

Нижегородский филиал
федерального государственного автономного образовательного учреждения
высшего образования «Национальный исследовательский университет
«Высшая школа экономики»

Международная конференция
«Машинное обучение и нелинейная динамика»

International Conference
«Machine Learning and Nonlinear Dynamics»

ML&ND'2025

Сборник трудов Международной научной конференции

15-18 октября 2025 года

Нижний Новгород
2025

УДК 517.9, 519.25, 530.182

ББК 16.63

М38

Машинное обучение и нелинейная динамика : Сборник трудов Международной
М38 научной конференции 15-18 октября 2025 года / Под ред. Н.В. Станкевич –
Н.Новгород: НИУ ВШЭ – Нижний Новгород, Саратов : Амирит, 2025. – 64 с.

ISBN 978-5-00279-008-1

В сборнике опубликованы материалы Международной научной конференции «Машинное обучение и нелинейная динамика», прошедшей в Нижегородском филиале Национального исследовательского университета «Высшая школа экономики» 15-18 октября 2025 года. Работы участников связаны с использованием методов машинного обучения для решения задач нелинейной динамики, применением методов нелинейной динамики для модернизации методов машинного обучения, а также с применением машинного обучения к прикладным задачам, в том числе задачам биологии и медицины. Для научных работников, преподавателей, аспирантов, студентов, специализирующихся в области нелинейной динамики, информационных технологий, машинного обучения, радиофизики, электроники, математическом моделировании.

Редакционная коллегия:

кандидат физ.-мат. наук Н.В. Станкевич (отв. редактор)

Рецензенты:

доктор физ.-мат. наук В.В. Клиньшов
(НИУ ВШЭ – Нижний Новгород)

кандидат физ.-мат. наук Н.И. Семенова
(Национальный исследовательский

Саратовский государственный университет им. Н.Г. Чернышевского)

Работа издана в авторской редакции

Конференция организована при финансовой поддержке
Национального исследовательского университета «Высшая школа экономики», программа
«Международное академическое сотрудничество»

УДК 517.9, 519.25, 530.182

ББК 16.63

© Авторы статей

© НИУ ВШЭ – Нижний Новгород, 2025

ISBN 978-5-00279-008-1



Подписано в печать 03.12.2025 г. Формат 60×90/16.

Гарнитура Times New Roman. Бумага офсетная.

Усл. печ. л. 4,13. Тираж 100 экз. Заказ № 6524-25.

Отпечатано в соответствии с предоставленными материалами
в ООО «Амирит», 410004, г. Саратов, ул. Чернышевского, 88.

Тел.: 8-800-700-86-33 | (845-2) 24-86-33 | Сайт: amirit.ru

Содержание

Секция 1 «Машинное обучение для решения задач нелинейной динамики».....5 Session 1 «Machine Learning for tasks of Nonlinear Dynamics»

Купцов П.В., Станкевич Н.В. Применение методов машинного обучения для реконструкции однопараметрического семейства динамических моделей по скалярному временному ряду.....	7
Каравасев А.С., Ишбулатов Ю.М., Боровкова Е.И., Безручко Б.П. Модели биосигналов для настройки параметров и верификации методов анализа данных.....	9
Бондаренко И.А., Гриднев В.В., Паленова А.В., Панюшев А.А., Посненкова О.М., Станкевич Н.В. Анализ фазовой структуры электрокардиограммы при нагрузке и в покое с использованием методов машинного обучения.....	11
Мяков Т.И., Станкевич Н.В. Влияние метода интегрирования дифференциального уравнения на этапе формирования обучающих датасетов на свойства нейроморфного отображения.....	14
Клименко К.Е., Оленин С.М., Леванова Т.А., Писарчик А.Н. Восстановление среднеполевой динамики нейрон-астроцитарного взаимодействия с использованием архитектуры dend-PLRNN.....	17
Panyushev A., Kuptsov P., Stankevich N. Machine learning for solving the problem of detecting multistable attractors.....	20
Gromov N., Levanova T., Smirnov L. On some properties of output matrices in reservoir computings.....	22
Багаутдинова Э.Р., Купцов П.В., Станкевич Н.В. Сравнительный анализ мультистабильных свойств модели типа Ходжкина-Хаксли и ее нейросетевого отображения.....	24

Секция 2 «Нелинейная динамика для задач машинного обучения».....27 Session 2 «Nonlinear Dynamics for Machine Learning »

Семенова Н.И., Колесников И.Д., Максимов Д.А., Москвитин В.М. Внутренние шумы в аппаратных нейронных сетях различных типов.....	29
Москвитин В.М., Семенова Н.И. Влияние внутреннего шумового воздействия на работу искусственной нейронной сети. Применение стратегии уменьшения шума.....	32
Kolesnikov I.D., Semenova N.I. Modeling and evaluation of the impact of additive and multiplicative noise on the performance of hardware neural networks.....	35
Соловьев И.А., Горюнов О.А., Клиньшов В.В. Обучение простой нейронной сети: компромисс между скоростью обучения и потреблением памяти.....	37

Секция 3 «Динамические системы».....41 Session 3 «Dynamical Systems »

Semenov V. Symmetry control for bistable systems: methods and applications.....	43
Gonchenko A. On discrete Lorez-like attractors in three-dimensional maps with axial symmetry	45
Шамсина А.Л., Станкевич Н.В. Мультистабильность в неавтономном генераторе квазипериодических колебаний.....	47

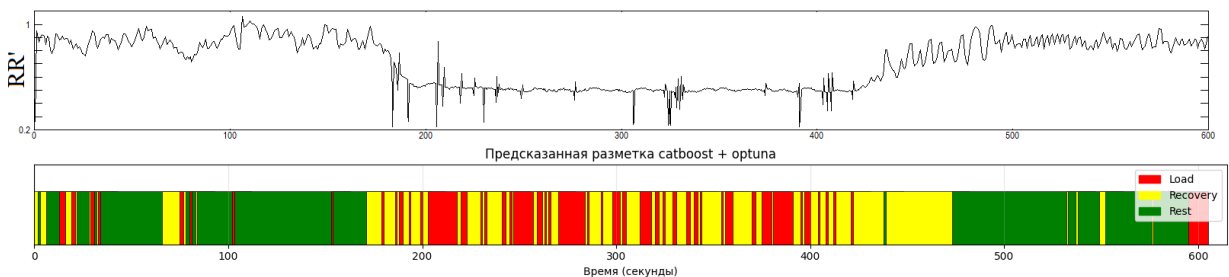
Gonchenko S., Safonov K. On mixed dynamics of reversible maps with symmetric homoclinic dynamics.....	49
--	----

Секция 4 « Машинное обучение в приложениях ».....	51
Session 4 « Machine learning in applications» - Workshop «AI in medicine»	

Семенова Н.И., Сергеев К.С., Слепнев А.В., Тучина Д.К., Тимошина П.А., Тучина В.В. Оценка гликирования тканей при диабете с использованием оптического просветления и нелинейной регрессии.....	53
Храбров К.А., Цыпин А.А., Кадулин А.А. Оптимизация с помощью графовых нейронных сетей.....	56
Смышляев А.П., Громов Н.В., Шарков А.А., Гребенюкова А.Д., Герасимова С.А., Лебедева А.В., Смиронов Л.А., Леванова Т.А., Писарчик А.Н. Детектирование сонных веретен у здоровых и пациентов с диагностированной эпилепсией с использованием алгоритмов искусственного интеллекта на основе гиперграфов.....	59
Программный комитет	62
Программа конференции	63

Секция 1

«Машинное обучение для решения задач нелинейной динамики»



Session 1

«Machine learning for tasks of nonlinear dynamics»

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РЕКОНСТРУКЦИИ ОДНОПАРАМЕТРИЧЕСКОГО СЕМЕЙСТВА ДИНАМИЧЕСКИХ МОДЕЛЕЙ ПО СКАЛЯРНОМУ ВРЕМЕННОМУ РЯДУ

Купцов Павел Владимирович
НИИ ВШЭ – Нижний Новгород
pkuptsov@hse.ru

Станкевич Наталия Владимировна
НИИ ВШЭ – Нижний Новгород
nstankevich@hse.ru

Аннотация. Исследование посвящено реконструкции динамических систем с помощью нейронных сетей на основе единственного скалярного временного ряда. Предложенный двухэтапный метод использует вариационный автоэнкодер и нейросетевое отображение для воссоздания семейства систем с одним управляющим параметром. Показано, что восстановленная модель способна воспроизводить различные динамические режимы исходной системы, в том числе и те, которые не наблюдались в процессе обучения.

Ключевые слова: нейросетевое отображение, реконструкция динамических систем, машинное обучение.

Важная характеристика нейронных сетей — свойство обобщения информации. Эта способность реализуется через выявление в данных скрытых закономерностей и взаимосвязей. Применительно к задаче реконструкции динамических систем, это свойство открывает значительные перспективы для построения моделей на основе экспериментальных данных [1, 2].

В настоящем исследовании анализируются возможности нейронных сетей по воспроизведению динамических режимов системы, соответствующих различным значениям управляющего параметра, при условии, что в распоряжении имеется лишь единственный скалярный временной ряд. Разработанная нейросетевая архитектура аппроксимирует целое семейство динамических систем с одним управляющим параметром. Это семейство адекватно отражает поведение исходной системы, воспроизводя её характерные режимы и бифуркационные переходы.

В качестве примера используется модель физиологического нейрона, построенная в рамках формализма Ходжкина-Хаксли и описываемая системой трёх обыкновенных дифференциальных уравнений. Исходные данные для реконструкции представляют собой одиночную запись значений одной из переменных указанной системы, полученную численным методом при фиксированных параметрах и постоянном временном шаге.

Процесс реконструкции осуществляется в два этапа. На первом этапе производится обучение вариационного автоэнкодера, задачей которого является преобразование исходного временного ряда в последовательность реконструированных векторов состояния. На втором этапе эти векторы используются для обучения модели, сформулированной в виде рекуррентного нейросетевого отображения. Обученное отображение функционирует как дискретная динамическая система с одним управляющим параметром. В работе показано, что реконструированная модель успешно воспроизводит

динамические режимы исходной системы, в том числе те, которые не были представлены в обучающей выборке.

Результаты получены в рамках проекта «Международное академическое сотрудничество» НИУ ВШЭ.

Библиографический список

1. *Купцов П. В., Станкевич Н. В. // Известия вузов. Прикладная нелинейная динамика, 2024 Т. 32, №1, С. 72-95.*
2. *Kuptsov P. V., Stankevich N. V. // Chaos, Solitons & Fractals. 2025, T. 191. С. 115818.*

МОДЕЛИ БИОСИГНАЛОВ ДЛЯ НАСТРОЙКИ ПАРАМЕТРОВ И ВЕРИФИКАЦИИ МЕТОДОВ АНАЛИЗА ДАННЫХ

*Караваяев Анатолий Сергеевич,
СГУ им. Н.Г. Чернышевского, Саратов
karavaevas@gmail.com*

*Ишбулатов Юрий Михайлович
СГУ им. Н.Г. Чернышевского, Саратов
ishbulatov95@mail.ru*

*Боровкова Екатерина Игоревна
СГУ им. Н.Г. Чернышевского, Саратов
rubanei@mail.ru*

*Безручко Борис Петрович
СГУ им. Н.Г. Чернышевского, Саратов
bezruchkobp@gmail.com*

Аннотация. В работе представлены математические модели электрокардиограммы и фотоплетизмограммы, позволяющие задавать известный паттерн фазовой синхронизации низкочастотных колебаний, отражающих автономный контроль кровообращения. Созданные модели адекватно воспроизводят спектральные и статистические свойства экспериментальных сигналов и используются для тестирования усовершенствованных алгоритмов детекции фазовой синхронизации. Результаты демонстрируют повышение чувствительности и специфичности метода диагностики синхронизации.

Ключевые слова: математическое моделирование, электрокардиограмма, фотоплетизмограмма, фазовая синхронизация, низкочастотные колебания, автономный контроль кровообращения, алгоритм детекции синхронизации, сердечно-сосудистые заболевания.

В работе предложены математические модели сигналов электрокардиограммы (ЭКГ) и фотоплетизмограммы (ФПГ), разработанные с возможностью задания заранее известного паттерна фазовой синхронизации низкочастотных колебаний, связанных с симпатическим контролем кровообращения. Такие модели воспроизводят свойства экспериментальных данных и позволяют настраивать параметры и тестировать методы детекции фазовой синхронизации.

Автономный контроль кровообращения играет ключевую роль в поддержании гомеостаза и функциональной стабильности сердечно-сосудистой системы. Нарушения в синхронизации низкочастотных осцилляций в вариабельности сердечного ритма и фотоплетизмограмме ассоциируются с развитием патологий, включая инфаркт миокарда и артериальную гипертензию. Это подчеркивает диагностическую значимость исследования физиологических биосигналов [1–4].

Целью работы является создание математических моделей ЭКГ и ФПГ, адекватно воспроизводящих экспериментальные сигналы с известным закономерным чередованием синхронных и асинхронных интервалов фазовой разности низкочастотных колебаний. Задачей является тестирование и улучшение алгоритма детекции фазовой синхронизации с использованием этих моделей.

Для моделирования разности мгновенных фаз используется модель фазовых осцилляторов с цветным шумом, параметры которого подобраны на основе анализа экспериментальных данных ЭКГ и ФПГ 10 здоровых добровольцев. Модели ЭКГ и ФПГ строятся на основе сумм гауссиан, описывающих основные пики и волновые формы сигналов [5, 6]. Алгоритм детекции синхронных участков основан на диагностике пологих участков разности мгновенных фаз [3, 4].

Разработанные модели качественно воспроизводят временные ряды и спектры реальных сигналов ЭКГ и ФПГ с характерной паттерном фазовой синхронизацией низкочастотных компонент RR-интервалов и ФПГ.

Благодаря тому, что участки фазовой синхронизации известны априорно, временные ряды моделей могут быть эффективно использованы для уточнения параметров методов диагностики синхронизации. В частности, уточнение параметров предложенного ранее метода [4] обеспечило улучшение его статистических свойств (чувствительность 0.69, специфичность 0.60, площадь под ROC-кривой 0.75) по сравнению с исходным методом.

Разработанные модели и алгоритмы представляют эффективный инструмент для объективного тестирования и совершенствования методов анализа фазовой синхронизации и связи биосигналов [7, 8]. Перспективно использование предложенных моделей для генерации обучающих выборок для классификаторов, построенных на основе нейросетевых методов. Результаты имеют значимость для фундаментальных исследований автономного контроля и практической медицины, в частности для диагностики сердечно-сосудистых заболеваний.

Работа выполнена при поддержке Российского научного фонда (грант № 23-12-00241).

Библиографический список

1. Gang Y. and Malik M. // Indian Pacing and Electrophysiology Journal. 2003. V. 3. Iss. 1. P. 34–40.
2. Shiogai Y. et al. // Physics Reports. 2010. V. 488. P. 51–110.
3. Karavaev A.S. et al. // CHAOS. 2019. V. 29. P. 121101.
4. Karavaev A.S. et al. // CHAOS. 2009. V. 19. P. 033112.
5. Tang Q. et al. // Sci Rep. 2020.
6. McSharry P.E. et al. // IEEE Trans Biomed Eng. 2003.
7. Mormann F. et al. // Physica D. 2000. V. 144. P. 358–369.
8. Rosenblum M.G. and Pikovsky A.S. // Phys. Rev. E. 2001. V. 64. P. 045202.

АНАЛИЗ ФАЗОВОЙ СТРУКТУРЫ ЭЛЕКТРОКАРДИОГРАММЫ ПРИ НАГРУЗКЕ И В ПОКОЕ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

*Бондаренко Илья Александрович,
НИУ ВШЭ – Нижний Новгород
iabondarenko_2@edu.hse.ru*

*Гриднев Владимир Иванович,
СГМУ им. В.И. Разумовского, Саратов
gridnev@cardio-it.ru*

*Паленова Анастасия Вячеславовна,
СГТУ им. Гагарина Ю.А., Саратов
anastasiapalenoval@gmail.com*

*Панюшев Антон Алексеевич,
НИУ ВШЭ – Нижний Новгород
aapanyushev@edu.hse.ru*

*Посненкова Ольга Михайловна,
СГМУ им. В.И. Разумовского, Саратов
posnenkova@cardio-it.ru*

*Станкевич Наталия Владимировна,
НИУ ВШЭ – Нижний Новгород
nstankevich@hse.ru*

Аннотация. Разработан классификатор ЭКГ с использованием машинного обучения. Предложенный классификатор в автоматическом режиме детектирует фазы покоя, нагрузки и восстановления по сигналам ЭКГ.

Ключевые слова: электрокардиограмма; машинное обучение; нагрузочные тесты.

В данной работе была исследована возможность автоматической классификации сигнала электрокардиограммы (ЭКГ) по фазам пробы с физической нагрузкой (покоя, нагрузки и восстановления) с применением алгоритмов машинного обучения и сравнение результата классификации с динамикой RR-интервалов.

Во время работы использовались ЭКГ 8 здоровых пациентов, представленные в виде двух отдельных файлов EDF, полученных с носимого кardiомонитора – кариофлешка ECG Dongle (ООО «Нордавинд», Дубна, Россия) общей длительностью около 14 минут с 6 каналами, соответствующими I, II и III стандартным отведениям, а также усиленным отведениям от конечностей aVR, aVL, aVF.

Для повышения общего объема выборки для обучения сигнал был разделён на перекрывающиеся окна длительностью 10 секунд при помощи метода скользящего окна с перекрытием 0.9. Такое разбиение обеспечивало достаточную детализацию для отслеживания переходных состояний между фазами при сохранении приемлемого объёма данных. Для каждого окна производилась детекция R-пику с помощью библиотеки NeuroKit2, после чего извлекались интервалы RR, временные промежутки между последовательными R-пиками. На основе полученных интервалов рассчитывались признаки вариабельности сердечного ритма, которые служили входными данными для

классификатора. Временные признаки включали в себя стандартные характеристики ЭКГ с подтвержденной множеством работ показательно отклонений у пациентов с хронической сердечной недостаточностью [1-3]: Mean_RR - среднее значение RR-интервалов, SDNN - стандартное отклонение RR-интервалов, RMSSD - квадратичное среднее последовательных разностей RR, pNN50 - доля пар RR, различающихся более чем на 50 мс, Mean_HR - средняя частота сердцебиения. Также рассчитывались частотные показатели HRV: LF power (0.04 – 0.15 Гц), HF power (0.15 – 0.4 Гц) и LF/HF ratio — их соотношение. Однако в ходе анализа было выявлено, что данные показатели в контексте заданной задачи не повышают качество классификации и имеют нулевой значение важности признака (рис.1), что связано с малым размером выборки и короткой длительностью окон.

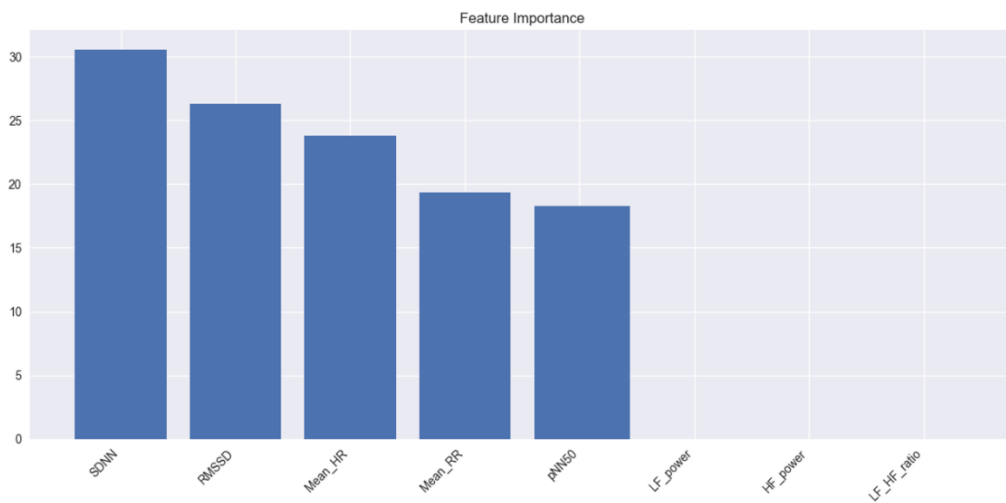


Рис.1. Значимость выбранных характеристик

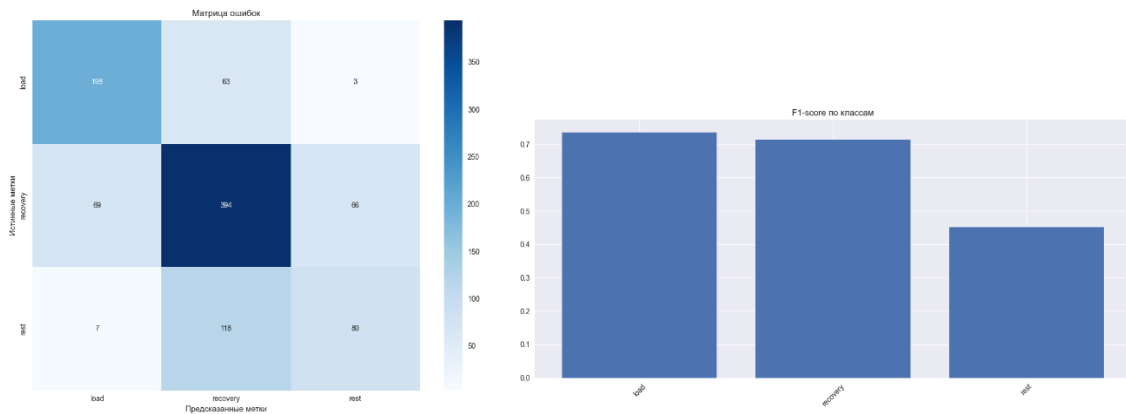


Рис.2. Метрики качества точности классификации обученной модели

Для построения модели классификации использовался алгоритм CatBoost, представляющий собой градиентный бустинг над решающими деревьями, хорошо работающий с табличными данными и устойчивый к переобучению на малых выборках. Подбор гиперпараметров выполнялся с помощью библиотеки Optuna, реализующей автоматическую оптимизацию на основе вероятностного поиска (TPE sampler). В ходе 50 итераций оптимизировались глубина деревьев, количество итераций (500-2000), скорость обучения (0.01-0.3) и коэффициент L2-регуляризации (1-10). Для повышения устойчивости модели использовались параметры `bootstrap_type = 'Bernoulli'` и `early_stopping_rounds = 50` с целевой метрикой: максимизация `accuracy`.

Метрики качества показали удовлетворительный уровень точности классификации при ограниченном объёме обучающих данных (рис.2). Для анализа вклада признаков использовались встроенные механизмы интерпретации CatBoost, что позволило оценить значимость выбранных временных характеристик (рис.1).

Для визуальной проверки корректности работы алгоритма были построены графики с наложением предсказаний модели на исходный сигнал и ручную разметку фаз. Сравнение показало, что предложенная модель в целом корректно воспроизводит структуру сигналов и адекватно выделяет переходные состояния между фазами, за исключением небольших промежутков сигнала.

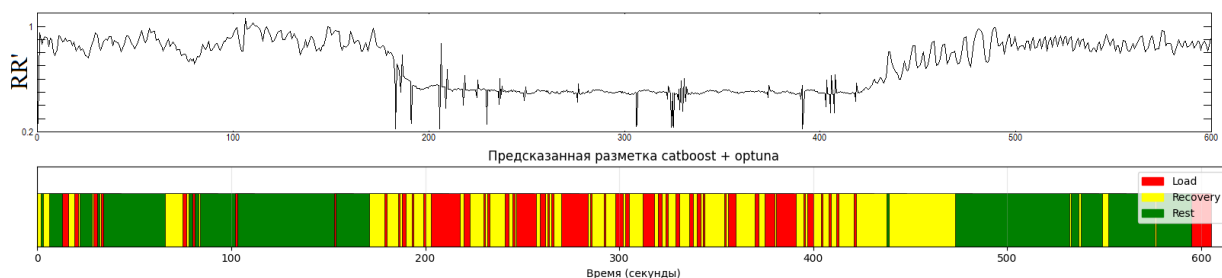


Рис.3. Сравнение динамики RR-интервалов сигнала ЭКГ и разметки моделью у испытуемого в фазе покоя, нагрузки и восстановления

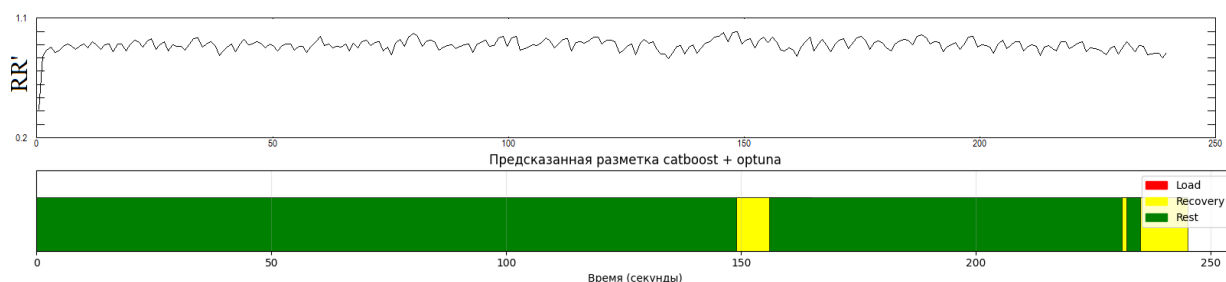


Рис.4. Сравнение динамики RR-интервалов сигнала ЭКГ и разметки моделью у испытуемого в фазе восстановления

Таким образом, разработанная методика, включающая комбинацию NeuroKit2 для детекции пиков, HRV-признаков и CatBoost с оптимизацией гиперпараметров в Optuna, позволяет реализовать необходимый подход к автоматическому определению фаз нагрузочной пробы в условиях ограниченного объема данных и высокой зашумленности сигнала ЭКГ в фазе нагрузки и начала периода восстановления.

Библиографический список

1. Wu L., Jiang Z., Li C., Shu M. // International Journal of Cardiology. 2014. Vol. 174, No. 3. P. 857–860.
2. Mulder M.B., Eidelson S.A., Buzzelli M.D., Gross K.R., Batchinsky A.I., Convertino V.A., Schulman C.I., Namias N., Proctor K.G. // Aerospace Medicine and Human Performance. 2019. Vol. 90, No. 12. P. 1009–1015.
3. Ayano Y.M., Geleto D.A., Jilcha K.D., Megersa S.M. // Diagnostics. 2022. Vol. 12, No. 12. P. 2971

ВЛИЯНИЕ МЕТОДА ИНТЕГРИРОВАНИЯ ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ НА ЭТАПЕ ФОРМИРОВАНИЯ ОБУЧАЮЩИХ ДАТАСЕТОВ НА СВОЙСТВА НЕЙРОМОРФНОГО ОТОБРАЖЕНИЯ

Мяков Тимофей Ильич,
НИУ ВШЭ – Нижний Новгород
timyakov@edu.hse.ru

Станкевич Наталия Владимировна,
НИУ ВШЭ – Нижний Новгород
nstankevich@hse.ru

Аннотация. В работе исследуется, как выбор численного метода интегрирования обыкновенных дифференциальных уравнений на этапе генерации обучающих данных влияет на точность нейросетевого аппроксиматора динамики. В качестве объекта изучения используются системы Рёсслера и Хиндмарша–Роуза. Датасеты формируются с помощью численных методов RK23, RK45 и DOP853, а модель представляет собой MLP, напрямую аппроксимирующую приращение состояния для фиксированного шага. Предложена функция потерь, согласованная с полужанной трапециевидной схемой, и введены метрики MSE и MRNE для сопоставления численных решений и предсказаний сети.

Ключевые слова: численное интегрирование; методы Рунге–Кутты; DOP853; RK23; RK45; система Рёсслера; система Хиндмарша–Роуза; нейросетевой аппроксиматор

Современные нейросетевые подходы позволяют аппроксимировать оператор потоковых динамических систем напрямую, минуя явное численное интегрирование на этапе применения модели [1-3]. При этом обучающие выборки для таких моделей обычно генерируются классическими интеграторами обыкновенных дифференциальных уравнений. В данной работе рассматривается вопрос: как выбор интегратора при генерации данных (его порядок, устойчивость, локальная ошибка) влияет на свойства обученного нейросетевого отображения.

В качестве объекта исследования рассматриваются две нелинейные системы: (i) система Рёсслера, демонстрирующая как устойчивые периодические автоколебания, так и хаотическое поведение; (ii) нейродинамическая система Хиндмарша–Роуза, моделирующая типичные осцилляции нейрона. Это позволяет сопоставить влияние интегратора на разных классах динамики.

Формирование датасетов. Выборки синтетически генерируются из параметрически заданных распределений по состояниям и параметрам системы; для построения таргетов применяется `solve_ivp` с методами RK23 (порядок 2), RK45 (порядок 4) и DOP853 (порядок 8) [4-8]. Таким образом, для одной и той же совокупности начальных условий получаются три альтернативных набора таргетов, отличающихся порядком локальной ошибки, гладкостью траекторий и свойствами накопления погрешности. Для корректного сравнения используется одинаковый шаг дискретизации при обучении модели и единая процедура нормализации признаков.

Модель. Применяется компактная MLP-архитектура с одним скрытым слоем на 128 нейронов (сигмоидальная активация), аппроксимирующая приращение состояния для фиксированного шага Δt : $u^{k+1} = u^k + \hat{a}$. На вход подаются нормализованные векторы состояния $u = [x, y, z]^T$ и параметров $p = [a, b, c]^T$. Такая постановка эквивалентна обучению

дискретного суррогат-оператора потока, что упрощает развертывание модели и последующее моделирование длительных траекторий.

Функция потерь и метрики. Лосс согласован с полуявной трапецевидной схемой: минимизируется квадратичная невязка между предсказанным приращением и правой частью ОДУ, усреднённой по началу и концу шага. Это поощряет динамическую согласованность предсказаний с физической моделью. Для оценки качества используются две метрики: MSE между траекторией численного интегратора и траекторией модели; MRNE — средняя относительная норма ошибки, нормализующая величину ошибки на масштаб входов. Такое сочетание позволяет выявить случаи, когда относительная и абсолютная ошибки ранжируют модели по-разному.

Дизайн экспериментов.

Режимы системы Рёсслера: (а) периодические автоколебания (например, $a = 0.1$, $b = 0.1$, $c = 6$) с отсечением переходного процесса; (б) хаотические автоколебания (например, $a = 0.15$, $b = 0.2$, $c = 10$) без отсечения. Для каждого режима обучаются три модели на трёх вариантах датасета (RK23/RK45/DOP853). Обучение в PyTorch с Adam и экспоненциальным scheduler; ранняя остановка по плато MRNE.

Система Хиндмарша–Роуза с параметрами $a = 1$, $b = 3$, $c = 1$, $d = 5$, $s = 4$, $\alpha = -1.6$. Для выбранных параметров формируются датасеты теми же интеграторами и обучаются модели той же архитектуры и с теми же настройками обучения, что обеспечивает сопоставимость результатов.

Ключевые наблюдения.

Периодические автоколебания в системе Рёсслера (рис.1): по абсолютной точности (MSE) наилучшей оказалась модель, обученная на данных порядка 4 (RK4/RK45). Датасет RK2 даёт несколько большую MSE, тогда как DOP853 формирует «слишком гладкие» таргеты, которые усваиваются хуже при фиксированной архитектуре и лоссе.

Хаотические автоколебания в системе Рёсслера: по средней MSE преимущество демонстрирует датасет RK2, в то время как MRNE остаётся близкой у всех трёх вариантов; интеграторы более высокого порядка не всегда выигрывают по MSE из-за чувствительности к начальным данным и быстрого накопления фазовой ошибки.

Система Хиндмарша–Роуза (рис.2): лучший результат показал датасет, сгенерированный RK4 (по MSE; MRNE остаётся сопоставимой или лучше относительно альтернатив), что указывает на пользу интегратора среднего порядка при осцилляторной нейронной динамике.

Общая тенденция: «слишком точные» таргеты (DOP853) не гарантируют лучшей обобщающей способности компактной MLP; умеренно точные схемы (RK4 в периодических режимах и RK2 в хаотическом Рёсслере) формируют обучающий сигнал, лучше согласованный с ёмкостью модели и выбранной функцией потерь.

Вывод. Порядок и свойства численного метода интегрирования на этапе формирования обучающего датасета существенно влияют на итоговую точность нейросетевого отображения. Интегратор фактически выступает скрытым гиперпараметром; его оптимальный выбор зависит от класса динамики, функции потерь и ёмкости модели. В представленных экспериментах это проявляется в преимуществе RK4 в устойчивом режиме Рёсслера и в системе Хиндмарша–Роуза, а также в преимуществе RK2 по средней MSE в хаотическом режиме Рёсслера при близких MRNE.

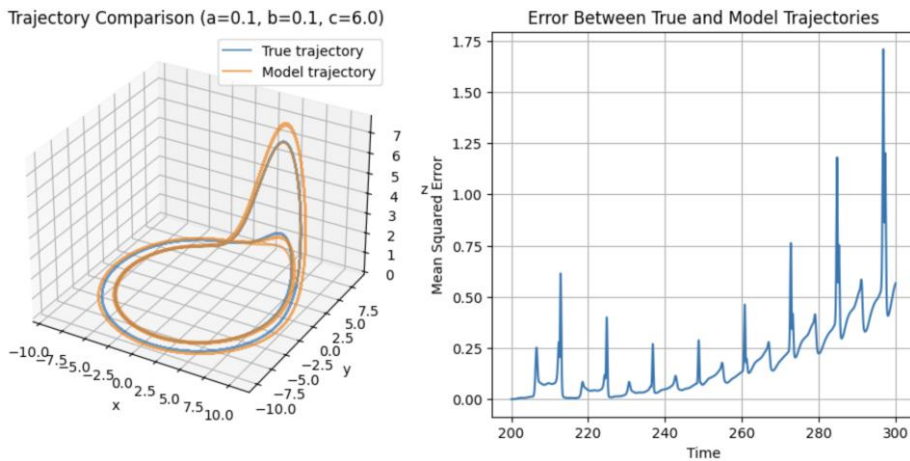


Рис.1. Сравнение траекторий системы Рёсслера, полученные численным методом РК4 и нейронной сетью, и соответствующая среднеквадратичная ошибка (MSE).

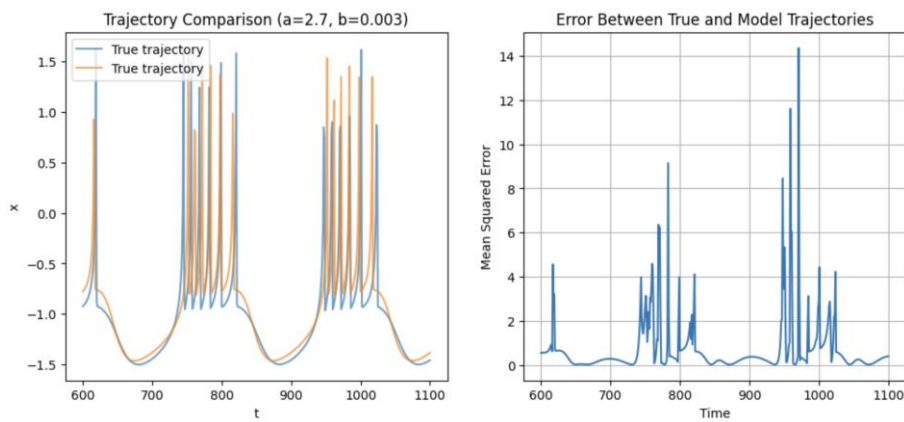


Рис.2. Сравнение траекторий системы Хиндмарша-Роуза, полученные численным методом РК4 и нейронной сетью, и соответствующая среднеквадратичная ошибка (MSE).

Библиографический список

1. Kingma D.P., Ba J. // arXiv preprint arXiv:1412.6980, 2017.
2. Kuptsov P.V., Kuptsova A.V., Stankevich N.V. // arXiv preprint arXiv:2104.05402, 2021
3. Kuptsov P. V., Stankevich N. V. // Chaos, Solitons & Fractals. 2025. V. 191. P. 115818.
4. Dormand J.R., Prince P.J. // Journal of Computational and Applied Mathematics. 1980.
5. Bogacki P., Shampine L.F. // Applied Mathematics Letters. 1989. V.3. No. 2.
6. Hairer E., Nørsett S.P., Wanner G. Solving Ordinary Differential Equations I: Nonstiff Problems. Springer, 1993.
7. Rössler O.E. // Physics Letters A. 1976. V. 57. No. 5. PP. 397-398.
8. Hindmarsh J.L., Rose R.M. // Proceedings of the Royal Society of London. Series B: Biological Sciences. 1984. V. 221. No. 1222. PP. 87-102.
9. Paszke A., Gross S., Massa F., Lerer A., Bradbury J., Chanan G., Killeen T. et al. // Advances in Neural Information Processing Systems. 2019. V.32.

ВОССТАНОВЛЕНИЕ СРЕДНЕПОЛЕВОЙ ДИНАМИКИ НЕЙРОН-АСТРОЦИТАРНОГО ВЗАИМОДЕЙСТВИЯ С ИСПОЛЬЗОВАНИЕМ АРХИТЕКТУРЫ dend-PLRNN

*Клименко Ксения Евгеньевна
ННГУ им. Н.И. Лобачевского – Нижний Новгород
kseniya.klimenko@unn.ru*

*Оленин Сергей Михайлович,
ННГУ им. Н.И. Лобачевского – Нижний Новгород
olenin@itmm.unn.ru*

*Леванова Татьяна Александровна
ННГУ им. Н.И. Лобачевского – Нижний Новгород
tatiana.levanova@itmm.unn.ru*

*Писарчик Александр Николаевич
Политехнический Университет Мадрида, Мадрид, Испания
alexander.pisarchik@upm.es*

Аннотация

В работе решается задача восстановления динамики феноменологической среднеполевой популяционной нейронной модели и её модификации с учетом астроцитарного влияния на основе временных рядов методами глубокого обучения. Для восстановления динамики была использована архитектура dend-PLRNN, которая представляет собой кусочно-линейную рекуррентную нейронную сеть с функциями активации ReLu и дендритными связями, которые позволяют эффективно уменьшить размер сети. В работе оценивается качество восстановления различных аттракторов исследуемых систем и изучается вопрос, как наличие в системе дополнительных элементов (астроцитов) влияет на качество восстановления динамики.

Ключевые слова: глубокое обучение; восстановление средне полевой динамики; архитектура dend-PLRNN; нейрон-астроцитарное взаимодействие.

Синаптическая передача играет ключевую роль в обработке информации в головном мозге и нервной системе. В последние десятилетия в ряде работ было показано, что ключевую роль в этом процессе играют не только нейроны, но и астроциты [1], которые взаимодействуют с пре- и постсинаптическими элементами, модулируя силу и пластичность нейронных связей.

Для моделирования нейрон-глиальных взаимодействий разработаны подходы на микро- и макроуровне. На микроуровне используются связанные сети, состоящие из слоев нейронов и астроцитов. На макроуровне применяются феноменологические среднеполевые модели [2, 3], которые модифицируют классические модели, добавляя уравнения для описания влияния астроцитов через глиотрансмиттеры, что позволяет описывать сложную динамику нейронных популяций и сопоставлять моделирование с экспериментальными данными.

Цель исследования — оценить, как присутствие астроцитов в системе влияет на качество восстановления на основе временных рядов аттракторов соответствующей динамической системы. Для достижения этой цели проведено сравнение результатов

восстановления аттракторов из данных временных рядов для классической модели Цодыкса-Маркграма (ЦМ) [4] и для модели Цодыкса-Маркграма с учетом астроцитарного влияния (ЦМА) [3, 5].

Восстановление динамики нейрон-астроцитарного взаимодействия выполнено с использованием архитектуры глубокого обучения dend-PLRNN [6], схематично изображенной на рис. 1. Архитектура представляет собой кусочно-линейную рекуррентную нейронную сеть, описываемую уравнением для M -мерного скрытого состояния:

$$z_t = Az_{t-1} + W\phi(z_{t-1}) + h + Cs_t,$$

которое описывает временную эволюцию M -мерного вектора скрытого состояния $z_t = (z_{1t}, z_{2t}, \dots, z_{Mt})$, A — диагональная матрица констант времени для каждого состояния, W — матрица весов связей, h — вектор смещений, C — веса внешних входов s_t . Функция активации задает дендритную нелинейность [8] и представляет собой линейные сплайны с наклонными дендритного входа/выхода $a_b \in R$ и порогами активации $h_b \in R^M$:

$$\phi(z_{t-1}) = \sum_{b=1}^B \alpha_b \max(0, -h_b) z_{t-1}$$

где B — число расширения базиса (basis expansion number), задает количество кусочно-линейных компонент, которые используются для аппроксимации нелинейной функции внутри модели. Увеличение параметра B позволяет более точно аппроксимировать нелинейные функции посредством разбиения на большее число линейных сегментов.

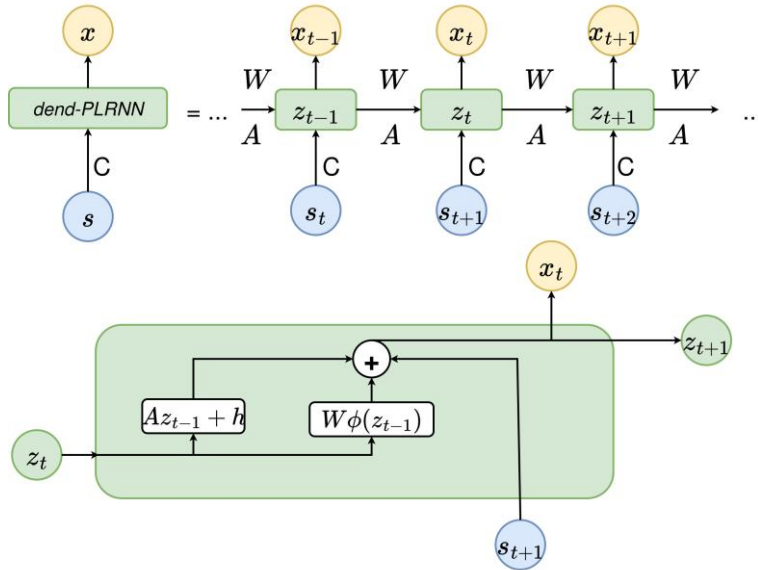


Рис. 1. Схема архитектуры нейронной сети dend-PLRNN: сверху показана классическая RNN, внизу — модифицированный блок dend-PLRNN

Обучение сети проводится методом обратного распространения ошибки во времени с частичной подстановкой истинных значений для стабилизации градиентов.

В качестве данных используются временные ряды (траектории), сгенерированные с использованием моделей ЦМ и ЦМА. В качестве метода интегрирования выбран метод Рунге-Кутты 4 порядка с постоянным шагом 0.01. Перед обучением данные стандартизируются для обеспечения корректного сравнения признаков и устойчивости процесса обучения.

Для анализа качества восстановления используется геометрическая метрика KLx [6], основанная на относительной энтропии. Эта метрика измеряет расхождение между распределениями состояний моделируемых систем в фазовом пространстве.

Показано, что используемая глубокая нейронная сеть dend-PLRNN успешно восстанавливает периодические, хаотические и стационарные режимы и соответствующие аттракторы для классической модели ЦМ и ЦМА, что объясняется тем, что по своей сути рассматриваемая архитектура dend-PLRNN сама является дискретной среднеполевой моделью, и таким образом, успешно улавливает особенности динамики других среднеполевых моделей. При этом периодические режимы хорошо восстанавливаются обеими моделями (рис. 2), а стационарный режим для модели ЦМА восстанавливается существенно точнее, чем для классической модели ЦМ. Таким образом, присутствие астроцитов в системе может не только не мешать, но и приводить к стабилизации предсказания.

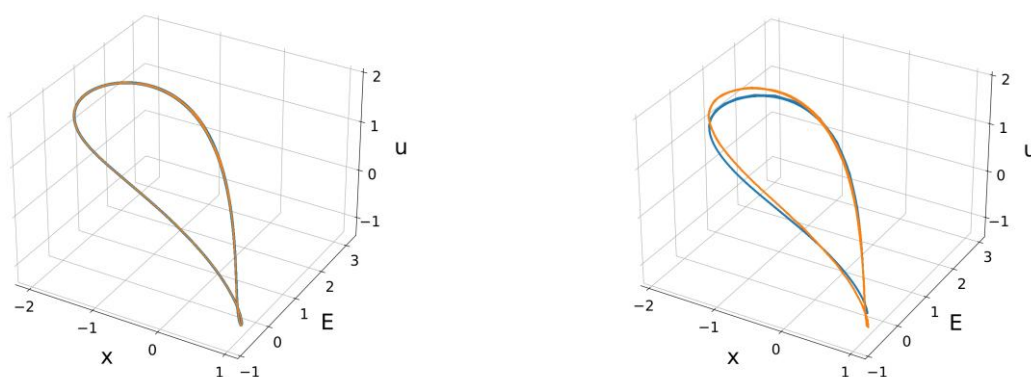


Рис. 2. Примеры восстановления динамики моделей. Рисунок слева соответствует модели ЦМ, $KLx = 0.039$, рисунок справа - модели ЦМА, $KLx = 15.870$. Траектории синего цвета соответствуют исходному аттрактору, оранжевые — восстановленному.

Работа выполнена при поддержке проекта РНФ № 25-22-00465

Библиографический список

1. Halassa M.M., Fellin T., Haydon P.G. Trends in molecular medicine. 2007. V. 13. №. 2. P. 54.
2. Cortes J.M., Desroches M., Rodrigues S., Veltz R., Muñoz M.A., Sejnowski T.J. Short-term synaptic plasticity in the deterministic Tsodyks–Markram model leads to unpredictable network dynamics // Proceedings of the National Academy of Sciences. 2013. V. 110. №. 41. P. 16610.
3. Olenin S.M., Levanova T.A., Stasenko S.V. Dynamics in the reduced mean-field model of neuron–glial interaction // Mathematics. 2023. V. 11. №. 9. P. 2143.
4. Mongillo G., Barak O., Tsodyks M. Synaptic theory of working memory // Science. 2008. V. 319. №. 5869. P. 1543-1546.
5. Barabash N.V., Levanova T.A., Smirnov L.A. Blue sky catastrophe in the phenomenological model of neuron–astrocyte interaction // Chaos: An Interdisciplinary Journal of Nonlinear Science. 2024. V. 34. №. 12.
6. Brenner M., Hess F., Mikhaeil J.M., Bereska L.F., Monfared Z., Kuo P.C., Durstewitz D. Tractable dendritic RNNs for reconstructing nonlinear dynamical systems // International conference on machine learning. – Pmlr. 2022. 3. 2292-2320.

MACHINE LEARNING FOR SOLVING THE PROBLEM OF DETECTING
MULTISTABLE ATTRACTORS

Panyushev Anton
HSE University, Nizhny Novgorod, Russia
apanyushev@hse.ru

Kuptsov Pavel
HSE University, Nizhny Novgorod, Russia
pkuptsov@hse.ru

Stankevich Nataliya
HSE University, Nizhny Novgorod, Russia
nstankevich@hse.ru

Abstract. A method for detecting multistability using neuromap is applied for multistable Chua model. It is shown that with a sufficient amount of data, the system can reproduce the dynamics on which it was trained, and with a small amount it does it roughly. Even with a small amount of data, the proposed algorithm found a hidden attractor. It is also noted that with fewer training pairs, the quality of reproducing the attractors decreases.

Keywords: neuromap, multistability, hidden attractors.

A series of neuromaps was trained on pairs (a point in phase space, a point on a trajectory passing through it through $t=0.01$) with a given number of pairs from the basins of attraction of each of the five attractors to reproduce the Chua system [1] with parameters $\alpha=8.4, \beta=12, \gamma=0.05, m_0=-1.2, m_1=-0.23$. Neuromaps was trained according to the algorithm proposed in [2] as a direct error propagation network with two hidden layers of 128 neurons each and *tanh* as an activation function, where the network input was normalized.

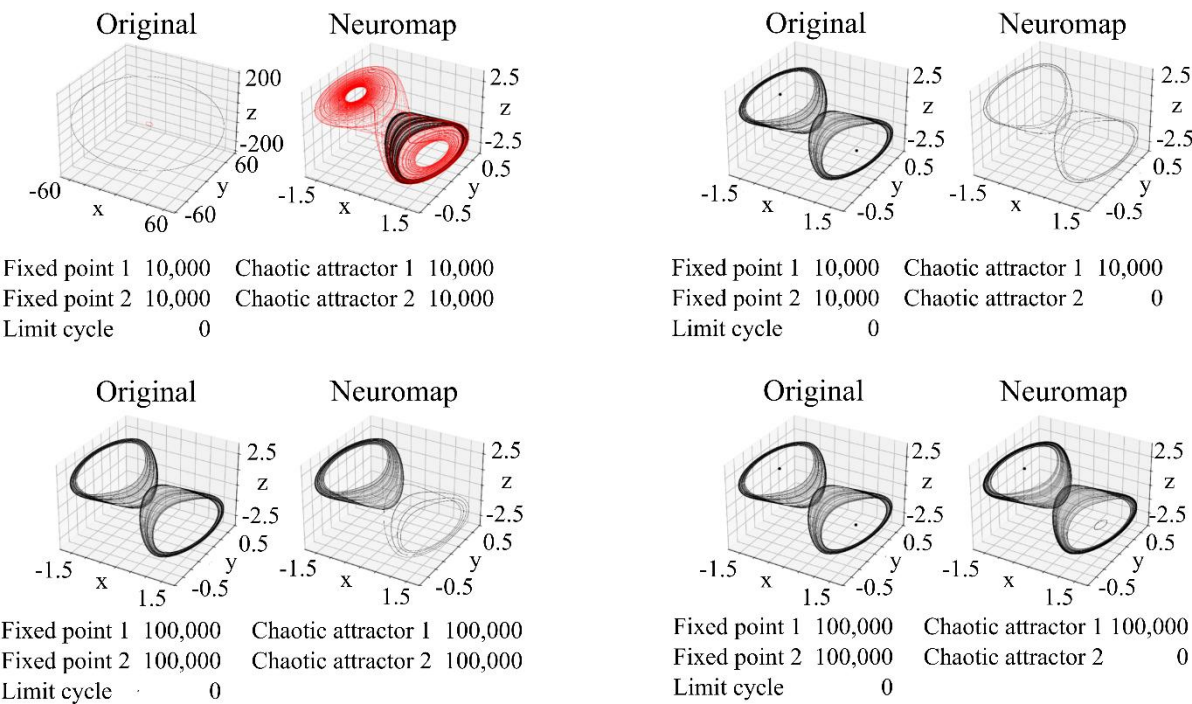


Fig.1. Comparison of the original attractor and the one reproduced by neuroimaging. The number of training pairs for each attractor is shown next to each experiment

Fig.1 demonstrates results of comparison solution of ODE (original) and corresponding trajectory of neuromap. It is shown that in the absence of a limit cycle in the dataset, the envelopes of the other attractors, neuroimaging finds it in a certain distorted form: its radius decreases. Chaotic attractors with insufficient data are found as similar limit cycles or as a merged chaotic attractor. The quality of the attractor reproduction depends on the amount of training data. Whereas the limit cycle was found in each of the four experiments.

The work carried out in the frame of project “International academic cooperation” HSE University.

References

1. *Chua L. O.* Chua circuit //Scholarpedia. – 2007. – Т. 2. – №. 10. – С. 1488.
2. Kuptsov P. V., Kuptsova A. V., Stankevich N. V. Artificial neural network as a universal model of nonlinear dynamical systems //arXiv preprint arXiv:2104.05402. – 2021.

ON SOME PROPERTIES OF OUTPUT MATRICES IN RESERVOIR COMPUTINGS

Gromov Nikolay

Lobachevsky University, Nizhny Novgorod, Russia

gromov@itmm.unn.ru

Levanova Tatiana

Lobachevsky University, Nizhny Novgorod, Russia

tatiana.levanova@itmm.unn.ru

Smirnov Lev

Lobachevsky University, Nizhny Novgorod, Russia

lev.smirnov@itmm.unn.ru

Abstract. The central hypothesis of this research is that increasing the number of distinct trajectories in the training set causes the trained output weight matrices to become more similar to each other, which in turn enhances the reliability and accuracy of predictions by preventing the reservoir from diverging to its own, incorrect dynamics.

Keywords: machine learning, reservoir computing, Li-Sprott System, Rössler System.

Reservoir Computing (RC) is a promising machine learning approach for this task due to its computational efficiency. The architecture typically consists of three layers: an input layer, a recurrent reservoir with fixed internal connections, and a trainable output layer. Only the output weights are adapted during training, while the reservoir and input weights are initialized randomly and remain fixed. This study investigates the training process of RC for predicting the trajectories of chaotic dynamical systems. A key challenge in data-driven forecasting is accurately reconstructing a system's dynamics, especially when it is chaotic or multistable (possessing multiple coexisting attractors), from limited observational data.

The research was conducted on two canonical chaotic systems: the Li-Sprott system and the Rössler system. The Li-Sprott system is a 4D multistable system where a torus attractor coexists with a symmetric pair of strange chaotic attractors [1]. The Rössler System is a 3D chaotic system with a lower Lyapunov exponent and a longer predictability time [2].

In this study, we employ a conventional RC architecture composed of three main components: an input layer, which receives time-series data and projects it into the high-dimensional reservoir space; a reservoir layer, consisting of 300 neurons (sparsely connected with fixed random weights); a trainable output layer, which reads from the reservoir state and is optimized to perform predictions based on the dynamic information encoded in the reservoir. The reservoir's state evolves according to a differential equation:

$$f(r_t, x_t) = -r_t + \tanh(Wr_t + GW_{in}x_t + B),$$

where x_t is an input sample of the system at the time t , r_t is a reservoir state at the time t , W is an adjustment matrix of the reservoir, W_{in} is an input matrix with strange G , B is a bias vector.

It is integrated numerically using a 4th-order Runge-Kutta scheme. The reservoir updates its internal state multiple times (25) before making a single prediction for the next point in the system's trajectory. The reservoir was trained on a set of different trajectories (from 1 to 10). For

the test trajectories, after a "warm-up" the RC operates autonomously: its own predicted output is fed back as its input for the next step.

When trained on only one trajectory, the RC often failed to generalize. The predicted trajectory initially follows the true path but then suddenly shifts to a dynamical regime not present in the original system. This shift is characterized by a strong zero-frequency component in the power spectrum. It was found that output matrices leading to these shifts were distant from each other in terms of the Frobenius norm.

The core hypothesis was confirmed: as the number of training trajectories increased, the number of predicted trajectories that exhibited shifts decreased. For the Li-Sprott system, this decline was sharp and clear. For the Rössler system, shifts were less frequent overall but were eliminated entirely when 10 trajectories were used for training.

The distance between different matrices trained on the same number of trajectories decreased as the number of training trajectories increased. Both the mean distance and the interquartile range shrank. This indicates that the training process converges to a more consistent and unique mapping from the reservoir state to the system's state.

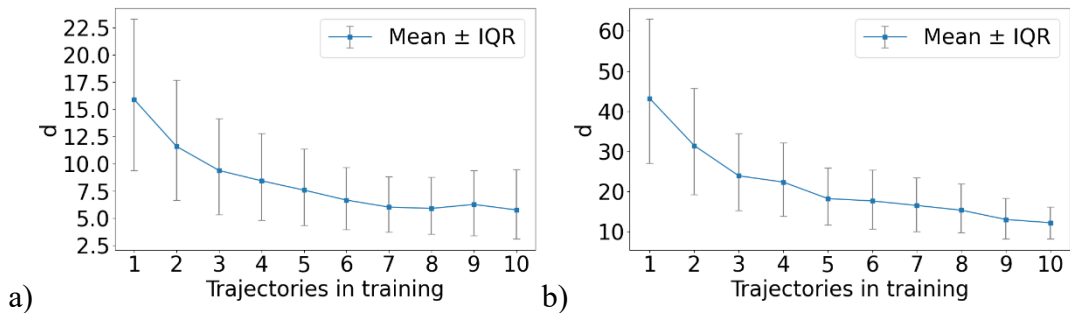


Fig. 2. Mean distance d between trained matrices with 10-th and 90-th quartile bands, depending on the number of trajectories in training n . (a) The Li-Sprott system. (b) The Rössler system.

This finding has practical importance for applications where data collection is expensive or difficult (e.g., neurophysiology). It provides a guideline for the optimal amount of data required for reliable forecasting, suggesting that a small set of distinct trajectories (e.g., ~ 10) can be sufficient for robust performance. The results may also be generalized to other types of recurrent neural networks, opening avenues for future research.

Funding for this project was provided by the Ministry of Science and Higher Education of the Russian Federation, Contract FSWR-2024-0005.

References

1. Li C., Sprott J. C. Coexisting hidden attractors in a 4-D simplified Lorenz system. *International Journal of Bifurcation and Chaos*. – 2014. – T. 24. – №. 03. – C. 1450034.
2. Rössler O. E. An equation for continuous chaos. *Physics Letters A*. – 1976. – T. 57. – №. 5. – C. 397-398.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МУЛЬТИСТАБИЛЬНЫХ СВОЙСТВ МОДЕЛИ ТИПА ХОДЖКИНА-ХАКСЛИ И ЕЁ НЕЙРОСЕТЕВОГО ОТОБРАЖЕНИЯ

Багаутдинова Эльмира Рафиковна,
НИИУ ВШЭ – Нижний Новгород
bagautdinovaerl@mail.ru

Купцов Павел Владимирович
НИИУ ВШЭ – Нижний Новгород
pkuptsov@hse.ru

Станкевич Наталия Владимировна
НИИУ ВШЭ – Нижний Новгород
nstankevich@hse.ru

Аннотация. В работе проводится сравнительный анализ мультистабильных свойств модифицированной модели Шермана–Ринцеля с дополнительным калиевым каналом типа Ходжкина–Хаксли, и её нейросетевого отображения. Мультистабильность описывается с помощью вероятностных характеристик по ансамблю случайных начальных условий, при этом нейросетевая модель аппроксимирует отображение исходной системы по шагу интегрирования. Показано, что нейросетевое отображение корректно воспроизводит расположение областей бистабильности и форму аттракторов, однако несколько завышает долю начальных условий, приводящих к равновесному режиму.

Ключевые слова: модели типа Ходжкина–Хаксли; мультистабильность; нейросетевое отображение.

Модели типа Ходжкина–Хаксли широко используются для описания возбудимых клеток (нейронов, кардиомиоцитов, β -клеток) и воспроизводят богатый набор режимов — покоя, спайков и берстов [1]. Одной из ключевых особенностей таких систем является мультистабильность, то есть сосуществование нескольких устойчивых режимов при фиксированных параметрах. Для β -клеток мультистабильность связана, в частности, с сосуществованием молчаливых и состояний автоколебаний в виде берстов, и может играть роль в регуляции секреции инсулина [2].

Численное исследование мультистабильности требует интегрирования большого числа траекторий для разных начальных условий и параметров, что становится вычислительно затратным при переходе к сетям осцилляторов. Это мотивирует использование искусственных нейронных сетей в качестве суррогатных моделей, аппроксимирующих динамику динамических систем и позволяющих ускорить параметрический анализ.

Рассматривается модифицированный вариант модели Шермана–Ринцеля [3] для β -клетки:

$$\begin{aligned}\tau \dot{V} &= -I_{Ca}(V) - I_k(V, n) - I_{K2}(V, n) - I_S(V, S), \\ \tau \dot{n} &= \sigma(n_\infty(V) - n), \\ \tau_S \dot{S} &= S_\infty(V) - S.\end{aligned}$$

Динамика мембранного потенциала V , быстрого калиевого канала, n и медленной переменной S задаётся системой уравнений типа Ходжкина–Хаксли с кальциевым, калиевым и медленным токами:

$$\begin{aligned}I_{Ca}(V) &= g_{Ca} m_\infty(V)(V - V_{Ca}), \\ I_K(V, n) &= g_K n_\infty(V)(V - V_K),\end{aligned}$$

$$I_{K2}(V, n) = g_{K2}p_{\infty}(V)(V - V_{K2}),$$

$$I_S(V, n) = g_S S_{\infty}(V)(V - V_S),$$

Функция, описывающая открытие ионных каналов в оригинальной модели, имеет следующий вид:

$$\omega_{\infty}(V) = [1 + \exp \frac{V_{\omega} - V}{\theta_{\omega}}]^{-1}, \quad \omega = m, n, S.$$

Мультистабильность обеспечивается введением дополнительного калиевого тока $IK2$ с немонотонной функцией активации, представленной разностью двух сигмоид:

$$p_{\infty}(V) = [\exp \frac{V - V_p}{\theta_p} + \exp \frac{V_p - V}{\theta_p}]^{-1}.$$

Такая локальная модификация порождает сосуществование равновесия и автоколебаний в некотором интервале параметров. В работе параметр проводимости дополнительного канала g_{K2} фиксируется на значении, при котором система демонстрирует берстовые колебания, а в качестве бифуркационного параметра используется сдвиг активационной кривой медленного тока V_S . Для ряда значений V_S наблюдается бистабильность «равновесие–берсты».

Для ускорения вычислений строится нейросетевое отображение, аппроксимирующее шаг интегрирования исходной системы [4]. Вход включает текущее состояние и параметр V_S , выход — состояние в следующий момент времени. Полносвязная сеть обучается на наборах пар $(x(t), x(t+\Delta t))$, полученных численным интегрированием исходной модели для различных начальных условий и параметров. После обучения нейросетевое отображение воспроизводит временные ряды и фазовые портреты, близкие к траекториям исходной системы.

Локализация области мультистабильности описывается вероятностным методом. Для фиксированного V_S генерируется ансамбль случайных начальных условий в ограниченной области фазового пространства. Каждая траектория интегрируется до выхода на установившийся режим, после чего по отсутствию колебаний производится классификация: равновесие или автоколебания. На этой основе оценивается вероятность попадания в равновесный режим:

$$P_{eq} = \frac{N_{eq}}{N_{tot}} \cdot 100\%,$$

где N_{eq} — число траекторий, приходящих к равновесию. Аналогичная процедура применяется к нейросетевому отображению.

На рисунке 1 полученные вероятностные характеристики P_{eq} для исходной модели и нейросетевого отображения демонстрируют хорошее совпадение. В обоих случаях область бистабильности по параметру V_S определяется интервалом, где вероятность равновесного режима мала, но отлична от нуля, что соответствует узкому бассейну притяжения равновесия при доминировании колебательного аттрактора.

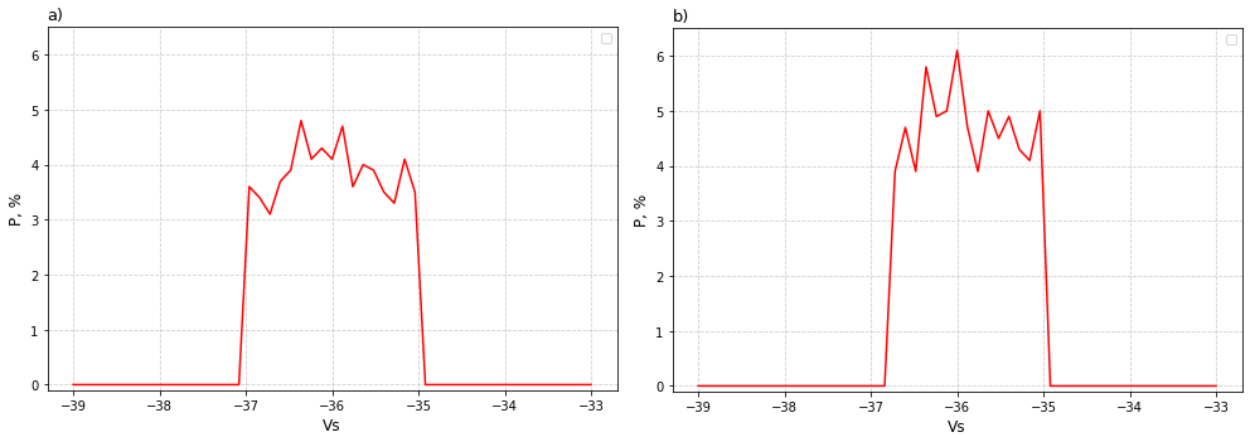


Рис.1. Вероятностные характеристики для а) математического моделирования, б) нейросетевого отображения

Для нейросетевого отображения уменьшается область бистабильности по V_s и форма зависимости P_{eq} , однако в ней значения вероятности увеличиваются только на 1.3% по сравнению с исходной моделью. Это можно интерпретировать как следствие аппроксимации сложных границ бассейнов притяжения нейросетью, приводящее к их «размыванию». При этом сравнение фазовых портретов показывает хорошее совпадение форм аттракторов и расположения равновесия.

Проведён сравнительный анализ мультистабильных свойств модифицированной модели типа Ходжкина–Хаксли и её нейросетевого отображения на основе вероятностных характеристик по начальным условиям. Показано, что нейросетевое отображение корректно воспроизводит расположение областей бистабильности и качественную структуру аттракторов, несмотря на небольшое завышение вероятности равновесного режима.

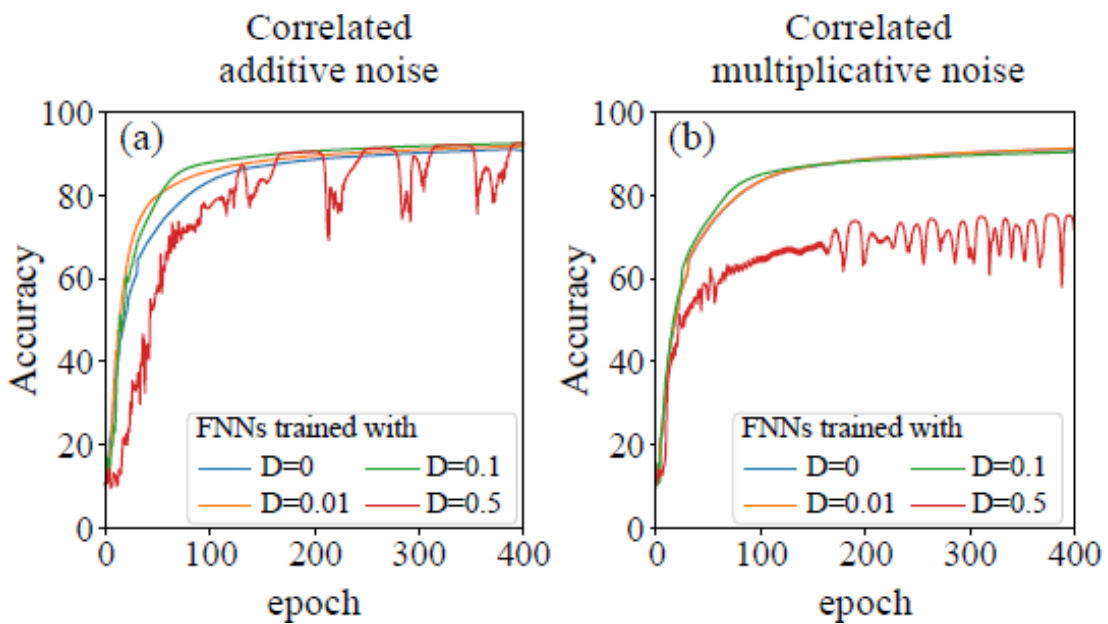
Работа выполнена в рамках проекта «Международное академическое сотрудничество» НИУ ВШЭ

Библиографический список

1. Pedersen M. G. et al. A biophysical model of electrical activity in human β -cells. *Biophys. J.* 2010. Vol. 99. P. 3200–3209.
2. Stankevich N., Mosekilde E. Coexistence between silent and bursting states in a model of pancreatic β -cells with an additional potassium current. *Chaos*. 2017. V. 27, №. 123101. P. 12.
3. Sherman A, Rinzel J, Keizer J. Emergence of organized bursting in clusters of pancreatic beta-cells by channel sharing. *Biophys J* 1988;54(3):411–25.
4. Kuptsov P.V., Stankevich N.V., Bagautdinova E.R. Discovering dynamical features of Hodgkin–Huxley-type model of physiological neuron using artificial neural network. *Chaos, Solitons and Fractals*. 2023. V. 167 No 113027

Секция 2

«Нелинейная динамика для задач машинного обучения»



Session 2

«Nonlinear dynamics for machine learning»

ВНУТРЕННИЕ ШУМЫ В АППАРАТНЫХ НЕЙРОННЫХ СЕТЯХ РАЗЛИЧНЫХ ТИПОВ

*Семенова Надежда Игоревна,
СГУ им. Н.Г. Чернышевского
semenovani@sgu.ru*

*Колесников Иван Дмитриевич,
СГУ им. Н.Г. Чернышевского
kole200@yandex.ru*

*Максимов Даниил Алексеевич,
СГУ им. Н.Г. Чернышевского
danilmax2003maksimov@gmail.com*

*Москвитин Виктор Максимович,
СГУ им. Н.Г. Чернышевского
vmoskvitin@gmail.com*

Аннотация. В последние годы появилось большое количество работ, посвященных аналоговой (аппаратной) реализации искусственных нейронных сетей, в которых нейроны и связи между ними основаны не на компьютерных вычислениях, а на физических принципах. Такие сети в некоторых случаях обеспечивают улучшенную энергоэффективность и масштабируемость, но могут быть подвержены влиянию внутренних шумов. В данной работе изучается влияние внутреннего шума на обученные и упрощенные глубокие, рекуррентные, свёрточные и спайковые нейронные сети. Белый гауссовский шум вводится четырьмя способами: аддитивным и мультипликативным (в зависимости от воздействия на каждый нейрон); коррелированным и некоррелированным (в зависимости от того, как он влияет на весь резервуар).

Ключевые слова: нейронные сети; искусственные нейронные сети; аппаратные нейронные сети; шум; шумовое воздействие; глубокие нейронные сети; рекуррентные нейронные сети; свёрточные нейронные сети; спайковые нейронные сети.

Физическая реализация искусственных нейронных сетей (ИНС) в реальном эксперименте – довольно новое направление в области машинного обучения [1,2]. Такие сети дают существенное преимущество в скорости работы, однако могут иметь внутренние шумы, исходящие из составляющих экспериментальной установки. При этом шум отдельных компонентов может распространяться по сети и потенциально накапливаться. Следовательно, такие сети могут быть уязвимы к воздействию шума, что является существенным препятствием в процессе обучения и функционирования. На данный момент опубликовано несколько работ, в которых показывается, что наличие шума внутри экспериментально реализованной ИНС мешает обучению или функционированию [3,4]. В смоделированных на компьютере сетях шум, как правило, добавлялся только во входные данные, что влияло на процесс работы и обучения [5]. Однако практически нет работ, которые анализировали бы распространение и накопление шума в ИНС с точки зрения шума внутри нейронов или связи.

Здесь будут представлены результаты, полученные для нескольких типов нейронных сетей. В качестве источников шумов рассматриваются источники белого гауссовского шума. В зависимости от способа воздействия шума на искусственные нейроны

использовалось аддитивное и мультипликативное шумовое воздействие, а в зависимости от влияния на слой нейронов рассматривалось коррелированное и некоррелированное воздействие. Для начала рассматривается влияние аддитивных и мультипликативных шумовых воздействий на упрощенные глубокую и рекуррентную сеть, когда нейроны имеют линейную функцию активации [6]. Некоррелированный шум в такой сети может быть уменьшен благодаря слабой связи между нейронами. В случае аналогичной сети, но с нейронами с нелинейной функцией активации (гиперболический тангенс и сигмоида) ситуация немного сложнее [7]. Для упрощенной нейронной сети с глобальной связью между нейронами было обнаружено, что накопление шума в целом ограничено, а добавление дополнительных слоев сети не ухудшает соотношение сигнал-шум ниже определенного предела. Для обученных нейронных сетей аналитическое предсказание уровня шума имеет некоторые особенности, но все еще может применяться. В качестве расширения поставленных задач были предложены способы уменьшения влияния различных типов шумов [8]. Для уменьшения влияния коррелированного аддитивного шумового воздействия была введена концепция нейронов-призраков (*ghost neurons*), а для некоррелированных шумов – концепция пулов.

Что касается рекуррентной сети, то рассматриваются сети с типичными матрицами связей резервуара: случайные равномерные и ленточные матрицы с различной связностью [9]. Все НЭС были обучены предсказывать хаотическую реализацию системы Макки–Гласса. Показано, что распространение шума в резервуаре в основном контролируется статистическими свойствами выходной матрицы связей, а именно ее средним и средним квадратом. Эти значения влияют на накопление определенных типов шумового воздействия. Также показывается, что существуют условия, при которых даже шума интенсивностью 10^{-20} уже достаточно, чтобы полностью потерять полезный сигнал. Также показывается, какие типы шума наиболее критичны для сетей с различными функциями активации (гиперболический тангенс, сигмоидальная и линейная) и если сеть является самозамкнутой.

Для свёрточной сети было показано, что аддитивные шумы (как коррелированный, так и некоррелированный) в сетях со свёрточным слоем накапливаются сильнее, чем в сети без него. Зависимости точности работы сети от интенсивности мультипликативного коррелированного шума выглядят одинаково для сетей со свёрточным слоем и без, а аналогичные зависимости для мультипликативного некоррелированного шума выглядят гораздо лучше для сети со свёрточным слоем. Также были рассмотрены слои дискретизации с *MaxPooling* и *MeanPooling*. Их использование позволяет существенно поднять точность при наличии аддитивного шума в свёрточном слое. Спадание точности при увеличении интенсивности мультипликативного коррелированного шума практически одинаковое для сетей со слоем субдискретизации и без него. Что же касается некоррелированного мультипликативного шума, то сети с *MaxPooling* менее устойчивы к его воздействию. Заключение. Показано, что аддитивный шум сильно ухудшает работу сети при наличии свёрточного слоя, однако его негативное воздействие можно уменьшить при наличии слоя субдискретизации сразу после слоя свёртки. Для мультипликативного шума выводы не такие однозначные. В большинстве случаев его влияние одинаково независимо от наличия слоев свёртки и субдискретизации. Однако использование *MaxPooling* в слое субдискретизации делает сеть менее устойчивой к мультипликативному некоррелированному шумовому воздействию. Кроме этого, была предложена аналитическая оценка влияния шумового воздействия на такие сети [10].

Что касается спайковых сетей, мы рассмотрели обученные сети, состоящие из моделей нейронов ФицХью-Нагумо в колебательном режиме [11]. Для таких сетей было рассмотрено некоррелированное аддитивное и мультипликативное шумовое воздействие, которое вводилось в первое уравнение системы:

$$\begin{aligned}\varepsilon \dot{x}_i &= x_i - \frac{x_i^3}{3} - y_i \cdot \left(1 + \sqrt{2D_M} \xi_M(i, t)\right) + \sqrt{2D_A} \xi(i, t) \\ \dot{y}_i &= x_i + a + I(i, t).\end{aligned}\quad (1)$$

И аналогичным образом для второго уравнения системы

$$\begin{aligned}\varepsilon \dot{x}_i &= x_i - \frac{x_i^3}{3} - y_i \\ \dot{y}_i &= x_i \cdot \left(1 + \sqrt{2D_M} \xi_M(i, t)\right) + a + I(i, t) + \sqrt{2D_A} \xi(i, t).\end{aligned}\quad (2)$$

Для обеих систем уравнений x_i и y_i – фазовые переменные системы ФитцХью—Нагумо, описывающие динамику i -го осциллятора, ε и a – параметры системы, которые были зафиксированы, $I(i, t)$ – внешнее воздействие, описывающее влияние предыдущего слоя обученной сети. Подробнее описание системы можно прочитать в работе [11].

Для таких сетей было показано, что внедрение обоих типов шумового воздействия приводит к существенной потере точности уже при интенсивности шума порядка 10^{-2} , однако для аддитивного воздействия в первом уравнении спад происходит значительно медленнее.

Работа выполнена за счет гранта Российского научного фонда № 25-72-10055, <https://rscf.ru/project/25-72-10055/>

Библиографический список

1. Tuma T., et al. // Nature Nanotechnology. 2016. V. 11, P. 693–699.
2. Bueno J., et al. // Optica. 2018. V. 5, № 6. P. 756–760.
3. Soriano M.C., et al. // IEEE transactions on neural networks and learning systems. 2015. V. 26. P. 388–393.
4. Moon S., Shin K., Jeon D. // IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2019. V. 27. № 6. P. 1455–1459.
5. Tait A. N., et al. // Scientific Reports. 2017. V. 7. № 1. P. 1–10.
6. Semenova N., Porte X., Andreoli L., Jacquot M., Larger L., Brunner D. // Chaos. 2019. V. 29, № 10. P. 103128.
7. Semenova N., Larger L., Brunner D. // Neural Networks. 2022. V. 146. P. 151–160.
8. Semenova N.I., Brunner D. // Chaos. 2024. V. 34, № 5. P. 051101.
9. Semenova N. // Eur. Phys. J. Spec. Top. 2025. doi: 10.1140/epjs/s11734-025-01697-7
10. Kolesnikov I.D., Semenova N. // Chaos. 2025. V. 35. № 6. P. 061104.
11. Bogatenko T.R., Sergeev K.S., Slepnev A.S., Kurths J., Semenova N.I. // Chaos. 2023. V. 33. № 7. P. 073122.

ВЛИЯНИЕ ВНУТРЕННЕГО ШУМОВОГО ВОЗДЕЙСТВИЯ НА РАБОТУ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ. ПРИМЕНЕНИЕ СТРАТЕГИИ УМЕНЬШЕНИЯ ШУМА

*Москвитин Виктор Максимович,
СГУ им. Н.Г. Чернышевского
vmmoskvitin@gmail.com*

*Семенова Надежда Игоревна,
СГУ им. Н.Г. Чернышевского
semenovani@sgu.ru*

Аннотация. В работе проводится анализ влияния внутреннего шумового воздействия на работу обученной нейронной сети решающей задачу классификации изображений, и оценка степени устойчивости сети к различным видам шумового воздействия, а также рассмотрение метода уменьшения влияния шумового воздействия на работу сети.

Ключевые слова: нейронные сети; аппаратные нейронные сети; внутренний шум; устойчивость; пул нейронов.

Искусственные нейронные сети (ИНС) в последние десятилетия стали основой решений значительного числа задач в области обработки информации, таких как распознавание образов, анализ временных рядов, автономное вождение и др. Однако дальнейшее увеличение их сложности и масштабов упирается в ограничения цифровых вычислительных систем, требующее всё большего роста энергозатрат. В качестве решения данной проблемы активно исследуется переход к аналоговым моделям, реализуемым на основе физических принципов — оптических, мемристивных и фотонных технологий. Их ключевое преимущество заключается в высокой скорости обработки данных и крайне низком энергопотреблении. Но с появлением аппаратных решений возникает новая фундаментальная проблема: влияние физических шумов, присущих любым аналоговым элементам. Эти шумы вызывают флуктуации сигналов нейронов, что ставит под вопрос стабильность и итоговую точность всей системы, особенно в таких ответственных задачах, как классификация.

Актуальность настоящей работы обусловлена активным развитием аппаратных нейросетевых технологий и необходимостью обеспечения их надёжности в условиях физически неизбежного внутреннего шума.

Шум в искусственных нейронных сетях представляет собой нежелательные случайные колебания сигнала, возникающие в процессе вычислений и передачи информации между элементами сети. Аппаратные нейросети подвержены внутреннему шуму, обусловленному флуктуациями напряжения, температурными дрейфами, нестабильностью компонентов и другими физическими источниками. Такой шум не связан с ошибками во входных данных и возникает независимо от обучающей выборки, влияя непосредственно на внутренние процессы обработки информации в нейронах.

С математической точки зрения, шум в нейроне можно описать как возмущение, накладываемое на выходное значение активационной функции. Пусть $f(x)$ — активационная функция, определяющая выходной сигнал нейрона в зависимости от суммы его входов x . При наличии аддитивного шума сигнал на выходе нейрона приобретает вид:

$$y_i^n = f(x_i^n) + \sqrt{2D_A} \xi_A(t) \quad (1)$$

где y_i^n – выход i -го нейрона в n -м слое, D_A – интенсивность аддитивного шума, а $\xi_A(t)$ реализация белого гауссовского шума.

Если шум является мультипликативным, он изменяет выход нейрона, масштабируя его значение. Тогда выражение для выходного сигнала принимает форму:

$$y_i^n = f(x_i^n) * (1 + \sqrt{2D_M} \xi_M(t)) \quad (2)$$

где D_M – интенсивность мультипликативного шума, а $\xi_M(t)$ – реализация белого гауссовского шума. В этом варианте флуктуации зависят от амплитуды самого сигнала, и, следовательно, более активные нейроны подвергаются более сильному искажению.

Независимо от типа, шум характеризуется параметром интенсивности D , от которого зависит его воздействие на сеть. При малых значениях этого параметра влияние шума может быть незначительным, однако при его увеличении точность нейросети заметно снижается, особенно если возмущение накладывается на скрытые слои. Это объясняется тем, что в скрытых слоях происходит формирование и передача представлений, критичных для корректной работы всей модели. Шум в этих слоях способен нарушить передачу признаков между уровнями, что в итоге приводит к потере классификационной способности сети.

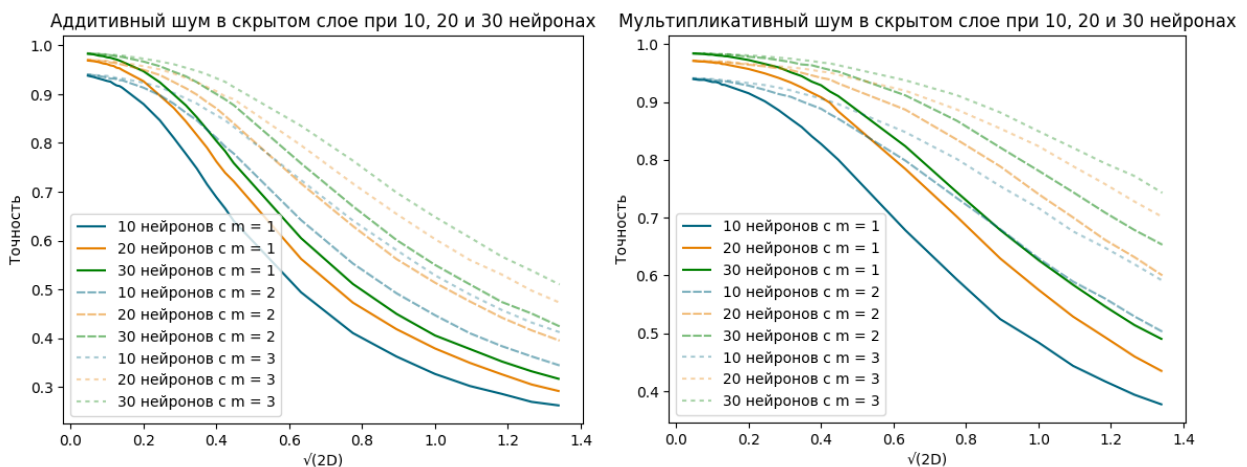


Рис.1. График зависимости точности работы сети от интенсивности шума

В качестве основной стратегии подавления шумового воздействия применяется техника нейронного пулинга, заключающаяся в формировании групп (пулов) из нескольких дублирующих нейронов, каждый из которых обрабатывает один и тот же входной сигнал и благодаря статистическому усреднению результатов работы этих нейронов, достигается эффект подавления шумовой компоненты. Стратегия была так названа из-за схожести с особенностями построения пулингового слоя (pooling layer в англоязычной литературе). Эта методика позволяет существенно повысить устойчивость сети к шуму, в частности к шумам, влияющим на скрытые слои.

Библиографический список

1. *Semenova N.I., Brunner D.* // *Chaos*. 2024. V. 34, № 5. P. 051101.
2. *Semenova N., Larger L., Brunner D.* // *Neural Networks*. 2022. V. 146. P. 151-160.
3. *Semenova N., et al.* // *Chaos*. 2019. V. 29, № 10. P. 103128.
4. *Lin X. et al.* // *Science*. 2018. V. 361. № 6406. P. 1004-1008.
5. *Mourgias-Alexandris G. et al.* // *Nature communications*. 2022. V. 13. № 1. P. 5572.
6. *Amato F. et al.* // *Journal of applied biomedicine*. 2013. V. 11. № 2. P. 47-58.
7. *Marabini R., Carazo J. M.* // *Biophysical Journal*. 1994. V. 66. № 6. P. 1804-1814.
8. *Samborska I. A. et al.* // *J. NanoPhotoBioSciences*. 2014. V. 2. № 14. P. 2347-7342.
9. *Wang X. et al.* // *Chaos*. 2024. V. 34. №. 3. P. 033108.

MODELING AND EVALUATION OF THE IMPACT OF ADDITIVE AND MULTIPLICATIVE NOISE ON THE PERFORMANCE OF HARDWARE NEURAL NETWORKS

Kolesnikov Ivan Dmitrievich

Saratov State University

kole200@yandex.ru

Semenova Nadezhda Igorevna

Saratov State University

semenovani@sgu.ru

Abstract. This work analyzes the effect of internal additive and multiplicative noise on the training process and robustness of a fully connected neural network. To this end, we compare accuracy under noise for a network trained with noise and for networks trained at different noise levels. We find that a small amount of noise during training increases robustness to interference, whereas excessive noise degrades performance.

Keywords: neural networks, noise, machine learning.

Internal noise in hardware neural networks

Internal noise is inevitable in hardware implementations of neural networks and can substantially alter a model's behavior in deployment. We study a deep feedforward network on the handwritten digit recognition task and compare the effects of additive and multiplicative noise, as well as their correlated and uncorrelated character across neurons within a layer. Noise was injected into the hidden layers during both training and testing. This made it possible to compare noise-free operation, training with noise, and their combinations, as well as to trace the dynamics of accuracy.

We consider a deep feedforward neural network for handwritten digit recognition. The network is trained on MNIST (60,000 training and 10,000 test 28×28 grayscale images); pixel values are normalized to the [0, 1] range and fed to the corresponding input neurons. The architecture includes three layers; the output layer contains 10 neurons with a softmax activation, forming a probability distribution over classes 0–9. The decision is determined by the index of the neuron with the maximum probability.

It was found that training without noise provides the best result on clean data, but such a model is less tolerant to interference. Adding a small amount of internal noise during training acts as mild regularization: under noisy conditions, accuracy decreases more slowly, and when test noise is reduced, it partially recovers. Excessive noise, by contrast, worsens accuracy.

Additive correlated noise introduced during training virtually removes the model's sensitivity to interference of the same type: quality depends little on the level of such noise at test time. For uncorrelated disturbances, improved robustness is achieved with small noise levels during training.

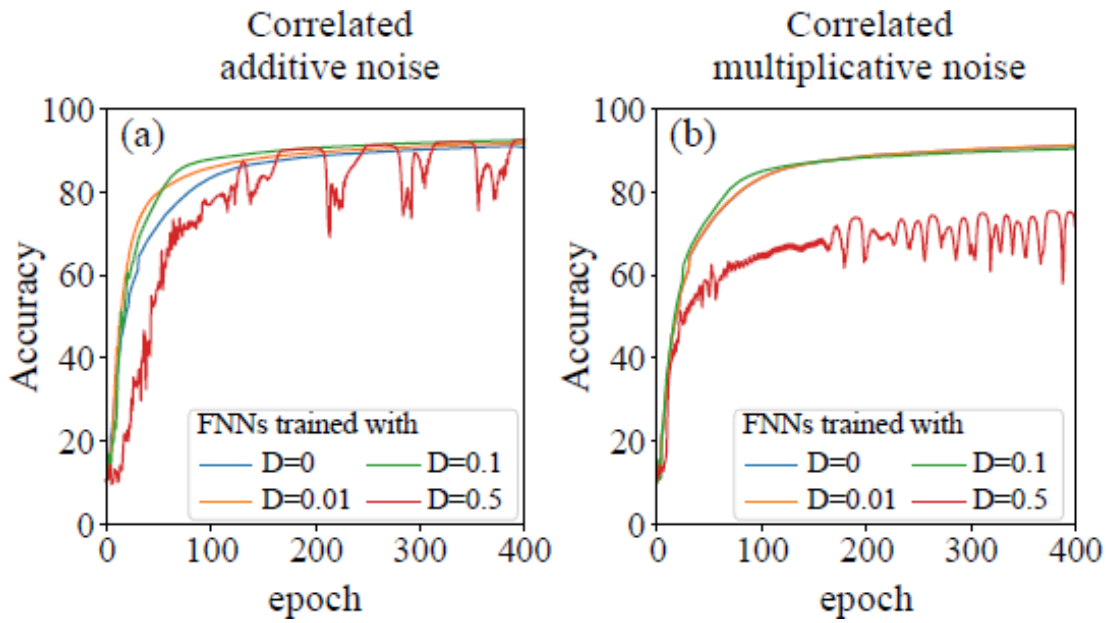


Fig. 1. Effect of correlated noise on an FNN.

The work is supported by the Russian Science Foundation, grant No. 25-72-10055, <https://rscf.ru/project/25-72-10055/>

References

1. Moskvitin V. M., Semenova N. I. Influence of noise on a recurrent neural network with nonlinear neurons. *Izvestiya VUZ: Applied Nonlinear Dynamics*. 2023;31(4).
2. LeCun Y., Cortes C., Burges C. J. C. The MNIST database of handwritten digits.
3. Mackey M. C., Glass L. Oscillation and chaos in physiological control systems. *Science*. 1977.

ОБУЧЕНИЕ ПРОСТОЙ НЕЙРОННОЙ СЕТИ: КОМПРОМИСС МЕЖДУ СКОРОСТЬЮ ОБУЧЕНИЯ И ПОТРЕБЛЕНИЕМ ПАМЯТИ

*Соловьев Игорь Александрович,
НИУ ВШЭ – Нижний Новгород
solocool46@gmail.com*

*Горюнов Олег Алексеевич,
НИУ ВШЭ – Нижний Новгород
oleggorunov2002@mail.ru*

*Клиньшов Владимир Викторович,
НИУ ВШЭ – Нижний Новгород
vladimir.klinshov@gmail.com*

Аннотация. В данном исследовании изучается компромисс между потреблением памяти и эффективностью оптимизации путём обобщения алгоритма Shampoo для поддержки матриц предобуславливания произвольного размера. Мы исследуем, как различные размеры предобуславливателей влияют на скорость обучения простой нейронной сети, решающей задачу регрессии. Результаты показывают сильную положительную корреляцию: более крупные матрицы неизменно приводят к более быстрому обучению.

Ключевые слова: предобусловленная оптимизация, градиентный спуск, мультистабильность.

Введение

Один из способов повысить эффективность классического градиентного спуска в задачах машинного обучения – это использовать так называемые предобуславливатели, представляющие собой особые матрицы, вычисляемые на каждом шаге обучения и используемые для изменения значений градиента. Классическими примерами такого подхода являются метод Ньютона [1], Adagrad [2] и другие.

Хотя предобуславливатель может радикально увеличить скорость обучения, использование полной матрицы может оказаться невозможным из-за ограничений по занимаемой памяти. Наиболее экономичными являются методы, использующие диагональные предобуславливающие матрицы, благодаря чему потребление памяти растёт линейно, а не квадратично с размером модели [3]. Другой способ уменьшения занимаемой памяти заключается в использовании блочно-диагональных матриц, блоки которых соответствуют различным тензорам модели [4].

В настоящей работе мы расширяем последний подход, основанный на малых матрицах предобуславливания, но выбираем их размеры произвольно, а не ограничиваясь размером тензоров.

Методы

Поскольку наша работа является расширением алгоритма оптимизации Shampoo, начнем с его краткого описания. Во-первых, предположим, что параметры модели образуют матрицу $W \in R^{m \times n}$, и мы стремимся минимизировать функцию потерь $f(W)$. Полноматричный предобуславливатель будет, очевидно, матрицей размерности $mn \times mn$, которая может быть чрезмерно большой. Вместо этого Shampoo использует две матрицы

меньшего размера $L_t \in R^{m \times m}$ и $R_t \in R^{n \times n}$, в результате чего их суммарный размер равен $m^2 + n^2$ вместо $m^2 n^2$.

Правило обновления весов формулируется следующим образом

$$\begin{aligned} L_t &= L_{t-1} + G_t G_t^T \\ R_t &= R_{t-1} + G_t^T G_t \\ W_{t+1} &= W_t - \eta L_t^{-\frac{1}{4}} G_t R_t^{-\frac{1}{4}} \end{aligned} \quad (1)$$

где t – номер шага, η – шаг обучения, G – вектор градиентов.

В настоящей работе мы развиваем подход, предложенный в [4] с использованием небольших предобуславливателей с формой, отличающейся от стандартной. Мы тестируем этот подход на перцептроне, в простой задаче аппроксимации функции $y = g(x) = \cos(\pi x)$. В модели $N = 100$ нейронов в скрытом слое, выход сети задается как:

$$\tilde{y} = \tilde{g}(x) = \sum_{i=1}^N W_i \sigma(w_i x + b_i) + B \quad (2)$$

где $\sigma(\)$ – сигмоидальная функция, W_i – выходные веса, w_i – входные веса, b_i – смещения скрытого слоя, и B – смещение выходного слоя. Они имеют размеры 1×100 , 1×100 , 1×100 и 1×1 соответственно. Хранение тех же самых параметров в тензорах другой формы позволяет использовать существенно более маленькие матрицы предобуславливания и, таким образом, экономить память. Различные возможные конфигурации тензоров и суммарный размер предобуславливателей представлены в таблице 1.

Таблица 1. Сравнение размера предобуславливателей для возможных форм тензора параметров сети	
Форма тензора	Размер предобуславливателя
1x100 (Adagrad)	301
1x100 (Shampoo)	30001
2x2x25	1900
2x5x10	388
4x5x5	199
2x2x5x5	175

Для всех конфигураций мы обучили сеть на 1000 эпох по алгоритму Shampoo с использованием функции ошибки MSE

$$C = \frac{1}{n} \sum_i^n (\tilde{g}(x_i) - g(x_i))^2 \quad (2)$$

где $n = 1000$ – размер всего датасета, а x_i – отдельный обучающий пример. Мы использовали стандартный шаг обучения $\eta = 0.1$ и повторили эксперимент 100 раз для каждой конфигурации весов с различных начальных условий.

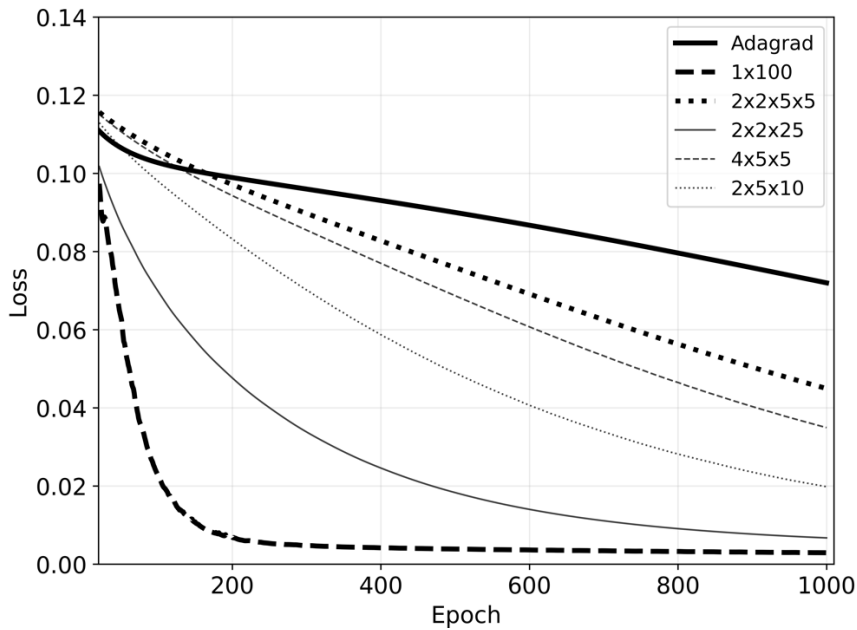


Рис.1. Кривые обучения (усредненные по 100 экспериментам) для разных конфигураций тензоров весов (указаны в легенде)

Результаты

На рисунке 1 показаны кривые обучения для различных конфигураций тензора весов, усредненные по 100 экспериментам с различными случайными начальными условиями. Для сравнения те же самые эксперименты были выполнены и со стандартным оптимизатором Adagrad. Из полученных результатов видно, что алгоритм оптимизации Shampoo демонстрирует более высокую скорость сходимости по сравнению с Adagrad в том числе и при других размерах предобуславливателя.

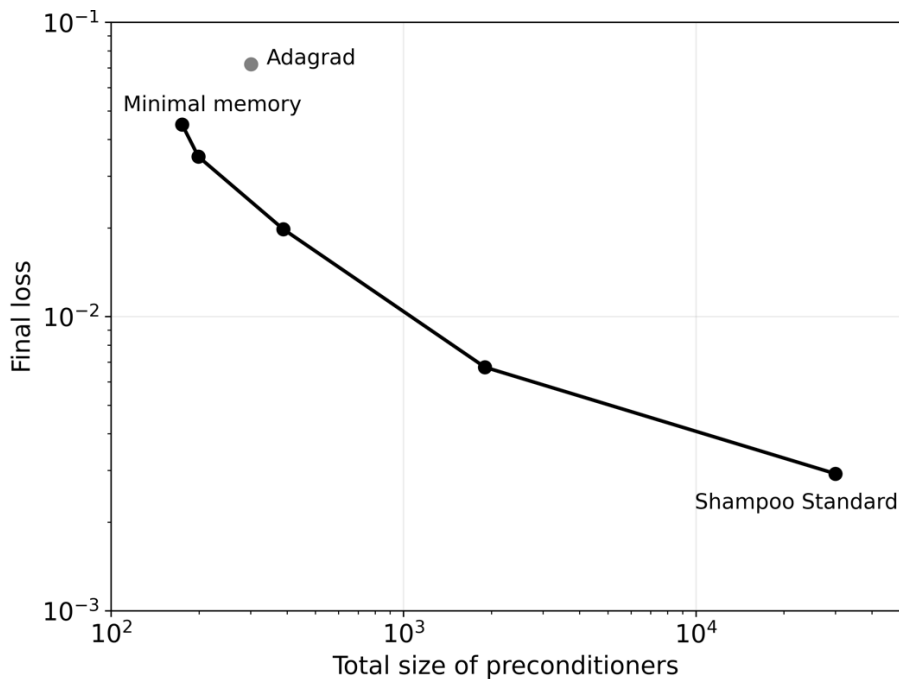


Рис.2. Итоговое значение функции ошибки после 1000 эпох (усредненной по 100 экспериментам) от общего размера предобуславливателей

Чтобы лучше проиллюстрировать эту тенденцию, была изучена также зависимость итогового значения функции ошибки после 1000 эпох (усреднённой по 100 экспериментам)

от общего размера предобуславливателей (общего количества элементов во всех предобуславливающих матрицах). Результаты представлены на рис. 2 и подтверждают корреляцию между скоростью оптимизации и общим размером предобуславливателей. Стоит обратить внимание, что производительность, по-видимому, растёт сублинейно с ростом объёма памяти: для достижения в 10 раз большей производительности требуется в 100 раз больше памяти.

Заключение

В настоящей работе была изучена производительность методов оптимизации, основанных на градиентных тензорах, в зависимости от размера используемых ими предобуславливателей. В качестве исследуемого оптимизатора был выбран Shampoo [4], но, в отличие от его оригинальной имплементации, использовались разные настраиваемые формы тензоров весов, а следовательно размеры предобуславливателей, что позволило существенно снизить объем используемой памяти. Результаты исследования показывают, что общее количество элементов в матрицах предобуславливания сильно коррелирует с производительностью: чем больше матрицы, тем быстрее происходит оптимизация.

Проверка наличия подобной корреляции в более крупных и сложных моделях машинного обучения — основная задача будущих исследований.

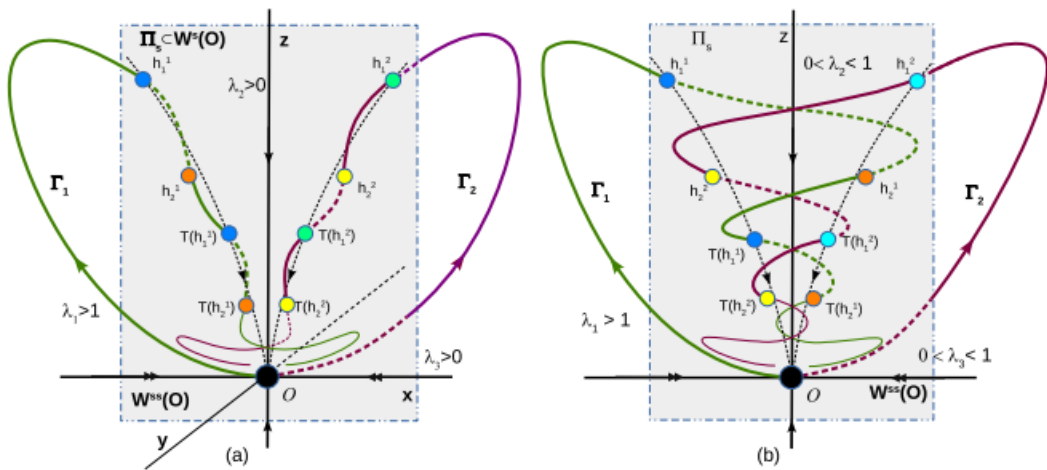
Работа выполнена в рамках проекта «Международное академическое сотрудничество» НИУ ВШЭ.

Библиографический список

1. P. E. Gill, W. Murray, M. H. Wright, Practical optimization. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2019.
2. J. Duchi, E. Hazan, and Y. Singer, “Adaptive subgradient methods for online learning and stochastic optimization”, J. Mach. Learn. Res., vol.12, pp. 257-269, 2011.
3. D. P. Kingma, J. Ba, “Adam: A method for stochastic optimization”, Int. Conf. Learning Representations, 2015.
4. V. Gupta, T. Koren, Y. Singer “Shampoo: Preconditioned stochastic tensor optimization”, Int. Conf. Machine Learning, pp. 1842-1850, 2018

Секция 3

«Динамические системы»



Session 3

«Dynamical systems»

SYMMETRY CONTROL FOR BISTABLE SYSTEMS: METHODS AND APPLICATIONS

Semenov Vladimir

Saratov State University, Saratov, Russia

semenov.v.v.ssu@gmail.com

Abstract. A manifold of methods for controlling the symmetry of single and coupled bistable oscillators is discussed. A list of considered approaches involves deterministic techniques, the action of noise and various coupling topologies. The associated applications such as the wavefront propagation control and spin-network implementations are also taken into discussion.

Keywords: bistability; symmetry; wavefront propagation; spin networks.

The nonlinearity's symmetry is a crucial factor for the dynamics of bistable oscillators. In particular, it is known that adjusting the symmetry properties enables enhancing and suppressing stochastic resonance [1], as well as controlling the velocity and direction of wavefront propagation [2]. A distinguishable application where the symmetry is principally important lies in the field of spin network (or spin-glass model) implementation. Over the past few decades, the hardware implementation of such spin-glass models have generated tremendous interest due to the prospects of solving NP-hard problems (problems with non-deterministic polynomial-time hardness). An important class of spin-glass models is the Ising machine that can be applied for a wide range of NP-hard problems. A key observation behind these methods is that optimization problems can be efficiently mapped onto specific classical Ising models [3]. Solving the specific optimization problem then translates into finding the ground state of the corresponding Ising Hamiltonian [4]. The Ising model describes an ensemble of binary spins. In the simplified case, the spin states are either in the spin up or the spin down, which represents two coexisting states of a bistable spin assumed to be a symmetric node. For this reason, it is especially important to reduce the spin's asymmetry when physically implementing a network of such systems.

The phenomenon being extremely sensitive to the presence of asymmetry in bistable systems is wavefront propagation: the larger is the asymmetry, the faster is the wavefront propagation. This principle works in bistable media [5,6], bistable delayed-feedback oscillators [7,8] and ensembles of coupled oscillators [9-11]. Thus, nonzero wavefront velocity indicates the presence of asymmetry. To achieve zero wavefront velocity in bistable systems, one can apply multiplicative [5,6,9] or additive [10,11] noise, vary the properties of interaction within networks such as the coupling strength or the coupling topology [9-11].

A universal approach for varying the velocity and direction of wavefront propagation through the symmetry control is proposed in paper [2]. It is based on the Taylor-series expansion of the system's nonlinearity function for further analyzing it as a polynomial series. In particular, the third-order Taylor series is applied to achieve the correspondence with the cubic nonlinearity. In such a case, the symmetric cubic nonlinearity is obtained at nonzero linear and cubic terms in the Taylor series, whereas other terms tend to zero. At the first stage, these conditions are analysed mathematically in order to derive the necessary expressions for the parameters of systems under study. After that, the obtained relationships and conditions are examined in the context of practical realizability. In particular, the symmetry must be achieved at adequate conditions such as an absolute temperature being higher than absolute zero, concentrations of substances within the range from 0 to 100%, positive values of light intensities, etc. The described methodology for the symmetry control was successfully applied to affect the propagating fronts on an example of an optical setup based on an optically-addressed spatial light modulator [2,12]. Meanwhile, the

possibilities of applying such method go beyond optics and wavefront propagation control. It must be noted, that the developed symmetry control scheme can be applied in the context of dynamical systems of different properties and nature, including networks of coupled oscillators where the nonlinearity of the network element can be expanded into the Taylor series.

This work is supported by the Russian Science Foundation (project No. 24-72-00054).

<https://rscf.ru/project/24-72-00054/>

References

1. *Li J.-h.* // Physical Review E. 2002. V. 66. №. 3. P. 031104.
2. *Semenov V.V., Porte X., Larger L., Brunner D.* // Physical Review B. 2023. V. 108. №. 2. P. 024307.
3. *Lucas A.* // Frontiers in Physics. 2014. V. 2. P. 5.
4. *Barahona F.* // Journal of Physics A. 1982. V. 15. №. 10. P. 3241.
5. *Engel A.* // Physics Letters A. 1985. V. 113. №. 3. P. 139.
6. *Méndez V., Llopis I., Campos D., Horsthemke W.* // Journal of Theoretical Biology. 2011. V. 281. №. 1. P. 31.
7. *Zakharova A., Semenov V.V.* // Neural Networks. 2025. V. 183. P. 106939.
8. *Semenov V.V.* // Chaos, Solitons and Fractals. 2025. V. 198. P. 116533.
9. *Semenov V.V., Jalan S., Zakharova A.* // Chaos, Solitons and Fractals. 2023. V. 173. P. 113656.
10. *Semenov V.V.* // Chaos, Solitons and Fractals. 2025. V. 198. P. 116533.
11. *Semenov V.V.* // Physics Letters A. 2025. V. 532. P. 130189.
12. *Semenov V., Porte X., Abdulhalim I., Larger L., Brunner D.* // Chaos. 2021. V. 31. №. 12. P. 121104.

ON DISCRETE LORENZ-LIKE ATTRACTORS IN THREE-DIMENSIONAL MAPS WITH AXIAL SYMMETRY

Gonchenko Alexander

Lobachevsky State University of Nizhny Novgorod, Mathematical

Center «Mathematics of Future Technologies», Russia;

HSE University, Nizhny Novgorod, Russia

e-mail: agonchenko@mail.ru

Abstract. We discuss two types of discrete Lorenz attractors, standard and twisted, in the case of three-dimensional quadratic maps with the axial symmetry and constant Jacobian.

Keywords: discrete Lorenz attractor, 3D map, axial symmetry, twisted attractor

We propose a new class of quadratic three-dimensional maps with the axial symmetry $S: x \rightarrow -x, y \rightarrow -y, z \rightarrow z$ and constant Jacobian that is additional to the well-known class of quadratic three-dimensional Hénon maps. We study dynamical properties and bifurcations in three-parameter families of these axial symmetric maps and show that such maps can have discrete Lorenz attractors, including those of new types in which one-dimensional unstable separatrices of a saddle fixed point are twisted. We also demonstrate typical scenarios that start with a stable fixed point and lead to the emergence of discrete Lorenz attractors of various types. Additionally, we show numerically that such attractors can be robustly chaotic, i.e. they are pseudohyperbolic and, in particular, preserve the maximal Lyapunov exponent to be positive for all smooth perturbations.

The corresponding three-parameter families are as follows

$$T_\alpha : \bar{x} = y, \bar{y} = \nu_1 x + \nu_2 y + yz, \bar{z} = \nu_3 z + \alpha y^2$$

where $\alpha = \pm 1$ and ν_1, ν_2, ν_3 are parameters. The Jacobian $J = -\nu_1 \nu_3$ of the map is constant. The point $O(0,0,0)$ is always the fixed point of the map and we consider the problem of studying so-called homoclinic attractors [1,2], which contain the point O .

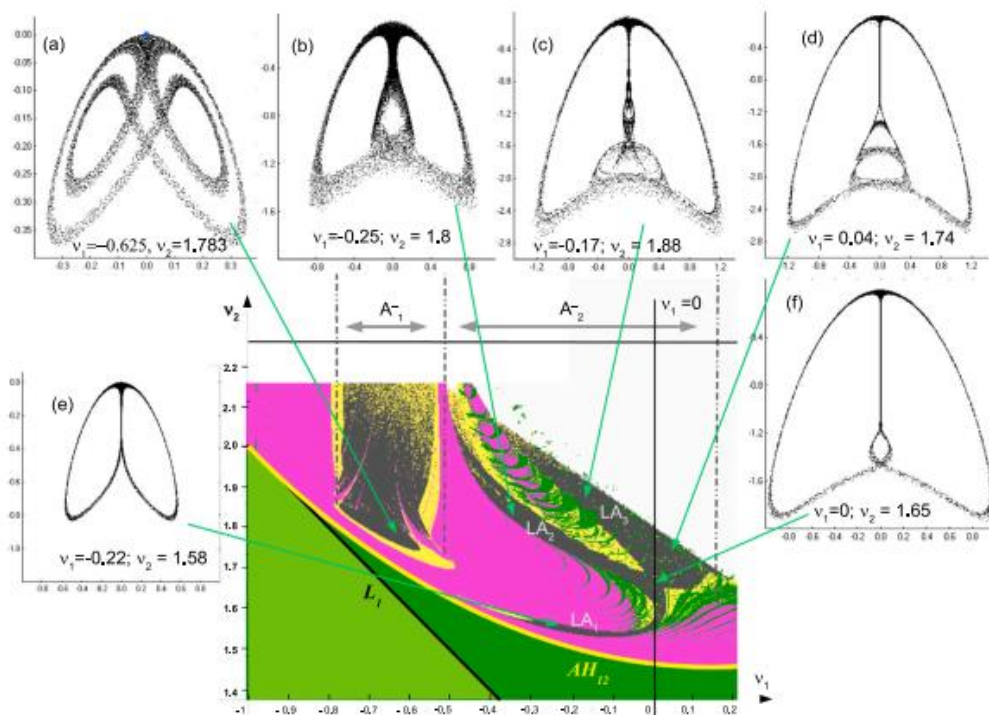


Fig.1.

We show that map T_α have very rich Lorenz-like dynamics. Namely, it can has both *standard discrete Lorenz attractors* as well as the so-called *twisted Lorenz attractors*, see Fig.1. In this figure main elements of the bifurcation diagram in the (v_1, v_2) -parameter plane for the map T in the case $\alpha=1$ with $v_3=0.8$ are shown accompanied by phase portraits of various discrete Lorenz-like attractors for the corresponding values of parameters (attractors in fig. (a) and (e) are standard, while others are twisted). As can see, phase portraits for standard and twisted discrete Lorenz-like attractors are different. We found that this difference is connected, in particular, with the fact that one dimensional unstable manifolds (separatrices) of the point O for such attractors have different structures as topological knots tied around the invariant z -axis, see Fig.2.

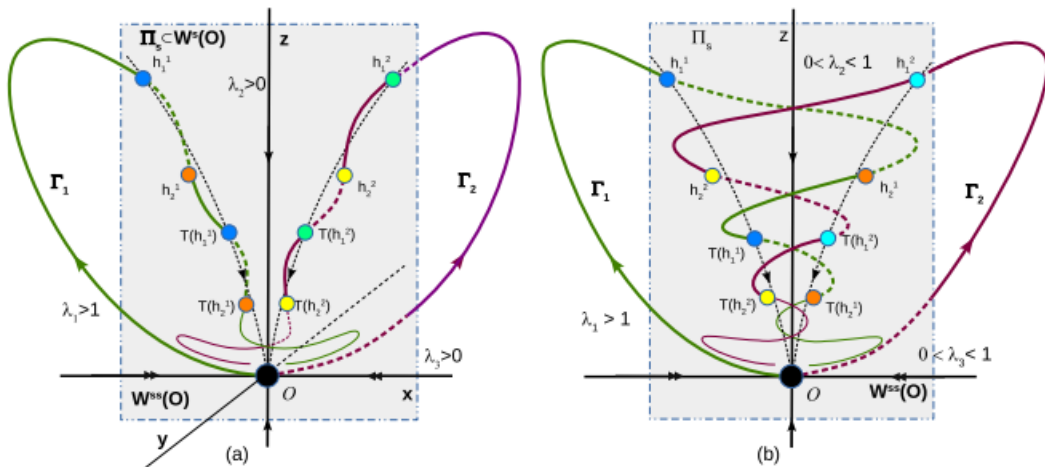


Fig. 2. . Illustrations for homoclinic butterfly structures in the cases when fixed point has multipliers $\lambda_1 > 1$, $0 < \lambda_3 < 1$ in the cases of a discrete Lorenz-like attractor (a) of standard type where separatrices are separated; (b) of twisted type where separatrices are tied in a knot

Constructing bifurcation diagrams, like in Fig.1, we trace also various bifurcation scenarios leading to the emergence of the Lorenz-like attractors including both orientable (for $J = -v_1 v_3 > 0$) and non-orientable (for $J = -v_1 v_3 < 0$) ones. We also shown that both the standard and twisted Lorenz-like attractors can be pseudohyperbolic, using the method of (ρ, φ) -cloud form [3,4]. In particular, all attractors shown in Fig.1 are pseudohyperbolic, except for the attractor of Fig.1(f) that is an attractor of map T with $J=0$ (for such 3D-endomorphisms the notion of pseudohyperbolicity is not applicable).

The work is supported by the Russian Science Foundation grant 24-11-00339.

References

1. Gonchenko A.S., Gonchenko S.V., Shilnikov L.P. // Rus J Nonlinear Dyn. 2012;8:3–28. [in Russian].
2. Gonchenko A., Gonchenko S. // Phys D. 2016;337:43–57.
3. Gonchenko A.S., Gonchenko S.V., Kazakov A.O., Kozlov A.D. // Int J Bifurcation Chaos. 2018;28(11):1830036.
4. Gonchenko S.V., Kazakov A.O., Turaev D. // Nonlinearity. 2021;34(4):2018–2047.

МУЛЬТИСТАБИЛЬНОСТЬ В НЕАВТОНОМНОМ ГЕНЕРАТОРЕ КВАЗИПЕРИОДИЧЕСКИХ КОЛЕБАНИЙ

Шамсина Алина Ленаровна
НИУ ВШЭ – Нижний Новгород
alshamsina@edu.hse.ru

Станкевич Наталия Владимировна,
ТМД НИУ ВШЭ – Нижний Новгород
nstankevich@hse.ru

Аннотация. В работе проводится анализ неавтономной модели генератора квазипериодических колебаний под периодическим импульсным воздействием. Локализованы области мультистабильности системы при различных значениях параметров. Обнаружена мультистабильность между различными торам, Построены бассейны притяжения аттракторов и карты показателей Ляпунова.

Ключевые слова: квазипериодические колебания, мультистабильность, динамические системы, показатели Ляпунова.

Мультистабильность представляет собой одно из важнейших явлений в теории нелинейных динамических систем, характеризующееся сосуществованием нескольких устойчивых динамических режимов при одних и тех же значениях управляющих параметров. Данное явление имеет фундаментальное значение для понимания сложного поведения нелинейных систем и находит широкое применение в различных областях науки и техники [1-2]. Возможны различные типы мультистабильности, когда сосуществуют периодические и непериодические аттракторы, или различные типы непериодических аттракторов. В рамках данной работы мы обсудим интересный тип мультистабильности, когда сосуществуют различные торы (в сечении Пуанкаре им отвечают инвариантные кривые), который был нами обнаружен в модели квазипериодического генератора под периодическим импульсным воздействием.

Модель неавтономного генератора квазипериодических колебаний записывается в виде следующей системы дифференциальных уравнений:

$$\begin{aligned} \ddot{x} - (\lambda + z + x^2 - \beta x^4) \dot{x} + \omega_0^2 x &= A \sum \delta(t - nT), \\ \dot{z} &= b(\varepsilon - z) - \kappa \dot{x}^2. \end{aligned} \quad (1)$$

Параметр λ определяет условия возникновения автоколебаний, ε характеризует энергетическое питание системы, ω_0 определяет одну из частотных компонент системы, параметр b связан с инерционными свойствами энергетической цепи, коэффициент κ отвечает за отбор энергии. Импульсное воздействие описывают следующие параметры: A - амплитуда импульсного воздействия, T - период, δ - дельта-функция Дирака.

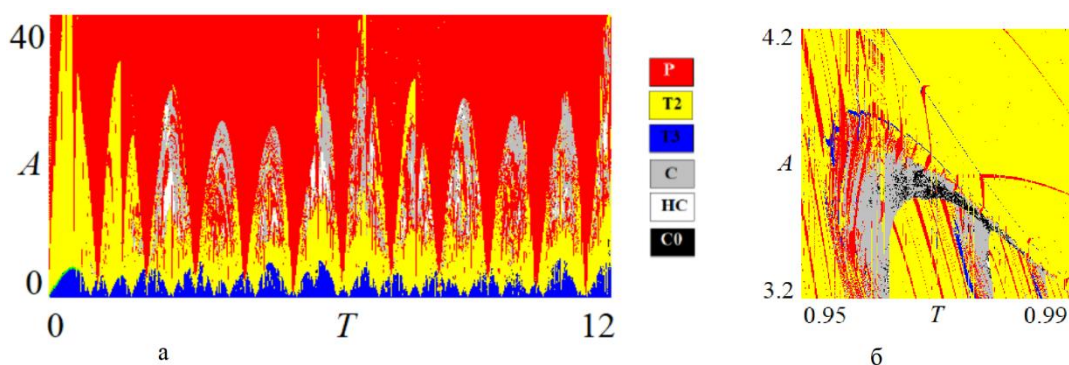


Рис.1. а) Карта показателей Ляпунова при $\lambda=1$, $\beta=1/18$, $\varepsilon=0$, $\omega_0=6$, $\kappa=0.02$;
б) Фрагмент карты показателей Ляпунова в окрестности области мультистабильности

Для исследования системы были построены карты показателей Ляпунова (рис.1), на которой было обнаружено, что в окрестности языков синхронизации формируется хаос по различным сценариям. В этих участках (рис.1б) карта показателей Ляпунова имеет «ребристую» структуру с эффектом «наложения». Такие резкие переходы свидетельствуют о наличии мультистабильности в системе. Действительно, если построить деревья бифуркаций и спектры показателей Ляпунова при данных параметрах с наследованием в двух направлениях сканирования, то можно увидеть, что система выходит на два разных аттрактора, один из которых отвечает тору, а второй удвоенному тору.

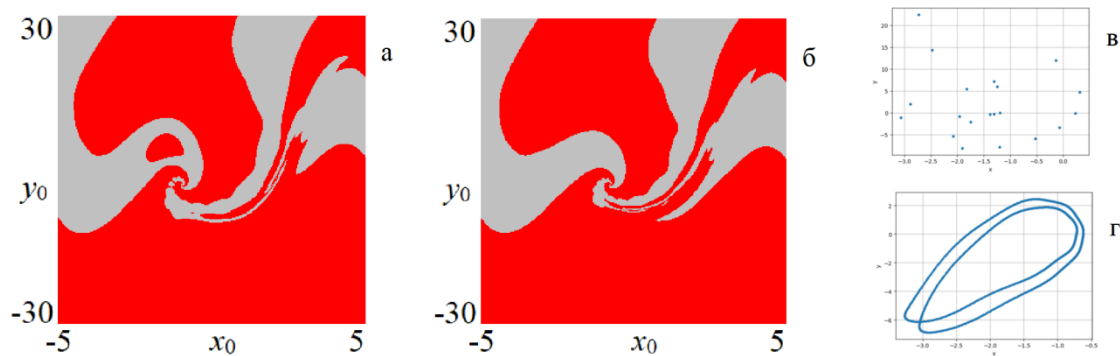


Рис.2. а) Бассейн притяжения при $z_0 = 1$, б) Бассейн притяжения при $z_0 = 0.5$; в) Сечение Пуанкаре красного аттрактора, г) Сечение Пуанкаре серого аттрактора.

Для периода $T = 0.97$ на рис.2а,б представлены изображения бассейнов притяжения: серым участкам соответствует аттрактор с удвоенной инвариантной кривой в сечении Пуанкаре (рис.2г), а красным участкам – режим, отвечающий резонансной однооборотной инвариантной кривой (в данном случае – это цикла периода 20) (рис.2в). Бассейны представляют собой слайсы трехмерного пространства, и рис.2а и рис.2б соответствуют различному выбору сечения. Бассейны имеют сложную структуру, небольшое изменение параметров системы приводит к переходу на новый аттрактор.

В работе будет обсуждаться, как возникают и исчезают сосуществующие аттракторы при изменении параметров в различных областях мультистабильности.

Работа выполнена в рамках проекта «Международное академическое сотрудничество» НИУ ВШЭ.

Библиографический список

1. Kloeden, Peter E., and Martin Rasmussen. Nonautonomous dynamical systems. No. 176. American Mathematical Soc., 2011.
2. Stefanovska, Aneta, and Peter VE McClintock, eds. Physics of biological oscillators: new insights into non-equilibrium and non-autonomous systems. Springer Nature, 2021.
3. Кузнецов А. П., Станкевич Н. В. Автономные системы с квазипериодической динамикой. Примеры и свойства: Обзор //Известия высших учебных заведений. Прикладная нелинейная динамика. – 2015. – Т. 23. – No. 3. – С. 71-93.

ON MIXED DYNAMICS OF REVERSIBLE MAPS WITH SYMMETRIC HOMOCLINIC DYNAMICS

Gonchenko Sergey

Lobachevsky State University of Nizhny Novgorod,
Mathematical Center «Mathematics of Future Technologies»,
HSE University, Nizhny Novgorod, Russia
sergey.gonchenko@mail.ru

Safonov Klim

HSE University, Nizhny Novgorod, Russia
safonov.klim@yandex.ru

Abstract: We discuss the so-called Reversible Mixed Dynamics conjecture (RMD-conjecture) and give its prove for one parameter families unfolding generally symmetric quadratic and cubic homoclinic tangencies.

Keywords: reversible map, symmetric orbit, homoclinic tangency, mixed dynamics, Newhouse region.

In the paper [1] it was proposed the following conjecture on mixed dynamics in the case of two-dimensional reversible maps.

RMD-conjecture. *Two-dimensional reversible maps with mixed dynamics are generic in Newhouse regions where maps with symmetric homoclinic and/or heteroclinic tangencies are dense.*

The existence of such Newhouse regions with mixed dynamics was proved in [2,3] for general one-parameter families containing maps with nontransversal heteroclinic contours containing points with the Jacobian less and greater than 1. This conjecture was also proved in [4,5] for various settings covering completely the so-called a priori nonconservative case including the case of a symmetric couple of quadratic homoclinic tangencies. In the present work we prove the conjecture for remaining two codimension-1 cases: (a) a symmetric quadratic homoclinic tangency and (b) a symmetric cubic homoclinic tangency. The prove in both these cases are quite similar and based on the following lemma.

Lemma. *Let f be a two-dimensional reversible map having a saddle fixed point O whose the stable and unstable invariant manifolds have a symmetric homoclinic tangency at the points of a homoclinic orbit G . Let f_m be a one parameter family which unfolds generally the initial tangency at $m = 0$. Then there exist an infinite sequence $\{m_k\}$ of values of m such that $m_k \rightarrow 0$ as $k \rightarrow \infty$ and the map $f(m_k)$ has a symmetric couple of quadratic homoclinic tangencies and these tangencies split generally when varying m .*

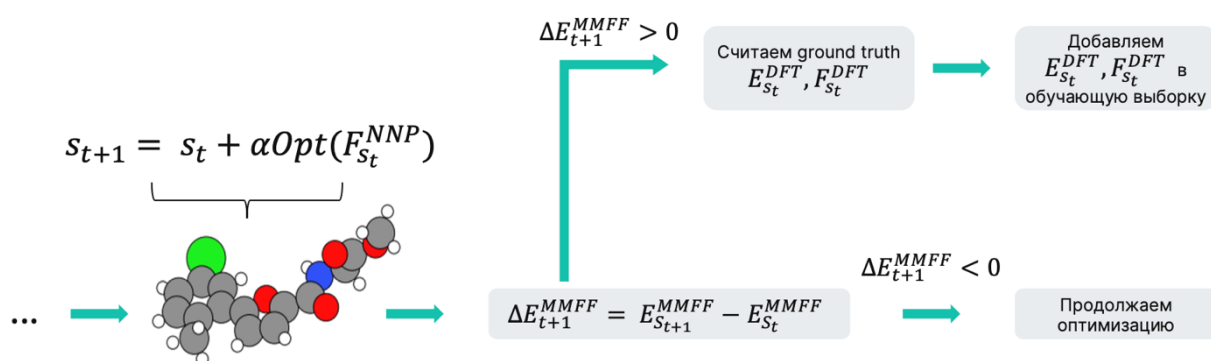
The work is supported by the Russian Science Foundation grant 24-11-00339.

References.

1. Delshams A., Gonchenko S.V., Gonchenko V.S., Lazaro J.T., Sten'kin O. // Nonlinearity, 26 (2013), 1—33.
2. Gonchenko S.V., Turaev D.V., Shilnikov L.P. // Steklov Math. Trans., 1997, v.216, pp. 76-125.
3. Lamb J.S.W., Stenkin O.V. // Nonlinearity, 17, No. 4, 1217–1244 (2004).
4. Delshams A., Gonchenko M., Gonchenko S., Lazaro J. // DCDS - A, 2018, 38 (9) : 4483-4507.
5. Gonchenko S., Gonchenko M., Synitskii I. // Izv. Math., 84:1 (2020), 23-51.

Секция 4

«Машинное обучение в приложениях»



Session 4

«Machine learning in applications»

ОЦЕНКА ГЛИКИРОВАНИЯ ТКАНЕЙ ПРИ ДИАБЕТЕ С ИСПОЛЬЗОВАНИЕМ ОПТИЧЕСКОГО ПРОСВЕТЛЕНИЯ И НЕЛИНЕЙНОЙ РЕГРЕССИИ

*Семенова Надежда Игоревна,
СГУ им. Н.Г. Чернышевского
semenovani@sgu.ru*

*Сергеев Константин Сергеевич,
СГУ им. Н.Г. Чернышевского
kssergeev@mail.ru*

*Слепнев Андрей Вячеславович,
СГУ им. Н.Г. Чернышевского
slepnevav@sgu.ru*

*Тучина Дарья Кирилловна,
СГУ им. Н.Г. Чернышевского
tuchinadk@mail.ru*

*Тимошина Полина Александровна,
СГУ им. Н.Г. Чернышевского
timoshina2906@mail.ru*

*Тучин Валерий Викторович,
СГУ им. Н.Г. Чернышевского
tuchinvv@mail.ru*

Аннотация. В данной работе были рассмотрены данные, в которых в кожу крыс вводился просветляющий агент. Данные разделялись на 3 основные группы: контроль (здоровые), 2-недельный диабет, 4-недельный диабет. Анализировалась динамика прозрачности образцов во времени для различных длин волн. В работе предлагаются методы нормирования полученных данных и применения нелинейной регрессии для их обработки. Статистика коэффициентов, полученные при помощи регрессии, позволяет однозначно разделить контрольные данные и данные для 4-недельного диабета.

Ключевые слова: оптическое просветление; фототерапия; спектральный анализ; диабет; диагностика; гиперосмотические агенты; глицерин; регрессия; машинное обучение; кожа; биофотоника.

Одним из новых и эффективных методов лечения рака кожи и других пролиферативных заболеваний, таких как псориаз, является фототерапия, при этом из-за многократного рассеяния воздействующих световых волн существует проблема ограничения проникновения излучения в глубину ткани. Эту проблему можно решить с помощью введения оптических просветляющих агентов, многие из которых являются гиперосмотическими. Простым и удобным методом повышения глубины проникновения оптического излучения является оптическое просветление биоткани, заключающееся во введении во внутритканевое пространство биосовместимых жидкостей (оптических просветляющих агентов) с достаточно высоким показателем преломления [1, 2]. Параллельно с этим было выявлено, что использование оптического просветления (введение просветляющих агентов) может быть применено для диагностики не только рака

кожи, но и других заболеваний. В данной работе этот эффект рассматривается в контексте диабета.

В данной работе были рассмотрены данные, в которых в кожу крыс вводился просветляющий агент (глицерин 70%). Данные разделялись на 3 основные группы: контроль (здоровые), 2-недельный диабет, 4-недельный диабет. Анализировалась динамика прозрачности образцов во времени для различных длин волн. Имеющиеся в распоряжении данные уже анализировались в работах [3,4].

Принципиальной особенностью данного исследования является то, что предлагается алгоритм автоматического определения наличия диабета, используя изменения спектральных характеристик во времени. Метод основан на применении машинного обучения (регрессии) для определения характерных особенностей и свойств законов, по которым спектральные характеристики стремятся к предельным значениям.

На Рис. 1 показана зависимость величин пропускания света длин волн 600, 700, 800, 900 нм от времени при четырех фиксированных длинах волн для кожи. Зеленый, синий и красный цвет на графике соответствует контрольным данным, 2- и 4-недельному диабету, соответственно. Для корректного отображения и возможности сравнения данные предварительно нормировались. Структура всех данных такова, что со временем значения спектра стремятся к определенному верхнему уровню S_n , на значение которого нормировалась вся реализация. Что же касается нормировки по времени, то для этого использовалась величина T_n , которая соответствовала моменту выхода на значение 95% от S_n .

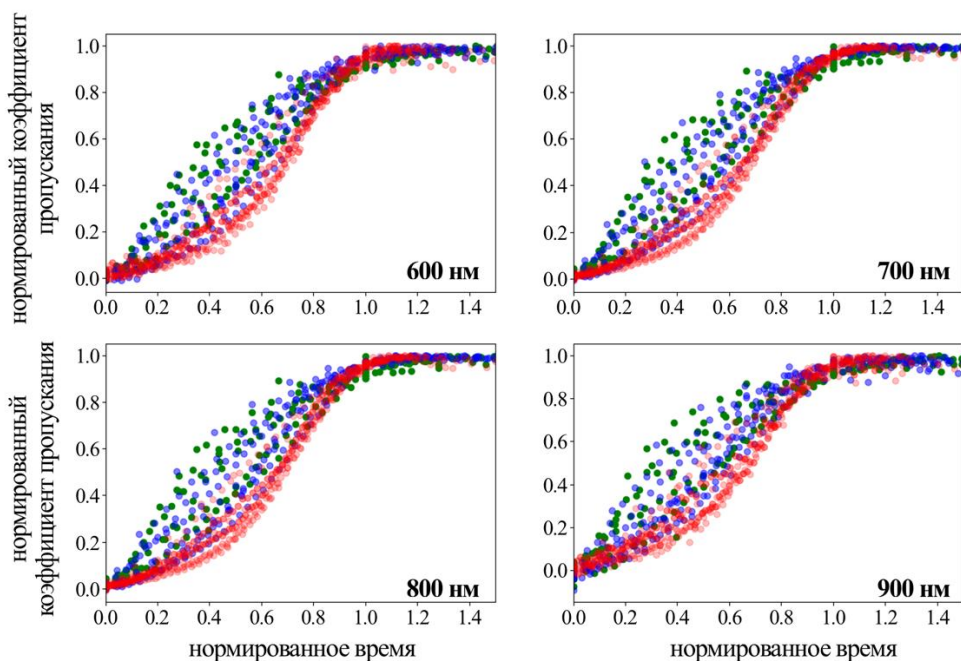


Рис.1. Зависимость спектральных характеристик от нормированного времени при фиксированной длине волны 600, 700, 800, 900 нм для контрольных данных (зеленый), 2-недельного диабета (синий) и 4-недельного диабета (красный).

Из графиков Рис. 1 видно, что качественно выход на предельное значение для контрольных данных и 4-недельного диабета довольно сильно отличается. Поэтому для аппроксимации этих участков была применена регрессия (аппроксимация полиномом) вида:

$$f(x) = C_0 + C_1x + C_2x^2 + C_3x^3. \quad (1)$$

На Рис. 2 (верхние графики) приведены результаты применения регрессии для всех трех групп данных для кожи при фиксированной длине волны 700 нм. Для каждого из образцов были получены значения коэффициентов C_0 , C_1 , C_2 , C_3 при соответствующих степенях x . Исходя из полученной статистики были построены диаграммы размаха (нижние графики Рис. 2) для каждого из коэффициентов.

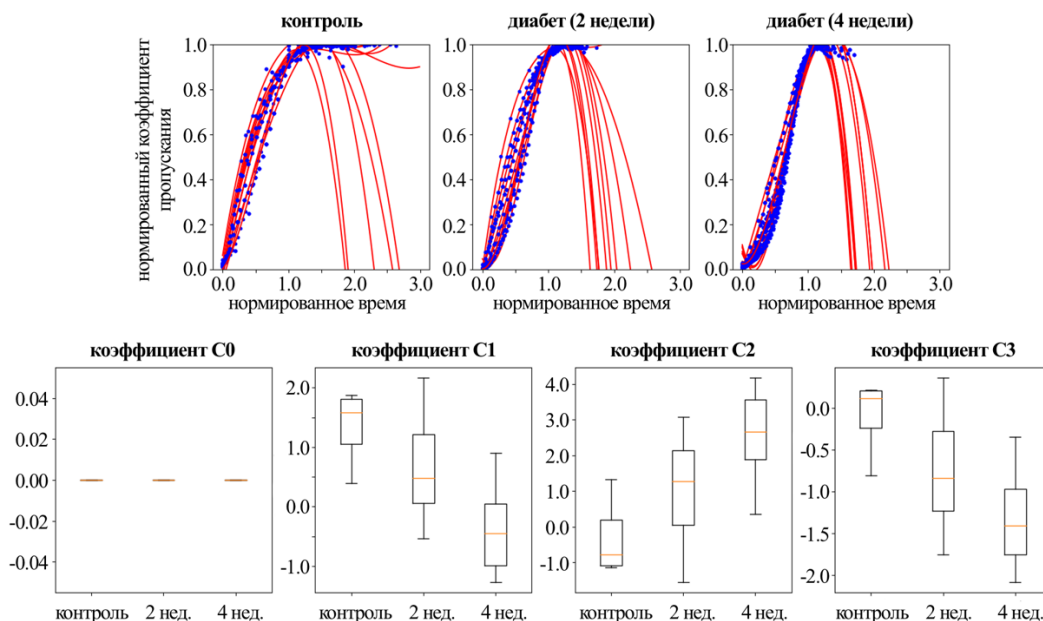


Рис.2. Точки на верхних графиков соответствуют тому же, что и Рис. 1 для 700 нм. Красными линиями показаны результаты аппроксимации (1). На нижних графиках приведены диаграммы размаха полученных коэффициентов C_0 , C_1 , C_2 , C_3 .

Из сравнения контрольных данных, а также данных для 2- и 4-недельного диабета видно, что при увеличении срока диабета происходит явное уменьшение коэффициентов C_1 и C_3 и рост коэффициента C_2 , при этом квантили распределений практически не перекрываются для контроля и 4-недельного диабета.

Работа выполнена при финансовой поддержке Минобрнауки России в рамках выполнения государственного задания (проект № FSRR-2023-0007).

Библиографический список

1. Тучина Д.К., Меерович И.Г., Синдеева О.А., Жердева В.В., Казачкина Н.И., Соловьев И.Д., Савицкий А.П., Богданов мл. А.А., Тучин В.В. // Квантовая электроника. 2021. Т. 51. № 2. С. 104–117.
2. Tuchin V.V., Zhu D., Genina E.A. Handbook of Tissue Optical Clearing: New Prospects in Optical Imaging. Boca Raton, FL : Taylor & Francis Group LLC ; CRC Press, 2022. 688 p.
3. Tuchina D.K., Bucharskaya A.B., Dyachenko, P.A., Dikht N.I., Terentyuk G.S., Tuchin V.V. Optical and Structural Properties of Biological Tissues under Simulated Diabetes Mellitus Conditions. Chapter 1 in “Biomedical Photonics for Diabetes Research”, CRC Press, 2022. Pp. 1–31.
4. Tuchina D.K., Bashkatov A.N., Bucharskaya A.B., Genina E.A., Tuchin V.V. // Journal of Biomedical Photonics & Engineering. 2017. V. 3. № 2. P. 020302.

ОПТИМИЗАЦИЯ С ПОМОЩЬЮ ГРАФОВЫХ НЕЙРОННЫХ СЕТЕЙ

Храбров Кузьма Агдреевич,

AIRI

khrabrov@airi.net

Цыпин Артем Андреевич,

AIRI

tsyпин@airi.net

Кадури́н Артем Аликович

AIRI

kadurin@airi.net

Аннотация. В работе приводится обзор методов оптимизации молекулярных конформаций. Приводится механизм ускорения классических численных методов с помощью графовых нейронных сетей. Приводятся методы, позволяющие снизить вычислительную стоимость сбора обучающей выборки для нейронных потенциалов, используемых для оптимизации.

Ключевые слова: графовые нейронные сети, нейронные потенциалы, теория функционала плотности.

1. Молекулярные конформации и энергия

Конформация — пространственное расположение атомов в молекуле; задаётся координатами $\mathbf{r}_i \in \mathbb{R}^3$ и типами атомов $z_i \in \mathbb{N}$.

Потенциальная энергия $E(s)$ определяет вероятность нахождения молекулы в данной конформации: $p(s) \propto \exp(-E(s)/kT)$.

Энергия и силы $F_i = -\partial E / \partial \mathbf{r}_i$ позволяют проводить молекулярную динамику (MD).

Конформация определяет свойства: дипольный момент, HOMO–LUMO gap, поляризуемость и др.

2. Методы расчёта энергии

Эмпирические поля сил (MMFF и др.) — быстрые, но неточные.

Квантово-химические методы (уравнение Шрёдингера) — точные, но вычислительно дорогие.

DFT — компромисс между точностью и скоростью, но неэффективен для систем >1000 атомов.

Более точные методы (например, CCSD(T)) ещё дороже и сложнее применимы для больших систем.

→ Нужны приближённые модели, способные предсказывать энергию и силы быстро и точно.

3. Графовые нейронные сети (GNN)

Молекула представляется как граф $G = (V, E)$, где вершины — атомы, а рёбра — связи.

Message Passing: нейросетевой подход, в котором на каждом шаге осуществляется обмен сообщениями (признаками) между соседними атомами. После n итераций этого процесса следует Readout-функция, предсказывающая целевые свойства. Readout-функция должна быть инвариантна (эквивариантна) к перестановке вершин.

GNN позволяют строить нейронные потенциалы (NNP) — аппроксимации поверхностей потенциальной энергии.

Пример NNP-архитектуры: SchNet

Три блока: Embedding → Interaction → Output.

Использует фильтры на основе расстояний между атомами и радиальных базисных функций.

Поддерживает периодические системы.

Обучается на энергии и силах, используя автодифференцирование.

4. Оптимизация конформаций

Традиционно: оптимизация по антиградиенту энергии, вычисленной DFT — дорого.

С NNP: энергия и силы аппроксимируются нейросетью → ускорение ~2000 раз.

Проблема — distribution shift: модель ошибается на новых конформациях.

5. Обучение и улучшение NNP

- Используются открытые датасеты: ∇^2 DFT (AIRI, 2024), SPICE (2023).
- Ошибки растут при оптимизации из-за смещения распределения.
- Решение: добавление DFT-траекторий в обучающую выборку.

→ Ошибка снижается при увеличении данных.

6. Подход GOLF (Gradual Optimization Learning Framework)

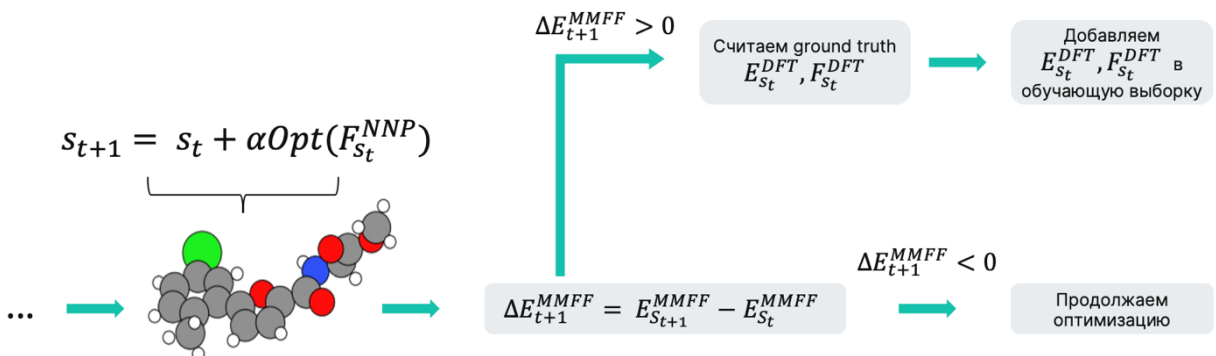


Рис.1. Схема оптимизации с помощью GOLF

- Итеративная оптимизация с нейронным потенциалом и суррогатным оракулом (например, MMFF).
- Если энергия снижается по мнению суррогатного оракула — продолжаем оптимизацию.
- Если растёт — добавляем текущую конформацию в обучающую выборку.
- Результат: снижение потребности в данных с 500 000 → 10 000 конформаций и экономия 9 CPU-лет.

7. Расширения и будущее (GOLF 2.0)

Если нет суррогатного оракула:

- Использовать промежуточные версии NNP как оракул.
- Использовать неуверенность модели для выбора, когда запускать дорогие расчёты.

8. Перспективы и дальнейшие направления

- Мультиагентное RL: каждый атом — агент, минимизирующий энергию (пример: MACS, 2025).

- Молекулярная динамика с GNN: графовые сети предсказывают силы и ускоряют MD.
- Перспектива — генеративные пропагаторы (GNN + диффузионные модели).

9. Выводы

1. Нейронные потенциалы — альтернатива DFT для задач оптимизации.
2. GNN потенциально способны масштабировать квантовую механику на миллионы шагов MD.
3. GOLF снижает стоимость обучения и повышает надёжность NNP.
4. В будущем — активное обучение, мультиагентные системы и гибридные DFT + NNP подходы.

Работа выполнена при поддержке института AIRI.

Библиографический список

1. Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., & Dahl, G. E. (2017, July). *Neural message passing for quantum chemistry*. In *International conference on machine learning* (pp. 1263-1272). PMLR.
2. Schütt, K., Kindermans, P. J., Sauceda Felix, H. E., Chmiela, S., Tkatchenko, A., & Müller, K. R. (2017). *Schnet: A continuous-filter convolutional neural network for modeling quantum interactions*. *Advances in neural information processing systems*, 30.
3. Schütt, Kristof, Oliver Unke, and Michael Gastegger. "Equivariant message passing for the prediction of tensorial properties and molecular spectra." *International Conference on Machine Learning*. PMLR, 2021
4. Smith, Daniel GA, et al. "PSI4 1.4: Open-source software for high-throughput quantum chemistry." *The Journal of chemical physics* 152.18 (2020).
5. Khrabrov, K., Ber, A., Tsypin, A., Ushenin, K., Rumiantsev, E., Telepov, A., ... & Kadurin, A. (2024). ∇^2 DFT: A Universal Quantum Chemistry Dataset of Drug-Like Molecules and a Benchmark for Neural Network Potentials. *arXiv preprint arXiv:2406.14347*.
6. Eastman, Peter, et al. "Spice, a dataset of drug-like molecules and peptides for training machine learning potentials." *Scientific Data* 10.1 (2023): 11.
7. Tsypin A. et al. *Gradual Optimization Learning for Conformational Energy Minimization* //The Twelfth International Conference on Learning Representations.

**ДЕТЕКТИРОВАНИЕ СОННЫХ ВЕРЕТЕН У ЗДОРОВЫХ СУБЪЕКТОВ И
ПАЦИЕНТОВ С ДИАГНОСТИРОВАННОЙ ЭПИЛЕПСИЕЙ С
ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
НА ОСНОВЕ ГИПЕРГРАФОВ**

*Смышляев Александр Павлович,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
aleksandr.smyshlaev@yandex.ru*

*Громов Николай Валерьевич,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
gromov@itmm.unn.ru*

*Шарков Артем Алексеевич,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
НИКИ педиатрии детской хирургии имени академика Ю.Е. Вельтищева – Москва
МГЦ Геномед – Москва
a.sharkov@pedklin.ru*

*Гребенюкова Анна Дмитриевна,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
Федеральный центр мозга и нейротехнологии ФМБА – Москва
МГЦ Геномед – Москва
grebenyukova@fccps.ru*

*Герасимова Светлана Александровна,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
gerasimova@neuro.nnov.ru*

*Лебедева Альбина Владимировна,
ННГУ им.Н.И.Лобачевского – Нижний Новгород,
ПИМУ Минздрава РФ - Нижний Новгород
lebedeva@neuro.nnov.ru*

*Смирнов Лев Александрович,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
lev.smirnov@itmm.unn.ru*

*Леванова Татьяна Александровна,
ННГУ им.Н.И.Лобачевского – Нижний Новгород
tatiana.levanova@itmm.unn.ru*

*Писарчик Александр Николаевич,
Политехнический Университет – Мадрид
alexander.pisarchik@ctb.upm.es*

Аннотация. В данной работе рассмотрено применение алгоритмов искусственного интеллекта для обнаружения сонных веретен в записях электроэнцефалографии (ЭЭГ) на основе гиперграфов, что позволило учесть сложные коллективные паттерны физиологических процессов. На наборе данных от видео-ЭЭГ лаборатории МГЦ «Геномед»

получено качество детектирования, сравнимое с последними результатами, описанными в литературе.

Ключевые слова: гиперграфы, ЭЭГ, веретена сна, машинное обучение, искусственный интеллект.

Материалы и методы. В качестве набора данных использовались 24 записи, предоставленные видео-ЭЭГ лабораторией МГЦ «Геномед» [1]. Каждая запись включает 19 биполярных отведений ЭЭГ, записанные с частотой дискретизации 199.5 Гц. В наборе данных 12 записей получены от пациентов с диагностированной эпилепсией и 12 от здоровых субъектов, включая следующие возрастные группы:

1. Ранний детский возраст (1–3 года): 2 пациента (возраст: 2 и 3 года),
2. Дошкольный возраст (4–6 лет): 3 пациента (возраст: 4, 5 и 6 лет),
3. Младший школьный возраст (7–11 лет): 6 пациентов (возраст: 7, 8, 8, 8, 8 и 10 лет),
4. Подростковый возраст (12–18 лет): 1 пациент (возраст: 12 лет).

Предобработка данных. С целью минимизации шума и выделения участков сонных веретен была выполнена предобработка, описанная в [1], включающая полосовую фильтрацию в диапазоне 10–16 Гц и последующую нормализацию амплитуд.

Анализ данных. Процедура анализа для каждой записи состояла из нескольких последовательных этапов.

1. *Сегментация сигнала и оконный анализ:* Для анализа динамики ЭЭГ-сигнал был сегментирован методом скользящего окна. Длина окна составляла 150 временных отсчетов (~ 0.75 с при частоте дискретизации 199.5 Гц), а перекрытие между соседними окнами — 120 отсчетов (~ 0.6 с). Данные значения длины окна и шага были определены экспериментально как гиперпараметры перебором по сетке.

2. *Оценка корреляции каналов:* Связь между парой сигналов $x_i(t)$ и $x_j(t)$ двух каналов в записи ЭЭГ в каждом временном окне рассчитывалась методом максимальной кросс-корреляции с временной задержкой, при этом максимальная временная задержка была выбрана равной $M=20$.

3. *Построение графов связности:* Для каждого временного окна матрица попарных кросс-корреляций между ЭЭГ-каналами бинаризовалась путем применения порогового значения (thr) для выделения наиболее сильных и информативных связей. На основе полученной матрицы смежности строился неориентированный граф, в котором вершины представляли каналы, а ребра — сильную корреляционную связь.

4. *Идентификация сетей высокого порядка и построение гиперграфа:* Для перехода от парных взаимодействий к групповым была построена гиперграфовая модель, в которой гиперребра формировались на основе максимальных клик, найденных с помощью алгоритма Брона-Кербоша в графе корреляций. В такой структуре каждое гиперребро объединяет структурно связанную группу ЭЭГ-каналов, демонстрирующих высокую степень корреляции в заданном временном окне.

5. *Извлечение признаков из гиперграфов:* Для количественной характеристики структурной роли каждой вершины (канала) в структуре гиперграфа были рассчитаны три метрики центральности высокого порядка (Hypergraph Centrality Coefficients, hcc) [2], а именно: коэффициент независимой связности, коэффициент кластеризации, коэффициент масштаба участия. Для каждого канала ЭЭГ в анализируемом окне был сформирован характеристический вектор.

6. *Формирование набора данных:* В результате выполнения данных шагов был сформирован итоговый набор данных, где каждый объект представлял собой временное окно, охарактеризованное вектором признаков. Этот набор данных послужил основой для последующего обучения и валидации моделей искусственного интеллекта.

Обучение классификатора. Для автоматической детекции сонных веретен на основе сформированных векторов гиперграфовых признаков был использован ансамблевый классификатор Случайный Лес (Random Forest, RF), зарекомендовавший себя в схожих задачах [2]. Этот же классификатор применялся на этапе предобработки для подбора оптимальных гиперпараметров.

Результаты. Оценка качества модели проводилась с помощью перекрестной проверки по методу исключения по одному пациенту (Leave-One-Patient-Out Cross-Validation, LOO-CV). Модель RF предсказывает вероятность наличия веретена в каждом временном окне; для получения бинарной метки был подобран оптимальный порог этой вероятности (RF-thr). Для учета временных зависимостей в данных в вектор признаков каждого окна были добавлены признаки двух соседних окон ($t-1$, $t+1$).

При использовании исключительно гиперграфовых признаков итоговая средняя пособытийная F1-мера составила 0.75. Для повышения точности классификации исходный набор признаков был расширен за счет добавления абсолютной спектральной мощности сигнала в диапазоне 10–16 Гц, рассчитанной для каждого канала методом Уэлча. Включение данного признака позволило значительно улучшить качество детекции, увеличив среднюю пособытийную F1-меру до 0.78.

Усредненная по всем пациентам F1-мера была снижена из-за низкой производительности модели на 3 из 24 записей. При исключении данных записей из расчета среднее значение F1-меры возрастает до 0.85 (Таблица 1), что подчеркивает высокий потенциал предложенного гибридного подхода на более однородной выборке.

	Только гиперграфовые признаки	Гиперграфовые признаки + дополнительный признак
Событийная F1-мера (24 записи)	0.75 (RF-thr = 0.75)	0.78 (RF-thr = 0.79)
Событийная F1-мера (21 запись)	0.81 (RF-thr = 0.75)	0.85 (RF-thr = 0.89)

Таблица 1. Событийная F1-мера

Исследование поддержано Минобрнауки России, Госзадание FSWR-2024-0005.

Библиографический список

1. Gromov N. V., Lebedeva A. V., Sharkov A. A., et al. Automated Sleep Spindle Analysis in Epilepsy EEG Using Deep Learning // Preprints. – 2025. – P. 2025092116.
2. Gao X., Zhu Y., Yang Y., et al. A seizure detection method based on hypergraph features and machine learning // Biomedical Signal Processing and Control. – 2022. – Vol. 77. – P. 103769.
3. Kwon H., Walsh K. G., Berja E. D., et al. Sleep spindles in the healthy brain from birth through 18 years // Sleep. – 2023. – Vol. 46, No. 4. – P. zsad017.

Программный комитет

Клинышов В.В. – НИУ ВШЭ, Нижний Новгород, Россия

Кузнецов Н.В. – СПбГУ, Санкт-Петербург, Россия

Купцов П.В. – НИУ ВШЭ, Нижний Новгород, Россия

Масловская А.Г. – Университет Иннополис, Иннополис, Россия

Починка О.В. – НИУ ВШЭ, Нижний Новгород, Россия

Семенова Н.И. – СГУ им. Н.Г. Чернышевского, Саратов, Россия

Станкевич Н.В. – НИУ ВШЭ, Нижний Новгород, Россия

Франович И. – Институт физики Белграда, Национальный институт
Республики Сербии, Белград, Сербия

International conference
“Machine learning and Nonlinear dynamics - 2025”
15.10.2025 – 18.10.2025

HSE University – Nizhniy Novgorod
25/12, Bolshaya Pecherskaya Street

15 October, Wednesday

14.00-14.45	Registration	Hall, 1st floor
14.45-15.00	Opening	305
15.00 – 18.00	Workshop «AI for medicine»	305
15.00-15.20	Kadurin A.A. AI for drug discovery	
15.20-15.40	Klinshov V.V. Deep learning for automatic spine radiograph labeling	
15.40-16.10	Bursov A.I. Experience in developing a medical device with AI in the ISP RAS	
16.10-16.25	Coffee-Break	
16.25-16.45	Gombolevskiy V.A. Medical AI: Algorithms that are changing practice	
16.45-17.05	Khrabrov K.A. Optimization of molecular conformations using graph neural networks	
17.05-17.25	Levanova T.A. Automated Sleep Spindle Analysis in Epilepsy EEG Using Deep Learning	
17.25-17.40	Semenova N.I. Assessment of tissue glycation in diabetes using optical clearing and nonlinear regression	
17.40-18.00	Lebedeva A.V. Development of a technology for restoring neuronal activity in the rat hippocampus using AI	

16 October, Thursday

14.45-15.00	Coffee-Break	
15.00 – 18.00	Section “Nonlinear dynamics for machine learning”	305
15.00-15.45	Semenova N.I. Internal noise in hardware neural networks of various types	
15.45-16.30 on-line	Kuptsov P.V. Application of machine learning methods to reconstruct a one-parameter family of dynamic models from a scalar time series	
16.30-17.00	Solovev I.A. Training a simple neural network: a tradeoff between speed and memory consumption	
17.00-17.30	Kolesnikov I.D. Modeling and evaluation of the impact of additive and multiplicative noise on the performance of hardware neural networks	
17.30-17.45	Myakov T.I. The influence of the differential equation integration method at the stage of training dataset formation on the properties of neuronetwork map	
17.45-18.00	Panyushev A.A. Machine learning for solving the problem of detecting multistable attractors	

International conference
“Machine learning and Nonlinear dynamics - 2025”
15.10.2025 – 18.10.2025

HSE University – Nizhniy Novgorod
25/12, Bolshaya Pecherskaya Street

17 October, Friday

14.45-15.00	Coffee-Break	
15.00 – 18.00	Section “Dynamical Systems”	305
15.00-15.45	Semenov V.V. Symmetry control for bistable systems: methods and applications	
15.45-16.15 on-line	Čubrović M. Semiclassical chaos and superdiffusion in Bose-Hubbard lattices	
16.15-16.45	Gonchenko A.S. On discrete Lorenz-like strange attractors in systems with axial symmetry	
16.45-17.00 on-line	Franović I. How ketogenic diet impacts multistability and allows for control of seizure-like activity: new insights from a next-generation neural mass approach	
17.00-18.00	Poster Session, Coffee-Break	Hall, 3rd floor

Poster session

Smyshlyaev A.P. Sleep spindles detection using hypergraph learning

Klimenko K.E. Reconstruction of the mean-field dynamics of neuron-astrocyte interaction using the dend-PLRNN architecture

Gromov N.V. On some properties of output matrices in reservoir computings

Bagautdinova E.R. Comparative analysis of multistable properties of the Hodgkin-Huxley model and its neural network map

Moskvitin V.M. The Impact of Internal Noise on the Performance of an Artificial Neural Network: Application of a Noise Reduction Strategy

Shamsina A.L. Multistability in non-autonomous generator of quasi-periodic oscillations

Qasim M. Classification of dynamical regimes in the Chialvo map with Machine Learning

Safonov K. On mixed dynamics of reversible maps with symmetric homoclinic dynamics

18 October, Saturday

14.45-15.00	Coffee-Break	
15.00 – 17.30	Section “Machine learning for tasks of nonlinear dynamics”	305
15.00-15.45	Klinshov V.V. Nonlinear dynamics and Machine learning: experience of interaction	
15.45-16.30	Karavaev A.S. Biosignal models for parameter tuning and verification of data analysis methods	
16.30-17.00	Posnenkova O.M., Bondarenko I.A. Analysis of the phase structure of the electrocardiogram under stress and at rest using machine learning methods	
17.00-17.30 on-line	Ishbulatov Yu.M. Neural network method for estimating the strength of directional interaction between self-oscillating systems	
17.30-18.00	Closing	