ИССЛЕДОВАНИЕ СОСТЯЗАТЕЛЬНЫХ АТАК НА РЕГРЕССИОННЫЕ МОДЕЛИ МАШИННОГО ОБУЧЕНИЯ В БЕСПРОВОДНЫХ СЕТЯХ 5G

Легашев Л. В.[1], Жигалов А. Ю.[2]

DOI: 10.21681/2311-3456-2024-3-61-67

Цель исследования: *Исследование влияния состязательных атак на метрики качества регрессионных моделей машинного обучения.*

Метод исследования: *Эмуляция данных распространения сигнала в MIMO системах, синтез состязательных примеров, выполнение состязательных атак на модели машинного обучения, обучение бинарных классификаторов для обнаружения состязательных аномалий в данных.*

Результат исследования: *В статье проведена генерация сценария и исследовательский анализ набора данных с помощью эмулятора DeepMIMO. Выполнена состязательная атака с максимизацией знака градиента методом FGSM. Выполнено экспериментальное сравнение бинарных классификаторов для обнаружения отравленных данных. Выполнен анализ динамики изменения метрик качества регрессионной модели в сценарии без состязательных атак, сценарии выполнения состязательной атаки и сценарии изоляции отравленных данных. Выполнение состязательной атаки FGSM с максимизацией знака градиента увеличивает значение метрики MSE в среднем на 33% и снижает значение метрики R^2 в среднем на 10%. Бинарный классификатор LightGBM с точностью в 98% успешно обнаруживает записи с состязательными аномалиями в табличных данных. Регрессионные модели машинного обучения уязвимы к состязательным атакам, при этом своевременный интеллектуальный анализ сетевого трафика и передаваемых по сети данных позволяет обнаруживать злонамеренную сетевую активность.*

Научная новизна: *исследованы методы выполнения состязательных атак на регрессионную модель для задачи прогнозирования комбинированных потерь пути распространения сигнала от базовой станции до конечных пользователей в эмулируемом сегменте беспроводных сетей последнего поколения.*

Ключевые слова: *состязательные атаки, беспроводные самоорганизующиеся сети, машинное обучение, регрессия, MIMO.*

RESEARCH ON ADVERSARIAL ATTACKS ON REGRESSION MACHINE LEARNING MODELS IN 5G WIRELESS NETWORKS

Legashev L. V.[3], Zhigalov A. Yu.[4]

The purpose of research: *Study the impact of adversarial attacks on the evaluation metrics of regression ML models.*

The methods of research: *Emulation of signal propagation data in MIMO systems, synthesis of adversarial samples, execution of adversarial attacks on machine learning models, training of binary classifiers to detect adversarial anomalies in data.*

1 Легашев Леонид Вячеславович, кандидат технических наук, ведущий научный сотрудник лаборатории цифровых решений и аналитики больших данных Оренбургского государственного университета, г. Оренбург, Россия. E-mail: silentgir@gmail.com. ORCID: 0000-0001-6351-404X.

2 Жигалов Артур Юрьевич, младший научный сотрудник лаборатории искусственного интеллекта и анализа данных Оренбургского государственного университета, г. Оренбург, Россия. E-mail: leroy137.artur @gmail.com. ORCID: 0000-0003-3208-1629.

3 Leonid V. Legashev, Ph.D., Leading Researcher, Laboratory of Digital Solutions and Big Data Analytics, Orenburg State University, Orenburg, Russia. E-mail: silentgir@gmail.com. ORCID: 0000-0001-6351-404X.

4 Artur Yu. Zhigalov, Junior Researcher, Laboratory of Digital Solutions and Big Data Analytics Orenburg., Orenburg, Russia. E-mail: leroy137.artur@gmail.com. ORCID: 0000-0003-2752-7198

Scientific novelty: *methods for performing adversarial attacks on a regression model for the problem of predicting the combined losses of the signal propagation path from the base station to end users in the emulated segment of the latest generation wireless networks have been studied.*

The result of research: *Scenario generation and exploratory analysis of a dataset using the DeepMIMO emulator carried out. An adversarial attack with gradient sign maximization using the FGSM method was performed. An experimental comparison of binary classifiers for detecting malicious data was performed. An analysis of the dynamics of changes in the evaluation metrics of a regression model was performed in a scenario without adversarial attacks, a scenario under adversarial attack, and a scenario with isolating compromised data. Performing an adversarial FGSM attack with gradient sign maximization increases the value of the MSE metric by an average of 33% and reduces the value of the R^2 metric by an average of 10%. The LightGBM binary classifier successfully detects records with adversarial anomalies in tabular data with 98% accuracy. Regression-based machine learning models are vulnerable to adversarial attacks, but timely intelligent analysis of network traffic and data transmitted over the network can detect malicious network activity.*

Keywords: *adversarial attacks, wireless ad hoc networks, machine learning, regression, MIMO.*

Введение и обзор современного состояния исследований

Повсеместное распространение беспроводных сетей последнего поколения, развитие технологий миллиметровых радиоволн (mmWave), антенных систем massive MIMO (massive Multiple Input Multiple Output) и, как следствие, возросший уровень передаваемых по сети данных от множества пользователей влечет за собой проблемы обеспечения сетевой безопасности. Автор публикации [1] посвящает свою работу анализу безопасности физического уровня для беспроводных сетей 5G/6G. Современные модели машинного обучения (МО) активно используются для анализа сетевого трафика и выявления злонамеренной сетевой активности, но при этом сами модели глубокого обучения могут быть уязвимы к состязательным атакам, цель которых заключается в компрометации эффективности таких моделей. Состязательные атаки вида «белый ящик» (white box attacks) характерны для случаев, в которых злоумышленник имеет прямой доступ к моделям машинного обучения с возможностью исследования исходного кода и архитектуры. Состязательные атаки вида «черный ящик» (black box attacks) характерны для случаев, в которых злоумышленник имеет возможность тестировать готовую модель. Намеренное добавление специально подготовленных состязательных возмущений в исходные данные может привести к компрометации качества модели машинного обучения.

Множество актуальных исследований состязательных атак посвящены проблеме классификации на основе табличных или графических данных, авторы статьи [2] выполняют комплексный анализ состязательных атак на системы машинного обучения и обсуждают методы их защиты. Следует отметить, что практически отсутствуют публикации по исследованию состязательных атак на задачи регрессии, в том числе в области беспроводных сетей, что подчеркивает актуальность настоящего исследования. Авторы публикации [3] выполняют состязательную атаку вида «белый ящик» на табличные данные, успешно обманывая нейронную сеть и снижая её производительность. В исследовании [4] проводится анализ устойчивости сильно параметризованных линейных моделей к состязательным атакам с целью максимизации ошибки прогнозирования. В статье [5] исследуется устойчивость коэффициентов регрессии к состязательным примерам, подготовленным для «отравления» исходных данных обучения модели МО. Публикация [6] посвящена анализу уязвимости регрессионных моделей многомерных временных рядов к состязательным атакам. Авторы показывают, что исследуемые модели уязвимы к проводимым атакам, что критически важно для безопасности. В исследовании [7] представлены два алгоритма для выполнения состязательных атак на модели регрессии. Авторы статьи [8] отмечают, что подготовленные состязательные примеры, сгенерированные для атаки «белого ящика» можно эффективно использовать для выполнения состязательной атаки на неизвестную злоумышленнику модель регрессии, то есть для выполнения атаки вида «черный ящик». В публикации [9] выполнялось исследование по обнаружению состязательных атак на модели прогнозирования LSTM и временной свёрточной сети на основе алгоритмов одноклассового метода опорных векторов и локального уровня выброса. Авторы статьи [10] описывают общий подход, основанный на анализе возмущений алгоритмов обучения для выполнения состязательных атак на регрессионные модели. В публикации [11] исследуются способы ослабления негативного влияния состязательных примеров на модель робастной непараметрической регрессии. Авторы исследования [12] отмечают важность обеспечения безопасности в автомобильных самоорганизующихся сетях и исследуют различные

варианты выполнения состязательных атак на модели регрессии и варианты защиты от них.

В этой статье будет проведено исследование влияния состязательных атак на метрики качества моделей машинного обучения, а также способы обнаружения таких атак в различных моделируемых сценариях распространения сигнала MIMO антенн.

1. Методы генерации состязательных примеров

Состязательная атака уклонения (dodging attack) в случае задачи классификации является атакой, при которой злоумышленник ставит задачу неправильной классификации объекта, при этом неважно, как именно будет классифицирован объект и к какому некорректному классу он будет отнесен. В случае задачи регрессии атака уклонения заключается в резком увеличении порога ошибки модели регрессии, предсказываемое значение должно быть как можно больше/меньше реального значения. Состязательная отравляющая атака (poisoning attack) – вид атаки, выполняемой в момент обучения моделей искусственного интеллекта, связанных с подмешиванием «отравленных» данных в тренировочный набор данных. В публикации [13] проводится анализ модификаций моделей машинного обучения посредством отравления данных для обучения с количественной оценкой рисков при разработке систем с искусственным интеллектом.

Рассмотрим базовые подходы для генерации состязательных образцов, которые могут быть применены для атаки моделей машинного обучения, построенных на основе табличных данных. Наиболее популярным подходом является использование метода быстрого знака градиента.

1.1. Алгоритм быстрого знака градиента (Fast Gradient Sign Method, FGSM)

Идея данного метода заключается в том, что он вычисляет градиенты функции потерь по отношению к исходным данным, а затем использует знак градиентов для создания нового «отравленного» изображения, которое максимизирует потери J модели машинного обучения:

$$x' = \varepsilon \cdot sign(\nabla_x J(\theta,x,y)), \quad (1)$$

где ε – минимальный уровень шума, θ – модель нейронной сети, $sign(\nabla_x J(\theta,x,y))$ – знак градиента, ∇_x – градиент, x – исходные данные, y – целевое значение для x.

Также следует отметить следующие алгоритмы атак на табличные данные:

1.2. Алгоритм атаки на расстоянии (Distance-based attack)

Данный метод состоит в том, чтобы минимизировать расстояние между объектом и синтетической записью с разными выходными метками. Особенность данного подхода состоит в предварительной группировке состязательных образцов в соответствии с квазиидентификаторами и выставлении соответствующего секретного признака как наиболее распространенное значение (моду). Для данного алгоритма правило обновления можно задать следующим образом:

$$y' = argmax_{|y'i-y|}\ min_{r\in R} \|(x'_i \mid t) - r\|_2, \quad (2)$$

где r – вектор возмущений значений признаков.

1.3. Алгоритм низкого профиля (Low Profile Algorithm)

Данный метод [14] состоит в том, чтобы минимизировать взвешенную норму вектора возмущений на признаках табличных данных при максимизации доли примеров $x \in X$, с ложными ответами на выходе. Для данного алгоритма правило обновления можно задать следующим образом:

$$x'_{i+1} = Clip\{x' + (r'_i + \alpha \cdot \cdot [-\nabla r_j(x'_i,t) + \lambda\|v \odot r\|_p]\}, i = 0,\ldots,N-1, \quad (3)$$

$$x' = argmin_{x'_i}\ d_v(x_i)$$

где λ – коэффициент компромисса, v – вектор важности признаков, N – максимальное количество итераций, α – коэффициент масштабирования

2. Генерация и исследование наборов данных массивных MIMO сетей

2.1. Генерация набора данных сценария «Boston5G_28»

Для генерации наборов данных массивных MIMO сетей на основе точной 3D-трассировки лучей Remcom мы использовали фреймворк DeepMIMO [15]. Рассмотрен сценарий «Boston5G_28» – сценарий на открытом пространстве, созданный на основе центра Бостона, со зданиями варьируемой высоты. На улице зафиксирована одна базовая станция (BS) на высоте 15 м, оборудованная всенаправленной антенной. В качестве массивов пользователей (UE) выступают две сетки антенн с общим количеством в 965 090 пользователей, расположенных на высоте 2 м, расстояние между пользователями – 37 см. Стандартная рабочая частота эмуляции – 28 ГГц. Каждый пользователь состоит из одной всенаправленной антенны. Расстояние между углами трассировки лучей составляет 0,25 градусов. Бетон и влажная земля используются в качестве материалов для зданий и местности соответственно. Модель распространения сигнала такова, что каждый путь канала может пройти максимум через 4 отражения, прежде чем сигнал базовой станции достигнет приемника (пользователя). Задана пропускная способность в 0.1 МГц. Общая схема расположения пользователей и базовой станции представлена на рисунке 1. Бóльшая часть пользователей отрезана от базовой станции

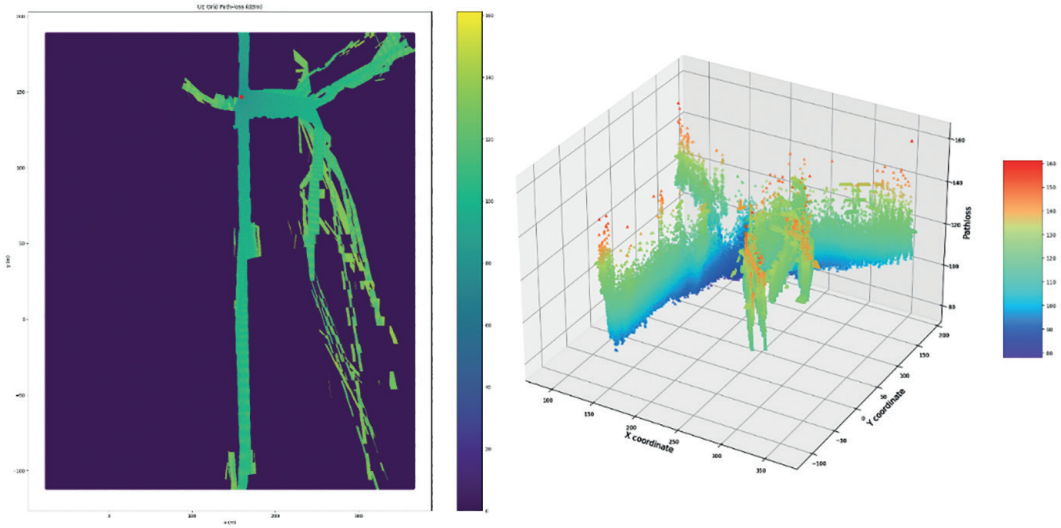

Рис.1. 2d- и 3d- карты городского пространства для сценария «Boston5G_28». Базовая станция отмечена красной точкой на 2d-карте. Цветовая схема соответствует комбинированным потерям сигнала на пути канала между пользователями и базовой станцией. Желтый/красный цвет – высокие потери, зеленый цвет – средние потери, темно-синий цвет – низкие.

в соответствии с топологией эмулируемого сегмента города. Можно динамически отслеживать изменение комбинированных потерь сигнала на пути распространения от источника (базовой станции) до конечных пользователей с учётом архитектуры эмулируемого сегмента города и отражений сигнала.

Для сгенерированного набора данных доступны координаты отправителя и получателей, матрица каналов отправителя и получателей, а также различные числовые характеристики распространении сигнала. В итоговый набор данных после выполнения расчетов сценария выбраны следующие признаки:

1. ***X coordinate*** – координата на оси X пользователя относительно эмулируемой области.
2. ***Y coordinate*** – координата на оси Y пользователя относительно эмулируемой области.
3. ***Distance*** – расстояние между базовой станцией и каждым пользователем, в метрах.
4. ***Pathloss*** – комбинированные потери на пути канала между отправителем и получателем («затухание» сигнала антенны), в децибелах относительно 1 милливатта.
5. ***DoA_phi*** – азимутальный угол прибытия, в градусах.
6. ***DoA_theta*** – зенитный угол прибытия, в градусах.
7. ***DoD_phi*** – азимутальный угол отправления, в градусах.
8. ***DoD_theta*** – зенитный угол отправления, в градусах.
9. ***Phase*** – фаза пути распространения сигнала, в градусах.
10. ***Power*** – сила сигнала при получении, в ватт.
11. ***Time of arrival*** – время получения сигнала, в секундах.
12. ***Line of Sight (LoS)*** – статус сигнала, принимаемый одно из трёх значений из {-1, 0, 1}.

($LoS = 1$): Путь прямой видимости существует. ($LoS = 0$): существуют только пути вне прямой видимости, при этом путь прямой видимости заблокирован. ($LoS = -1$): между передатчиком и приемником нет путей (полная блокировка).

2.2. Исследование сгенерированного набора данных

Итоговый набор данных содержит 105 842 записей, при этом 40 387 пользователей находятся в зоне прямой видимости базовой станции ($LoS = 1$), а 65 455 пользователей находятся вне зоны прямой видимости базовой станции ($LoS = 0$). Метрика ***pathloss*** – комбинированные потери на пути канала – является одной из ключевых метрик оценки качества беспроводных сетей последнего поколения и указывает насколько эффективной является действующая сетевая топология. Значение ***pathloss*** можно прогнозировать на основе имеющихся данных о состоянии сети при передаче сигнала между базовой станцией и большим массивом пользователей. Состязательная атака на модель регрессии должна резко увеличивать или уменьшать предсказываемое значения по отношению к оригинальному значению целевого столбца. Для выполнения состязательной атаки на модель прогнозирования потерь сигнала злоумышленнику выгодно резко увеличивать прогнозируемое значение. Злоумышленники могут выполнять атаку на регрессионные модели машинного обучения путем отравления исходных данных, в результате чего комбинированные потери на пути канала резко

возрастут, и пользователи могут потерять доступ к базовой станции в соответствии с действующими протоколами маршрутизации. В текущем исследовании сфокусируемся на задаче генерации, обнаружения и противодействия таким состязательным атакам.

На рисунке 2 представлены графики рассеяния и гистограммы для комбинированных потерь на пути канала, а также значимость признаков для прогнозирования.

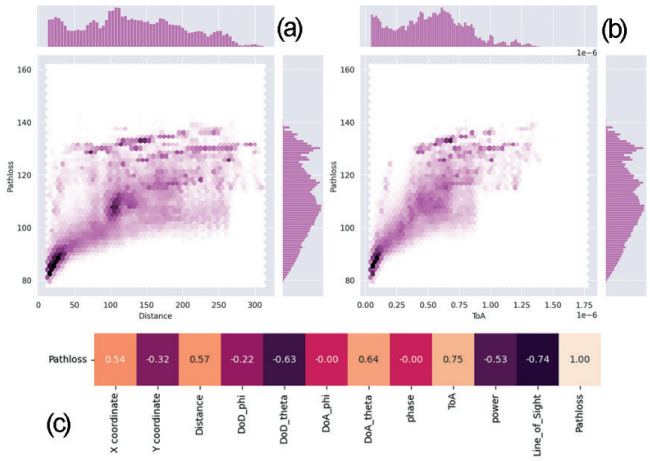

Рис. 2. Гистограммы распределения признака Pathloss в зависимости от расстояния пользователя до базовой станции (a) и в зависимости от времени прибытия сигнала (b), фрагмент матрицы корреляции (c).

Из рисунков 2(a) и 2(b) мы можем визуально выделить три пика высоких потерь сигнала в зависимости от расстояния пользователя до базовой станции и в зависимости от времени прибытия сигнала. На рисунке 2(c) представлен фрагмент матрицы корреляции, показывающий сильную прямую зависимость признака ***Pathloss*** от признаков ***Time of arrival***, ***DoA_theta*** и ***Distance*** и сильную обратную зависимость от признаков ***Line of sight***, ***DoA_theta*** и ***Power***. Действительно, увеличение времени получения сигнала приводит к увеличению комбинированных потерь на пути канала. Для пользователей, находящихся в зоне прямой видимости базовой станции, комбинированные потери на пути канала уменьшаются ввиду отсутствия отражений сигнала по пути его распространения.

3. Исследование состязательных атак на регрессионные модели машинного обучения

Полученный в разделе 2 набор данных разбит в соотношении 40:40:20 на тренировочную выборку для обучения регрессионной модели, выборку для отравления данных при выполнении состязательной атаки и тестовую выборку для валидации данных. Варьирование всех элементов знака градиента позволяет контролировать «направление» ошибки. На рисунке 3 показано, как изменяется прогнозируемое значение комбинированных потерь сигнала pathloss в зависимости от знака градиента.

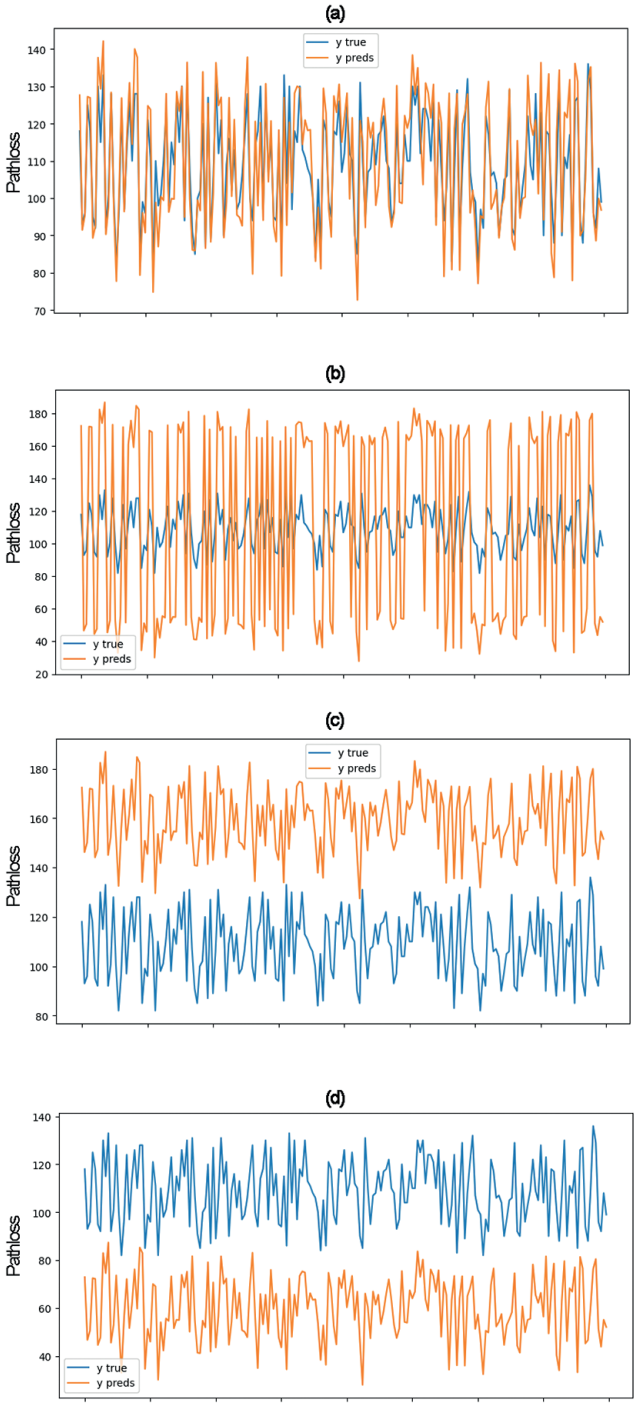

Рис. 3. Фрагмент тестового набора данных в различных сценариях: (a) – обученная модель линейной регрессии, (b) – атака FGSM с флуктуацией знака градиента, (c) – атака FGSM с максимизацией знака градиента, (d) – атака FGSM с минимизацией знака градиента

В текущем исследовании рассмотрены три основных сценария исследования состязательных атак на табличные данные:

1. Сценарий обучения модели регрессии без сторонних вмешательств (Undefended Model). Выполним обучение регрессора LinearRegressor для задачи прогнозирования комбинированных потерь ***pathloss*** по метрике оценки качества Mean Squared Error (MSE) и R^2. Линейная регрессионная модель показала хорошую точность (см. рисунок 4(а)) при решении задачи прогнозирования показателя ***pathloss*** на основе других признаков.

При построении архитектуры нейронной сети градиентный спуск сходится в локальной точке экстремума, поэтому общий алгоритм обучения модели регрессии выглядит следующим образом:

1.1 Выполнено обучение линейной регрессии из библиотеки sklearn на основе метода наименьших квадратов.

1.2 Полученные веса и свободный коэффициент (сдвиг) использованы при инициализации нейронной сети с одним линейным слоем и без функции активации с помощью библиотеки pytorch.

1.3 Выполнено тестирование построенной нейронной сети.

1.4 Подсчитаны метрики качества регрессионной модели.

2. Сценарий отравления исходных данных для обучения на основе генеративно-состязательных сетей (Attacked Model). Выполним состязательную атаку FGSM с варьированием показателя окрестности $\varepsilon = [1^{-10}, 1^{-9}, 1^{-8}, 1^{-7}]$ и доли атакуемых данных $fract = [0.2, 0.4, 0.6, 0.8, 0.95, 0.99999]$.

На рисунке 4 показана зависимость метрик качества от размера ε окрестности для обученной модели линейной регрессии. Значение $\varepsilon = 1^{-7}$ и выше приводит к резкому росту значений метрики MSE и снижению значений метрики R^2, что является нецелесообразным при проведении состязательной атаки, т.к. очень сильное отклонение модели будет расцениваться как выброс (outlier) или аномалия в данных.

В результате исследований по варьированию параметров FGSM можем сделать вывод о том, что модель линейной регрессии наиболее уязвима к атаке FGSM с максимизацией знака градиента с параметрами ε = 1^{-10} и fract = 0.99999, в остальных конфигурациях отклонения в метриках незначительны.

3. Сценарий обнаружения и противодействия состязательным атакам на исходные данные (Secured Model). Полученные во втором сценарии отравленные наборы данных использованы для обучения классификаторов LightGBM, CatBoost и XGBoost для решения задачи бинарной классификации: обычные данные (benign data) с меткой «0» или отравленные данные (malicious data) с меткой «1». Оптимальные параметры классификаторов подобраны с помощью

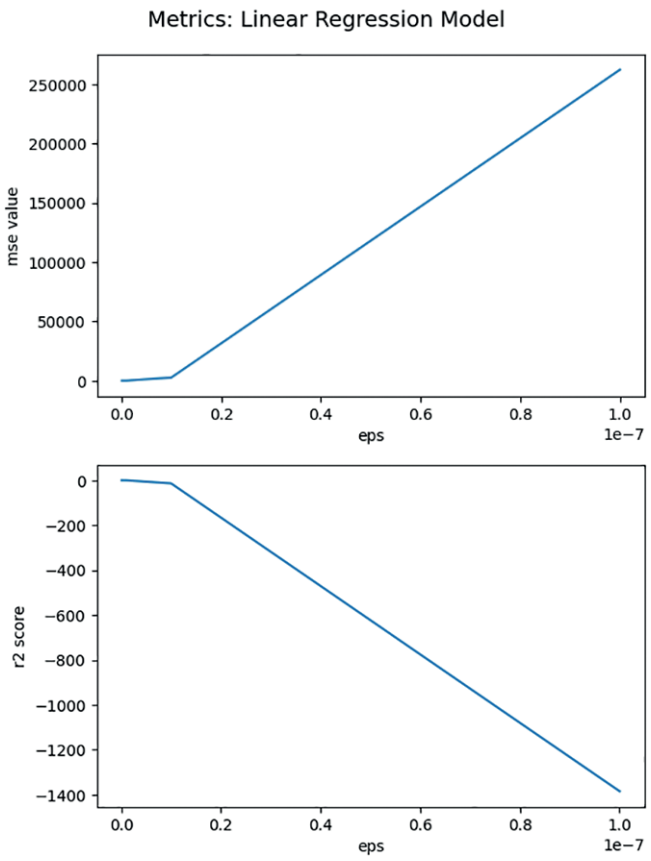

Рис. 4. Зависимость значений метрик MSE и R^2 score от размера ε окрестности

инструмента GridSearchCV. В таблице 1 представлены результаты сравнения трёх классификаторов. Для обучения классификаторов случайным образом выбран набор данных, в котором отравленные и обычные данные пропорционально сбалансированы.

Таблица 1

Сравнение бинарных классификаторов по обнаружению аномалий

Классификатор	$\varepsilon = 1^{-10}$, $fract = 0.6$		
	Precision	**Recall**	**F1-score**
LGBMClassifier (max_depth=20, n_estimators=500, num_leaves=20, subsample=0.7)	0.9835	0.9833	0.9834
CatBoost (depth'=4, 'learning_rate'=0.02, 'iterations'=100)	0.9777	0.9646	0.9703
XGBoost (n_estimators=500)	0.9828	0.9816	0.9822

Лучшие результаты по обнаружению состязательных аномалий показывает классификатор LightGBM с параметрами max_depth=20, n_estimators=500, num_leaves=20, и subsample=0.7. По результатам

работы классификатора на тестовых данных удалим из набора данных обнаруженные состязательные примеры и получим сокращенный набор данных из 5029 записей, на котором повторно выполним оценку качества регрессионной модели.

4. Обсуждение и выводы

На каждом из трёх этапов выполнялся подсчет основных метрик качества регрессионных моделей. В таблице 2 представлена динамика изменения метрик качества линейной регрессионной модели в зависимости от исследуемого сценария. Выполнение состязательной атаки FGSM с максимизацией знака градиента и параметрами показателя окрестности $\varepsilon = 1^{-10}$ и доли атакуемых данных $fract = 0.99999$ увеличивает значение метрики MSE в среднем на 33% и снижает значение метрики R^2 в среднем на 10%. Бинарный классификатор LightGBM с подобранными оптимальными гиперпараметрами с точностью в 98% успешно обнаруживает записи с состязательными аномалиями в табличных данных, изоляция которых позволяет восстановить метрики регрессионной модели до исходных значений.

В рамках проведенного исследования выполнена генерация табличных данных сценария сегмента беспроводной сети на базе эмулятора DeepMIMO; выполнено построение состязательных примеров с целью максимизации прогнозируемого значения комбинированных потерь сигнала от базовой станции до конечных пользователей; выполнено обучение бинарного классификатора по распознаванию отравленных данных; показана динамика изменения метрик качества линейной регрессионной модели в приложениях беспроводных сетей поколения 6G. Регрессионные модели машинного обучения уязвимы к состязательным атакам, своевременный интеллектуальный анализ сетевого трафика и передаваемых по сети данных может обнаруживать злонамеренную сетевую активность в сегменте беспроводной сети последнего поколения.

Таблица 2

Динамика изменения метрик качества линейной регрессионной модели

Сценарий	$\varepsilon = 1^{-10}$, $fract = 0.99999$	
	MSE	R2
Undefended Model {Linear Regression}	38.51	0.80
Attacked Model {FGSM}	51.40 ↑	0.72 ↓
Secured Model {LightGBM}	37.55 ↓	0.80 ↑

Исследование выполнено за счет гранта Российского научного фонда (проект № 22-71-10124).

Литература

1. *Петров И. А. Безопасность физического уровня для сетей 5G/6G // Вопросы кибербезопасности. – 2023. – №. 3. – С. 55.*
2. *Котенко И. В. и др. Атаки и методы защиты в системах машинного обучения: анализ современных исследований // Вопросы кибербезопасности. – 2024. – №. 1. – С. 59.*
3. *Gupta K. et al. An adversarial attacker for neural networks in regression problems // IJCAI Workshop on Artificial Intelligence Safety (AI Safety). – 2021.*
4. *Ribeiro A. H., Schön T. B. Overparameterized linear regression under adversarial attacks // IEEE Transactions on Signal Processing. – 2023. – V. 71. – P. 601–614.*
5. *Li F., Lai L., Cui S. On the adversarial robustness of linear regression // 2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP). – IEEE, 2020. – P. 1–6.*
6. *Mode G. R., Hoque K. A. Adversarial examples in deep learning for multivariate time series regression // 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR). – IEEE, 2020. – P. 1–10.*
7. *Kong X., Ge Z. Adversarial Attacks on Regression Systems via Gradient Optimization // IEEE Transactions on Systems, Man, and Cybernetics: Systems. – 2023.*
8. *Meng L. et al. White-box target attack for EEG-based BCI regression problems // Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part I 26. – Springer International Publishing, 2019. – P. 476–488.*
9. *Santana E. J. et al. Detecting and mitigating adversarial examples in regression tasks: A photovoltaic power generation forecasting case study // Information. – 2021. – V. 12. – №. 10. – P. 394.*
10. *Balda E. R., Behboodi A., Mathar R. Perturbation analysis of learning algorithms: Generation of adversarial examples from classification to regression // IEEE Transactions on Signal Processing. – 2019. – V. 67. – №. 23. – P. 6078–6091.*
11. *Zhao P., Wan Z. Robust nonparametric regression under poisoning attack // Proceedings of the AAAI Conference on Artificial Intelligence. – 2024. – V. 38. – №. 15. – P. 17007–17015.*
12. *Deng Y. et al. An analysis of adversarial attacks and defenses on autonomous driving models // 2020 IEEE international conference on pervasive computing and communications (PerCom). – IEEE, 2020. – P. 1–10.*
13. *Костогрызов А. И., Нистратов А. А. Анализ угроз злоумышленной модификации модели машинного обучения для систем с искусственным интелектом // Вопросы кибербезопасности. – 2023. – №. 5. – С. 9.*
14. *Ballet V. et al. Imperceptible adversarial attacks on tabular data // arXiv preprint arXiv:1911.03274. – 2019. DOI: https://doi.org/10.48550/arXiv.1911.03274*
15. *Alkhateeb A. DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications // arXiv preprint arXiv:1902.06435. – 2019. DOI: https://doi.org/10.48550/arXiv.1902.06435*