

УДК 656.2:519.7

Идентификация нештатных ситуаций железнодорожного транспорта России на основе кластерного анализа данных

Макеева В. В. ✉, Вивчарь Р. М., Березкин А. А.

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича
Санкт-Петербург, 193232, Российская Федерация

Постановка задачи: идентификация нештатных ситуаций на железной дороге для обеспечения безопасности и оптимизации процессов включает в себя наблюдение за состоянием инфраструктуры, движением поездов, погодными условиями и другими факторами, способными привести к чрезвычайным ситуациям. **Целью работы** является разработка специализированного программного обеспечения с использованием кластерного анализа данных для выявления причин аварий на железнодорожном транспорте России, что критически важно из-за его значительного влияния на общественную безопасность и экономику страны, а также из-за потенциальных угроз, связанных с возможными чрезвычайными ситуациями, особенно при перевозке опасных грузов. **Используемые методы:** кластерный анализ для систематизации разнообразных видов аномалий, возникающих на железнодорожном транспорте, с учетом таких параметров, как источники возникновения проблем, географическое расположение и временные характеристики. Они позволят выявить ключевые группы ситуаций и лучше понять их природу, а также оценить эффективность предотвращения чрезвычайных ситуаций. **Результат:** с использованием метода кластерного анализа данных осуществляется выявление основных закономерностей, лежащих в основе непривычных обстоятельств в функционировании железнодорожного транспорта. Это позволяет прийти к существенным выводам, включая идентификацию различных категорий нештатных ситуаций, оптимизацию мер безопасности и повышение результативности железнодорожной системы. Элементами **новизны** представленного решения является применение кластерного анализа для выявления тенденций, закономерностей и потенциальных сценариев в данных, связанных с авариями на железной дороге. **Теоретическая/Практическая значимость** состоит в выработке рекомендаций по улучшению инфраструктуры, повышению мер безопасности и в конечном итоге повышению общей эффективности транспортной системы в долгосрочной перспективе.

Ключевые слова: нештатные ситуации, чрезвычайные ситуации, кластерный анализ, железнодорожный транспорт, оптимизация процессов, тенденции, мониторинг, безопасность, Data Mining (анализ данных), разработка программного обеспечения и рекомендаций

Источник финансирования: статья подготовлена в рамках прикладных научных исследований СПбГУТ, регистрационный номер 1023031600087-9 в ЕГИСУ НИОКТР.

Актуальность исследования

В настоящее время железнодорожный транспорт является важной составляющей инфраструктуры России и играет ключевую роль в ее транспортной системе, перемещая как пассажиров, так и товары на огромные расстояния. Несмотря на различные факторы, такие как погодные условия, технические сбои, человеческие ошибки и другие непредвиденные обстоятельства, которые могут привести к возникновению нештатных ситуаций, железнодорожная отрасль стремится обеспечить безопасное и

Библиографическая ссылка на статью:

Макеева В. В., Вивчарь Р. М., Березкин А. А. Идентификация нештатных ситуаций железнодорожного транспорта России на основе кластерного анализа данных // Вестник СПбГУТ. 2024. Т. 2. № 2. С. 1. EDN: KSEHFQ

Reference for citation:

Makeeva V., Vivchar R., Berezkin A. Identification of Railroad Transport Emergency Situations in Russia Based on Data Clustering Analysis // Herald of SPbSUT. 2024. Vol. 2. Iss. 2. P. 1. EDN: KSEHFQ

бесперебойное движение поездов. Эти ситуации могут вызывать задержки, аварии, повреждения грузов и угрозу безопасности пассажиров и персонала. Идентификация и анализ нештатных ситуаций имеет решающее значение для обеспечения безопасности и эффективности железнодорожного транспорта. Одним из методов, который может быть использован для анализа данных и выявления скрытых закономерностей, является кластерный анализ.

Кластерный анализ — статистический метод сегментации данных, позволяющий классифицировать объекты по группам (кластерам) на основе их сходства. Эта процедура помогает систематизировать информацию в соответствии с ее характерными признаками и выявлять закономерности в данных.

В контексте анализа нештатных ситуаций на железнодорожном транспорте кластеризация позволяет группировать различные виды инцидентов на основе их общих характеристик, что помогает глубже понять причины и факторы, которые приводят к возникновению таких ситуаций.

Посредством анализа больших объемов данных, включая информацию о расписаниях поездов, технических характеристиках локомотивов, метеорологических условиях и других факторах, можно выявить структуру и взаимосвязи между различными формами нештатных ситуаций. Применение методов кластерного анализа также позволяет создать систему мониторинга, которая осуществляет постоянный анализ данных и предупреждает о потенциально опасных ситуациях заранее. Это позволяет снизить риск аварий и оптимизировать планирование ресурсов железнодорожного транспорта.

Таким образом, идентификация нештатных ситуаций на основе кластерного анализа представляет собой мощный инструмент для повышения безопасности, эффективности и надежности железнодорожного транспорта России. В работе мы рассмотрим примеры применения кластерного анализа для выявления нештатных ситуаций, их анализа и разработки мер по их предотвращению.

Оценка современной экологической ситуации на железнодорожном транспорте России и необходимость автоматизации анализа нештатных ситуаций

Экологическая ситуация в мире находится в критическом положении из-за дестабилизации баланса взаимодействия с природой при отсутствии хорошо разработанных национальных стратегий экологической безопасности. В связи с этим с каждым годом все больше материальных и финансовых ресурсов вкладывается в ликвидацию природных и техногенных катастроф. Одной из наиболее тревожных проблем является перевозка опасных/токсичных материалов по железной дороге.

Железнодорожный транспорт играет важную роль в экологической системе России, обеспечивая половину всех перевозок в стране, включая опасные грузы [1]. С увеличением объемов грузоперевозок возрастает необходимость анализа информации о возможных чрезвычайных ситуациях.

Поскольку данные о нештатных ситуациях на железнодорожном транспорте зачастую неоднозначны и не структурированы, для их изучения и обработки целесообразно использовать методы интеллектуального анализа данных с учетом специфики исследуемого объекта. Таким образом, важность использования информационных технологий для исследования нештатных ситуаций на железной дороге состоит в возможности оперативного и эффективного управления подобными ситуациями, анализе возможных последствий и принятии обоснованных управленческих решений.

Способы и средства обнаружения нештатных ситуаций на железнодорожных магистралях

На современных железнодорожных системах используются различные методы и технологии для обнаружения экстренных ситуаций, таких как аварии, сбои в работе оборудования и другие опасные сценарии. Благодаря прогрессу в области разработки технических решений уровень безопасности на железнодорожном транспорте значительно повышается.

Среди методов обработки больших объемов информации на железнодорожных системах можно выделить Data Mining (анализ данных) — процесс извлечения ценной информации из больших объемов данных для поиска скрытых закономерностей, трендов, аномалий и других полезных сведений [2]. В зависимости от постановки задачи все алгоритмы анализа данных разделяются на контролируемое обучение (обучение с учителем), предполагающее поэтапное построение моделей данных и обучение модели до достижения желаемых результатов, и неконтролируемое обучение (обучение без учителя), использующееся в случае отсутствия предварительных знаний о данных.

Основные задачи Data Mining — классификация, регрессия, поиск ассоциативных правил и кластеризация [3]. Рассмотрим их применение на примере анализа аварийных ситуаций на железной дороге.

Задача *классификации* [4] заключается в определении класса экстренной ситуации на основе ее характеристик; ее недостатками являются необходимость обширного обучающего набора данных и проблемы переобучения.

Регрессия позволяет определить значение параметра для известных характеристик экстренных ситуаций на железной дороге, который представляется набором вещественных чисел.

Ассоциативный анализ, или *поиск ассоциативных правил*, может помочь выявить комбинации факторов, таких как погодные условия, техническое состояние оборудования и наличие особенностей маршрута, которые могут увеличивать вероятность возникновения нештатных ситуаций.

Кластеризация [4] используется для группировки объектов на основе их сходства. В отличие от классификации при кластеризации не предполагается наличие заранее заданных меток для групп, и модель сама определяет структуру данных. Целью кластеризации является разделение набора данных на группы (кластеры), где объекты внутри одного кластера более похожи друг на друга, чем на объекты из других групп. Например, кластерный анализ может помочь выделить категории аварийных ситуаций, связанных с техническими дефектами, человеческим фактором или внешними воздействиями, что в дальнейшем может быть полезно для разработки специализированных мер по предотвращению подобных ситуаций.

Каждый из этих методов имеет свои преимущества и может применяться в различных сферах, включая идентификацию нештатных ситуаций на железнодорожном транспорте, где классификация может помочь в определении типа аварии, регрессия — в прогнозе ее параметров, поиск ассоциативных правил — в выявлении нестандартных ситуаций и аномалий (обнаруживая скрытые отношения в данных), а кластеризация — в группировке ситуаций с похожими особенностями для выявления закономерностей.

Оценка результатов выявления нештатных ситуаций на железной дороге

Нештатные ситуации на железной дороге могут иметь серьезные последствия, такие как аварии, простои в движении поездов, ущерб для окружающей среды и, что самое важное, угроза для жизни и здоровья людей. Эффективное выявление и оценка нештатных ситуаций играют ключевую роль в обеспечении безопасности и надежности железнодорожного транспорта.

Процесс выявления чрезвычайных ситуаций включает в себя несколько этапов:

- 1) *мониторинг*: постоянное наблюдение за состоянием инфраструктуры, поездов, погодными условиями и другими факторами, которые могут привести к возникновению чрезвычайных ситуаций;
- 2) *обнаружение*: использование современных технологий, таких как сенсоры, видеонаблюдение, системы дистанционного мониторинга для выявления возможных проблемных ситуаций;
- 3) *оценка*: анализ полученных данных и оценка их потенциального воздействия на безопасность движения поездов и пассажиров.

Применение кластерного анализа для оценки результатов выявления нештатных ситуаций на железной дороге представлено ниже пошагово.

Шаг 1. Группировка аварийных ситуаций осуществляется для объединения различных видов аварий в зависимости от их уникальных характеристик (факторы возникновения, место событий, временные рамки и другие параметры); способствует выявлению ключевых групп аварийных ситуаций и более глубокому пониманию их природы.

Шаг 2. Оценка эффективности профилактики аварийных ситуаций предоставляет возможность оценить степень успешности предотвращения нештатных ситуаций в различных группах. Например, постоянное возникновение аварий в конкретном кластере с высоким уровнем риска может свидетельствовать о необходимости совершенствования механизмов предупреждения в данной категории событий.

Шаг 3. Идентификация тенденций и прогнозирование помогают выявить тенденции и паттерны в данных о нештатных ситуациях, что позволяет более точно прогнозировать возможные сценарии и разрабатывать меры предотвращения заранее.

Реализация специализированного программного обеспечения с использованием кластерного анализа данных для выявления причин аварий на железнодорожном транспорте России на языке программирования Python представлена на рисунке 1.

```
import pandas as pd
import numpy as np
import folium

from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from folium.plugins import MarkerCluster

# Загрузка данных о нештатных ситуациях
incident_data = {
    'incident_id': [1, 2, 3, 4, 5],
    'location': ['Москва', 'Санкт-Петербург', 'Екатеринбург', 'Новосибирск', 'Красноярск'],
    'description': ['Сход поезда с рельсов', 'Пожар на станции', 'Обрыв контактной сети', 'ДТП на пе-
реезде', 'Техническая неисправность поезда'],
    'date': ['2024-03-15', '2024-03-16', '2024-03-16', '2024-03-17', '2024-03-17']
}
df_incidents = pd.DataFrame(incident_data)

# Загрузка географических данных
geo_data = {
    'station_name': ['Москва', 'Санкт-Петербург', 'Новосибирск', 'Екатеринбург', 'Красноярск'],
    'latitude': [55.7558, 59.9343, 55.0084, 56.8389, 56.0106],
    'longitude': [37.6176, 30.3351, 82.9357, 60.6057, 92.8524],
    'railway_type': ['Высокоскоростная', 'Пассажирская', 'Грузовая', 'Смешанная', 'Пассажирская'],
    'track_condition': ['Хорошее', 'Удовлетворительное', 'Плохое', 'Хорошее', 'Удовлетворительное']
}
df_geo = pd.DataFrame(geo_data)

# Обработка данных (Объединение данных или добавление новых признаков)
df_combined = pd.merge(df_incidents, df_geo, left_on='location', right_on='station_name', how='left')
df_combined = df_combined.drop('station_name', axis=1)

# Вывод обработанных данных
print(df_combined)

# Выбираем признаки для кластеризации (широта и долгота)
features = df_combined[['latitude', 'longitude']]

# Нормализуем данные
scaler = StandardScaler()
features_scaled = scaler.fit_transform(features)

# Применяем алгоритм KMeans для кластеризации
kmeans = KMeans(n_clusters=3)

# Задаем количество кластеров
clusters = kmeans.fit_predict(features_scaled)

# Добавим информацию о кластерах в исходные данные
df_combined['cluster'] = clusters

# Вывод результата кластеризации
print(df_combined[['location', 'latitude', 'longitude', 'cluster']])

# Создаем карту с центрированием и начальным масштабом
map_clusters = folium.Map(location=[55.7558, 37.6176], zoom_start=5)

# Добавляем точки нештатных ситуаций на карту в каждый кластер
for cluster_num in df_combined['cluster'].unique():
    cluster_data = df_combined[df_combined['cluster'] == cluster_num]
    cluster_map = MarkerCluster().add_to(map_clusters)
    for idx, row in cluster_data.iterrows():
        folium.Marker([row['latitude'], row['longitude']], popup=f"Location: {row['location']}",
            icon=folium.Icon(color='red')).add_to(cluster_map)

# Сохраняем интерактивную карту в файл
map_clusters.save('cluster_map.html')

# Группируем данные по кластерам
cluster_groups = df_combined.groupby('cluster')

# Проходим по каждому кластеру для анализа характеристик
for cluster_num, cluster_data in cluster_groups:
    print(f"Cluster {cluster_num}:")
    print(cluster_data[['location', 'description', 'railway_type', 'track_condition']].describe())
    print("-----")
```

Рис. 1. Программа на языке Python для идентификации нештатных ситуаций на железнодорожном транспорте России, реализующая алгоритм кластерного анализа

Повышение эффективности идентификации нештатных ситуаций на железнодорожном транспорте

Эффективная идентификация и управление нештатными ситуациями играют ключевую роль в обеспечении безопасности и надежности железнодорожного транспорта. В данном исследовании анализ эффективности идентификации нештатных ситуаций на железнодорожном транспорте представлен на основе числовых данных [5]. Анализируются такие показатели, как общее число и типы нештатных ситуаций, среднее время реакции на них, а также процент успешно разрешенных случаев.

Цель анализа — определение уровня готовности и эффективности системы идентификации нештатных ситуаций, а также выявление областей, требующих дополнительного внимания и улучшений. Анализ основан на данных, собранных за определенный период, и предоставляет ценную информацию для разработки стратегий по улучшению системы безопасности и более эффективному реагированию на потенциальные угрозы.

Общее количество и типы нештатных ситуаций, которые возникали на российском железнодорожном транспорте за 2023 год, представлены в таблице 1.

Таблица 1. Число и типы нештатных ситуаций

Тип нештатной ситуации	Число случаев за год
Сход с рельсов	15
Пожар на станции	7
Отключение электропитания	10
Происшествия с пассажирами	5

Примечания:

- 1) сход с рельсов (15 случаев за год): важно обратить внимание на возможные причины сходов и принять меры по их предотвращению, так как это одна из наиболее серьезных и опасных ситуаций;
- 2) пожар на станции (7 случаев за год): необходимо провести анализ причин возникновения пожаров и улучшить системы пожарной безопасности на станциях;
- 3) отключение электропитания (10 случаев за год): может быть связано с техническими проблемами или внешними воздействиями; необходимо улучшить системы электроснабжения для предотвращения подобных ситуаций;
- 4) происшествия с пассажирами (5 случаев за год): требует дополнительного анализа для выявления причин и принятия мер по обеспечению безопасности пассажиров.

Эффективность идентификации нештатных ситуаций также зависит от скорости реакции на них. Данные о среднем времени реакции представлены в таблице 2.

Таблица 2. Среднее время реакции

Тип нештатной ситуации	Среднее время реакции (в минутах)
Сход с рельсов	20
Пожар на станции	15
Отключение электропитания	10
Происшествия с пассажирами	25

Примечания:

- 1) сход с рельсов (20 минут): необходимо сократить время реакции для быстрого восстановления движения поездов и минимизации последствий;
- 2) пожар на станции (15 минут): важно улучшить системы обнаружения и тушения пожаров для быстрой локализации и ликвидации возгораний;
- 3) отключение электропитания (10 минут): требуется оптимизация систем мониторинга и реагирования для быстрого восстановления электроснабжения;
- 4) происшествия с пассажирами (25 минут): необходимо улучшить процедуры эвакуации и медицинской помощи для быстрой реакции на происшествия с пассажирами.

Критическим показателем эффективности идентификации нештатных ситуаций является процент успешно решенных проблем. Данные за 2023 год представлены в таблице 3.

Идентификация нештатных ситуаций на железнодорожном транспорте также осуществляется с использованием современных технологий, таких как системы мониторинга и датчики. Данные о степени использования таких технологий представлены в таблице 4 и представляют собой ценную информацию для разработки стратегий улучшения системы предотвращения и реагирования на возможные угрозы на железнодорожном транспорте.

Таблица 3. Количество успешно решенных случаев

Тип нештатной ситуации	Количество успешно решенных проблем (в %)
Сход с рельсов	80
Пожар на станции	90
Отключение электропитания	85
Происшествия с пассажирами	75

Примечания:

- 1) сход с рельсов (80 % успешно решенных случаев): показатель успеха в решении данного типа нештатных ситуаций достаточно высок, однако требуется постоянное обновление и совершенствование процедур для его улучшения;
- 2) пожар на станции (90 % успешно решенных случаев): очень высокий процент успешно решенных случаев говорит об эффективности систем пожаротушения и организации оперативных мероприятий;
- 3) отключение электропитания (85 % успешно решенных случаев): высокий процент успешно решенных случаев свидетельствует о хорошей готовности и реагировании на подобные ситуации;
- 4) происшествия с пассажирами (75 % успешно решенных случаев): необходимо улучшить процедуры и обучение персонала для повышения эффективности реагирования на происшествия с пассажирами.

Таблица 4. Степень использования технологий

Технология	Степень использования (в %)
Системы видеонаблюдения	95
Датчики состояния путей	65
Автоматизированные системы управления	60

На основе проведенного анализа были выделены ключевые направления для повышения эффективности идентификации нештатных ситуаций. В частности, важно развивать системы машинного обучения и искусственного интеллекта для более точного и быстрого обнаружения аномалий на железнодорожных маршрутах. Также следует уделять внимание интеграции различных источников данных, включая информацию о погодных условиях, состоянии инфраструктуры и поведении пассажиров.

Реализация методов кластеризации для анализа нештатных ситуаций на железнодорожном транспорте с использованием языка программирования Python

В данном разделе рассмотрим основанную на алгоритме *k*-means (*k*-средних) программу с использованием языка программирования Python и различных библиотек (см. рисунок 1), разработанную для анализа и предотвращения аварийных ситуаций на железнодорожных системах.

Программа начинает свою работу с загрузки данных о нештатных ситуациях и географических параметров. Затем данные объединяются на основе совпадения местоположения для последующего анализа. Далее выбираются показатели широты и долготы для кластеризации, данные нормализуются, и применяется алгоритм *k*-means для разделения точек по координатам на карте. Результаты кластеризации визуализируются на интерактивной карте для наглядного представления.

Принципы работы программы лежат в объединении и обработке данных, нормализации для корректной кластеризации, применении *k*-means для разделения точек на кластеры на основе географического расположения, а также визуализации результатов на карте. Результаты работы программы позволяют группировать точки ситуаций на карте, определять центры кластеров для выявления областей повышенного риска и помогают в принятии решений по улучшению безопасности транспорта.

Интерпретация выводов программы представляет собой карту с кластерами, указывающими на области, предрасположенные к нештатным ситуациям, и обеспечивает анализ характеристик кластеров для выявления особенностей каждой группы событий, что ценно для специалистов по безопасности. Образец карты с кластерами представлен на рисунке 2.

Алгоритм *k*-means, в свою очередь, обладает временной сложностью:

$$O = nkid,$$

где n, k, i, d — количество элементов в 1-м, 2-м, 3-м и 4-м наборах данных, соответственно. Это подразумевает высокую эффективность на малых и средних наборах данных, хотя в больших могут возникнуть проблемы с производительностью из-за итеративного обновления центров кластеров.

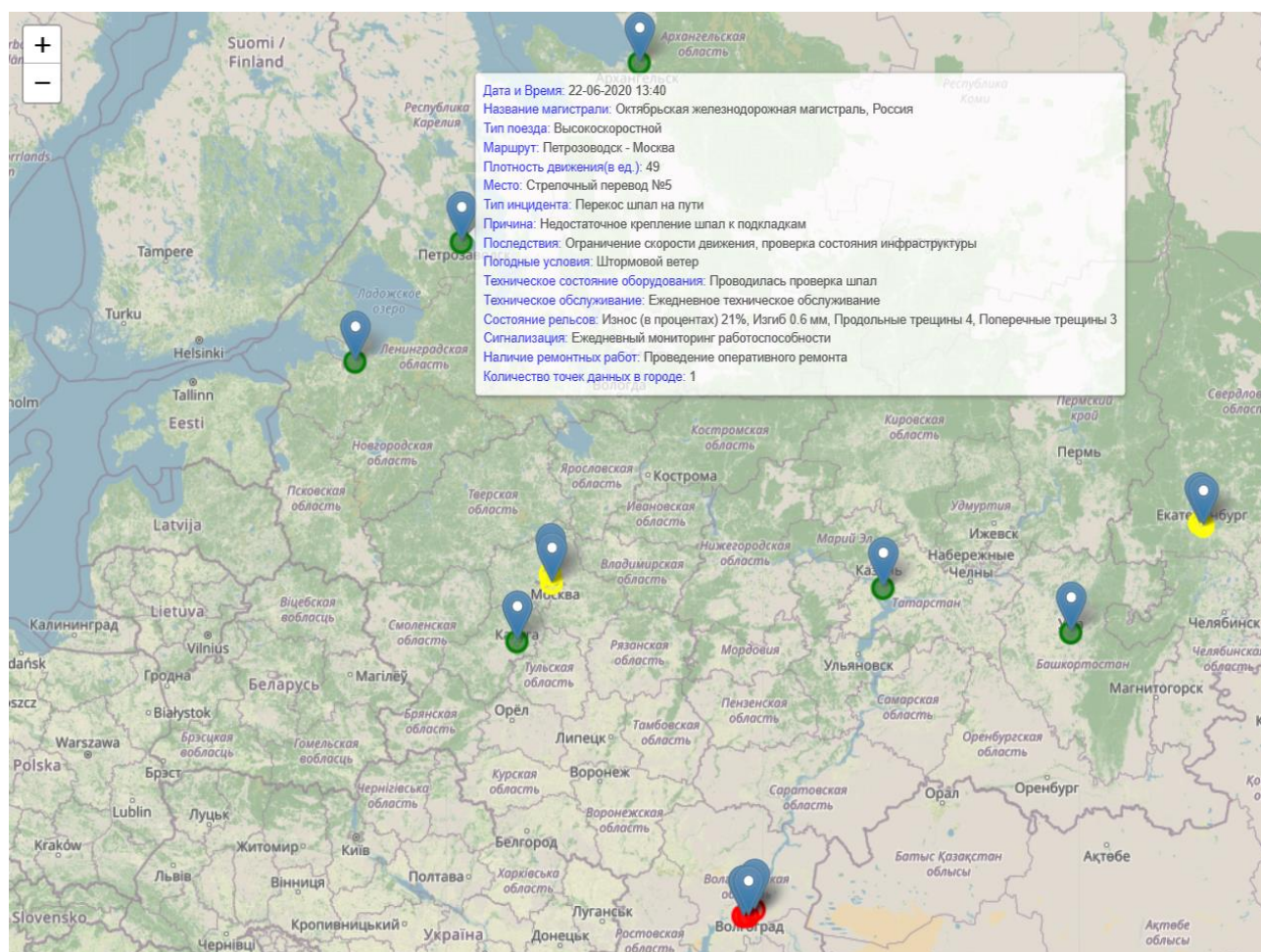


Рис. 2. Образец карты с кластерами

Выбор Python для реализации программы обусловлен его популярностью и богатой системой библиотек, таких как Pandas, NumPy, Scikit-learn и Folium, что делает его мощным инструментом для анализа данных, кластеризации и визуализации на карте. Таким образом, использование методов кластеризации в анализе нештатных ситуаций на железнодорожных системах с помощью Python и указанных библиотек позволяет эффективно выявлять и анализировать области с повышенным риском аварий.

Выводы

Исследование нештатных ситуаций на железнодорожном транспорте России при помощи кластерного анализа данных представляет собой эффективный подход к выявлению закономерностей и паттернов, которые раскрывают особенности возникновения подобных ситуаций:

- кластерный анализ позволяет выделить различные типы нештатных ситуаций на железной дороге, что в свою очередь способствует оперативному реагированию на них и принятию необходимых мер;
- посредством анализа данных о нештатных ситуациях и их классификации на кластеры можно улучшить процессы обеспечения безопасности на железнодорожном транспорте, повысив прогностическую способность и оперативность реагирования;
- результаты исследования могут применяться для разработки рекомендаций по совершенствованию инфраструктуры и процессов на железнодорожном транспорте, что поможет в будущем повышению эффективности всей транспортной системы.

Разработанные рекомендации [6] по совершенствованию инфраструктуры и процессов железнодорожного транспорта таковы:

- модернизация и расширение существующих железнодорожных путей для увеличения пропускной способности и сокращения времени в пути;
- строительство новых высокоскоростных железнодорожных линий для сокращения времени следования между крупными городами;
- внедрение системы мониторинга состояния путей и оборудования для оперативного выявления и устранения дефектов;
- улучшение системы безопасности на станциях и в поездах, включая видеонаблюдение, датчики движения и аварийные кнопки;
- внедрение системы автоматического управления поездами (АТС, *аббр. от англ. Automatic Train Control*) для предотвращения столкновений и обеспечения безопасной работы поездов;
- внедрение системы умного управления движением поездов для оптимизации расписаний, увеличения частоты поездок и сокращения времени ожидания;
- переход на использование электрической тяги и внедрение современных энергосберегающих технологий для снижения выбросов вредных веществ;
- стимулирование развития грузовых перевозок по железной дороге путем создания специальных тарифных программ и снижения административных барьеров;
- модернизация инфраструктуры для грузовых перевозок, включая строительство современных терминалов и складских комплексов;
- разработка мобильных приложений для пассажиров с информацией о расписаниях, задержках и услугах на станциях;
- разработка и внедрение комплексных программ развития железнодорожной инфраструктуры, основанных на взаимодействии государственных органов, частных компаний, академических институтов и общественных организаций.

Внедрение кластерного анализа данных в систему мониторинга и управления безопасностью железнодорожного транспорта в России является неотъемлемым шагом для повышения уровня безопасности, эффективности и надежности транспортной системы.

Литература

1. Грузооборот и пассажирооборот на железнодорожном транспорте // ТК Капро. URL: <https://cargotk.com/news/gruzooborot-i-passazhirooborot-na-zheleznodorozhnom-transporte/> (дата обращения 13.04.2024)
2. Leskovec J., Rajaraman A., Ullman J. Mining of Massive Datasets. Cambridge: Cambridge University Press, 2014. 467 p.
3. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Mining: учебное пособие. СПб.: БХВ-Петербург, 2004. 336 с. EDN: QMNOPT
4. Шубинский И. Б., Замышляев А. М., Проневич О. Б., Игнатов А. Н., Платонов Е. Н. Применение методов машинного обучения для прогнозирования опасных отказов объектов железнодорожного пути // Надежность. 2020. Т. 20. № 2. С. 43–53. DOI: 10.21683/1729-2646-2020-20-2-43-53. EDN: PEEIGO
5. Методы и средства анализа данных // Bourabai Research: Технологии XXI века. URL: <http://www.bourabai.ru/tpoi/analysis.htm> (дата обращения 13.04.2024)
6. Отчет МСЖД 2023: статистика происшествий на железных дорогах // Zdmira. 23.03.2024. URL: <https://zdmira.com/articles/otchet-mszhd-2023-statistika-proisshествij-na-zhele-znykh-dorogakh> (дата обращения 13.04.2024)
7. Правила безопасности и порядок ликвидации аварийных ситуаций с опасными грузами при перевозке их по железным дорогам, утв. МПС РФ 25.11.1996 г. № ЦМ-407, МЧС РФ от 31.10.1996 г. № 9/733/3-2. URL: <https://normativ.kontur.ru/document?moduleId=1&documentId=263129> (дата обращения 13.04.2024)

Статья поступила 23 апреля 2024 г.
Одобрена после рецензирования 29 апреля 2024 г.
Принята к публикации 2 мая 2024 г.

Информация об авторах

Makeeva Виктория Витальевна — студент 4-го курса факультета инфокоммуникационных сетей и систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича. E-mail: vika.makeeva2015@yandex.ru

Vivchar Роман Михайлович — кандидат технических наук, доцент, доцент кафедры программной инженерии и вычислительной техники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича. E-mail: rvivcar12@sut.ru

Berezkin Александр Александрович — кандидат технических наук, доцент, доцент кафедры программной инженерии и вычислительной техники Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича. E-mail: berezkin.aa@sut.ru

Identification of Railroad Transport Emergency Situations in Russia Based on Data Clustering Analysis

V. Makeeva ✉, R. Vivchar, A. Berezkin

The Bonch-Bruevich Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Purpose. Identification of emergency situations on the railway to ensure safety and optimize processes. This includes monitoring the state of infrastructure, train movements, weather conditions and other factors that can lead to emergencies. **The purpose** of the work is to develop specialized software using cluster data analysis to identify the causes of accidents in Russian railway transport. This is critically important because of the significant role of rail transport on public safety and the economy of the country, as well as because of the potential threats associated with possible emergencies, especially during the transportation of dangerous goods. **Methods.** Cluster analysis to systematize various types of anomalies occurring in the railway infrastructure, taking into account parameters such as sources of problems, geographical location and time characteristics. This will help identify key groups of situations and better understand their nature, as well as evaluate the effectiveness of emergency prevention in various groups. **Results.** Using the method of cluster data analysis, the main patterns underlying unusual circumstances in the functioning of railway transport are identified. This allows us to come to significant conclusions, including the identification of various categories of emergency situations, optimization of safety measures and improvement of the effectiveness of the railway system. **Novelty.** The novelty of the proposed solution lies in applying cluster analysis to identify trends, patterns, and potential scenarios in data related to railway accidents. **Practical relevance.** This research aims to provide recommendations for improving infrastructure, enhancing safety measures, and ultimately contributing to the overall efficiency of the transportation system in the long term perspective.

Key words: emergency situations, emergencies, cluster analysis, railway transport, process optimization, trends, safety, Data Mining software development

Information about Author

Makeeva Victoria — a 4th Year Student of Faculty of Information and Communication Networks and Systems (The Bonch-Bruevich Saint Petersburg State University of Telecommunications).
E-mail: vika.makeeva2015@yandex.ru

Vivchar Roman — Ph. D. of Engineering Sciences, Associate Professor, Associate Professor at the Department of Software Engineering and Computer Engineering (The Bonch-Bruevich Saint Petersburg State University of Telecommunications). E-mail: rvivcar12@sut.ru

Berezkin Aleksander — Ph. D. of Engineering Sciences, Associate Professor, Associate Professor at the Department of Software Engineering and Computer Engineering (The Bonch-Bruevich Saint Petersburg State University of Telecommunications). E-mail: berezkin.aa@sut.ru