также схем из набора ISCAS85 [20] с помощью открытого маршрута проектирования OpenROAD.
2. Генерация множества вариантов реальных схем (функционирующих схем), на основе синтезированных, путём изменения резистивной сети питания по принципу удаления некоторых шин питания. Удаление некоторого ограниченного числа вертикальных и горизонтальных шин питания верхнего уровня в схеме не приведёт к функциональным сбоям, что позволяет использовать данную схему для обучения.
3. Генерация случайного числа источников напряжения в схеме с расположением их в допустимых координатах на топологии, что также позволяет считать новые схемы реальными.

4. Получение SPICE нетлистов для сгенерированного набора схем и их моделирование для получения тепловых карт со значениями IR drop в каждом узле схемы в виде «.csv» файла: ir_drop_map.csv.

На примере одной из схем набора ISCAS85 на рис. 5 показаны топологические представления до модернизации сетки питания и после, а на рис. 6 показаны соответствующие им карты падения напряжения.

Таким образом, в результате генерации данных были получены наборы схем, где каждая схема описывается в виде четырех «.csv» файлов, необходимых для обучения ML-модели нейронной сети.

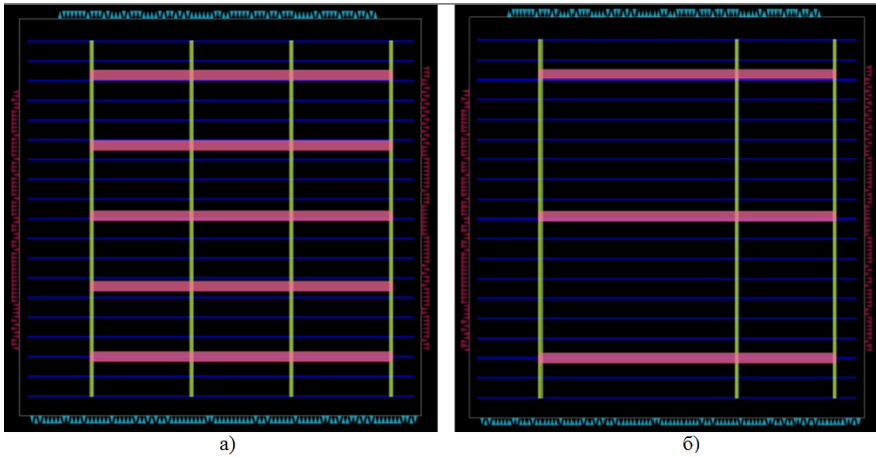


Рис. 5. Топологическое представление схемы c7552 из набора тестовых схем ISCAS85. Изображение до преобразований сети питания (а), изображение после удаления некоторых шин питания верхнего уровня (б).

Fig. 5. Topological representation of the c7552 circuit from the ISCAS85 test circuit set. Image before power supply network conversion (a), image after removing some of the top-level power supply rails(b).

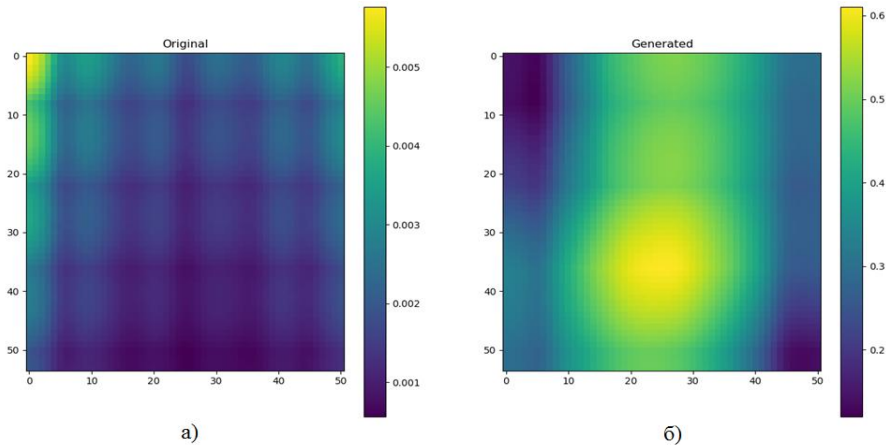


Рис. 6. IR drop для оригинальной сети питания схемы c7552 (а) и для модифицированной сети питания схемы c7552 (б).

Fig.6. IR drop for the original c7552 circuit power supply network (a) and for the modified c7552 circuit power supply network (b).

7. Нейросетевая модель

Изначально для решения задачи необходимо было определиться с архитектурой нейронной сети. Были протестированы Unet [11] и FPN [21] архитектуры с разными кодировщиками. В качестве кодировщиков использовали свёрточные нейронные сети из семейства EfficientNet [22], а также Multi-Axis Vision Transformers – MaxVit [23]. Выбор этих архитектур определялся доступностью кода для обучения, высокими метриками при решении задач классификации и предыдущим опытом успешного использования при решении задач сегментации. Сравнение на валидационном наборе данных показало, что MaxVit имеет выше метрику MAE, чем EfficientNetB5. Также требовалось выбрать размер входа для нейронной сети. Изменение размеров исходных изображений приводит к потере информации, поэтому изображения должны подаваться пиксель в пиксель. Но не обязательно подавать всё изображение, а достаточно подавать какую-то небольшую часть. Максимальный размер схемы для обучения в пикселях был 930 X 930. Современные нейронные сети для классификации изображений на базе трансформеров редко принимают на вход изображения более 512 X 512 пикселей, а типовой размер входа обычно 224 X 224 пикселей. Поэтому мы пробовали разные размеры изображений от 128 X 128 до 512 X 512. В наших экспериментах максимальное качество достигалось на больших размерах. По всей видимости, это связано с тем, что таким образом нейронная сеть имеет больше информации для предсказания значений IR drop. Если изображение меньше указанного размера, оно дополняется нулями до заданного размера.

В качестве аугментаций данных использовалось:

- 1) выбор случайного патча нужного размера из изображения;
- 2) случайный поворот на 0/90/180/270 градусов;
- 3) случайное отражение по вертикали/горизонтали.

На этапе использования (инференса) при анализе заданной схемы, она сначала преобразовывалась в изображение и затем методом плавающего окна анализировалась на предмет значения IR drop. Окно имело размер 512 X 512 пикселей, то есть равное входу нейронной сети, шаг можно было варьировать. При этом, чем меньше шаг, тем выше точность, но ниже скорость работы. Также для улучшения точности был использован метод Test time Augmentation [24]. Изображение вращается на 90 градусов и отражается во всех 8 возможных положениях, и результат усредняется. Можно использовать только часть положений из 8, что тоже дает возможность находить некоторый баланс между точностью и скоростью работы.

Во время обучения мы столкнулись с проблемой, что метрика F1 для тестовых примеров иногда равна нулю, но при этом значение метрики MAE низкое. Это связано с особенностью расчёта метрики F1. Если значение 90% от максимума равно N, то если мы предскажем значение меньше N, то получим False Negative, а если больше N, то True Positive и тем самым улучшим метрику F1. Однако с точки зрения MAE метрики ничего не поменяется. Для больших значений IR drop выгоднее предсказывать значение IR drop с некоторым запасом в большую сторону и избегать заниженных значений IR drop. Поэтому loss-функция была модифицирована следующим образом:

если $pred < true$, то $loss = K * MAE$, если $pred \geq true$, то $loss = MAE$,

коэффициент K в нашем случае был равен 2.

Дополнительно на этапе инференса значения более чем 90% от максимального IR drop домножались на поправочный коэффициент 1.05. Это также позволило улучшить метрику F1 на валидации и при этом почти не ухудшить метрику MAE.

Изначально мы столкнулись с проблемой, что обучение не сходится, если подавать значения в абсолютных величинах. По всей видимости, это связано с тем, что погрешность вычислений начинает превышать искомые значения или же с использованием смешанной точности (mixed

precision) во время обучения. Поэтому мы сделали масштабирование значений как входных данных, так и искомого IR drop: источники тока домножались на 10^8 , карта эффективных расстояний на 100 и карта плотности PDN на 100. Значения IR drop домножались на 10^5 .

8. Результаты

Для тестирования скорости и точности работы были использованы данные, предоставленные организаторами ICCAD 2023. Это 10 реальных схем (см. табл. 1). Сравнивалась скорость и точность работы SPICE симулятора NGSPICE [25] и предложенной нейросетевой модели.

Табл. 1. Размеры тестовых схем

Table 1. Test circuit sizes

Название тестовой схемы	Размер, μm^2	Число резисторов
testcase1	298 X 298	24087
testcase2	298 X 298	22327
testcase3	930 X 930	228439
testcase4	930 X 930	214889
testcase5	641 X 641	110465
testcase6	641 X 641	102525
testcase11	204 X 204	10859
testcase12	204 X 204	10407
testcase17	566 X 566	84044
testcase18	566 X 566	80038

Поскольку обучение нейронной сети проводилось с использованием K-Fold валидации со значением параметра $k=5$, то по результатам было получено 5 моделей. Были предсказаны значения по схеме Out of Fold, то есть для каждой из пяти моделей были предсказаны значения для схем, которые не принимали участия в тренировке. Результаты тестирования приведены в табл. 2. Из приведённой выше таблицы, время расчёта падения напряжения с использованием нейронной сети даже для больших схем меньше секунды. Основное время занимает преобразование SPICE описания в 2D-матрицу. Эта операция может быть оптимизирована, что даст дополнительный прирост по скорости. В данной работе оптимизация не проводилась. Результаты сравнения реальной карты IR drop с предсказанной для нескольких тестовых схем приведены на рис. 7-9.

Предложенное решение попало в тройку лидеров конкурса ICCAD Contest 2023 (Problem C).

9. Выводы

Из результатов исследования видно, что метод использования свёрточных нейросетей для предсказания IR drop имеет высокую точность (средняя абсолютная ошибка имеет значения, не превышающие $15 \cdot 10^{-5}$). Для больших схем метод с использованием нейронных сетей в десятки раз (10-30 раз) превосходит по скорости классические методы на базе моделирования. С ростом размера схем отрыв по скорости увеличивается. Дополнительно

данный метод в будущем можно легко адаптировать для предсказания эффекта электромиграции. Однако существует ряд проблем, которые ещё нужно решить для применения метода на практике:

1. В настоящее время разработанная модель поддерживает только одну технологию – Nangate45 и скорее всего не справится с расчетом IR drop на других технологиях. Для решения этой проблемы требуется добавлять тестовые данные из других технологий. В этом случае может быть получено универсальное решение для этой задачи.
2. Требуется расширять публичный набор данных для решения этой задачи. Нужен либо набор реальных тестовых схем, либо сразу результаты физического синтеза (генерации PDN) для них в рамках различных технологий.
3. Как правило, современные СБИС состоят не только из стандартных ячеек, но и из множества макроблоков, соединяющихся между собой. В данном решении никак не учитывается возможность обработки схем, содержащих в себе макроблоки.
4. В данном решении поддерживается только один алгоритм генерации PDN, встроенный в маршрут OpenROAD.

Табл. 2. Результаты сравнения скорости NGSPICE и нейросетевой модели
Table 2. Results of speed comparison between NGSPICE and neural network model

Название тестовой схемы	Время, с			Ускорение по сравнению с NGSPICE (х раз)	Метрика MAE * 10^{-5}	Метрика F1
	NgSPICE	Преобразование SPICE нетлиста в матрицы	Расчёт нейросети			
testcase1	5.8	0.9	0.2	5.2	9.30	0.62
testcase2	4.7	0.9	0.1	4.7	14.31	0.87
testcase3	1230.3	45.4	0.8	26.6	4.84	0.75
testcase4	966.9	45.2	0.8	21	10.18	0.63
testcase5	170.3	11.9	0.4	13.8	6.10	0.81
testcase6	125.9	11.9	0.4	10.1	10.21	0.60
testcase11	1.1	0.5	0.4	1.2	10.59	0.52
testcase12	1.1	0.5	0.4	1.2	14.27	0.62
testcase17	83.6	7.3	0.6	10.5	3.41	0.69
testcase18	75.8	7.4	0.7	9.3	4.39	0.79

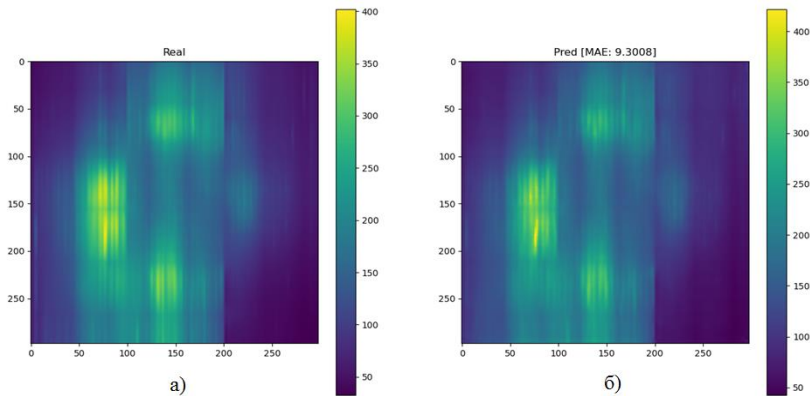


Рис. 7. Тестовая схема 1. Реальный IR drop (а), предсказанный IR drop (б)
(значения IR drop нужно домножить на 10^{-5}).

Fig. 7. Testcase 1. Real IR drop (a), predicted IR drop (b) (IR drop values need to be multiplied by 10^{-5}).

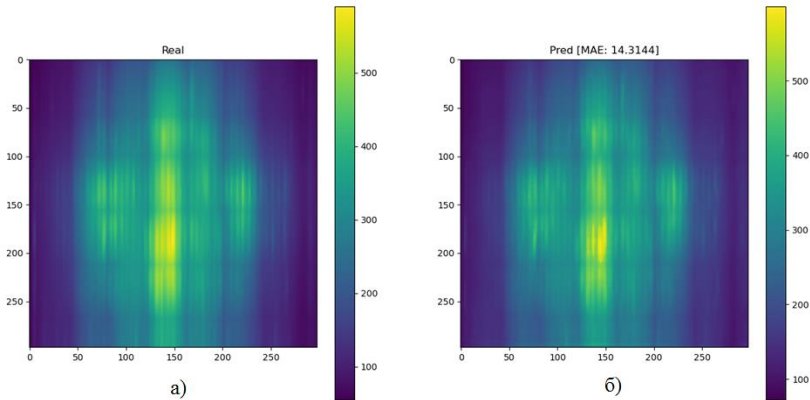


Рис. 8. Тестовая схема 2. Реальный IR drop (а), предсказанный IR drop (б)
(значения IR drop нужно домножить на 10^{-5}).

Fig. 8. Testcase 2. Real IR drop (a), predicted IR drop (b) (IR drop values need to be multiplied by 10^{-5}).

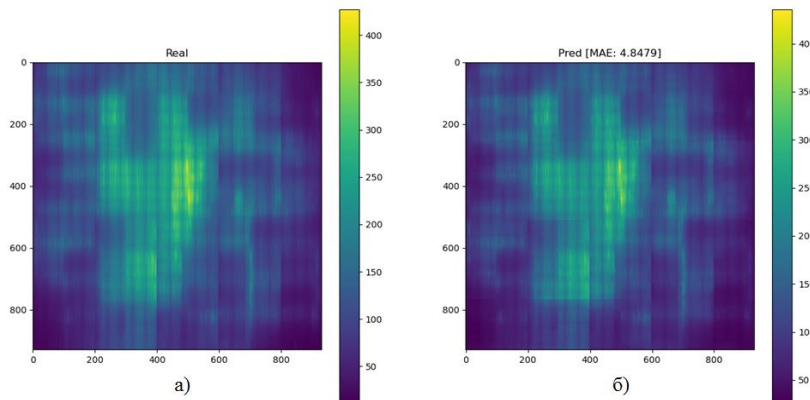


Рис. 9. Тестовая схема 3. Реальный IR drop (а), предсказанный IR drop (б)
(значения IR drop нужно домножить на 10^{-5}).

Fig. 9. Testcase 3. Real IR drop (a), predicted IR drop (b) (IR drop values need to be multiplied by 10^{-5}).

Список литературы / References

- [1]. Интернет-ресурс <https://siliconvlsi.com/what-is-ir-drop/> (дата обращения – 27.11.2023).
- [2]. Интернет-ресурс https://vlsi-backend-adventure.com/ir_analysis.html (дата обращения – 27.11.2023).
- [3]. Интернет-ресурс <https://teamvlsi.com/2020/07/ir-analysis-in-asic-design-effects-and.html> (дата обращения – 27.11.2023).
- [4]. Chhabria V. A., Ahuja V., Prabhu A., Patil N., Jain P., and Sapatnekar S. S. Thermal and IR Drop Analysis Using Convolutional Encoder-Decoder Networks. *Proceedings of Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2021.
- [5]. Chia-Tung Ho and Andrew B Kahng IncPIRD: Fast Learning Based Prediction of Incremental IR Drop. *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, 2019.
- [6]. Zhiyao Xie, Haoxing Ren, Brucek Khailany, Ye Sheng, Santosh Santosh, Jiang Hu, and Yiran Chen PowerNet: Transferable Dynamic IR Drop Estimation via Maximum Convolutional Neural Network. *Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2020.
- [7]. Chi-Hsien Pao, An-Yu Su, and Yu-Min Lee XGBIR: An xgboost-based IR drop predictor for power delivery network. *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 2020
- [8]. Chhabria V. A., Zhang Y., Ren H., Keller B., Khailany B., and Sapatnekar S. S. MAVIREC: ML-Aided Vectored IR-Drop Estimation and Classification. *Proceedings of Design, Automation, and Test in Europe (DATE)*, 2021.
- [9]. Интернет-ресурс <https://2023.iccad.com/> (дата обращения – 27.11.2023).
- [10]. Интернет-ресурс <http://iccad-contest.org/> (дата обращения – 27.11.2023).
- [11]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.
- [12]. Интернет-ресурс <https://skine.ru/articles/533011/> (дата обращения – 27.11.2023).
- [13]. Интернет-ресурс <https://academy.yandex.ru/handbook/ml/article/metriki-klassifikacii-i-regressii> (дата обращения – 27.11.2023).
- [14]. Интернет-ресурс <https://eda.ncsu.edu/freepdk/freepdk45/> (дата обращения – 27.11.2023).
- [15]. Интернет-ресурс <https://github.com/UMN-EDA/BeGAN-benchmarks> (дата обращения – 27.11.2023).
- [16]. Chhabria V. A. et al. BeGAN: Power grid benchmark generation using a process-portable GAN-based methodology. *2021 IEEE/ACM International Conference on Computer Aided Design (ICCAD)*. – IEEE, 2021. – С. 1-8.
- [17]. Интернет-ресурс <https://github.com/The-OpenROAD-Project> (дата обращения – 27.11.2023).
- [18]. Ajayi T., Blaauw D. Openroad: Toward a self-driving, open-source digital layout implementation tool chain. *Proceedings of Government Microcircuit Applications and Critical Technology Conference*. – 2019.
- [19]. Интернет-ресурс <https://electroandi.ru/toe/metod/metod-uzlovykh-potentsialov.html> (дата обращения 27.11.2023).
- [20]. Интернет-ресурс http://vscripts.ru/w/Схемы_ISCAS85 (дата обращения 27.11.2023).
- [21]. Kirillov, A., He, K., Girshick, R., & Dollár, P. (2017). A unified architecture for instance and semantic segmentation. In *CVPR*.
- [22]. Tan M., Le Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International conference on machine learning*. – PMLR, 2019. – С. 6105-6114.
- [23]. Tu, Z., Talebi, H., Zhang, H., Yang, F., Milanfar, P., Bovik, A., & Li, Y. (2022, October). Maxvit: Multi-axis vision transformer. *European conference on computer vision* (pp. 459-479). Cham: Springer Nature Switzerland.
- [24]. Shanmugam D. et al. Better aggregation in test-time augmentation // *Proceedings of the IEEE/CVF international conference on computer vision*. – 2021. – С. 1214-1223.
- [25]. Интернет-ресурс <https://ngspice.sourceforge.io/> (дата обращения 27.11.2023).

Информация об авторах / Information about authors

Роман Александрович СОЛОВЬЁВ – чл.-кор. РАН, доктор технических наук, главный научный сотрудник Института Проблем Проектирования в Микроэлектронике РАН. Сфера научных интересов: автоматизация проектирования цифровых СБИС, нейронные сети и машинное обучение, система остаточных классов, физический синтез интегральных схем.

Roman Aleksandrovich SOLOVYEV – Dr. Sci. (Tech.), member-corr. RAS, chief researcher at the Institute of Design Problems in Microelectronics RAS. Area of scientific interests: automation of digital VLSI design, neural networks and machine learning, residue number system, physical synthesis of integrated circuits.

Дмитрий Владимирович ТЕЛЬПУХОВ – доктор технических наук, заместитель директора по научной работе Института Проблем Проектирования в Микроэлектронике РАН. Сфера научных интересов: автоматизация проектирования цифровых СБИС, логический синтез, радиационно-стойкое проектирования, машинное обучение и нейронные сети, система остаточных классов.

Dmitry Vladimirovich TELPUKHOV – Dr. Sci. (Tech.), Deputy Director for Scientific Work of the Institute of Design Problems in Microelectronics RAS. Research interests: automation of digital VLSI design, logic synthesis, radiation-resistant design, machine learning and neural networks, residual class system.

Евгений Денисович ДЕМИДОВ – аспирант 2 курса «Национального исследовательского университета «МИЭТ», инженер-разработчик отдела разработки средств автоматизированного проектирования СБИС в компании ООО «Альфачип». Тема научно-исследовательской работы: «Разработка средств анализа IR drop эффекта с применением методов машинного обучения». Сфера научных интересов: автоматизация проектирования цифровых СБИС, специальные виды анализа СБИС, машинное обучение и нейронные сети.

Evgeny Denisovich DEMIDOV – 2nd year postgraduate student of National Research University "MIET", development engineer computer-aided design tools of VLSI development department at Alphachip Company. Theme of scientific research work: «Development of tools for analyzing IR drop effect using machine learning methods». Research interests: automation of digital VLSI design, special types of VLSI analysis, machine learning and neural networks.

Илья Ильич ШАФЕЕВ – студент 2 курса магистратуры «Национального исследовательского университета «МИЭТ», инженер-исследователь в Институте Проблем Проектирования в Микроэлектронике РАН. Сфера научных интересов: автоматизация проектирования цифровых СБИС, специальные виды анализа СБИС, машинное обучение и нейронные сети.

Ilya Ilyich SHAFEEV – 2nd year master's student of National Research University "MIET", research-engineer of the Institute of Design Problems in Microelectronics RAS. Research interests: automation of digital VLSI design, special types of VLSI analysis, machine learning and neural networks.

