

УДК 004.89

Михалев Кирилл Александрович  
ФГБОУ ВО "Российский государственный социальный университет"  
г. Москва, Россия  
Студент

## **ПРИМЕНЕНИЕ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ГОРОДСКОЙ ТРАНСПОРТНОЙ ИНФРАСТРУКТУРЫ (НА ПРИМЕРЕ ГКУ ЦОДД)**

*Аннотация. В статье рассматривается проблема оптимизации управления интеллектуальной транспортной системой (ИТС) мегаполиса с помощью методов интеллектуального анализа данных (Data Mining). На примере ГКУ ЦОДД Правительства Москвы разработаны алгоритмы автоматического выявления аномалий в работе комплексов фотовидеофиксации и пространственной кластеризации дорожных нарушений. Предложена целевая ИТ-архитектура на базе Data Lake. Обоснован экономический эффект от внедрения разработанных моделей, позволяющий предотвращать финансовые потери и оптимизировать маршрутизацию экипажей.*

*Ключевые слова: интеллектуальный анализ данных, ГКУ ЦОДД, Data Mining, машинное обучение, интеллектуальная транспортная система, выявление аномалий, пространственный анализ, бизнес-информатика.*

Mikhalev Kirill Alexandrovich  
Russian State Social University  
Moscow, Russia  
Student

**APPLICATION OF DATA MINING METHODS TO IMPROVE THE EFFICIENCY OF URBAN TRANSPORT INFRASTRUCTURE (ON THE EXAMPLE OF THE STATE INSTITUTION CENTER FOR TRAFFIC MANAGEMENT)**

*Abstract. The article discusses the problem of optimizing the management of a megacity's intelligent transport system (ITS) using data mining methods. Using the example of the State Institution "Center for Traffic Management" (TSODD) of the Moscow Government, algorithms for automatically detecting anomalies in the operation of photo and video recording complexes and spatial clustering of traffic violations have been developed. A target IT architecture based on a Data Lake is proposed. The economic effect of implementing the developed models is justified, making it possible to prevent financial losses and optimize crew routing.*

*Keywords: data mining, TSODD, machine learning, intelligent transport system, anomaly detection, spatial analysis, business informatics.*

## **Введение**

Современные мегаполисы сталкиваются с беспрецедентной нагрузкой на транспортную инфраструктуру. В этих условиях ключевым инструментом управления дорожным движением становится интеллектуальная транспортная система (ИТС), аккумулирующая колоссальные объемы данных с камер фотовидеофиксации, светофорных объектов и детекторов транспорта<sup>1</sup>. Переход к проактивному управлению транспортными потоками и внедрение искусственного интеллекта в транспортную отрасль является одним из приоритетов государственной политики Российской Федерации<sup>2</sup>.

На сегодняшний день ядром московской ИТС выступает ГКУ «Центр организации дорожного движения» (ЦОДД). Парк организации насчитывает тысячи комплексов фиксации, генерирующих петабайты информации. Однако, как отмечает руководство организации, камеры

---

<sup>1</sup> ГОСТ Р 56829-2015. Интеллектуальные транспортные системы. Термины и определения. – М. : Стандартинформ, 2016. – 26 с.

<sup>2</sup> О развитии искусственного интеллекта в Российской Федерации : Указ Президента РФ от 10.10.2019 № 490 // СПС КонсультантПлюс. – URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_335184/](https://www.consultant.ru/document/cons_doc_LAW_335184/) (дата обращения: 16.05.2026).

нуждаются в дополнительном непрерывном контроле со стороны нейросетевых алгоритмов и систем аналитики для исключения аппаратных сбоев и снижения нагрузки, связанной с ручным разбором инцидентов инспекторами<sup>3</sup>. Выявление инфраструктурных сбоев (например, неисправностей оборудования) и планирование маршрутов мобильных комплексов зачастую осуществляются в ручном режиме на основе ретроспективного анализа. Следствием этого становится неизбежный временной лаг между моментом возникновения неисправности или "зоны риска" и управленческой реакцией, что ведет к прямым финансовым потерям городского бюджета.

Целью данного исследования является разработка и апробация аналитических модулей на базе методов интеллектуального анализа данных (Data Mining) для повышения операционной и экономической эффективности ГКУ ЦОДД. Для достижения поставленной цели решаются следующие задачи: алгоритмизация поиска аномалий в работе оборудования, разработка модуля пространственного анализа нарушений, а также проектирование целевой ИТ-архитектуры.

### **Методы и исследования**

Исследование проводилось на основе анализа сырых данных из колоночной СУБД ClickHouse (таблицы фиксации и нарушений), а также выгрузок реестров оборудования и статусов отключений. Для решения выявленных проблем были спроектированы два ключевых аналитических алгоритма, реализованных на языке Python.

Одной из критических проблем является состояние «скрытого отказа», когда комплекс фотовидеофиксации технически исправен и фиксирует проезды транспортных средств, но не генерирует

---

<sup>3</sup> ИИ в дорожных камерах Москвы научили самопроверке для исключения ошибочных штрафов // Хабр. – URL: <https://habr.com/ru/news/1020416/> (дата обращения: 16.05.2026).

постановления о нарушениях. Для решения этой задачи нами применен метод поиска пересечений множеств с последующей классификацией отклонений на базе алгоритмов семейства Isolation Forest<sup>4</sup>.

Логика разработанного алгоритма детектирования аномалий включает:

1. Загрузку оперативных логов фиксации и нарушений за контрольный период, а также реестра плановых отключений оборудования.
2. Трансформацию данных и сопоставление потоков по базовому идентификатору камеры.
3. Изоляцию аномальных паттернов: фильтрацию пулов данных, где показатель числа фиксации больше нуля, а количество сгенерированных нарушений равно нулю (или имеет значение NaN). Фрагмент реализации данного этапа представлен на рисунке 1.
4. Верификацию аномалии путем сопоставления с реестром сервисных работ. В случае отсутствия планового ремонта комплексу автоматически присваивается статус, требующий оперативной проверки.

Python

```
# Поиск аномалий: фиксации > 0, нарушения == 0 или NaN  
  
anomaly_df = df_pivot[  
  
    (df_pivot['фиксации'] > 0) &  
  
    ((df_pivot['нарушения'].isna()) | (df_pivot['нарушения'] == 0))  
  
].copy()
```

---

<sup>4</sup> Liu F. T., Ting K. M., Zhou Z.-H. Isolation Forest // Proceedings of the 8th IEEE International Conference on Data Mining (ICDM). – Pisa, 2008. – P. 413–422.

```

# Сопоставление с интервалами отключений

merged = anomaly_df.merge(df_google[['camera_id_clean', 'Статус', ...]],
on='camera_id_clean', how='left')

mask_interval = (merged['Дата отключения_dt'].isna() | \
((merged['дата_dt'] >= merged['Дата отключения_dt']) & \
(merged['дата_dt'] <= merged['Дата включения_dt'])))

merged_filtered = merged[mask_interval]

```

Рисунок 1 — Фрагмент программного кода выявления аномалий в работе комплексов фотовидеофиксации (Python).

Источник: составлено автором.

Для оптимизации маршрутизации мобильных комплексов (экипажей фотовидеофиксации) разработан алгоритм построения динамических тепловых карт. На вход подаются массивы геокодированных данных об исторических нарушениях (полученные через API геокодеров). С использованием логарифмической шкалы нормализации строится карта плотности, позволяющая кластеризовать участки улично-дорожной сети с наивысшей концентрацией правонарушений для последующего перенаправления туда патрульных экипажей<sup>5</sup> (рисунок 2).

<sup>5</sup> Дюк В. А., Флегонтов А. В. Применение технологий интеллектуального анализа данных в естественнонаучных, технических и гуманитарных областях : монография. – Санкт-Петербург : Изд-во РГПУ им. А. И. Герцена, 2019. – 256 с.

## Фиксации (log scale)

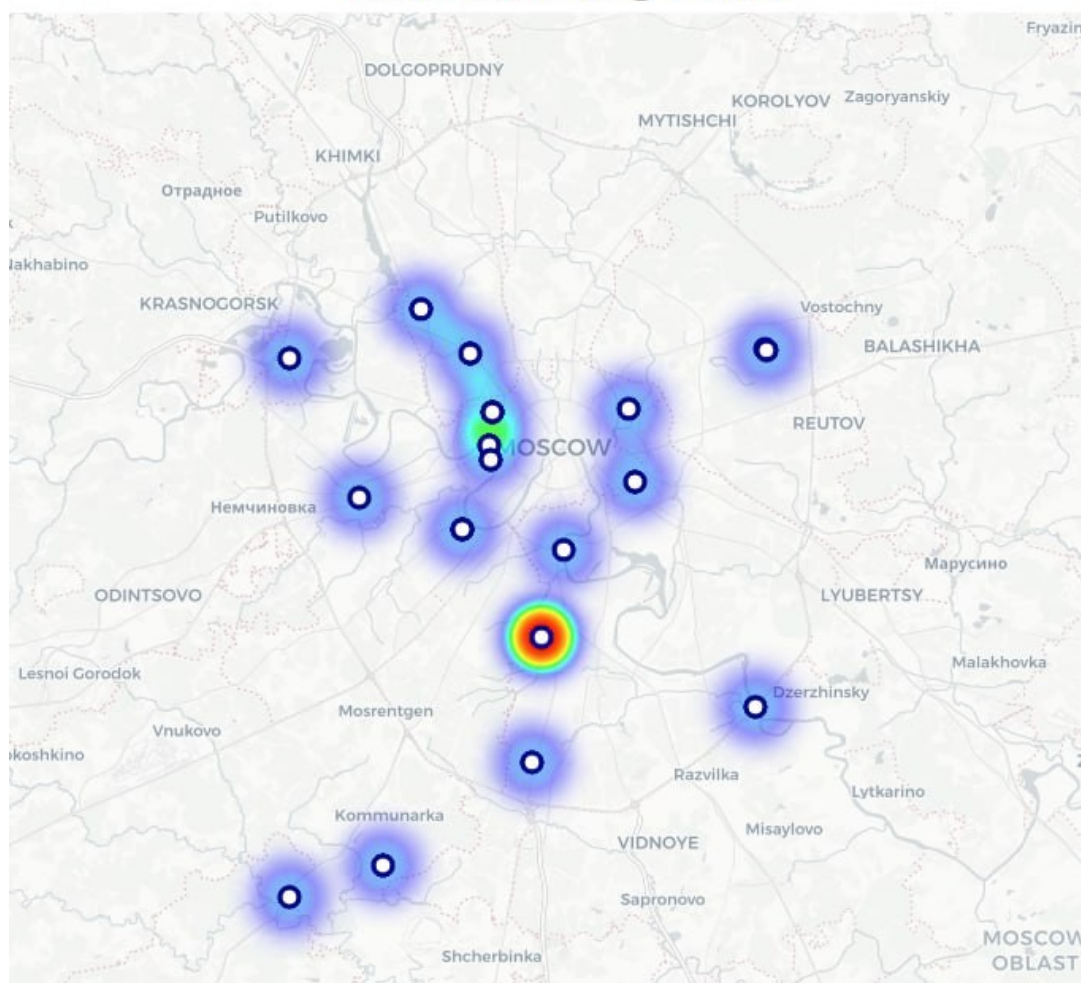


Рисунок 2 — Тепловая карта концентрации нарушений, построенная модулем.

Источник: составлено автором.

Дополнительно в рамках исследования были разработаны скрипты автоматизированной оценки результативности ручной корректировки нечитаемых государственных регистрационных знаков (ГРЗ) и модули пакетного сжатия графических материалов.

### **Результаты оригинального авторского исследования**

Апробация разработанных алгоритмов на реальных данных ГКУ ЦОДД подтвердила их высокую результативность. Модуль детектирования аномалий на тестовой выборке из более чем 3000 комплексов позволил выявить около 3% скрытых аппаратных и программных сбоев. Точность классификации составила 95% (рисунки 3.1 и 3.2).

camera_id	дата	фиксации	нарушения	Состояние_реестр	Утвержденные_нарушения
123	2026 Апрель 21	69419		Утвержден (действующий)	107, 2, 77
1234	2026 Апрель 21	25683		Утвержден (действующий)	156
12345	2026 Апрель 22	23850		Утвержден (действующий)	107, 2, 8, 80
1534	2026 Апрель 21	23582		Утвержден (действующий)	138, 157
15888	2026 Апрель 22	21588		Утвержден (действующий)	138, 157
242245	2026 Апрель 21	20689		Утвержден (действующий)	138, 157

Рисунок 3.1 — Фрагмент отчёта о выявлении аномалий в работе камер (1 часть).

Источник: составлено автором.

Статус	Причина отключения	Статус_работы	Причина_аномалии
Открыто	Ремонт УДС	Не работает	Отключена (статус "Открыто", причина: Ремонт УДС)
Открыто	Замена асфальта	Не работает	Отключена (статус "Открыто", причина: Замена асфальта)
Открыто	Изменение ОДД	Не работает	Отключена (статус "Открыто", причина: Изменение ОДД)
		Работает	⚠ Работает, но нет нарушений (требуется проверка)
		Работает	⚠ Работает, но нет нарушений (требуется проверка)
		Работает	⚠ Работает, но нет нарушений (требуется проверка)

Рисунок 3.2 — Фрагмент отчёта о выявлении аномалий в работе камер (2 часть).

Источник: составлено автором.

Автоматизация данного процесса сокращает время выявления неисправного комплекса с нескольких суток (до формирования планового отчета) до 24 часов. С учетом того, что средний комплекс генерирует около 50 постановлений в сутки при среднем штрафе 500 руб., сокращение простоя одной единицы оборудования на трое суток предотвращает потерю бюджетных поступлений в размере 75 000 руб.

Применение пространственного анализа показало, что около 20% мобильных экипажей патрулируют зоны с низкой результативностью. Реаллокация маршрутов 10 комплексов в выявленные «горячие точки» (зоны высокой концентрации нарушений), по минимальной оценке, обеспечивает прирост фиксаций на 15–20%. В финансовом эквиваленте это генерирует дополнительный экономический эффект в размере 2–4 млн руб. в год.

Внедрение вспомогательного алгоритма сжатия визуальных данных (с коэффициентом качества 50%) позволило снизить объем хранимого архива в 3–4 раза без потери аналитической ценности изображений, что существенно снижает операционные затраты на закупку серверного

оборудования. Агрегированная консервативная оценка годового экономического эффекта от внедрения предложенного комплекса алгоритмов составляет 10–15 млн руб.

Интеграция разработанных аналитических скриптов в единую корпоративную среду требует системного пересмотра ИТ-архитектуры организации. На смену модели с ручным сбором данных предложена целевая архитектура (TO-BE), функционирующая по принципу автоматического конвейера. В качестве ключевой рекомендации предлагается создание единого озера данных (Data Lake) на базе СУБД ClickHouse или PostgreSQL для консолидации разрозненных потоков. В дальнейшем целесообразна контейнеризация алгоритмов (например, с использованием Docker) и их запуск по расписанию через оркестраторы (Apache Airflow) для полного устранения рутинных операций. Интеграция витрин данных с отечественной BI-платформой Visiology обеспечит предоставление аналитических дашбордов руководству в режиме реального времени.

### **Заключение**

Проведенное исследование доказало высокую эффективность применения методов интеллектуального анализа данных в управлении транспортной инфраструктурой. Разработанные алгоритмы поиска аномалий и пространственной кластеризации позволили устранить ограничения ретроспективной аналитики и выявить неочевидные финансовые потери в операционной деятельности ГКУ ЦОДД.

Разработанный подход позволил формализовать признак «скрытого отказа» оборудования и создать механизм его автоматической верификации, что является решением научной и практической проблемы в области эксплуатации комплексов фотовидеофиксации. Переход к

проактивной целевой модели управления на основе концепции Data Lake обладает быстрой окупаемостью и масштабируемостью, что делает разработанные решения применимыми для внедрения в интеллектуальные транспортные системы других регионов в рамках дальнейшего развития транспортной аналитики.

### **Использованные источники**

1. ГОСТ Р 56829-2015. Интеллектуальные транспортные системы. Термины и определения. – М. : Стандартинформ, 2016. – 26 с.
2. О развитии искусственного интеллекта в Российской Федерации : Указ Президента РФ от 10.10.2019 № 490 // СПС КонсультантПлюс. – URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_335184/](https://www.consultant.ru/document/cons_doc_LAW_335184/) (дата обращения: 16.05.2026).
3. ИИ в дорожных камерах Москвы научили самопроверке для исключения ошибочных штрафов // Habr. – URL: <https://habr.com/ru/news/1020416/> (дата обращения: 16.05.2026).
4. Liu F. T., Ting K. M., Zhou Z.-H. Isolation Forest // Proceedings of the 8th IEEE International Conference on Data Mining (ICDM). – Pisa, 2008. – P. 413–422.
5. Дюк В. А., Флегонтов А. В. Применение технологий интеллектуального анализа данных в естественнонаучных, технических и гуманитарных областях : монография. – Санкт-Петербург : Изд-во РГПУ им. А. И. Герцена, 2019. – 256 с.

### **List of sources used**

1. GOST R 56829-2015. Intelligent transport systems. Terms and definitions. – Moscow : Standartinform, 2016. – 26 p.

2. On the Development of Artificial Intelligence in the Russian Federation : Decree of the President of the Russian Federation No. 490 of October 10, 2019 // ConsultantPlus Legal System. – URL: [https://www.consultant.ru/document/cons\\_doc\\_LAW\\_335184/](https://www.consultant.ru/document/cons_doc_LAW_335184/) (accessed: 16.05.2026).

3. AI in Moscow traffic cameras taught to self-check to eliminate erroneous fines // Habr. – URL: <https://habr.com/ru/news/1020416/> (accessed: 16.05.2026).

4. Liu F. T., Ting K. M., Zhou Z.-H. Isolation Forest // Proceedings of the 8th IEEE International Conference on Data Mining (ICDM). – Pisa, 2008. – P. 413–422.

5. Dyuk V. A., Flegontov A. V. Application of data mining technologies in natural science, technical and humanitarian fields : monograph. – St. Petersburg : Herzen State Pedagogical University of Russia Publ., 2019. – 256 p.