

УТВЕРЖДАЮ

Первый проректор – проректор по научной деятельности
федерального государственного автономного
образовательного учреждения высшего образования
«Казанский (Приволжский) федеральный университет»
Доктор физико-математических наук, профессор


Д.А. Тажурский
«...» апреля 2026 года


ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

Федерального государственного автономного образовательного учреждения высшего образования «Казанский (Приволжский) федеральный университет» на диссертационную работу ЮРТИНА Алексея Артемьевича «Нейросетевые методы восстановления потоковых данных», представленную на соискание ученой степени кандидата физико-математических наук по специальности 2.3.5 – Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Актуальность темы диссертации

Выполненное диссертационное исследование направлено на разработку новых методов восстановления потоковых данных, представленных в форме временных рядов. Актуальность этой тематики определяется стремительным ростом объемов потоковых данных, наблюдаемым в настоящее время, и необходимостью их интеллектуального анализа. Пропуски в потоковых данных, возникающие вследствие технических сбоев и ряда других объективных причин, снижают достоверность аналитических выводов и требуют разработки эффективных методов восстановления пропущенных данных. Важным направлением исследования, выполненного диссертантом, является повышение точности нейросетевых моделей восстановления данных за счет построения подходящей функции потерь, соответствующей рассматриваемой задаче, и разработки методов прогнозирования точности и времени обучения нейросетевых моделей.

Цель и задачи диссертации

Цель диссертационного исследования А.А. Юртина заключалась в разработке и исследовании новых методов восстановления потоковых данных, представленных в форме многомерных временных рядов, на основе совместного использования нейросетевых моделей и поведенческих шаблонов. Для достижения поставленной цели соискателем были поставлены и успешно решены следующие задачи.

1. Разработаны следующие нейросетевые модели, методы и алгоритмы, предназначенные для восстановления потоковых данных:

- нейросетевые методы восстановления многомерных временных рядов в офлайн и онлайн режимах;
- построена функция потерь для обучения нейросетевых моделей восстановления временных рядов;
- предложен метод прогнозирования ошибки и времени обучения нейросетевых моделей восстановления временных рядов.

2. Проведены вычислительные эксперименты, результаты которых продемонстрировали эффективность предложенных нейросетевых моделей, методов и алгоритмов.

Содержание диссертации

Диссертация состоит из введения, четырех глав, заключения, списка литературы и трех приложений. Объем диссертации составляет 193 страницы, объем библиографии 168 наименований.

В первой главе рассмотрены основные этапы жизненного цикла нейросетевых моделей, применяемых для восстановления потоковых данных, представленных в форме временных рядов. Проведен анализ существующих методов и инструментальных средств, используемых для восстановления пропущенных значений временных рядов. Систематизированы современные подходы к решению этой задачи. Рассмотрены применяемые функции потерь и методы предварительной оценки качества нейросетевых моделей. На основе этого анализа выявлены ключевые ограничения существующих решений и определены направления проведения дальнейших исследований.

Во второй главе представлены новые нейросетевые методы восстановления потоковых данных, предназначенные для различных режимов обработки временных рядов. Предложен оригинальный метод SANNI для онлайн

восстановления, обеспечивающий заполнение пропусков на основе предыдущих наблюдений. Разработан метод SAETI для офлайн восстановления, позволяющий обрабатывать пропущенные значения во временных рядах независимо от их расположения. Описаны архитектуры нейросетевых моделей, реализующих ключевые этапы обработки данных в рамках предложенных методов.

В третьей главе представлены новые инструментальные средства, предназначенные для оценки и повышения точности нейросетевых моделей восстановления потоковых данных, представленных в форме временных рядов. Предложена новая функция потерь MPDE, учитывающая поведенческие особенности при сравнении подпоследовательностей временного ряда в процессе обучения нейросетевых моделей. Разработан параллельный алгоритм вычисления новой функции потерь. Представлен метод tsGAP прогнозирования ошибки и времени обучения, предназначенный для предварительной оценки точности восстановления.

Четвертая глава содержит результаты вычислительных экспериментов, проведенных соискателем. Представленные эксперименты подтвердили эффективность предложенных методов и функции потерь. Сравнительный анализ, представленный в этой главе, показал, что метод SANNI обеспечивает высокую точность онлайн восстановления временных рядов и превосходит существующие аналоги, а метод SAETI демонстрирует преимущество при офлайн восстановлении данных. Нейросетевые модели, обученные с использованием функции потерь MPDE, обеспечили более высокую точность восстановления по сравнению с моделями, обученными с другими аналогичными функциями потерь. Метод tsGAP продемонстрировал более точное прогнозирование ошибки и времени обучения нейросетевых моделей восстановления потоковых данных по сравнению с существующими методами.

Автореферат диссертации корректно и в полном объеме отражает основные положения, содержание и результаты диссертационного исследования.

Соответствие специальности. Содержание и результаты диссертации соответствуют паспорту специальности 2.3.5 — «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных се-

тей»: п. 4. Интеллектуальные системы машинного обучения, управления базами данных и знаний, инструментальные средства разработки цифровых продуктов.

Научная новизна диссертационного исследования А.А. Юртина заключается в следующих результатах.

1. Разработаны новые нейросетевые методы восстановления потоковых данных, представленных в форме временных рядов, предназначенные для работы в онлайн и офлайн режимах и впервые использующие поведенческие шаблоны.

2. Предложена новая функция потерь для обучения нейросетевых моделей восстановления временных рядов, учитывающая поведенческое сходство подпоследовательностей в процессе обучения. Разработан эффективный параллельный алгоритм ее вычисления, обеспечивающий возможность применения предложенной функции потерь в современных фреймворках глубокого обучения.

3. Разработан новый метод прогнозирования ошибки и времени обучения нейросетевых моделей восстановления потоковых данных, представленных в форме временных рядов, основанный на представлении архитектуры модели в виде направленного ациклического графа и позволяющий выполнять предварительную оценку точности моделей без необходимости их непосредственного обучения.

Теоретическая значимость диссертационной работы состоит в следующем.

- В нейросетевых методах восстановления потоковых данных, разработанных соискателем, использованы оригинальные архитектуры нейросетевых моделей, позволяющие применять шаблоны временных рядов в качестве дополнительного источника информации.

- Предложенная функция потерь существенно учитывает поведенческую схожесть временных рядов в процессе обучения.

Практическая значимость диссертационной работы заключается в том, что разработанные нейросетевые модели и алгоритмы могут применяться для восстановления временных рядов в интеллектуальных системах обработки потоковых данных, а использование предложенных методов позволит автоматизировать подготовку потоковых данных, содержащих пропуски.

Обоснованность и достоверность результатов, содержащихся в диссертации, подтверждены проведенными вычислительными экспериментами на реальных и синтетических наборах данных. Эти эксперименты выполнены в соответствии с общепринятыми методологическими подходами и подробно описаны. Соискателем проведено сравнение полученных результатов с известными аналогами. Для обеспечения воспроизводимости полученные результаты реализации алгоритмов и экспериментальные данные размещены в открытых интернет-репозиториях.

Публикации и апробация

Результаты диссертационного исследования опубликованы в пяти научных работах в журналах категорий K1 и K2 Перечня ВАК. Из них 4 статьи опубликованы в российских научных журналах, входящих в Ядро РИНЦ, в том числе 1 статья в зарубежном журнале, приравненном к журналам категории K1 Перечня, индексируемом в квартиле Q2 базы данных Scopus. Опубликованные работы полно отражают основные положения и научные результаты диссертации. Материалы исследования прошли апробацию на 5 международных и всероссийских научных конференциях.

В качестве замечаний, относящихся прежде всего к оформлению диссертации и не затрагивающих содержание и значимость основных ее результатов, можно отметить следующие.

1. Новый метод SANNI восстановления временного ряда работает в режиме онлайн и при длительных сериях пропусков использует предыдущие восстановленные значения. Это может приводить к накоплению ошибки. Необходимо пояснить, почему предложенный метод является устойчивым к длительным сериям пропусков и каков механизм контроля надежности восстановленных значений.

2. Новая функция потерь MPDE позволяет обучать нейросетевые модели с учетом поведенческого сходства подпоследовательностей временного ряда. Почему же в вычислительных экспериментах качество нейросетевых моделей, обученных с использованием MPDE, оценивается с помощью другой, традиционной метрики RMSE, которая не использовалась в процессе обучения.

3. Некоторые части текста дублируют друг друга, особенно когда идет перечень того, что будет обсуждено или рассмотрено ниже. Все такие инфор-

мативные части (где сказано, что и где будет в диссертации далее) есть во введении, они достаточно подробно составлены и, как нам представляется, не требуют повторений. Например, вводный текст в раздел на стр. 32 без ущерба можно было перенести на стр. 22 или вообще исключить. То же относится к информативному материалу на стр. 68.

4. В диссертации (глава 3) есть Лемма 1 и несколько утверждений. Эта лемма представляется тривиальной и доказательства не требует (речь идет о простом выводе простого числового неравенства). Последующие утверждения как математические объекты также являются простыми и получаются несложными выкладками. И лемма, и названные утверждения нисколько не повышают математический уровень и значимость диссертации и без ущерба могли бы быть просто переведены в технические выкладки.

5. В тексте достаточно часто встречаются банальные или, наоборот, совсем непонятные фразы, которые требуют многократного прочтения для понимания и, скорее всего, переписывания. Например, «... поскольку квадрат любого действительного числа неотрицателен» (стр. 87); «уменьшаемое второго множителя правой части формулы ...», «в соответствии с цепным правилом ...» (стр. 92); «для наивного (прямого) вычисления вклада» (стр. 93); «частная производная MPDE по переменной вычисляется как среднее значение производных расстояний между окнами» (стр. 94); «как сумма разностей между точками восстановленного и истинного временного ряда, деленных на евклидово расстояние соответствующих окон ...» (стр. 95); «входными данными метода являются архитектура целевой нейросетевой модели и вектор параметров ее обучения» (стр. 101). Текст диссертации, в целом написанный грамотно и хорошим языком, выиграл бы, если бы таких фраз было меньше.

Указанные замечания не влияют на общую положительную оценку диссертационного исследования А.А. Юртина.

ЗАКЛЮЧЕНИЕ

Диссертация А.А. Юртина представляет собой самостоятельную и законченную научно-квалификационную работу, содержащую результаты, обладающие научной новизной. Соискателем разработаны новые нейросетевые методы восстановления многомерных временных рядов, обладающие значимой практической ценностью для решения задач обработки потоковых данных. По уровню решаемых задач, содержанию и полученным результатам диссертационная работа соответствует требованиям, установленным Положением о порядке присуждения ученых степеней, в том числе пункту 9.

А.А. Юртин заслуживает присуждения ученой степени кандидата физико-математических наук по научной специальности 2.3.5 – «Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей».

Диссертация и отзыв обсуждены на заседании кафедры цифровой аналитики и технологий искусственного интеллекта Института информационных технологий и интеллектуальных систем (ИТИС) федерального государственного автономного образовательного учреждения высшего образования «Казанский (Приволжский) федеральный университет» (КФУ), протокол № 4 от «7» апреля 2026 г.

Отзыв составил

Профессор кафедры цифровой аналитики и технологий искусственного интеллекта Института ИТИС КФУ, д. ф.-м. н., профессор, Заслуженный деятель науки Российской Федерации, Заслуженный деятель науки Республики Татарстан

Елизаров Александр Михайлович

Тел.: +7(937) 6151553; e-mail: amelizarov@gmail.com

420008, г. Казань, ул. Кремлевская, д. 35

Наименование организации, предоставившей отзыв:

Федеральное государственное автономное образовательное учреждение высшего образования «Казанский (Приволжский) федеральный университет»

Адрес: 420008, Россия, Республика Татарстан, г. Казань, ул. Кремлевская, д. 18

Телефон: + 7 (843) 292-72-16

Email: public.mail@kpfu.ru

www: <https://kpfu.ru/>

