Инженерный вестник Дона, №7 (2025)
ivdon.ru/ru/magazine/archive/n7y2025/10227

Методы предиктивной аналитики для построения проактивной системы мониторинга сети

И.А. Клычков

Московский государственный технологический университет «СТАНКИН»

Аннотация: Современные методы предиктивной аналитики позволяют значительно расширить возможности систем сетевого мониторинга, обеспечивая раннее обнаружение аномалий и потенциальных отказов. В данной статье представлены результаты исследования подходов к построению проактивной системы мониторинга сети с применением методов машинного обучения и статистического анализа. Показано, что использование комбинированных моделей на основе рекуррентных нейронных сетей и авторегрессионных моделей обеспечивает наиболее точное прогнозирование сетевого трафика с горизонтом предсказания до 10 временных интервалов. Практическая реализация предложенного подхода позволяет сократить количество незапланированных простоев на 27% и снизить время реагирования на инциденты на 35% по сравнению с традиционными реактивными системами мониторинга.
Ключевые слова: предиктивная аналитика, сетевой мониторинг, машинное обучение, статистический анализ, обнаружение аномалий, прогнозирование трафика, рекуррентные нейронные сети, авторегрессионные модели, проактивные системы, отказоустойчивость.

Введение

В условиях непрерывного роста объемов и сложности сетевого трафика традиционные реактивные системы мониторинга становятся неэффективными для обеспечения стабильной работы современных телекоммуникационных сетей. Согласно данным исследований, незапланированные простои сетевой инфраструктуры приводят к значительным финансовым потерям для бизнеса [1, 2]. Так, системы инфраструктурного мониторинга, которые лишь фиксируют уже произошедшие сбои, не позволяют своевременно предотвращать проблемы и минимизировать их последствия.

Предиктивная аналитика представляет собой подход, который позволяет значительно расширить горизонт времени для прогнозирования возникновения дефектов и аномалий в работе сети. Проактивный подход к мониторингу сетевой инфраструктуры не только обеспечивает большой запас времени для реагирования, но и существенно снижает как последствия сбоев,

так и связанный с ними ущерб [3]. В современных условиях разработка и внедрение проактивных систем мониторинга становится критически важной задачей для обеспечения надежной работы сетевых инфраструктур.

Научная значимость исследования заключается в разработке комплексного подхода к построению проактивной системы мониторинга сети на основе интеграции методов предиктивной аналитики. В отличие от существующих исследований, сосредоточенных на отдельных аспектах прогнозирования сетевого трафика [4, 5], данная работа предлагает целостную методологию создания системы мониторинга, способной не только выявлять аномалии, но и предсказывать потенциальные проблемы до их возникновения.

Предлагаемый подход позволяет преодолеть ограничения традиционных моделей мониторинга, связанные с их применимостью только к сетям простой структуры со стабильными во времени объемами трафика [6]. Научная новизна исследования заключается в разработке гибридных моделей прогнозирования, учитывающих как кратковременные, так и долговременные зависимости в сетевом трафике, а также в методике определения оптимального горизонта прогнозирования для различных типов сетей.

Обзор литературы

Мониторинг и анализ сетевого трафика представляют ключевой компонент управления сетью, особенно для обеспечения правильной работы крупномасштабных сетей, таких как Интернет [7]. По мере увеличения сложности интернет-сервисов и объема трафика становится всё сложнее разрабатывать масштабируемые приложения для мониторинга и анализа сетевого трафика. Приложениям для классификации трафика и регулирования требуются подходы, работающие в режиме реального времени и обладающие высокой масштабируемостью.

Традиционно прогнозирование трафика строилось на основе статистических авторегрессионных моделей, которые устанавливали корреляционные связи между текущими объемами трафика и его возможными будущими значениями [6]. Крюков и Чернягин предложили модель ARIMA с минимальным числом параметров для прогнозирования сетевого трафика [5]. В своей работе авторы показали, что адекватный прогноз значений сетевого трафика можно осуществить с помощью моделей: ARIMA(0,1,1) на 2 шага вперед, ARIMA(1,1,0) – на 6 шагов вперед, ARIMA(1,1,1) на 50 шагов вперед.

Однако статистические модели имеют существенные ограничения, в частности, большинство из них применимы только к стационарным статистическим данным, что сужает их область применения до сетей простой структуры со стабильными во времени объемами потребляемого трафика. Большинство современных сетей имеют сложную структуру и подвержены быстрым изменениям потоков трафика.

Для преодоления ограничений традиционных моделей и улучшения прогнозирования в последнее время предпринимаются активные усилия по использованию технологий машинного обучения. Митрохин и Башков в своей работе представили методологический подход к прогнозированию сетевого трафика в мобильных сетях, в частности в сетях LTE, с использованием рекуррентных нейронных сетей долгой краткосрочной памяти (Long Short-Term Memory, LSTM) [5]. Авторы отмечают, что сети LSTM превосходят другие подходы машинного обучения для анализа временных рядов в прогнозировании трафика.

В исследовании Li et al. [8] предлагается подход на основе глубокого обучения для надежного выявления аномальных событий из системных журналов. Авторы провели эмпирическое исследование с использованием 18

различных моделей для прогнозирования анализа в виртуализации сетевых функций.

Согласно данным журнала «Информатизация и системы управления в промышленности» существует проактивный подход к мониторингу, который позволяет значительно расширить горизонт времени для прогнозирования появления дефектов [3]. В этом подходе первично моделируется объект как набор алгоритмов и соответствующих правил, создавая своего рода «цифровой двойник» оборудования, повторяющий поведение реального объекта. Далее реальные данные объектов мониторинга поступают в «цифровой двойник», и в режиме реального времени ведется анализ расхождений и разнообразных аномалий. За счет этого удается прогнозировать их состояние и выявлять неисправности на ранней стадии их возникновения.

Naumen отмечает, что предиктивная аналитика позволяет не только прогнозировать события по историческим данным об инцидентах, но и выявлять аномалии в работе системы до того, как пользователи столкнутся с проблемами [1].

В исследованиях по предиктивной аналитике для оценки качества моделей прогнозирования используются различные метрики. Крюков и Чернягин использовали метрики средней абсолютной процентной погрешности (Mean Absolute Percentage Error, MAPE) и коэффициент стандартной погрешности (Standard Error Ratio, SER) для оценки качества прогнозных моделей [5]. Авторы показали, что оценка MAPE модели ARIMA(3,1,4) минимальна и находится между 10% и 20%, что соответствует хорошему прогнозу для двух шагов вперед.

Трошин отмечает, что для оценки эффективности предиктивных моделей в задачах прогнозирования сетевого трафика используются такие метрики, как среднеквадратическая ошибка (Root Mean Square Error, RMSE),

средняя абсолютная ошибка (Mean Absolute Error, MAE) и коэффициент детерминации (R^2) [6].

Методология исследования

В рамках данного исследования предлагается методология построения проактивной системы мониторинга сети, основанная на интеграции методов предиктивной аналитики. Концепция системы предполагает многоуровневую архитектуру, включающую:

1. Уровень сбора и предварительной обработки данных
2. Уровень анализа и прогнозирования
3. Уровень визуализации и принятия решений

Для реализации уровня сбора данных используется комплексный подход, предложенный в работе о мониторинге и предиктивной аналитике технологического оборудования на базе промышленного интернета вещей [9]. Система мониторинга состоит из устройств сбора данных, беспроводной сенсорной сети и облачного сервера.

Ключевым элементом предлагаемой методологии является уровень анализа и прогнозирования, который реализует гибридный подход к предсказанию возможных аномалий и отказов в сети.

В данном исследовании использовались данные сетевого трафика, собранные с помощью систем мониторинга в различных сегментах корпоративной сети. Сбор данных осуществлялся в течение 3 месяцев с 5-минутными интервалами. Массив данных включал следующие метрики:

- Объем входящего и исходящего трафика
- Задержка в передаче пакетов
- Процент потери пакетов
- Загрузка процессора и памяти сетевого оборудования
- Количество сетевых соединений

Для предварительной обработки данных были применены методы нормализации, устранения выбросов и заполнения пропусков. Особое внимание уделялось выявлению сезонных паттернов в данных, поскольку сетевой трафик обычно демонстрирует ярко выраженную суточную и недельную периодичность [4, 10].

В исследовании были разработаны и протестированы следующие модели прогнозирования:

1. Статистические модели:
    - ARIMA(p,d,q) с различными параметрами
    - Экспоненциальное сглаживание (ETS)
2. Модели машинного обучения:
    - Рекуррентные нейронные сети LSTM
    - Сверточные нейронные сети (Convolutional neural networks, CNN)
    - Гибридные модели CNN-LSTM
    - Модели градиентного бустинга (XGBoost, LightGBM)
3. Ансамблевые модели:
    - Стекинг на основе ARIMA и LSTM
    - Взвешенное усреднение прогнозов различных моделей

Для обучения моделей использовался метод скользящего окна с различными размерами окна и горизонта прогнозирования. Оптимизация гиперпараметров моделей проводилась с использованием метода поиска по сетке и байесовской оптимизации.

Для оценки эффективности разработанных моделей использовались следующие метрики:

- Средняя абсолютная процентная ошибка (MAPE)
- Среднеквадратическая ошибка (RMSE)
- Коэффициент детерминации (R^2)

- F1-мера для задачи классификации аномалий

Кроме того, для оценки практической ценности проактивной системы мониторинга были введены следующие метрики:

- Время упреждения (интервал времени между прогнозом аномалии и ее фактическим возникновением)
- Процент предотвращенных инцидентов
- Снижение времени реагирования на инциденты

Целью данного исследования является разработка и экспериментальная верификация методологии построения проактивной системы мониторинга сети на основе методов предиктивной аналитики, обеспечивающей раннее обнаружение аномалий и потенциальных отказов в работе сетевой инфраструктуры.

Для достижения поставленной цели были сформулированы следующие задачи:

1. Анализ существующих методов предиктивной аналитики в контексте их применимости для мониторинга сетевого трафика.
2. Разработка архитектуры проактивной системы мониторинга сети, интегрирующей методы предиктивной аналитики.
3. Создание и оптимизация комплекса моделей машинного обучения для прогнозирования различных аспектов сетевого трафика.
4. Определение оптимального горизонта прогнозирования для различных типов сетей и характеристик трафика.
5. Экспериментальная оценка эффективности разработанных моделей и системы в целом.
6. Формирование рекомендаций по внедрению предложенного подхода в практику управления сетевой инфраструктурой.

Результаты

В ходе исследования было проведено сравнение различных моделей предиктивной аналитики для прогнозирования сетевого трафика. Результаты оценки моделей с горизонтом прогнозирования от 1 до 10 временных интервалов представлены в таблице 1.

Таблица №1

Сравнение моделей прогнозирования сетевого трафика

Модель	MAPE (1 шаг)	MAPE (2 шага)	MAPE (5 шагов)	MAPE (10 шагов)
ARIMA(1,1,0)	8.3%	12.4%	15.7%	21.2%
ARIMA(0,1,1)	7.9%	14.6%	18.9%	24.5%
ARIMA(1,1,1)	6.8%	13.2%	17.5%	22.8%
ARIMA(3,1,4)	6.2%	9.8%	16.9%	23.7%
LSTM	5.7%	9.2%	13.8%	18.4%
CNN	6.9%	10.8%	15.2%	21.9%
CNN-LSTM	5.4%	8.7%	12.5%	17.3%
XGBoost	7.1%	11.5%	16.2%	22.4%
Ансамблевая модель	5.2%	8.3%	11.9%	16.5%

Как видно из представленных результатов, наиболее точные прогнозы обеспечивают ансамблевые модели и гибридные нейросетевые архитектуры CNN-LSTM. При этом для краткосрочного прогнозирования (1-2 шага) разница между различными моделями относительно невелика, однако с увеличением горизонта прогнозирования преимущество нейросетевых моделей становится более выраженным. Это согласуется с выводами Митрохина и Башкова о том, что сети LSTM превосходят другие подходы машинного обучения для анализа временных рядов в прогнозировании трафика [3].

Для оценки способности моделей выявлять аномалии в сетевом трафике был проведен эксперимент, в котором в тестовый набор данных были добавлены искусственно созданные аномалии различных типов

(внезапные пики трафика, медленные утечки, периодические колебания). Результаты представлены в таблице 2.

Таблица №2

Эффективность моделей в обнаружении аномалий

Модель	Точность	Полнота	F1-мера	Время упреждения (мин)
ARIMA	0.72	0.68	0.70	15.3
LSTM	0.85	0.79	0.82	22.7
CNN-LSTM	0.87	0.82	0.84	26.5
Ансамблевая модель	0.89	0.85	0.87	28.2

Результаты показывают, что нейросетевые и ансамблевые модели обеспечивают более высокую эффективность обнаружения аномалий и больший интервал упреждения по сравнению со статистическими моделями.

На основе результатов экспериментов была разработана архитектура проактивной системы мониторинга сети. Система включает следующие ключевые компоненты:

1. Модуль сбора данных: осуществляет сбор метрик с различных сетевых устройств и их предварительную обработку.

2. Модуль предиктивной аналитики: включает набор моделей для прогнозирования различных характеристик сетевого трафика и выявления потенциальных аномалий.

3. Модуль анализа рисков: оценивает вероятность и потенциальный ущерб от прогнозируемых аномалий.

4. Система принятия решений: формирует рекомендации для администраторов сети на основе результатов прогнозирования и оценки рисков.

5. Модуль визуализации: обеспечивает наглядное представление результатов мониторинга и прогнозирования.

Особенностью разработанной архитектуры является использование "цифрового двойника" сетевой инфраструктуры, который позволяет моделировать поведение сети в различных условиях и прогнозировать последствия потенциальных аномалий. Это соответствует подходу, где "цифровой двойник" оборудования используется для анализа расхождений и выявления аномалий [3].

Разработанная система была внедрена в инфраструктуру корпоративной сети крупной телекоммуникационной компании и тестировалась в течение 3 месяцев. За этот период были получены следующие результаты:

- Снижение количества незапланированных простоев на 27% по сравнению с предыдущим периодом.
- Сокращение времени реагирования на инциденты на 35%.
- Повышение эффективности использования полосы пропускания на 12%.
- Увеличение среднего времени между отказами (Mean time between failures, MTBF) на 19%.

Эти результаты подтверждают эффективность предложенного подхода к построению проактивной системы мониторинга сети на основе методов предиктивной аналитики.

Полученные результаты согласуются с выводами других исследователей о преимуществах методов глубокого обучения для прогнозирования сетевого трафика. Так, в работе Митрохина и Башкова отмечается, что сети LSTM превосходят другие подходы машинного обучения для анализа временных рядов в прогнозировании трафика [4]. В нашем исследовании это подтверждается более низкими значениями ошибки MAPE для моделей на основе LSTM по сравнению со статистическими моделями ARIMA.

В то же время текущие результаты расширяют выводы Крюкова и Чернягина, которые показали, что для краткосрочного прогнозирования (до 2 шагов) статистические модели типа ARIMA(3,1,4) обеспечивают хорошую точность прогноза [5]. В нашем исследовании показано, что для более длительных горизонтов прогнозирования (5-10 шагов) нейросетевые модели имеют значительное преимущество.

Важным отличием текущего подхода от существующих работ является использование ансамблевых методов, которые позволяют объединить преимущества различных моделей и повысить общую точность прогнозирования. Это согласуется с современными тенденциями в области машинного обучения, где ансамблевые методы показывают высокую эффективность в различных задачах прогнозирования.

Несмотря на полученные положительные результаты, исследование имеет ряд ограничений, которые следует учитывать при интерпретации результатов:

1. Тестирование системы проводилось в основном на данных корпоративной сети, что может ограничивать обобщаемость результатов на другие типы сетей (например, мобильные сети или сети интернет-провайдеров).

2. Не учитывались некоторые специфические факторы, влияющие на сетевой трафик, такие как внешние атаки или массовые отказы оборудования.

3. Временной горизонт тестирования (3 месяца) может быть недостаточным для оценки долгосрочной эффективности системы.

Выводы

В рамках данного исследования была разработана и экспериментально верифицирована методология построения проактивной системы мониторинга

сети на основе методов предиктивной аналитики. Основные результаты работы можно сформулировать следующим образом:

1. Проведен сравнительный анализ различных моделей предиктивной аналитики для прогнозирования сетевого трафика. Показано, что наилучшие результаты обеспечивают гибридные нейросетевые архитектуры CNN-LSTM и ансамблевые модели, особенно для горизонтов прогнозирования более 5 временных интервалов.

2. Разработана архитектура проактивной системы мониторинга сети, включающая модули сбора данных, предиктивной аналитики, анализа рисков, принятия решений и визуализации. Особенностью архитектуры является использование "цифрового двойника" сетевой инфраструктуры для моделирования сети и прогнозирования последствий аномалий.

3. Экспериментально показано, что внедрение разработанной системы позволяет снизить количество незапланированных простоев на 27%, сократить время реагирования на инциденты на 35% и повысить эффективность использования полосы пропускания на 12%.

4. Определены оптимальные горизонты прогнозирования для различных типов моделей и характеристик сетевого трафика. Для краткосрочного прогнозирования (1-2 шага) статистические модели ARIMA показывают сопоставимые с нейросетевыми моделями результаты, однако для более длительных горизонтов прогнозирования предпочтительно использование нейросетевых и ансамблевых моделей.

5. Сформулированы рекомендации по внедрению предложенного подхода в практику управления сетевой инфраструктурой, включая выбор оптимальных моделей прогнозирования в зависимости от характеристик сети и требуемого горизонта прогнозирования.

Таким образом, результаты исследования подтверждают эффективность применения методов предиктивной аналитики для

построения проактивной системы мониторинга сети и могут служить основой для практической реализации таких систем в различных типах сетевых инфраструктур.

Литература

1. Предиктивная аналитика в ИТ-мониторинге. URL: blog.naumen.ru/predictive-analytics-monitoring/ (дата обращения: 11.05.2025).

2. Прогноз для ИТ-инфраструктуры: как работает предиктивная аналитика. URL: blog.naumen.ru/howto-predictive-analytics/ (дата обращения: 11.05.2025).

3. Предиктивная аналитика и «цифровая зрелость». URL: isup.ru/articles/1/16123/ (дата обращения: 11.05.2025).

4. Митрохин В. Е., Башков И. Н. Прогнозирование мобильного трафика из необработанных данных с использованием сетей LSTM // Радиотехника, электроника и связь. 2019. С. 117–123.

5. Крюков Ю. А., Чернягин Д. В. ARIMA – модель прогнозирования значений трафика // Информационные технологии и вычислительные системы. 2011. № 2. С. 41–49.

6. Трошин А. В. Машинное обучение для прогнозирования трафика // Инфокоммуникационные технологии. 2019. Т. 17. № 4. С. 400–407.

7. Casas P., D'Alconzo A., Fiore M., Barlet-Ros P. et. al. A Survey on Big Data for Network Traffic Monitoring and Analysis URL: arxiv.org/abs/2003.01648 (дата обращения: 12.05.2025).

8. Li Z., Wang X., Chen Y., Xu K., Li Y. Predictive Analysis in Network Function Virtualization // Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies. 2018. pp. 120–133.

9. Добрынин С. Л., Бурковский В. Л. Мониторинг и предиктивная аналитика технологического оборудования на базе промышленного

интернета вещей // Вестник Воронежского государственного технического университета. 2020. Т. 16. № 5. С. 7-12.

10. Калытюк И. С., Французова Г. А., Гунько А. В. К вопросу выбора методов предиктивного анализа данных социальных медиа // Автоматика и программная инженерия. 2019. №4 (30). С. 9-17.

References

1. Prediktivnaja analitika v IT-monitoringe [Predictive analytics in IT monitoring]. URL: blog.naumen.ru/predictive-analytics-monitoring/ (accessed 11.05.2025).

2. Prognoz dlja IT-infrastruktury: kak rabotaet prediktivnaja analitika [Forecast for IT infrastructure: how predictive analytics works]. URL: blog.naumen.ru/howto-predictive-analytics/ (accessed 11.05.2025).

3. Prediktivnaja analitika i «cifrovaja zrelost'» [Predictive analytics and digital maturity]. URL: isup.ru/upload/pdf-zhurnala/2020/6-89-2020/050_052_IndaSoft_3.pdf (accessed 11.05.2025).

4. Mitrokhin V. E., Bashkov I. N. Radiotekhnika, elektronika i svyaz'. 2019. pp. 117–123.

5. Kryukov Yu. A., Chernyagin D. V. Informatsionnye tekhnologii i vychislitel'nye sistemy. 2011. № 2. pp. 41–49.

6. Troshin A. V. Infokommunikatsionnye tekhnologii. 2019. V. 17. № 4. pp. 400–407.

7. Casas P., D'Alconzo A., Fiore M., Barlet-Ros P. et. al. A Survey on Big Data for Network Traffic Monitoring and Analysis URL: arxiv.org/abs/2003.01648 (accessed 12.05.2025).

8. Li Z., Wang X., Chen Y., Xu K., Li Y. Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies. 2018. P. 120–133.

9. Dobrynin S. L., Burkovskiy V. L. Vestnik Voronezhskogo gosudarstvennogo tekhnicheskogo universiteta. 2020. V. 16. № 5. pp. 7-12.

10. Kalytyuk I. S., Frantsuzova G. A., Gun'ko A. V. Avtomatika i programmnaya inzheneriya. 2019. №4 (30). pp. 9-17.

Дата поступления: 16.05.2025

Дата публикации: 26.06.2025