

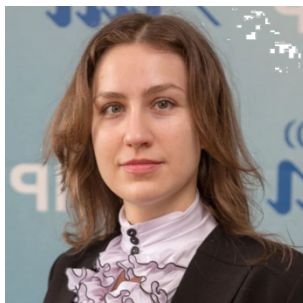
УДК 621.391.7: 004.9

ПРИМЕНЕНИЕ МЕТОДА МАШИННОГО ОБУЧЕНИЯ KNN ДЛЯ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ИНТЕНСИВНОСТИ ТРАФИКА В СЕТЯХ 5G (НА ПРИМЕРЕ ВИДЕОКОНФЕРЕНЦИЙ)



З. Джафаров

Заведующий кафедрой
компьютерных технологий
Азербайджанского
технического университета,
кандидат технических наук,
доцент
zafar.cafarov@aztu.edu.az



О.В. Бойправ

Заведующий кафедрой защиты
информации учреждения
образования «Белорусский
государственный университет
информатики и
радиоэлектроники», кандидат
технических наук, доцент
smi@bsuir.by



Э. Имамалиев

Доцент кафедры
компьютерных технологий
Азербайджанского
технического университета,
кандидат технических наук,
доцент
elman.imameliyev@aztu.edu.az

З. Джафаров

Окончил Азербайджанский институт нефти и химии (ныне Азербайджанский государственный нефтегазовый университет). Область научных интересов включает искусственный интеллект и машинное обучение, анализ данных и тестирование на основе моделей.

О.В. Бойправ

Окончила учреждение образования «Белорусский государственный университет информатики и радиоэлектроники». Области научных интересов – электромагнитное экранирование, кибербезопасность.

Э. Имамалиев

Окончил Азербайджанский технический университет. Область научных интересов охватывает современные сетевые технологии, архитектуры компьютерных сетей, методы анализа сетевого трафика и математические модели сетевых систем.

Аннотация. В работе представлены результаты исследования применения метода машинного обучения k -ближайших соседей (KNN) для краткосрочного прогнозирования трафика видеоконференций в сетях 5G. Актуальность исследования обусловлена тем, что трафик видеоконференций имеет нестационарный и всплесковый характер, что требует наличия точных прогнозов для управления ресурсами сетей в реальном времени и минимизации задержек. В ходе проведения исследования выполнено моделирование трафика, передаваемого по сети 5G, оценено влияние значения k на производительность с использованием показателей RMSE и MAE, на основании чего определено наиболее оптимальное значение k . На основе полученных результатов исследования установлено следующее: 1) метод KNN характеризуется простой реализацией; 2) с помощью метода KNN можно выполнять краткосрочное прогнозирование интенсивности трафика в сетях 5G с высокой точностью, что позволяет оптимизировать управление ресурсами указанных сетей и сокращать задержки трафика видеоконференций.

Ключевые слова: сеть 5G, видеоконференция, прогнозирование трафика, метод ближайших соседей, машинное обучение.

Введение. В последние годы сети 5G играют важную роль в развитии беспроводной связи с высокой скоростью, низкой задержкой и широкополосным доступом. Особенно в приложениях реального времени, таких как видеоконференции, интенсивность трафика быстро меняется и носит нестационарный характер. Это требует краткосрочного прогнозирования трафика для эффективного управления сетевыми ресурсами и оптимизации пользовательского опыта. Хотя традиционные статистические методы (например, ARIMA и Холта-Винтерса) являются точными, они имеют ограничения в прогнозировании трафика с всплесковыми паттернами, в частности трафика, формируемого в рамках видеоконференций. Поэтому методы машинного обучения, особенно метод k -ближайших соседей (KNN), позволяют более точно прогнозировать будущие значения трафика путем анализа паттернов трафика на уровне соседних узлов. В рамках представляемой работы выполнено исследование применения метода KNN для краткосрочного прогнозирования трафика видеоконференций в сетях 5G. Цель работы состояла в установлении зависимости точности и производительности указанного метода от значений k . Для достижения цели были решены следующие задачи:

- построена модель KNN для трафика, формируемого в рамках видеоконференций;
- выполнена оценка точности прогнозирования трафика с использованием метода при различных значениях k ;
- выполнена оценка производительности метода при различных значениях k с помощью показателей ошибок RMSE и MAE [1, 2];
- установлено наиболее оптимальное значение k для решения задач прогнозирования трафика, формируемого в рамках видеоконференций в сетях 5G.

Ключевые аспекты применения машинного обучения в сетях 5G. Прогнозирование краткосрочного трафика давно является предметом исследований в области управления сетями и оптимизации ресурсов. Традиционные статистические методы, такие как ARIMA и Холт-Винтерс, обеспечивают определенную точность при использовании данных временных рядов. Однако эти методы имеют ограничения в прогнозировании нестационарного трафика с всплесками, например, трафика видеоконференций.

В последние годы методы машинного обучения приобрели важное значение в прогнозировании трафика. Такие методы, как случайный лес, машины опорных векторов (SVM) и долговременная кратковременная память (LSTM), позволяют более точно изучать нелинейные и сложные закономерности трафика.

С другой стороны, метод KNN – это простой, легко интерпретируемый метод прогнозирования, который особенно интересен для приложений реального времени. Предыдущие исследования показали, что KNN хорошо работает с краткосрочным трафиком и что выбор гиперпараметров (значений k) напрямую влияет на точность прогнозирования. Однако исследования трафика видеоконференций в сетях 5G все еще ограничены. Существующие работы, как правило, проводились на трафике LTE или 4G, и применение KNN и других методов машинного обучения для управления ресурсами в реальном времени не получило широкого распространения. Поэтому тема применения KNN к трафику видеоконференций в сетях 5G является актуальной и инновационной.

В данной работе данные о трафике получены как на основе данных из реальных журналов событий сети 5G, так и на основе смоделированного набора данных. В ходе моделирования характеристики реального трафика имитировались путем добавления базового трафика, всплесков и случайного шума. Значения трафика за три предыдущих интервала ($lag1$, $lag2$, $lag3$) использовались в качестве входных переменных модели.

Набор данных был разделен на 80 % обучающей и 20 % тестовой частей путем деления на обучающую и тестовую выборки. Модель KNN обеспечивает прогнозирование на основе соседних узлов, то есть значение трафика в следующем интервале рассчитывается как среднее значение k ближайших соседей. Были протестированы различные значения k (1–20), и производительность оценивалась на основе RMSE и MAE. Было определено наиболее

оптимальное значение k , и окончательные прогнозы были сделаны с использованием этого параметра.

Точность прогнозов была показана с помощью визуальных графиков, а фактический и прогнозируемый трафик сравнивались. Этот подход демонстрирует как эффективность метода KNN для краткосрочных прогнозов, так и влияние выбора гиперпараметров на производительность. Результаты показывают, что KNN обеспечивает надежные прогнозы для трафика, сформированного в рамках видеоконференции в сети 5G, при этом оставаясь простым и легким для интерпретации.

Предложенный метод KNN был протестирован на смоделированном трафике, сформированном в рамках видеоконференции в сети 5G (рисунок 1). В качестве входных признаков при тестировании использовались значения трафика за три последних интервала ($lag1$, $lag2$, $lag3$). Набор данных был разделен на 80 % обучающей и 20 % тестовой частей, а производительность оценивалась для различных значений k (от 1 до 20).

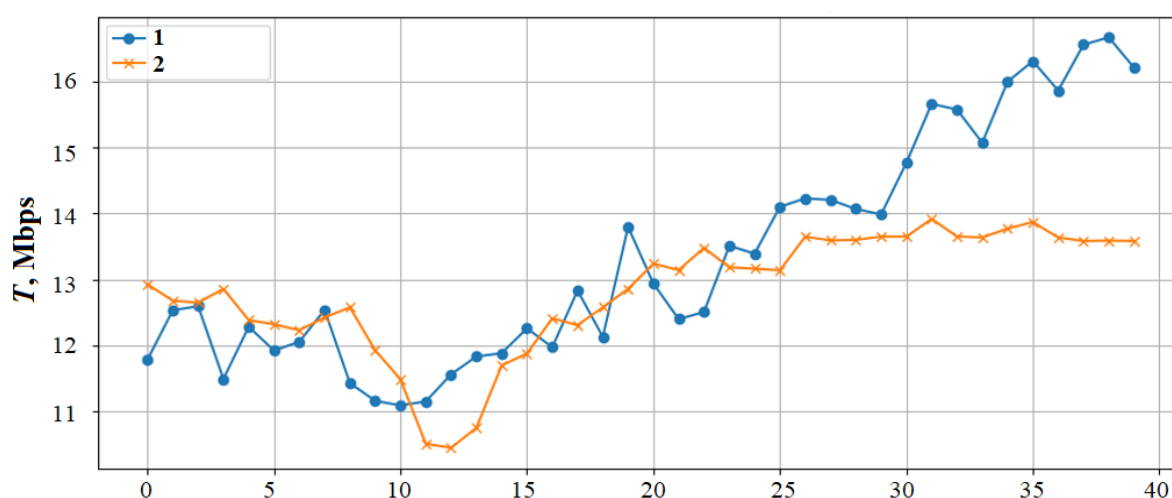


Рисунок 1. Графические представления трафика, формируемого в рамках видеоконференции в сети 5G, полученные на основе реальных данных (кривая 1) и результатов прогнозирования с использованием метода KNN (кривая 2)

Прогноз KNN в целом следует тренду реального трафика, но отстает, особенно в часы пик. Например, в интервале 25–38 наблюдается увеличение реального трафика, но KNN остается относительно стабильным (~ 13–14 Мбит/с). KNN прогнозирует более усредненное поведение и плохо улавливает внезапные всплески и спады. Это основное ограничение KNN для краткосрочного прогнозирования пикового трафика. В интервале 0–20 KNN относительно хорошо соответствует реальному трафику.

Ошибка прогнозирования (ER) для каждого k измерялась с помощью показателей RMSE и MAE. На основе анализа полученных результатов измерений (рисунок 2) установлено, что слишком малое или слишком большое значение k приводит к увеличению ошибок. Результаты также показали, что метод KNN, будучи простым и легким для интерпретации, обеспечивает высокую точность краткосрочных прогнозов. Экспериментальный анализ показывает, что как выбор гиперпараметров, так и количество прошлых интервалов, оказывают существенное влияние на точность прогноза.

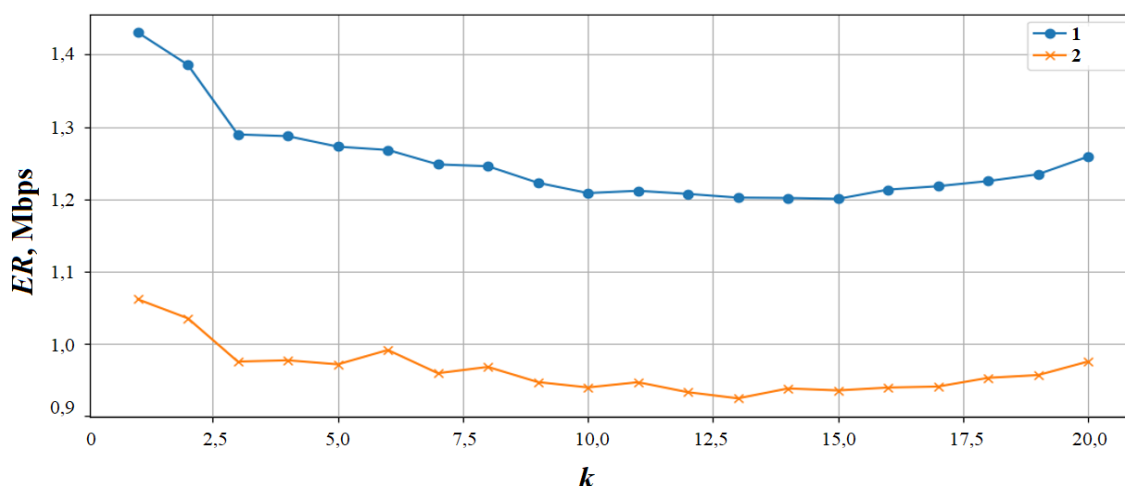


Рисунок 2. Зависимости ошибки прогнозирования от значения k , полученные с помощью показателей RMSE (кривая 1) и MAE (кривая 2)

Полученные результаты указывают на то, что метод KNN может эффективно применяться для управления ресурсами в реальном времени и оптимизации трафика, формируемого в рамках видеоконференций в сетях 5G. Кроме того, результаты показывают, что метод KNN обеспечивает высокую точность краткосрочных прогнозов трафика видеоконференций в сетях 5G. Оптимальное значение k , равное 5, минимизирует как RMSE (1,273 Мбит/с), так и MAE (0,971 Мбит/с), обеспечивая высокое соответствие между реальным трафиком и прогнозом. Слишком малое значение k делает прогнозы чувствительными к шуму, в то время как слишком большое значение k приводит к задержке прогнозов. Установлено, что метод KNN эффективно обучается как базовым, так и импульсным паттернам. Выбор количества прошлых интервалов (lag1–lag3) оказывает существенное влияние на точность прогнозирования; слишком мало информации об интервалах не обеспечивает достаточного контекста, в то время как слишком много интервалов замедляет работу модели сверх реального уровня. Простота и интерпретируемость метода делают его подходящим для управления сетью в реальном времени. Однако KNN имеет некоторые ограничения в обработке нестационарного трафика и трафика с очень высокой дисперсией.

Поэтому в будущих исследованиях можно достичь более высокой производительности с помощью гибридных моделей, таких как интеграция KNN-LSTM или онлайн-подходы KNN.

Кроме того, установлено, что метод KNN подходит как для экспериментов, так и для исследований, и может эффективно применяться для оптимизации ресурсов и сокращения задержек в трафике, формируемом в рамках видеоконференций в сетях 5G. Экспериментальная часть исследования проводилась на основе среды программирования Python. Для обработки данных, построения моделей и визуализации результатов использовался ряд научных библиотек экосистемы Python [3].

В ходе эксперимента данные временных рядов о трафике видеоконференций сначала обрабатывались с помощью Pandas, а затем обучались с помощью модели KNN, реализованной в библиотеке Scikit-learn.

После обучения модели полученные результаты прогнозирования сравнивались с реальными значениями трафика и представлялись в графической форме с использованием библиотеки Matplotlib.

Эта программная среда обеспечивала возможность повторения экспериментов, сравнительного анализа моделей и визуальной интерпретации результатов.

Заключение. Данное исследование показывает, что метод KNN может использоваться в качестве эффективного и практичного инструмента прогнозирования трафика видеоконференций в сетях 5G, а также имеет потенциал применения для оптимизации ресурсов и снижения задержки. В будущих исследованиях предлагается применять гибридные модели, например, интегрировать KNN с LSTM или тестировать онлайн-подходы KNN. Также перспективными направлениями исследований являются тестирование предложенного метода на реальных журналах событий сети 5G, адаптация модели для различных приложений видеоконференций и автоматизация оптимизации гиперпараметров.

Список литературы

- [1] Chen, M., Mao, S., & Liu, Y. (2014). *Big Data: A Survey*. *Mobile Networks and Applications*, 19(2), 171–209.
[2] Alaa A. Hussien, Heba Nashaat & Rehab F. Abdel-Kader. Machine Learning Techniques for Network Traffic Prediction in 5G Networks: Discover Applied Sciences Aims and scope. Volume 7, article number 1047, (2025)
[2] Scikit-learn: Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.

Авторский вклад

Джафаров Зафар – разработка концепции и методологии исследования, проведение моделирования и экспериментов, выполнение анализа результатов, подготовка первоначального варианта статьи.

Бойправ Ольга Владимировна – обзор литературы и подготовка введения, редактирование описания методологии и экспериментальных результатов, подготовка графиков и таблиц, окончательное редактирование статьи.

Имамалиев Эльман – разработка концепции и структуры исследования, описание методологии, проведение моделирования и экспериментов, анализ результатов, подготовка первоначального варианта статьи.

APPLICATION OF THE KNN MACHINE LEARNING METHOD FOR SHORT-TERM FORECASTING OF TRAFFIC INTENSITY IN 5G NETWORKS (ON THE EXAMPLE OF VIDEO CONFERENCES)

Z. Cafarov

*Head, Department of
Computer Technologies,
Azerbaijan Technical
University, PhD Associate
Professor*

O.V. Boiprav

*Head, Information Protection
Department of Computer Technologies,
Educational Establishment “Belarusian
State University of Informatics and
Radioelectronics”, PhD, Associate
Professor*

E. Imameliyev

*Associate Professor,
Department of Computer
Technologies, Azerbaijan
Technical University, PhD
Associate Professor*

Abstract. The paper presents the results of a study on the application of the k -nearest neighbors (KNN) machine learning method for short-term forecasting of videoconferencing traffic in 5G networks. The relevance of the study is due to the fact that videoconferencing traffic is non-stationary and bursty, which requires accurate forecasts for real-time network resource management and minimizing latency. During the study, traffic transmitted over a 5G network was modeled, the impact of the k value on performance was assessed using RMSE and MAE metrics, and the optimal k value was determined. Based on the obtained results, the following conclusions were established: 1) the KNN method is characterized by simple implementation; 2) using the KNN method, it is possible to perform short-term traffic intensity forecasting in 5G networks with high accuracy, which allows for optimized resource management of these networks and reduction of videoconferencing traffic latency.

Keywords: 5G network, video conferencing, traffic forecasting, nearest neighbor method, machine learning.