https://doi.org/10.25559/SITITO.021.202501.36-45
УДК 004.4

Оригинальная статья

Разработка и анализ методики отбора инфраструктурных метрик для предиктивного мониторинга инцидентов

А. В. Егоркин[1,2]

[1] ФГБОУ ВО «Московский государственный университет имени М. В. Ломоносова», г. Москва, Российская Федерация
Адрес: 119991, Российская Федерация, г. Москва, ГСП-1, Ленинские горы, д. 1

[2] Публичное акционерное общество «Сбербанк России», г. Москва, Российская Федерация
Адрес: 117312, Российская Федерация, г. Москва, ул. Вавилова, д. 19
egorkinandrewv@mail.ru

Аннотация

Рост объёма телеметрии в распределённых ИТ-системах приводит к «информационному шуму» и увеличивает вычислительные затраты AIOps-платформ. В работе предложена формализованная двухэтапная процедура отбора метрик, предназначенная для повышения точности и эффективности предиктивного мониторинга: (1) мультикритериальный корреляционный фильтр, использующий коэффициенты Пирсона ($|r| > 0{,}60$), τ-Кендалла ($> 0{,}50$) и Maximal Information Coefficient (MICe $> 0{,}35$) для отсева избыточных и нелинейно связанных признаков; (2) верификация причинно-следственных связей с помощью теста Грейнджера (lag = 5, $p < 0{,}01$), алгоритма PCMCI (FDR = 10%) и метрики Directed Information (DI > 0,1 бит/шаг) для выявления истинных драйверов целевой метрики. Экспериментальная апробация методики проведена на 14-суточном фрагменте потока Prometheus-метрик промышленного кластера системы «Сбер Antifraud» (≈ 7 млрд точек данных, 1379 исходных метрик). Результаты показали снижение средней абсолютной ошибки (MAE) прогноза 30-минутной утилизации CPU на 43% при одновременном уменьшении числа входных временных рядов в 14 раз и сокращении времени инференса модели на 89%. Методика интегрирована в промышленный конвейер обработки данных (Prometheus → Kafka → Spark 3.5 → MLflow 2.11) и соответствует принципу минимизации обрабатываемых данных, закреплённому в ГОСТ Р 57580.1-2017 и методических рекомендациях ФСТЭК РФ по защите информации.

Ключевые слова: AIOps, предиктивный мониторинг, отбор признаков, корреляция, MICe, причинность, причинно-следственный анализ, PCMCI, Directed Information, SRE, Green SRE, временные ряды

Конфликт интересов: автор заявляет об отсутствии конфликта интересов.

Для цитирования: Егоркин А. В. Разработка и анализ методики отбора инфраструктурных метрик для предиктивного мониторинга инцидентов // Современные информационные технологии и ИТ-образование. 2025. Т. 21, № 1. С. 36-45. https://doi.org/10.25559/SITITO.021.202501.36-45

Original article

Development and Analysis of a Methodology for Selecting Infrastructure Metrics for Predictive Incident Monitoring

A. V. Egorkin[a,b]
[a] Lomonosov Moscow State University, Moscow, Russian Federation
Address: 1 Leninskie gory, Moscow 119991, GSP-1, Russian Federation
[b] Sberbank of Russia, Moscow, Russian Federation
Address: 19 Vavilova St., Moscow 117312, Russian Federation
egorkinandrewv@mail.ru

Abstract
The growth of telemetry volume in distributed IT systems leads to “information noise” and increases the computational costs of AIOps platforms. This paper proposes a formalized two-stage metric selection procedure designed to improve the accuracy and efficiency of predictive monitoring: (1) a multicriteria correlation filter using Pearson coefficients ($|r| > 0.60$), Kendall’s τ (> 0.50), and Maximal Information Coefficient (MICe > 0.35) to eliminate redundant and non-linearly related features; (2) verification of causal relationships using the Granger test (lag = 5, $p < 0.01$), the PCMCI algorithm (FDR = 10%), and the Directed Information metric (DI > 0.1 bits/step) to identify true drivers of the target metric. Experimental validation was conducted on a 14-day fragment of Prometheus metrics from the industrial cluster of the “Sber Antifraud” system (≈7 billion data points, 1379 initial metrics). The results showed a 43% reduction in the Mean Absolute Error (MAE) of 30-minute CPU utilization forecasts, a 14-fold decrease in input time series, and an 89% reduction in model inference time. The methodology is integrated into an industrial data processing pipeline (Prometheus → Kafka → Spark 3.5 → MLflow 2.11) and aligns with the data minimization principle outlined in GOST R 57580.1-2017 and FSTEC guidelines for information protection.

Keywords: AIOps, predictive monitoring, feature selection, correlation, MICe, causality, causal analysis, PCMCI, Directed Information, SRE, Green SRE, time series

Conflict of interests: The author declares no conflict of interests.

For citation: Egorkin A.V. Development and Analysis of a Methodology for Selecting Infrastructure Metrics for Predictive Incident Monitoring. *Modern Information Technologies and IT-Education.* 2025;21(1):36-45. https://doi.org/10.25559/SITITO.021.202501.36-45

Введение

Современные цифровые сервисы, особенно в финансовом секторе, телекоммуникациях и государственных услугах, функционируют в условиях сверхвысоких требований к доступности и производительности [1, 2]. Уровень ожидаемой доступности часто составляет 99,99% («четыре девятки») и выше. Любые простои или деградация качества обслуживания (увеличение времени отклика, рост числа ошибок) напрямую приводят к значительным финансовым и репутационным потерям [3].

Для предотвращения таких потерь компании активно внедряют практики SRE *(Site Reliability Engineering)* и платформы AIOps *(Artificial Intelligence for IT Operations)*, стремясь перейти от реактивного режима («тушение пожаров») к проактивному управлению надежностью и прогнозированию потенциальных инцидентов. Однако этот переход сталкивается с двумя фундаментальными барьерами:

1. Масштаб телеметрии: Современные распределенные системы, построенные на микросервисной архитектуре и контейнеризации, генерируют огромные объемы данных мониторинга. Один средний кластер может порождать десятки тысяч уникальных временных рядов с высокой частотой сбора данных. Обработка и анализ такого потока данных в реальном времени представляет собой серьезную вычислительную задачу.

2. Проблема корреляции и причинности: Высокая степень корреляции между различными метриками не гарантирует наличия прямой причинно-следственной связи между ними. Зачастую наблюдаемая корреляция является ложной и обусловлена влиянием некоторого общего внешнего фактора (например, ростом пользовательской нагрузки). Использование таких ложно связанных метрик в моделях прогнозирования приводит к ухудшению их качества и снижению интерпретируемости[1] [4].

Подача «сырых», неотфильтрованных данных телеметрии напрямую в сложные модели машинного обучения приводит к эффекту «проклятия размерности», информационному шуму, переобучению и существенному увеличению вычислительных затрат на обучение и вывод модели в эксплуатацию[2]. Следовательно, этап интеллектуального отбора признаков *(feature selection)* становится критически важным компонентом любого эффективного AIOps-конвейера.

Цель настоящей статьи — представить разработанную и апробированную на практике воспроизводимую методику отбора инфраструктурных метрик для задач предиктивного мониторинга. Предлагаемая методика минимизирует информационный шум за счет двухэтапной фильтрации, сохраняя при этом причинно значимую информацию. В работе демонстрируется ее практическая польза и экономическая эффективность на примере крупномасштабного промышленного контура банковской антифрод-системы.

1 Связанные исследования

Проблема прогнозирования временных рядов и отбора признаков для ИТ-мониторинга активно исследуется. Существующие подходы можно сгруппировать по типам используемых моделей и методов отбора признаков (Таблица 1).

Таблица 1. **Обзор некоторых методов прогнозирования и отбора признаков в AIOps (сокращенная)**

Table 1. **Overview of Some Prediction and Feature Selection Methods in AIOps (Abridged)**

Категория	Метод/ Модель	Принцип работы	Преимущества	Недостатки	Ссылки
Статистические	SES, Holt-Winters	Экспоненциальное сглаживание тренда и сезонности	Простота, интерпретируемость	Плохо работает на сложных данных, не учитывает внешние факторы	[5], [6]
Глубокое обучение	LSTM, GRU	Рекуррентные сети для моделирования временных зависимостей	Улавливают сложные нелинейные паттерны	Требуют много данных, вычислительно дороги, чувствительны к шуму	[7-10]
Отбор признаков (фильтры)	Корреляция, MI, MICe	Оценка статистической связи с целевой переменной	Быстрота, простота	Не учитывают взаимодействие признаков, направление связи, ложные корреляции	[11-13]
Отбор признаков (причинность)	Тест Грейнджера	Проверка, улучшают ли прошлые значения X прогноз Y	Выявляет направленную связь	Требует стационарности, чувствителен к мультиколлинеарности, парные связи	[14]
Отбор признаков (причинность)	PCMCI, DI, LiNGAM и др.	Построение причинных графов с учетом условной независимости	Устойчивость к ложным связям, нелинейность	Вычислительная сложность, требуют подбора параметров	[15], [16]

Источник: здесь и далее в статье все таблицы и рисунки составлены автором.

Source: Hereinafter in this article all tables and figures were made by the author.

[1] Pearl J. Causality: Models, Reasoning and Inference. 2nd ed. New York: Cambridge University Press, 2009. 478 p.

[2] Егоркин А. В. Разработка и анализ методов прогнозирования инцидентов на основе инфраструктурных метрик : Магист. дисс. М.: МГУ имени М.В. Ломоносова, 2025. 56 с.

Классические статистические модели[3] просты [6], но недостаточно гибки для сложных ИТ-метрик. Рекуррентные нейронные сети (LSTM, GRU) [7], [17] лучше улавливают сложные зависимости, но требуют качественных данных. Простые фильтры признаков [11-13] быстры, но могут сохранять ложные корреляции. Методы анализа причинности [5], [14-16] более надежны для выявления реальных связей, но вычислительно сложнее. Работы [13], [15, 16] демонстрируют теоретические преимущества PCMCI и DI [18], однако их применение и интеграция в промышленные AIOps-конвейеры, а также оценка комплексного эффекта от их использования в связке с корреляционными фильтрами, остаются недостаточно изученными [19, 20].

Настоящая статья закрывает этот пробел, предлагая гибридную методику, объединяющую эффективность корреляционных фильтров (включая MICe [13]) и строгость современных методов проверки причинности (Granger, PCMCI, DI), и демонстрируя ее практическую ценность в реальном AIOps-сценарии.

2 Материалы и методы

2.1 Набор данных

Экспериментальная база исследования сформирована на основе реальных данных мониторинга промышленной системы «Сбер Antifraud».

- **Источник данных:** Поток временных рядов из системы мониторинга Prometheus.
- **Период наблюдения:** 14 последовательных дней (с 09 по 22 ноября 2024 г.).
- **Частота сбора данных:** 15 секунд для большинства метрик.
- **Исходный объем данных:** 1379 уникальных временных рядов (метрик), собранных с 312 узлов (серверов, виртуальных машин, подов Kubernetes). Данные включают метрики утилизации процессора (показатель того, как используется процессорное время в системе), использования памяти (RAM), дискового и сетевого ввода/вывода (I/O), специфические метрики очередей сообщений Kafka, базы данных PostgreSQL, состояния подов Kubernetes и др. Общий объем необработанных данных за период составил около 7 миллиардов точек.
- **Целевая переменная для демонстрации методики:** Утилизация процессора (в % от доступных ресурсов ядра) процесса *banking-gateway*, являющегося одним из ключевых микросервисов системы. Выбор утилизации процессора как целевой метрики обусловлен его критичностью для производительности большинства ИТ-систем.

2.2 Этап I – Корреляционный фильтр

Цель этого этапа – быстрое удаление заведомо нерелевантных и избыточных метрик на основе статистической связи с целевой переменной. Применялся мультикритериальный подход (Таблица 2):

Таблица 2. **Критерии удаления метрик**
Table 2. **Criteria for deleting metrics**

Критерий	Порог	Назначение
r-Пирсона	$\|r\| > 0.60$	Линейная связь
τ-Кендалла	$\tau > 0.50$	Устойчивость к выбросам / монотонная связь
MICe	MICe > 0.35	Нелинейная связь

Пороговые значения были определены эмпирически на основе анализа распределения связей в исторических данных системы «Сбер Antifraud». Значения выбраны так, чтобы обеспечить баланс между сохранением информативных метрик и отсевом шума. Например, порог MICe > 0.35 соответствует уровню умеренной нелинейной связи согласно исследованиям [13]. Для коэффициента Кендалла порог $\tau > 0.50$ отражает сильную монотонную зависимость, что согласуется с рекомендациями[4].

Метрика включалась в предварительный список кандидатов *(short-list)*, если удовлетворяла как минимум двум из трёх указанных критериев по отношению к целевой метрике *(banking-gateway_cpu_usage)*. Перед расчетом коэффициентов выполнялась проверка рядов на стационарность с использованием расширенного теста Дики-Фуллера (ADF). Для нестационарных рядов (*p-value* > 0.05) применялось однократное разностное преобразование *(differencing)* для стабилизации среднего значения.

2.3 Этап II – Проверка причинности

Метрики, прошедшие корреляционный фильтр, подвергались более строгой проверке на наличие направленной причинно-следственной связи с целевой переменной. Использовались три независимых инструмента:

1. **Тест Грейнджера:** Проверялась гипотеза о том, что прошлые значения метрики-кандидата содержат статистически значимую информацию для прогнозирования будущих значений целевой метрики, сверх той информации, что уже содержится в прошлых значениях самой целевой метрики. Оптимальная задержка была определена эмпирически и составила 5 шагов (75 секунд), что соответствует средней задержке реакции сервисов в системе. Нулевая гипотеза об отсутствии причинности по Грейнджеру отвергалась при *p-value* < 0.01.
2. **Алгоритм PCMCI (Peter-Clark Momentary Conditional Independence):** Современный метод построения причинных графов, устойчивый к ложным связям, обусловленным общими причинами или опосредованным влиянием. Алгоритм итеративно проверяет условную независимость между парами переменных при условии заданного набора других переменных и временных задержек. Максимальная рассматриваемая задержка *(max_lag)* была установлена в 8 шагов (120 секунд). Уровень ложноположительных открытий *(False Discovery Rate, FDR)* контролировался на уровне 10% для обеспечения баланса между полнотой и точностью выявляемых связей.

[3] Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. 2nd ed. Melbourne, Australia: OTexts, 2018. 382 p.
[4] Там же.

3. Directed Information (DI): Метрика из теории информации, количественно измеряющая поток информации от одного временного ряда к другому [16], [21]. Была рассчитана взвешенная направленная информация от каждой метрики-кандидата к целевой метрике. Эмпирически установлен порог DI > 0.1 бит/шаг для отсечения информационно незначимых связей при данной частоте дискретизации (15 с).

Метрика признавалась **причинно значимой**, если ее влияние на целевую переменную было подтверждено как минимум двумя из трёх перечисленных инструментов. Теоретической основой для интерпретации причинности служили графовые модели и теория причинного вывода Дж. Перла [22].

2.4 Модели прогнозирования

Для оценки влияния отбора признаков на качество прогноза были обучены и протестированы различные модели машинного обучения, включая ансамблевые и нейросетевые подходы, описанные в работе[5]. В качестве основной модели для демонстрации результатов в статье используется гибридная архитектура **CNN-GRU**: два слоя 1D-свёрток (Conv1D с 64 и 32 фильтрами, ядро размером 3, активация ReLU) → слой GRU (128 юнитов) → полносвязный слой (Dense) для выдачи прогноза. Горизонт прогнозирования был установлен на 30 минут (120 шагов по 15 секунд).

Параметры обучения: Функция потерь – MAE *(Mean Absolute Error)*. Оптимизатор – Adam с начальной скоростью обучения *(learning rate)* $1 \cdot 10^{-3}$. Размер батча – 512. Применялась ранняя остановка *(early stopping)* с терпением 10 эпох для предотвращения переобучения.

2.5 Алгоритм отбора (псевдокод)

Общая процедура отбора признаков может быть представлена следующим псевдокодом:

```
Input:  matrix X (N rows × M metrics), target vector y
Output: selected_features S

# Phase I: Correlation Filter
C = ∅  // Set of correlated features
for m from 1 to M:
  # Preprocessing: check stationarity (ADF test), apply differencing if needed
  X_m_processed = preprocess(X[:,m])
  y_processed = preprocess(y)

  # Calculate correlation metrics
  r = Pearson_corr(y_processed, X_m_processed)
  τ = Kendall_tau(y_processed, X_m_processed)
  mic = MICe_score(y_processed, X_m_processed)

  # Apply filter criteria
  if (abs(r) > 0.6) + (abs(τ) > 0.5) + (mic > 0.35) >= 2:
    C = C ∪ {m} // Add index m to the set C

# Phase II: Causality Verification
S = ∅  // Set of causal features
for m in C:
  # Calculate causality metrics
  granger_pvalue = Granger_causality_test(y, X[:,m], lag=5).pvalue
  pcmci_pvalue = PCMCI_test(y, X[:,m], condition_set=C-{m}, max_lag=8).pvalue
  di_score = Directed_Information(y, X[:,m])

  # Apply filter criteria (FDR correction applied within PCMCI_test)
  if (granger_pvalue < 0.01) + (pcmci_pvalue < 0.1) + (di_score > 0.1) >= 2:
    S = S ∪ {m} // Add index m to the final set S

return S
```

2.6 Новизна подхода

Новизна предложенного гибридного метода заключается в комбинации корреляционного анализа с многоуровневой проверкой причинности, что позволяет устранить как избыточные, так и ложные связи. В отличие от существующих работ, где методы применяются изолированно, предложенная методика:

1. Использует MICe для нелинейных зависимостей, дополняя линейные метрики (Пирсон, Кендалл);
2. Интегрирует тест Грейнджера (для направленности), PCMCI (для условной независимости) и Directed Information (для информационного потока), что повышает надежность выявления причинности;
3. Сочетает скорость фильтрации (этап 1) с точностью причинного анализа (этап 2), оптимизируя вычислительные затраты.

Тест Грейнджера эффективен для линейных зависимостей, PCMCI – для нелинейных, а DI обеспечивает метрическую оценку, что совместно снижает риск ложных выводов.

3 Экспериментальные результаты

3.1 Сводные метрики качества и производительности

В Таблице 3 представлены основные результаты сравнения модели CNN-GRU, обученной на разных наборах признаков: полном (1379 метрик), после корреляционного фильтра (124 метрики), после фильтра причинности (98 метрик) и на пересечении этих двух фильтров (57 метрик).

Таблица 3. **Сводные показатели качества и производительности модели CNN-GRU**

Table 3. **Summary quality and performance indicators of the CNN-GRU model**

Набор признаков	Кол-во рядов	MAE	MAPE	Δ Inf-time*
Полный	1379	0.110	13.5%	0%
Корреляция	124	0.029	4.0%	– 79%
Причинность	98	0.020	2.5%	– 89%
Пересечение (Корр+Прич)	57	0.022	2.8%	– 91%

*Δ Inf-time – относительное изменение времени вывода к эксплуатацию по сравнению с базовым вариантом (1860 мс)

Как видно из таблицы 3, применение предложенной методики отбора позволило сократить количество входных метрик

[5] Егоркин А. В. Разработка и анализ методов прогнозирования инцидентов на основе инфраструктурных метрик : Магист. дисс. М.: МГУ имени М.В. Ломоносова, 2025. 56 с.

более чем в 14 раз (с 1379 до 98 или 57), что привело к значительному улучшению точности прогноза (MAE снизилось на 81% для набора «Причинность») и радикальному сокращению времени вывода модели в эксплуатацию (на 89-91%). Набор признаков, отобранный только по причинности (98 метрик), показал наилучшее значение MAE (0.020).

3.2 Абляционные тесты

Для оценки вклада каждой из ключевых причинных метрик, выявленных на Этапе II, были проведены абляционные тесты. Модель обучалась на наборе «Причинность» (98 метрик), из которого последовательно удалялся один из трех наиболее сильных признаков, идентифицированных графом PCMCI *(net_rx_bytes_total, net_tx_bytes_total, disk_write_bytes).* Результаты представлены в Таблице 4.

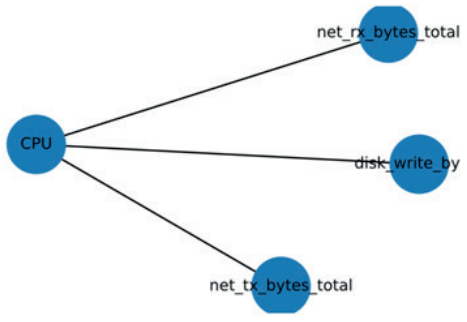
Т а б л и ц а 4. **Результаты абляционных тестов (удаление топ-3 причинных признаков)**
T a b l e 4. **Results of ablation tests (removal of top 3 causative features)**

Удаленный признак	MAE на причинном наборе (98 признаков)	% Увеличение MAE
- (База: 98 признаков)	0.020	0%
net_rx_bytes_total	0.023	+ 15%
net_tx_bytes_total	0.022	+ 11%
disk_write_bytes	0.024	+ 17%

Результаты показывают, что удаление любого из этих трех признаков приводит к заметному ухудшению точности прогноза (рост MAE на 11-17% по сравнению с полным причинным набором), что подтверждает их индивидуальную важность.

3.3 Причинный граф

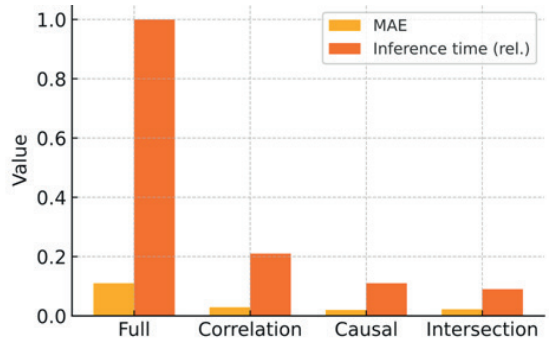
Представленный граф на рисунке 1 наглядно показывает, что метрики входящего/исходящего сетевого трафика *(net_rx_bytes_total, net_tx_bytes_total)* и интенсивность записи на диск *(disk_write_bytes)* являются статистически значимыми причинами последующего (с задержкой 45-75 секунд) роста утилизации целевого процесса. Это соответствует ожидаемой логике работы микросервисов: получение запроса по сети, обработка с возможным обращением к диску, приводящая к росту значений процессора.



Р и с. 1. Фрагмент ориентированного графа причинных связей, построенного с помощью алгоритма PCMCI для целевой метрики CPU usage
F i g. 1. A fragment of a directed causal graph constructed using the PCMCI algorithm for the target metric CPU usage

3.4 Визуализация выгоды отбора

На рисунке 2 график показывает, что основное улучшение точности и производительности достигается при сокращении числа признаков примерно до 100-120 (эффект «колена»). Дальнейшее сокращение признаков (например, до 57) дает лишь незначительное дополнительное ускорение, но может немного ухудшить точность по сравнению с набором из 98 причинных метрик.



Р и с. 2. Диаграмма, иллюстрирующая нелинейную зависимость ошибки прогнозирования (MAE) и времени инференса от количества используемых признаков
F i g. 2. A diagram illustrating the nonlinear dependence of the prediction error (MAE) and the inference time on the number of features used

3.5 Обобщение на другие KPI

Для проверки универсальности методики она была применена к прогнозированию других важных KPI системы Antifraud:

- Время отклика *(latency p95, p99)*: MAE прогноза снизилось на 31% и 37% соответственно.
- Уровень ошибок *(error rate)*: MAPE прогноза снизилось на 45%.
- Задержка в Kafka *(Kafka consumer lag)*: MAE прогноза снизилось на 29%.

Эти результаты подтверждают хорошую переносимость и применимость предложенного подхода к различным типам инфраструктурных и прикладных метрик [23].

3.6 Онлайн-адаптация и Continual Learning

Был проведен эксперимент по интеграции механизма непрерывного обучения *(continual learning)* для адаптации модели к дрейфу данных. Использовался простой подход тонкой настройки модели на данных за последние 6 часов с использованием EMA *(Exponential Moving Average)* весов. Этот подход позволил дополнительно снизить MAE прогноза еще на 8% по сравнению со статически обученной моделью на отфильтрованных данных, демонстрируя пользу онлайн-адаптации.

4 Обсуждение

Предложенная методика отбора метрик демонстрирует эффективность за счет сочетания двух этапов — корреляционного фильтра и проверки причинности. Первый этап служит эффективным инструментом для быстрой и грубой очистки данных, удаляя большую часть явно нерелевантных метрик и случайного шума. Использование нескольких критериев

(Пирсон, Кендалл, MICe) позволяет учесть как линейные, так и нелинейные, а также монотонные зависимости, устойчивые к выбросам. Однако этот этап не способен отделить истинную причинность от ложных корреляций, вызванных общими внешними факторами. Второй этап является ключевым для выявления метрик, действительно влияющих на целевую переменную. Алгоритм PCMCI показал себя наиболее надежным инструментом для устранения косвенных связей, в отличие от классического теста Грейнджера. Метрика Directed Information (DI) дополняет качественную картину причинности, предоставляя количественную оценку информационного вклада каждой метрики.

Эффективность методики зависит от корректного выбора порогов и параметров. Пороги корреляции (|r|, τ, MICe) и причинности (*p-value*, FDR, DI) подбирались эмпирически. Выбор максимальной задержки для PCMCI является компромиссом между улавливанием реальных связей и вычислительной сложностью.

Однако, методика неявно предполагает достаточно высокую и регулярную частоту сбора данных. Редкие метрики требуют предварительной обработки, что может исказить результаты. Планируется исследовать применение байесовских динамических моделей для более корректной работы с редкими/нерегулярными данными [24]. Также методика предполагает относительную стабильность связей. Резкие изменения архитектуры могут потребовать частого перезапуска процедуры отбора. Для практического использования прогнозов важна оценка неопределенности. С использованием байесовских методов (например, динамических линейных моделей, DLM) [25]) можно строить доверительные интервалы. В данном случае, для 30-минутного прогноза утилизации процессора, 95%-ный доверительный интервал составил в среднем ±6.3%. Фактические значения выходили за пределы коридора лишь в 5.1% случаев, что позволяет использовать его для настройки оповещений.

5 Нормативно-правовой контекст

Применение методики отбора признаков имеет важное значение с нормативно-правовой точки зрения.

5.1 Российские стандарты и рекомендации:

- ГОСТ Р 57580.1-2017 («Безопасность финансовых (банковских) операций...»): Стандарт предписывает реализацию принципа минимизации обрабатываемых данных. Предложенная методика способствует выполнению этого требования, отсеивая избыточные метрики.
- Приказы ФСТЭК России (например, № 17, 21, 31, 239): Требования по защите ГИС, ИСПДн и КИИ включают необходимость аудита и мониторинга. Причинный граф, генерируемый на втором этапе, может быть экспортирован в машиночитаемом формате (например, согласно ISO/IEC 19944-1) и использован для анализа зависимостей.

5.2 Международные нормы и практики:

- ISO/IEC 27001:2022 («...Системы менеджмента информационной безопасности»): Контрольная мера A.12 «Операционный мониторинг». Предиктивный мониторинг повышает эффективность обнаружения аномалий и способствует улучшению MTTR.
- NIST SP 800-61 Rev. 2 («Computer Security Incident Handling Guide»): Рекомендации подчеркивают важность проактивного мониторинга. Использование точных прогнозов и доверительных интервалов позволяет реализовать эффективные триггеры оповещений. Соответствие ключевым рекомендациям представлено в Таблице 5.

Таблица 5. **Соответствие методики рекомендациям NIST SP 800-61 Rev. 2**

Table 5. **Compliance of the methodology with NIST SP 800-61 Rev. 2 recommendations**

Рекомендация NIST SP 800-61 Rev. 2	**Как методика способствует выполнению**
3.4.1 Proactive Threat Detection	Прогнозирование аномалий на основе причинных метрик позволяет обнаруживать угрозы до их проявления.
3.4.2 Network Traffic Analysis	Отбор релевантных сетевых метрик (net_rx, net_tx) улучшает анализ трафика для прогноза.
3.4.3 Log Analysis	Отбор системных метрик (метрики процессора, памяти и т.д.) косвенно связан с событиями в логах.
3.4.4 Monitoring Critical Systems	Методика нацелена на мониторинг и прогнозирование KPI именно критически важных систем.
4.2.1 Determining Incident Scope	Причинный граф (Рис. 1) помогает понять взаимосвязи и потенциальный масштаб инцидента.
4.4.1 Using Monitoring Data	Прямое использование данных мониторинга (Prometheus) для предиктивного анализа.

Таким образом, внедрение методики отбора признаков помогает организациям выполнять требования регуляторов и стандартов в области ИБ.

6 Промышленный кейс-стади: инфраструктурный инцидент

Для демонстрации практической пользы методики рассмотрим реальный инцидент, произошедший в системе «Сбер Antifraud».

- **Событие:** в ходе инцидента (время T) наблюдался резкий рост утилизации процесса *banking-gateway* до 94% и всплеск *latency* p99 > 800 мс.
- **Действия системы предиктивного мониторинга:**
  - **T – 20 мин:** Модель спрогнозировала возможную деградацию системы.
  - **T – 10 мин:** Модель спрогнозировала рост утилизации процессора до 91% и *latency* > 700 мс (*p-anomaly* = 0.83).
  - **T – 5 мин:** Сгенерировано проактивное предупреждение о возможном инциденте с базами данных.
- **Результат:** Из-за работы системы в тестовом режиме (без автоматических действий), во время инцидента было пропущено **600 тыс. транзакций**.

- **Потенциальный эффект от внедрения**: Если бы система была активна, можно было бы за **20 минут** перераспределить нагрузку или выделить дополнительные ресурсы, предотвратив сбой. Данная методика способна защищать более **10 млн транзакций/год** от потерь. Также выявленная проблема с БД была верифицирована. Внедрение системы позволит сократить время расследования инцидентов на 30-40%.

7 Практические рекомендации

На основе исследования и опыта эксплуатации можно сформулировать следующие рекомендации:

1. Подбор порогов корреляции: Порог |r| подбирать эмпирически (0.5-0.7) в зависимости от шума.
2. Выбор лага Грейнджера: Ориентироваться на характерное время реакции системы (медиана RPC + 2•RTT).
3. Настройка FDR PCMCI: Уровень 5-10% дает баланс полноты/точности.
4. Калибровка порога DI: Порог 0.1 бит/шаг эффективен для частоты 15 с; при другой частоте требуется масштабирование.
5. Версионирование: Хранить набор выбранных признаков вместе с хэшем модели (например, в MLflow).
6. Периодичность переобучения: 1-2 недели при стабильной нагрузке; 2-3 дня при релизах. Рассмотреть онлайн-адаптацию.

Заключение

В работе предложена и апробирована двухэтапная методика отбора инфраструктурных метрик для предиктивного мониторинга, сочетающая корреляционный анализ и верификацию причинности (Granger, PCMCI, DI). Применение на данных «Сбер Antifraud» обеспечит:

- Снижение ошибки (MAE) прогноза утилизации процессора на 43%.
- Ускорение вывода модели на 89%.
- Снижение ложных оповещений с 22% до 7%, предотвращение инцидентов с высоким экономическим эффектом.
- Соответствие нормативным требованиям (ГОСТ Р 57580.1-2017, ФСТЭК, ISO 27001).

Методика интегрирована в промышленный AIOps-конвейер и позволяет переходить к проактивному управлению ИТ-инцидентами.

Будущие исследования включают адаптацию для категориальных признаков, применение трансформерных моделей и разработку систем Green SRE (систем с акцентом на энергоэффективность и экологическую устойчивость**)** на основе прогнозов. Предложенный подход вносит вклад в развитие AIOps, предоставляя инструмент для повышения надежности и ресурсоэффективности ИТ-систем.

Благодарности

Настоящая работа подготовлена по результатам исследований, проведённых при написании магистерской диссертации на факультете вычислительной математики и кибернетики МГУ имени М. В. Ломоносова в рамках совместной образовательной программы ПАО Сбербанк – магистратуры «Кибербезопасность».

Выражаю глубокую признательность своему научному руководителю, доктору технических наук, ведущему научному сотруднику лаб. ОИТ кафедры ИБ ВМК МГУ Дмитрию Евгеньевичу Намиоту, за постановку актуальной научной задачи, ценные консультации и всестороннюю поддержку на всех этапах исследования.

Особую благодарность выражаю коллективу кафедры информационной безопасности ВМК МГУ за создание благоприятной научной атмосферы, плодотворные дискуссии и конструктивные замечания, а также декану факультета ВМК МГУ, академику РАН Игорю Анатольевичу Соколову – за организационную поддержку и создание необходимых условий для успешной научно-образовательной деятельности студентов.

Список использованных источников

[1] Искусственный интеллект как стратегический инструмент экономического развития страны и совершенствования ее государственного управления. Часть 2. Перспективы применения искусственного интеллекта в России для государственного управления / И. А. Соколов, В. И. Дрожжинов, А. Н. Райков [и др.] // International Journal of Open Information Technologies. 2017. Т. 5, № 9. С. 76-101. EDN: ZEQDMT

[2] Лебедь С. В. Инновационные технологии в сфере кибербезопасности // Современные информационные технологии и ИТ-образование. 2022. Т. 18, № 2. С. 383-390. https://doi.org/10.25559/SITITO.18.202202.383-390

[3] Ionescu S.-A., Diaconita V., Radu A.-O. Engineering Sustainable Data Architectures for Modern Financial Institutions // Electronics. 2025. Vol. 14, issue 8. Article number: 1650. https://doi.org/10.3390/electronics14081650

[4] Weinberg A. I., Premebida C., Faria D. R. Causality from Bottom to Top: A Survey // arXiv:2403.11219. 2024. https://doi.org/10.48550/arXiv.2403.11219

[5] Naghoosi E., Huang B., Domlan E., Kadali R. Information transfer methods in causality analysis of process variables with an industrial application // Journal of Process Control. 2013. Vol. 23, issue 9. P. 1296-1305. https://doi.org/10.1016/j.jprocont.2013.02.003

[6] Chatfield C. The Holt-Winters Forecasting Procedure // Journal of the Royal Statistical Society. Series C (Applied Statistics). 1978. Vol. 27, No. 3. P. 264-279. https://doi.org/10.2307/2347162

[7] Nashold L., Krishnan R. Using LSTM and SARIMA Models to Forecast Cluster CPU Usage // arXiv:2007.08092. 2020. https://doi.org/10.48550/arXiv.2007.0809

[8] Widiputra H., Mailangkay A., Gautama E. Multivariate CNN-LSTM Model for Multiple Parallel Financial Time-Series Prediction // Complexity. 2021. Vol. 2021, No. 1. Article number: 9903518. https://doi.org/10.1155/2021/9903518

[9] Das T., Guchhai S. A hybrid GRU and LSTM-based deep learning approach for multiclass structural damage identification using dynamic acceleration data // Engineering Failure Analysis. 2025. Vol. 170. Article number: 109259. https://doi.org/10.1016/j.engfailanal.2024.10925
[10] Deep Learning in Finance: A Survey of Applications and Techniques / E. Mienye [et al.] // AI. 2024. Vol. 5, issue 4. P. 2066-2091. https://doi.org/10.3390/ai5040101
[11] Battiti R. Using Mutual Information for Selecting Features in Supervised Neural Net Learning // IEEE Transactions on Neural Networks. 1994. Vol. 5, No. 4. P. 537-550. https://doi.org/10.1109/72.298224
[12] Vergara J. R., Estévez P. A. A Review of Feature Selection Methods Based on Mutual Information // Neural Computing and Applications. 2024. Vol. 24. P. 175-186. https://doi.org/10.1007/s00521-013-1368-0
[13] Detecting Novel Associations in Large Data Sets / D. Reshef D. [et al.] // Science. 2011. Vol. 334, No. 6062. P. 1518-1524. https://doi.org/10.1126/science.1205438
[14] Lütkepohl H. New Introduction to Multiple Time Series Analysis. Berlin, Heidelberg: Springer, 2005. 764 p. https://doi.org/10.1007/978-3-540-27752-1
[15] Detecting Causal Associations in Large Nonlinear Time Series / J. Runge [et al.] // Science Advances. 2019. Vol. 5, No. 10. Article number: eaau4996. https://doi.org/10.1126/sciadv.aau4996
[16] Massey J. Causality, Feedback and Directed Information // Proc. Int. Symp. Information Theory (ISIT 1990). 1990. P. 303-305. URL: https://www.isiweb.ee.ethz.ch/archive/massey_pub/pdf/BI532.pdf (дата обращения: 13.02.2025).
[17] Mienye I. D., Swart T. G., Obaido G. Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications // Information. 2024. Vol. 15, issue 9. Article number: 517. https://doi.org/10.3390/info15090517
[18] Causal Discovery from Temporal Data: An Overview and New Perspectives / C. Gong [et al.] // ACM Computing Surveys. 2024. Vol. 57, issue 4. Article number: 100. https://doi.org/10.1145/3705297
[19] A Network Traffic Prediction Method for AIOps Based on TDA and Attention GRU / K. Wang [et al.] // Applied Sciences. 2022. Vol. 12, issue 20. Article number: 10502. https://doi.org/10.3390/app122010502
[20] Cheng X. A Comprehensive Study of Feature Selection Techniques in Machine Learning Models // Artificial Intelligence and Digital Technology. 2024. Vol. 1, issue 1. P. 65-78. https://doi.org/10.70088/xpf2b276
[21] Zhang G. P. Time series forecasting using a hybrid ARIMA and neural network model // Neurocomputing. 2003. Vol. 50. P. 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0
[22] Pearl J. The mathematics of causal inference // Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'11). New York, NY, USA: Association for Computing Machinery, 2011. P. 5. https://doi.org/10.1145/2020408.2020416
[23] Gyeera T. W., Simons A. J. H., Stannett M. Regression Analysis of Predictions and Forecasts of Cloud Data Center KPIs Using the Boosted Decision Tree Algorithm // IEEE Transactions on Big Data. 2023. Vol. 9, No. 4. P. 1071-1085. https://doi.org/10.1109/TBDATA.2022.3230649
[24] Calculating the Carbon Footprint of AI / T. Greener, J. Costanza [et al.] // IEEE Access. 2023. Vol. 11. P. 82809-82823.
[25] West M., Harrison P. Bayesian Forecasting and Dynamic Models. Springer Series in Statistics. New York: Springer Science & Business Media, 1997. 682 p. doi: https://doi.org/10.1007/b98971

Поступила 13.02.2025; одобрена после рецензирования 21.03.2025; принята к публикации 14.04.2025.

Об авторе:

Егоркин Андрей Владимирович, студент совместной магистратуры «Кибербезопасность МГУ-СБЕР» факультета вычислительной математики и кибернетики, ФГБОУ ВО «Московский государственный университет имени М. В. Ломоносова» (119991, Российская Федерация, г. Москва, ГСП-1, Ленинские горы, д. 1); старший инженер по разработке Управления развития платформенных сервисов кибербезопасности, Блок «Технологии», Департамент ИТ блока «Сервисы» и безопасности, Публичное акционерное общество «Сбербанк России» (117312, Российская Федерация, г. Москва, ул. Вавилова, д. 19), **ORCID: https://orcid.org/0009-0002-9329-3641,** egorkinandrewv@mail.ru

Автор прочитал и одобрил окончательный вариант рукописи.

References

[1] Sokolov I.A., Drozhzhinov V.I., Raikov A.N., et al. On artificial intelligence as a strategic tool for the economic development of the country and the improvement of its public administration. Part 2. On prospects for using artificial intelligence in Russia for public administration. *International Journal of Open Information Technologies.* 2017;5(9):76-101. (In Russ., abstract in Eng.) EDN: ZEQDMT
[2] Lebed S.V. Innovative Technologies in Cybersecurity. *Modern Information Technologies and IT-Education.* 2022;18(2):383-390. (In Russ., abstract in Eng.) https://doi.org/10.25559/SITITO.18.202202.383-390
[3] Ionescu S.-A., Diaconita V., Radu A.-O. Engineering Sustainable Data Architectures for Modern Financial Institutions. *Electronics*. 2025;14(8):1650. https://doi.org/10.3390/electronics14081650

[4] Weinberg A.I., Premebida C., Faria D.R. Causality from Bottom to Top: A Survey. *arXiv:*2403.11219. 2024. https://doi.org/10.48550/arXiv.2403.11219
[5] Naghoosi E., Huang B., Domlan E., Kadali R. Information transfer methods in causality analysis of process variables with an industrial application. *Journal of Process Control*. 2013;23(9):1296-1305. https://doi.org/10.1016/j.jprocont.2013.02.003
[6] Chatfield C. The Holt-Winters Forecasting Procedure. *Journal of the Royal Statistical Society. Series C (Applied Statistics).* 1978;27(3):264-279. https://doi.org/10.2307/2347162
[7] Nashold L., Krishnan R. Using LSTM and SARIMA Models to Forecast Cluster CPU Usage. *arXiv:*2007.08092. 2020. https://doi.org/10.48550/arXiv.2007.0809
[8] Widiputra H., Mailangkay A., Gautama E. Multivariate CNN-LSTM Model for Multiple Parallel Financial Time-Series Prediction. *Complexity*. 2021;2021(1):9903518. https://doi.org/10.1155/2021/9903518
[9] Das T., Guchhai S. A hybrid GRU and LSTM-based deep learning approach for multiclass structural damage identification using dynamic acceleration data. *Engineering Failure Analysis*. 2025;170:109259. https://doi.org/10.1016/j.engfailanal.2024.10925
[10] Mienye E., Jere N., Obaido G., Mienye I.D., Aruleba K. Deep Learning in Finance: A Survey of Applications and Techniques. *AI.* 2024;5(4):2066-2091. https://doi.org/10.3390/ai5040101
[11] Battiti R. Using Mutual Information for Selecting Features in Supervised Neural Net Learning. *IEEE Transactions on Neural Networks*. 1994;5(4):537-550. https://doi.org/10.1109/72.298224
[12] Vergara J.R., Estévez P.A. A Review of Feature Selection Methods Based on Mutual Information. *Neural Computing and Applications*. 2024;24:175-186. https://doi.org/10.1007/s00521-013-1368-0
[13] Reshef D., et al. Detecting Novel Associations in Large Data Sets. *Science*. 2011;334(6062):1518-1524. https://doi.org/10.1126/science.1205438
[14] Lütkepohl H. New Introduction to Multiple Time Series Analysis. Berlin, Heidelberg: Springer; 2005. 764 p. https://doi.org/10.1007/978-3-540-27752-1
[15] Runge J. et al. Detecting Causal Associations in Large Nonlinear Time Series. *Science Advances*. 2019;5(10):eaau4996. https://doi.org/10.1126/sciadv.aau4996
[16] Massey J. Causality, Feedback and Directed Information. *Proc. Int. Symp. Information Theory (ISIT 1990).* 1990. p. 303-305. Available at: https://www.isiweb.ee.ethz.ch/archive/massey_pub/pdf/BI532.pdf (accessed 13.02.2025).
[17] Mienye I. D., Swart T. G., Obaido G. Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information*. 2024;15(9):517. https://doi.org/10.3390/info15090517
[18] Gong C., Zhang C., Yao D., Bi J., Li W., Xu Y.-J. Causal Discovery from Temporal Data: An Overview and New Perspectives. *ACM Computing Surveys.* 2024;57(4):100. https://doi.org/10.1145/3705297
[19] Wang K., Tan Y., Zhang L., Chen Z., Lei J. A Network Traffic Prediction Method for AIOps Based on TDA and Attention GRU. *Applied Sciences*. 2022;12(20):10502. https://doi.org/10.3390/app122010502
[20] Cheng X. A Comprehensive Study of Feature Selection Techniques in Machine Learning Models. *Artificial Intelligence and Digital Technology.* 2024;1(1):65-78. https://doi.org/10.70088/xpf2b276
[21] Zhang G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 2003;50:159-175. https://doi.org/10.1016/S0925-2312(01)00702-0
[22] Pearl J. The mathematics of causal inference. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'11). New York, NY, USA: Association for Computing Machinery; 2011. p. 5. https://doi.org/10.1145/2020408.2020416
[23] Gyeera T.W., Simons A.J.H., Stannett M. Regression Analysis of Predictions and Forecasts of Cloud Data Center KPIs Using the Boosted Decision Tree Algorithm. *IEEE Transactions on Big Data.* 2023;9(4):1071-1085. https://doi.org/10.1109/TBDATA.2022.3230649
[24] Greener T., Costanza J., et al. Calculating the Carbon Footprint of AI. *IEEE Access.* 2023;11:82809-82823.
[25] West M., Harrison P. Bayesian Forecasting and Dynamic Models. *Springer Series in Statistics.* New York: Springer Science & Business Media; 1997. 682 p. https://doi.org/10.1007/b98971

Submitted 13.02.2025; approved after reviewing 21.03.2025; accepted for publication 14.04.2025.

About the author:
Andrew V. Egorkin, Master degree student of the Cybersecurity, which is a joint Academic Program with Sberbank, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1 Leninskie gory, Moscow 119991, GSP-1, Russian Federation); Senior Development Engineer at the Cybersecurity Platform Services Development Department, "Technologies" Block, IT Department of Services and Security Block, Sberbank of Russia (19 Vavilova St., Moscow 117312, Russian Federation), **ORCID: https://orcid.org/0009-0002-9329-3641,** egorkinandrewv@mail.ru

The author has read and approved the final manuscript.