ПРИМЕНЕНИЕ ЦИФРОВЫХ ДВОЙНИКОВ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Яровой Р. В.[1], Карганов В. В.[2], Лукашенок В. И.[3]

DOI:10.21681/3034-4050-2026-1-42-51

Ключевые слова: *цифровой двойник, телекоммуникационная сеть, обнаружение аномалий, синтетические данные, машинное обучение, телеметрия, предиктивное обслуживание, искусственный интеллект.*

Аннотация

Цель работы: *анализ возможности применения технологии цифровых двойников для обнаружения аномалий в телекоммуникационных сети с применением технологий искусственного интеллекта и обучением на синтетических данных.*

Метод исследования: *основан на применении математического моделирования, который предполагает созидание цифрового двойника телекоммуникационной сети посредством применения синтетически сгенерированных данных, которые имитируют поведение сети, а для обнаружения аномалий используется машинное обучение – обучение LSTM-автоэнкодера, с последующей оценкой качества обнаружения аномалий на основе метрик (precision, recall и F1-score).*

Результаты исследования: *включают разработку модели цифрового двойника телекоммуникационной сети, а также генерацию датасета на основе синтетических данных. В ходе исследования продемонстрировано, что может быть создан цифровой двойник телекоммуникационной сети, синтетически сгенерированы данные для обучения модели машинного обучения, при помощи которой могут быть обнаружены различные аномалии. В дальнейшем модель машинного обучения может быть использована в реальных телекоммуникационных сетях связи для обнаружения разного рода аномалий.*

Научная новизна: *заключается в разработке методики генерации синтетических данных с аномалиями, которые адаптированы под телекоммуникационные сети, а также в использовании цифрового двойника как инструмента для тестирования алгоритмов машинного обучения для обнаружения аномалий.*

Введение

Телекоммуникационные сети, которые используются в настоящее время для обмена информацией, охватывают все большее количество территорий, трансформируются и модифицируются с течением времени. Они предоставляют возможность огромному числу пользователей обмениваться разного рода информацией (звук, изображения, видео), являются основой для функционирования от мобильных сетей связи до интернета вещей, от автономного транспортного сообщения до услуг телеприсутствия в сфере здравоохранения. С учетом всеобъемлющего использования в подавляющем числе сфер жизни необходимость автоматизации наблюдения за различными параметрами сети, а также управления ими и адаптации становится важным элементом для предоставления качественных услуг связи.

С учетом того, что традиционные методы обнаружения аномалий [1], с ростом объема трафика в телекоммуникационных сетях, становятся менее эффективными, необходим качественно новый переход от традиционных методов к более модернизированным. Более важным является вопрос обнаружения разного рода аномалий, таких как DDoS-атаки, перегрузка каналов, технические сбои, которые впоследствии могут привести к ухудшению качества предоставляемых услуг связи или к временному выходу из строя оборудования.

С учетом вышеизложенного одним из возможных решений, которое может повысить устойчивость и безотказность системы, является технология цифрового двойника.

1 Яровой Роберт Владимирович, научный сотрудник научно-исследовательского центра Военной академии связи, г. Санкт-Петербург, Россия. E-mail: Nadzar@yandex.ru

2 Карганов Виталий Вячеславович, кандидат технических наук, доцент, старший научный сотрудник научно-исследовательского центра Военной академии связи, г. Санкт-Петербург, Россия. E-mail: vitalik210277@mail.ru

3 Лукашенок Василий Ильич, младший научный сотрудник научно-исследовательского центра Военной академии связи, г. Санкт-Петербург, Россия. E-mail: lukasenokvasilij@gmail.com

1. Теоретические основы цифровых двойников в телекоммуникациях

1.1. Понятие и определение цифрового двойника

Впервые концепция цифрового двойника, как виртуальной копии физического объекта, была предложена в 2002 году Майклом Гривсом. Цифровой двойник может быть использован для исследования влияния различных воздействий на моделируемый объект, не подвергая риску реальный объект. На данный момент концепция активно развивается и стандартизируется в различных сферах [2].

Согласно стандарту, BS ISO/IEC 30173, цифровой двойник – это «виртуальное представление физического объекта или системы, которое используется для понимания, прогнозирования и оптимизации характеристик и поведения физического объекта».

В телекоммуникационных системах цифровой двойник представляет из себя виртуальную модель сети, которая соединена с физической сетью и получает актуальные данные при помощи телеметрии и позволяет следующее [3]:

- моделировать поведение сети при различном влиянии на сеть;
- прогнозировать сбои и оптимизировать ресурсы сети;
- тестировать алгоритмы управления без ущерба для физической сети;
- автоматизировать процессы мониторинга сети.

1.2. Архитектура цифрового двойника телекоммуникационной сети

Архитектура цифрового двойника телекоммуникационной сети в большинстве случаев состоит из следующих частей [4]:

1. Физический уровень, который содержит реальные компоненты сети:

- базовые станции;
- устройства пользователей;
- ядро сети;
- различное сетевое оборудование: коммутаторы, маршрутизаторы, серверы.

2. Цифровой уровень (модель), который содержит виртуальное представление физических объектов:

- модели базовых станций, каналов связи, пользователей;
- алгоритмы расчета качества сигнала, пропускной способности, задержки;
- модели поведения пользователей (к примеру активность и мобильность).

3. Аналитический уровень, который содержит обработку данных и система принятия решений:

- сбор и хранение телеметрии;
- анализ в реальном времени;
- управление сетью (настройка параметров телекоммуникационной сети для предоставления услуг связи должного уровня).

1.3. Ключевые метрики телекоммуникационных сетей

Для создания цифрового двойника в целях предиктивной защиты от различных аномалий в телекоммуникационной сети необходимо подобрать набор соответствующих параметров [5, 6]:

1. Active UE: отражает число устройств, находящихся в зоне покрытия сети и ведущие обмен данными в настоящий момент времени. Данный параметр влияет на нагрузку телекоммуникационной сети, на число потребляемых ресурсов, а также на качество обслуживания.

2. DL Traffic: обозначает объем данных, передаваемых от сети к устройствам (пользователям сети). Данный параметр особо важен для планирования пропускной способности сети и распределения нагрузки между передающим оборудованием.

3. RSRP (Received Signal Reference Power, измеряется в дБм): данный параметр обозначает мощность сигнала, принимаемого устройством от базовой станции и принадлежит к перечню ключевых показателей качества радиоканала.

4. Active Sessions: данный параметр означает количество активных сессий в сети. Данный параметр близок по значению к Active UE, но может отличаться от него в определенных случаях (к примеру, когда пользователь в данный момент времени одновременно ведет разговор и пользуется интернетом).

1.4. LSTM-автоэнкодеры, как инструмент обнаружения аномалий

Для детектирования аномалий в телекоммуникационных сетях анализируют временные ряды с высокой размерностью и сложными зависимостями. В качестве инструмента для решения данной задачи могут быть использованы методы машинного обучения, которые

способны обучиться нормальным паттернам и определять отклонения от них для выявления аномалий в сети [7].

Можно выделить следующие достоинства автоэнкодеров:

- возможно обучение без учителя (не требуются размеченные данные для всех типов аномалий);
- способность обрабатывать многомерные данные (могут одновременно учитывать UE, трафик, RSRP и сессии в процессе мониторинга);
- обладают высокой эффективностью при грамотном обучении и настройке.

LSTM выбран исходя из следующих причин:

- работа с последовательностями, т.е. обладает способностью учета временных зависимостей (например, рост трафика за последний период времени);
- устойчивость к шуму (благодаря внутренним механизмам (забывание, вход, выход));
- предрасположенность к работе с длинными последовательностями (важно для телекоммуникационных сетей, в моментах аномалии могут проявляться в течении времени.

1.5. Методология обнаружения аномалий на основе ошибки восстановления

Основная идея метода – сначала обучить автоэнкодер на нормальных данных, для того чтобы он мог точно восстанавливать их. После обучения на нормальных данных, провести обучения на тестовых данных (включая аномалии), где вычисляется ошибка восстановления (MSE) [8], а если ошибка превышает порог – считается, что произошла аномалия.

Пусть $X \in R^{T \times D}$ – входная последовательность длиной T с D признаками.

Автоэнкодер предсказывает $\hat{X} = f_\theta(X)$.

Ошибка восстановления может быть определена по следующей формуле:

$$MSE = \frac{1}{T \times D} \sum_{t=1}^{T} \sum_{d=1}^{D} (X_{t,d} - \hat{X}_{t,d})^2, \quad (1)$$

где X – исходная последовательность, $\hat{X}$ – восстановленная последовательность, T – длина последовательности, D – число признаков.

Порог τ определяется как q-й перцентиль ошибки на нормальных данных по следующей формуле:

$$\tau = percentile(MSE_{normal}, q). \quad (2)$$

Аномалия в телекоммуникационной сети обнаруживается, если $MSE > \tau$.

2. Практическая реализация цифрового двойника телекоммуникационной сети

2.1. Генерация синтетических телеметрических данных

В целях создания цифрового двойника была разработан процедура генерации синтетических данных, которые имитируют работу телекоммуникационной сети в течение суток (24 часов) с шагом 10 секунд. Общее количество временных меток равно 8640, что соответствует 24 часам × 60 минут × 6 (интервалов в минуту).

Основой для создания нормальной нагрузки сети стал суточный профиль, который характерен для городских сетей, определяется по следующей формуле:

$$load\_profile(t) = 0{,}3 + 0{,}7 \cdot \cdot \left[\exp\left(-\frac{(h(t) - 8)^2}{8}\right) + \exp\left(-\frac{(h(t) - 20)^2}{10}\right)\right], \quad (3)$$

где $h(t) = [t/360] \bmod 24$ – текущий час суток, t – номер временного интервала.

На основе данного профиля были сгенерированы четыре ключевые метрики, которые представлены в таблице 1.

Таблица 1.

Ключевые метрики

№ п/п	Метрика	Диапазон
1	Active UE	10–300
2	DL Traffic Mbps	0–150
3	RSRP dBm	(–120) – (–70)
4	Active Sessions	5–400

Здесь ε – случайный шум, имитирующий флуктуации реальной сети.

Физический смысл метрик:

- Active UE – означает число активных пользовательских устройств, которое напрямую связано с нагрузкой;
- DL Traffic – означает объем передаваемых данных и важен для планирования ресурсов телекоммуникационной сети;
- RSRP – означает уровень сигнала и определяет качество связи;
- Active Sessions – означает число активных сессий и коррелирует с UE, но может отличаться от UE из-за мульти-сессий на устройствах пользователей.

2.2. Внедрение аномалий

В целях проверки эффективности и работоспособности цифрового двойника в синтетические данные были искусственно внедрены аномалии, характерные для реальных телекоммуникационных сетей [9–12]. Первая аномалия – технический сбой – в данном случае моделируется ситуация при которой происходит намеренное ухудшение качества радиоканала вследствие отказа оборудования или внешних помех. Данная аномалия была симулирована в интервале с 12:00 до 12:10, когда уровень сигнала (показатель RSRP) искусственно понижался до диапазона от –115 до –105 дБм, что соответствует критично низкому качеству связи. В тоже время, когда было симулировано снижение качества связи, в процессе мониторинга наблюдалось снижение загрузочного объема трафика на 70 %, что можно обусловить ухудшением качества соединения с активными пользователями, число которых также демонстрировало некоторое снижение. Вышеописанная ситуация отражает обычный сценарий сбоя в сети, при котором падение качества сигнала приводит к ухудшению сервиса и, следовательно, к потере пользователей. Вторая аномалия – DDoS-данный тип атаки имитирует злонамеренную нагрузку на сеть, при которой объем передаваемых данных стремительно возрастает, но при этом не уваливается число активных устройств пользователей [13, 14]. Аномалия с имитацией DDoS-атаки была симулирована во временной период с 18:30 до 18:45, когда DL Traffic искусственно доводился до уровня 120–200 Мбит/с, что значительно превышает нормальные значения для данного времени суток, при этом количество активных пользователей в сети оставалось неизменным. Состояние системы в данный момент времени с подобными параметрами соответствует стандартным признакам DDoS-атаки, т.к. рост трафика не сопровождается ростом пользовательской активности, что соответственно указывает на искусственное увеличение нагрузки на систему. Соответственно две вышеописанные аномалии были отмечены метками – «FAILURE» для технического сбоя и «DDOS» для атаки – что позволит в дальнейшем использовать их для оценки качества разработанной модели для обнаружения аномалий в сети.

2.3. Визуализация синтетических данных

На рисунке 1 представлены графики трех ключевых метрик за 24 часа: DL Traffic, RSRP и Active UE. Красные полупрозрачные области выделяют зоны аномалий:

Результаты анализа рисунка 1 позволили сделать следующие выводы:

- в зоне технического сбоя (с 12:00 до 12:10) наблюдается резкое падение показателей RSRP и DL Traffic, что соответственно сочетается с понижением количества пользователей;
- во временном промежутке с DDoS-атакой (18:30–18:45) наблюдается резкое повышение объема трафика без роста количества активных пользователей, что является маркером намеренного искусственного повышения нагрузки;

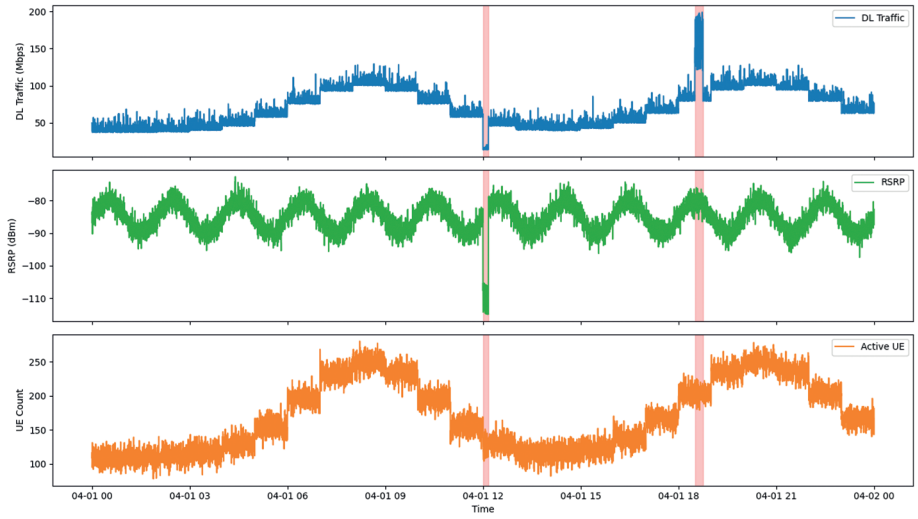

Рис. 1. Визуализация синтетических телеметрических данных телекоммуникационной сети с внедренными аномалиями

- нормальные параметры в телекоммуникационной сети в соответствии с суточным профилем с увеличением нагрузки в 8:00 и 20:00 соответствуют человеческой активности.

2.4. Подготовка данных для обучения модели

В целях качественного обучения модели автоэнкодера были подготовлены данные следующим образом:

- выбраны признаки, на основе которых формируется результат (Active_UE, DL_Traffic_Mbps, RSRP_dBm, Active_Sessions);
- проведена нормализация признаков: использован MinMaxScaler для масштабирования всех признаков в диапазоне от 0 до 1 включительно;
- формирование временных интервалов, где для работы с LSTM использованы последовательности длиной 60 шагов (10 минут).

2.5. Архитектура и обучение LSTM-автоэнкодера

Для обнаружения аномалий была выбрана архитектура LSTM-автоэнкодера, которая хорошо работает с временными рядами и способна учиться нормальным паттернам, выявляя отклонения.

Модель создана с использование библиотеки TensorFlow [15] и построена на следующей архитектуре из четырех блоков:

- Encoder: LSTM слой с 32 нейронами – предназначен для сжатия входной последовательности в скрытый массив данных;
- RepeatVector: создает копию скрытого массива данных для каждого временного интервала;
- Decoder: LSTM с 32 нейронами – восстанавливает исходную последовательность;
- TimeDistributed Dense: выходной слой, предназначенный для предсказывания значения определенных признаков на каждом временном интервале.

На основе синтетически сгенерированных данных [16] модель была обучена на нормальных данных (8431 последовательностей) в течение 15 эпох с размером батча равным 32 и размером валидационной выборки равным 10 % (от датасета), что в свою очередь соответствует рисунку 2.

На графике (рис. 2) можно обнаружить, что функция потерь на тренировочной и валидационной выборке имеет тенденцию стабильного снижения, следовательно, модель имеет хорошую сходимость. После 15 эпох функция потерь достигает значения ~0.005, что в свою очередь свидетельствует о том, что модель научилась восстанавливать нормальные паттерны.

2.6. Обнаружение аномалий на полном датасете

После обучения модель была применена к полному датасету (включая аномалии). Для каждой последовательности вычислялась ошибка восстановления (MSE) согласно формуле 1. Порог для определения аномалии был установлен как 95-й перцентиль ошибки на нормальных данных на основе формулы 2. В данному случае аномалия обнаруживается, если $MSE > \tau$.

2.7. Оценка качества обнаружения аномалий

Для оценки эффективности модели были рассчитаны стандартные метрики Precision, Recall, F1 [17] согласно формулам 4, 5, 6 соответственно:

$$\text{Precision} = \frac{TP}{TP + FP}; \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN}\text{TP/(TP+FN)}; \quad (5)$$

$$\text{F1} = 2 \times \frac{Precision \times Recall}{Precision + Recall}, \quad (6)$$

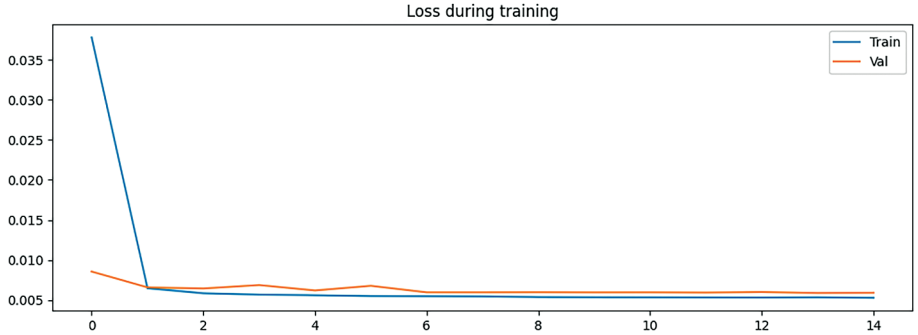

Рис. 2. График потерь

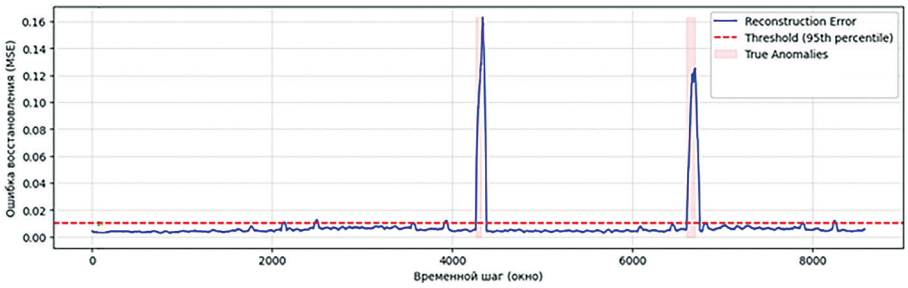

Рис. 3. Обнаружение аномалий с помощью цифрового двойника

где TP – истинно положительные, FP – ложно положительные, FN – ложно отрицательные.

В итоге были получены следующие результаты:

- Precision 89,2 % – означает, что из всех обнаруженных аномалий, 89,2 % действительно являются аномалиями;
- Recall 94,1% – означает, что модель обнаружила 94,1 % всех реальных аномалий;
- F1-score 91,6% – означает, что модель эффективно определяет истинные случаи, с низкой вероятностью ошибки.

2.8. Визуализация результатов обнаружения аномалий

На рисунке 3 представлена динамика ошибки восстановления для всех временных диапазонов, которая получена в результате работы LSTM-автоэнкодера. На основе визуального анализа графика можно сделать вывод о способности модели обнаруживать отклонения от нормального поведения сети на основе анализа различных временных последовательностей на основе ряда параметров.

На основе анализа графика (рис. 3) можно сделать следующие выводы:

1. На подавляющей части временного диапазона ошибка восстановления имеет низкое значение и не превышает 0,01, что также свидетельствует о том, что модель была качественно обучена на нормальных данных и способна восстанавливать нормальные паттерны телеметрии.

2. На временном шаге ~4200 можно наблюдать резкий скачок ошибки восстановления, который в пике достигает ~0,16, что в более чем 10 раз больше порога, который равен ~0,012, при этом данный скачок совпадает с розовой областью, которая обозначает аномалию и соответствует временному диапазону технического сбоя, который симулирован во временном диапазоне с 12:00 до 12:10.

3. Следующий рост ошибки $MSE = \sim 0{,}13$ восстановления наблюдается на временном шаге ~6700, который соответствует временному диапазону симулирования DDoS-атаки с 18:30 до 18:45. Пик роста ошибки восстановления совпадает с областью аномалии и также значительно превышает пороговое значение.

4. Для баланса между чувствительностью и количеством ложных срабатываний был выбран порог на уровне 95-го перцентиля, следствием чего модель адекватно реагирует на небольшие естественные всплески активности, а при значительных всплесках правильно детектирует аномалии.

5. Возможность вывода актуального состояния телекоммуникационной сети в визуальном формате (в качестве примеру можно привести график, который отображен на рисунке 3) предоставляет оператору сети возможность своевременно реагировать на возникающие аномалии с учетом возможности записи времени начала и идентификации типа атаки.

Объединяя вышеизложенное можно сделать вывод, что предложенный подход на основе LSTM-автоэнкодера, обученного на синтетических данных, обладает возможностью надежно детектировать как технические сбои, так и DDoS-атаки, путем анализа временных паттернов телеметрии.

2.9. Выводы по практической реализации

В итоге цифровой двойник позволил:

- генерировать правдоподобные синтетические данные с подконтрольными аномалиями;
- смоделировать и обучить LSTM-автоэнкодер для обнаружения аномалий;
- достигнуть показателей качества по метрике F1-score = 0,916, что является показателем качества модели;
- графически отобразить работу модели на примере двух типов аномалий (технического сбоя и DDoS-атаки).

Выше представленные результаты подтверждают возможность применения цифровых двойников в целях обучения, тестирования и внедрения алгоритмов обнаружения аномалий в телекоммуникационных сетях.

3. Анализ результатов и обсуждение

3.1. Интерпретация метрик качества

Ссылаясь на раздел 2.7., модель продемонстрировала следующие показатели качества:

- Precision: 0,892;
- Recall: 0,941;
- F1-score: 0,916.

Основываясь на показателях перечисленных метрик, модель обладает высокой эффективности модели в задаче обнаружения аномалий. Рассмотрим каждый показатель подробнее.

Precision (точность) = 0,892. Точность отражает количество истинно аномальных событий среди тех, которые модель классифицировала как аномалии. Значение 89,2 % означает, что на каждые 100 обнаруженных аномалий 89 действительно истинные, а 11 – соответственно ложные срабатывания. Достигнутый уровень считается приемлемым для систем мониторинга, в которых лишние оповещения не влияют на работу сети и не выводят ее из строя.

Recall (полнота) = 0,941 – данная метрика отражает способность модели обнаруживать все реальные аномалии. Значение 94,1 % соответствует тому, что модель успешно обнаруживает 94 из 100 реальных аномалий, что действительно важно в телекоммуникационных сетях, где какое-либо пропущенное негативное влияние может отрицательно сказаться на работоспособности сети или полностью вывести ее из строя.

F1-score = 0,916. F1-мера, отражающая гармоническое среднее между точностью и полнотой и отражает баланс между данными характеристиками. Значение 0,916 подтверждает тот факт, что был достигнут консенсус между точностью и полнотой, т.е. модель имеет небольшое число ложных срабатываний и в тоже время правильно выявляет большинство аномалий.

Исходя из удовлетворительных значений перечисленных метрик можно сделать вывод, что модель может быть использована в системах мониторинга телекоммуникационных сетей.

3.2. Анализ причин ложных срабатываний и пропусков

С учетом достаточно высоких показателей метрик, модель все же допускает ошибки. На графике (рис. 3) присутствуют участки, на которых ошибка восстановления превышает заданный порог, но не детектируется как истинная аномалия (розовая область). Данные случаи могут трактоваться следующими факторами:

1. Переходные процессы и шум.

В начале и в конце временного диапазона могут наблюдаться небольшие скачки, которые могут быть связаны с процессами запуска, остановки и шумом в данных. Примером является промежуток от 2000 до 2500 шага, где наблюдается небольшой рост ошибки восстановления, которая превышает порог и не совпадает с истинными аномалиями. Данный случай может быть вызван случайными процессами, которые модель интерпретирует как отклонение от нормального состояния.

2. Пороговое значение.

С учетом того, что порог установлен на уровне 95 перцентиля, для определенного перечня аномалий, которые могут происходить в телекоммуникационных сетях, он (порог) может быть завышен, что приведет к тому, что аномалия не будет своевременно обнаружена. В обучающих данных также могут присутствовать выбросы, которые в итоге «приподнимают» порог вверх, что способствует пропуску незначительных аномалий. Для решения данного случая может быть использован динамический порог, который основан на статических методах (например, $\mathrm{mean} + 3\sigma$).

3. Ограничения модели.

При использовании LSTM-автоэнкодера, который обучен на нормальных данных, автоэнкодер будет пытаться восстановить их, при этом, если аномалия будет иметь форму, которая близка к нормальному паттерну (в качестве примера можно привести медленный рост числа активных пользователей), модель может не обнаружить аномалию. Решением данного случая может быть использование добавочных признаков (скорость изменения RSRP) а также задействование ансамблевых методов.

3.3. Влияние параметров на качество модели

На успех детектирования аномалий влияет следующий перечень параметров:

1. Длина временного окна.

Грамотный выбор временного диапазона (60 шагов равные по длительности 10 минутам) позволяет модели учитывать краткосрочные зависимости без излишней нагрузки на вычислительные ресурсы оборудования, но для обнаружения аномалий, которые по длительности меньше временного диапазона окна, необходим подбор соответствующей длины временного окна.

2. Количество нейронов в LSTM.

Каждый LSTM-слой в архитектуре модели использует 32 нейрона, что обеспечивает необходимую емкость для корректного обучения, если увеличить число нейронов до 64 или до 128 при грамотном обучении позволит повысить точность модели (но увеличивается время обучения и риск переобучения).

3. Количество эпох.

В данной исследовании модель обучалась в течении 15 эпох, чего достаточно для ее сходимости. Дальнейшее увеличение числа функция потерь продолжит снижаться, но незначительно, что указывает на достижение плато (рис. 2), т.е. дальнейшие обучение модели без изменения архитектуры и увеличения датасета не позволит увеличить качество модели.

Заключение

В рамках проведенного анализа было выявлено, что подход с применением цифрового двойника и LSTM-автоэнкодера обладает требуемыми качествами для обнаружения аномалий. Разработанная модель показывает высокие показатели по раду метрик и с высокой вероятностью обнаруживает технические сбои и DDoS-атаки. При этом каждое «свое решение» модель основывает не только на основе текущего момента времени, но также на основе предыдущих параметров, что делает ее более подходящим вариантом для телекоммуникационных сетей, чем традиционные методы.

Однако, учитывая ранее изложенные результаты исследований в камках данной статьи, модель имеет незначительные ограничения, которые связаны с чувствительностью к различным шумам, конечному набору признаков и заданному уровню порога. В целях повышения качества модели рекомендуется:

- применять более сложные архитектуры (к примеру, CNN-LSTM и Transformer);
- использовать при оценке качества модели дополнительные метрики (SINR, Latency, Handover Count);
- расширять количество параметров, на основе которых производится многомерный анализ, до перечня параметров в реальной сети.

Перечисленные выше рекомендации позволят создать эффективный инструмент для автоматизированного мониторинга телекоммуникационных сетей, который будет своевременно реагировать на разного рода угрозы и предотвращать их прежде чем они нанесут деструктивное воздействие, которое снизит качество предоставляемых услуг связи или выведет оборудование из строя.

Литература

1. Попов А. А. Методы обнаружения аномалий в сетевом трафике // В сборнике: Наука, инновации, образование: актуальные вопросы XXI века. Сборник статей XIV Международной научно-практической конференции. Пенза, 2025. С. 30–33.
2. Шпак П. С., Сычева Е.Г., Меринская Е. Е. Концепция цифровых двойников как современная тенденция цифровой экономики // Вестник Омского университета. Серия: Экономика. 2020. Т. 18. № 1. С. 57–68.
3. Канаев А. К., Степанова А. Р. Цифровые двойники в телекоммуникациях // Научно-техническая конференция Санкт-Петербургского НТО РЭС им. А. С. Попова, посвященная Дню радио. 2024. № 1(79). С. 228–230.
4. Афонин И. Г., Кучерявый Е. А., Осипов Д. В., Морозов А. В. Алгоритмы ML для построения цифровых двойников в 5G // Вестник связи. 2024. № 12. С. 29–30.
5. Талибаева А. И., Рамазанов М. Б., Сайкен Д. Р., Хисамутдинов Р. М. Исследование технологических параметров телекоммуникационных сетей связи: методы оценки производительности беспроводных сетей // В сборнике: Global Challenges – Scientific Solutions II. proceedings. Antwerp, 2020. С. 225–227.
6. Новиковский К. В. Анализ параметров и модели качества в современных телекоммуникационных системах // Вестник Воронежского института высоких технологий. 2025. № 2(53).

7. Адамовский Е. Р., Богуш Р. П., Чертков В. М. Исследование эффективности LSTM нейронных сетей для прогнозирования занятости канальных ресурсов на основе данных карты радиосреды когнитивной системы связи // В сборнике: Информатика: проблемы, методы, технологии. Материалы XXIV Международной научно-практической конференции им. Э. К. Алгазинова. Воронеж, 2024. С. 148–153.
8. Кукурхоев А. М. Функция потерь MSE и MAE // В сборнике: Наука, образование, инновации: актуальные вопросы и современные аспекты. сборник статей XV Международной научно-практической конференции в 2 частях. Пенза, 2022. С. 85–86.
9. Амосов О. С., Амосова С. Г., Иванов Ю. С., Жиганов В. С. В. Использование глубоких нейронных сетей для распознавания аномалий сетевого трафика в информационно-телекоммуникационных системах предприятий // В сборнике: Управление развитием крупномасштабных систем MLSD'2019. Материалы двенадцатой международной конференции Научное электронное издание. Под общей ред. С. Н. Васильева, А. Д. Цвиркуна. 2019. С. 968–971.
10. Протасова М. А. Нейросетевой классификатор аномалий телекоммуникационной сети // В сборнике: Нейроинформатика–2015. XVII Всероссийская научно-техническая конференция с международным участием: сборник научных трудов. Ответственный редактор А. Г. Трофимов. 2015. С. 138–148.
11. Живодерников А. Ю., Ковайкин Ю. В., Лебедев П. В. Анализ источников сетевых аномалий в системах управления телекоммуникационными сетями // В сборнике: Проблемы технического обеспечения войск в современных условиях. Труды IV межвузовской научно-практической конференции. 2019. С. 298-301.
12. Акыев Г. А. Современные методы и алгоритмы мониторинга телекоммуникационных сетей // Инновационная наука. 2025. № 2-2. С. 37–39.
13. Яровой Р. В., Рябов Г. А., Карганов В. В. Кибербезопасность в мире инфотелекоммуникаций: вызовы и стратегии защиты // В сборнике: Инновационная деятельность в Вооруженных Силах Российской Федерации. Труды всеармейской научно-практической конференции. Санкт-Петербург, 2023. С. 373–377.
14. Зеленский М. Д. DDoS-атаки: типы атак, устранение DDoS-атак // В сборнике: Студенческая наука для развития информационного общества. сборник материалов IV Всероссийской научно-технической конференции: в 2-х томах. 2016. С. 241–243.
15. Эдель Г. Е. Глубокое обучение с использованием библиотеки TensorFlow // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. 2020. № 1-2. С. 162–164.
16. Чайкин Г. А. Создание синтетических данных пользовательской активности на основе вопросно-ответных текстовых данных // Процессы управления и устойчивость. 2025. Т. 12. № 1. С. 417–421.
17. Никкель К. Е., Сосунов А. А. Классификация данных: метрики оценки качества моделей и алгоритмы // В сборнике: Цифровизация: новые тренды и опыт внедрения. Сборник статей Международной научно-практической конференции. Уфа, 2024. С. 39–44.

APPLICATION OF DIGITAL TWINS IN TELECOMMUNICATIONS SYSTEMS

Yarovoy R. V.[4], Karganov V. V.[5], Lukashenok V. I.[6]

Keywords: *digital twin, telecommunications network, anomaly detection, synthetic data, machine learning, telemetry, predictive maintenance, artificial intelligence.*

Abstract

The aim of the work *is to analyze the possibility of using digital twin technology to detect anomalies in telecommunication networks using artificial intelligence technologies and training on synthetic data.*

Research method: *is based on the use of mathematical modeling, which involves the creation of a digital twin of a telecommunications network through the use of synthetically generated data that simulate the behavior of the network, and machine learning is used to detect anomalies – training of an LSTM autoencoder, with subsequent assessment of the quality of anomaly detection based on metrics (precision, recall and F1-score).*

4 Robert V. Yarovoy, Researcher, Research Center, Military Academy of Communications, St. Petersburg, Russia. E-mail: Nadzar@yandex.ru

5 Vitaly V. Karganov, Ph.D. of Technical Sciences, Associate Professor, Senior Researcher of the Research Center of the Military Academy of Communications, St. Petersburg, Russia. E-mail: vitalik210277@mail.ru

6 Vasily I. Lukashenok, Junior Researcher, Research Center, Military Academy of Communications, St. Petersburg, Russia. E-mail: lukasenokvasilij@gmail.com

Results of the study: *include the development of a digital twin model of a telecommunications network, as well as the generation of a dataset based on synthetic data. In the course of the study, it was demonstrated that a digital twin of a telecommunications network can be created, data can be synthetically generated to train a machine learning model, with the help of which various anomalies can be detected. telecommunication networks to detect various kinds of anomalies.*

Scientific novelty: *lies in the development of a methodology for generating synthetic data with anomalies that are adapted to telecommunication networks, as well as in the use of a digital twin as a tool for testing machine learning algorithms to detect anomalies.*

References

1. Popov A. A. Metody obnaruzhenija anomalij v setevom trafike // V sbornike: Nauka, innovacii, obrazovanie: aktual'nye voprosy XXI veka. Sbornik statej XIV Mezhdunarodnoj nauchno-prakticheskoj konferencii. Penza, 2025. S. 30–33.
2. Shpak P. S., Sycheva E. G., Merinskaja E. E. Koncepcija cifrovyh dvojnikov kak sovremennaja tendencija cifrovoj jekonomiki // Vestnik Omskogo universiteta. Serija: Jekonomika. 2020. T. 18. № 1. S. 57–68.
3. Kanaev A. K., Stepanova A. R. Cifrovye dvojniki v telekommunikacijah // Nauchno-tehnicheskaja konferencija Sankt-Peterburgskogo NTO RJeS im. A.S. Popova, posvjashhennaja Dnju radio. 2024. № 1(79). S. 228–230.
4. Afonin I. G., Kucherjavyj E. A., Osipov D. V., Morozov A. V. Algoritmy ML dlja postroenija cifrovyh dvojnikov v 5G // Vestnik svjazi. 2024. № 12. S. 29–30.
5. Talibaeva A. I., Ramazanov M. B., Sajken D. R., Hisamutdinov R. M. Issledovanie tehnologicheskih parametrov telekommunikacionnyh setej svjazi: metody ocenki proizvoditel'nosti besprovodnyh setej // V sbornike: Global Challenges – Scientific Solutions II. proceedings. Antwerp, 2020. S. 225–227.
6. Novikovskij K. V. Analiz parametrov i modeli kachestva v sovremennyh telekommunikacionnyh sistemah // Vestnik Voronezhskogo instituta vysokih tehnologij. 2025. № 2(53).
7. Adamovskij E. R., Bogush R. P., Chertkov V. M. Issledovanie jeffektivnosti LSTM nejronnyh setej dlja prognozirovanija zanjatosti kanal'nyh resursov na osnove dannyh karty radiosredy kognitivnoj sistemy svjazi // V sbornike: INFORMATIKA: PROBLEMY, METODY, TEHNOLOGII. Materialy XXIV Mezhdunarodnoj nauchno-prakticheskoj konferencii im. Je.K. Algazinova. Voronezh, 2024. S. 148–153.
8. Kukurhoev A. M. Funkcija poter' MSE i MAE // V sbornike: Nauka, obrazovanie, innovacii: aktual'nye voprosy i sovremennye aspekty. sbornik statej XV Mezhdunarodnoj nauchno-prakticheskoj konferencii v 2 chastjah. Penza, 2022. S. 85–86.
9. Amosov O. S., Amosova S. G., Ivanov Ju. S., Zhiganov V. S. V. Ispol'zovanie glubokih nejronnyh setej dlja raspoznavanija anomalij setevogo trafika v informacionno-telekommunikacionnyh sistemah predprijatij // V knige: Upravlenie razvitiem krupnomasshtabnyh sistem MLSD'2019. Materialy dvenadcatoj mezhdunarodnoj konferencii Nauchnoe jelektronnoe izdanie. Pod obshhej red. S. N. Vasil'eva, A. D. Cvirkuna. 2019. S. 968–971.
10. Protasova M. A. Nejrosetevoj klassifikator anomalij telekommunikacionnoj seti // V sbornike: Nejroinformatika–2015. XVII Vserossijskaja nauchno-tehnicheskaja konferencija s mezhdunarodnym uchastiem: sbornik nauchnyh trudov. Otvetstvennyj redaktor A. G. Trofimov. 2015. S. 138–148.
11. Zhivodernikov A. Ju., Kovajkin Ju. V., Lebedev P. V. Analiz istochnikov setevyh anomalij v sistemah upravlenija telekommunikacionnymi setjami // V sbornike: Problemy tehnicheskogo obespechenija vojsk v sovremennyh uslovijah. Trudy IV mezhvuzovskoj nauchno-prakticheskoj konferencii. 2019. S. 298–301.
12. Akyev G. A. Sovremennye metody i algoritmy monitoringa telekommunikacionnyh setej // Innovacionnaja nauka. 2025. № 2-2. S. 37–39.
13. Jarovoj R. V., Rjabov G. A., Karganov V. V. Kiberbezopasnost' v mire infotelekommunikacij: vyzovy i strategii zashhity // V sbornike: Innovacionnaja dejatel'nost' v Vooruzhennyh Silah Rossijskoj Federacii. Trudy vsearmejskoj nauchno- prakticheskoj konferencii. Sankt-Peterburg, 2023. S. 373–377.
14. Zelenskij M. D. DDoS-ataki: tipy atak, ustranenie DDoS-atak // V sbornike: Studencheskaja nauka dlja razvitija informacionnogo obshhestva. sbornik materialov IV Vserossijskoj nauchno-tehnicheskoj konferencii: v 2-h tomah. 2016. S. 241–243.
15. Jedel' G. E. Glubokoe obuchenie s ispol'zovaniem biblioteki TensorFlow // Jelektronnye sredstva i sistemy upravlenija. Materialy dokladov Mezhdunarodnoj nauchno-prakticheskoj konferencii. 2020. № 1-2. S. 162–164.
16. Chajkin G. A. Sozdanie sinteticheskih dannyh pol'zovatel'skoj aktivnosti na osnove voprosno-otvetnyh tekstovyh dannyh // Processy upravlenija i ustojchivost'. 2025. T. 12. № 1. S. 417–421.
17. Nikkel' K. E., Sosunov A. A. Klassifikacija dannyh: metriki ocenki kachestva modelej i algoritmy // V sbornike: Cifrovizacija: novye trendy i opyt vnedrenija. Sbornik statej Mezhdunarodnoj nauchno-prakticheskoj konferencii. Ufa, 2024. S. 39–44.