

УДК 656.021.5

## Особенности использования big data при исследовании грузопотоков на железнодорожном транспорте

Ф. А. Ярмолинский<sup>1</sup>, О. Д. Покровская<sup>1</sup>, Е. Д. Пасечник<sup>1</sup>, Е. В. Пакулина<sup>1</sup>,  
А. А. Трапезников<sup>2</sup>

<sup>1</sup> Петербургский государственный университет путей сообщения Императора Александра I, Россия, 190031, Санкт-Петербург, Московский пр., 9

<sup>2</sup> Военный институт железнодорожных войск и военных сообщений ВА МТО им. генерала армии А. В. Хрулёва

**Для цитирования:** Ярмолинский Ф. А., Покровская О. Д., Пасечник Е. Д. и др. Особенности использования big data при исследовании грузопотоков на железнодорожном транспорте // Бюллетень результатов научных исследований. — 2024. — Вып. 1. — С. 107–122. DOI: 10.20295/2223-9987-2024-01-107-122

### Аннотация

**Цель:** охарактеризовать особенности и перспективы использования инструментария и технологий больших данных в управлении перевозочным процессом на железных дорогах. **Методы:** нейросетевое моделирование, системный анализ, прогнозирование, программирование, большие данные, предиктивная аналитика. **Результаты:** предложена даталогическая модель сущностей для хранения актуальных данных по грузопотокам, предложена структура построения системы для накопления информации. Кроме того, в работе изучаются прикладные вопросы решения проблем хранения, получения и обработки данных при использовании методов big data. **Практическая значимость:** улучшение процесса управления процессами перевозок на железнодорожном транспорте в условиях цифровой трансформации в части получения более точных прогнозов.

**Ключевые слова:** big data, грузопотоки, предиктивная аналитика, цифровизация, прогноз, нейросетевая модель.

### Введение

Структура железнодорожного транспорта России отражает масштабность и сложность непрерывно обрабатываемых данных в управлении процессами перевозок. Структура же и количество сведений по информационным объектам могут значительно отличаться, но при этом дополнять друг друга по различным запросам [1, 2].

Процесс обработки и накопления данных ведется непрерывно из различных источников. В рамках анализа грузоперевозок наиболее интересны данные по таким информационным объектам, как:

- сети различных полигонов РЖД со станциями, участками;
- подъездные пути предприятий и сами предприятия;
- поезда, вагоны, контейнеры и грузы [3].

При исследовании информационного пространства, работающего в экосистеме «Российских железных дорог», в частности в условиях множества инфор-

мационных систем (ИС), предлагается унифицированный источник информирования, содержащий объем сведений, достаточный для аналитики [4]. В этом случае все данные по различным информационным объектам (вагон, контейнер и др.) будут накапливаться в АСОУП (Автоматизированная система оперативного управления перевозками), которая содержит:

- 1) поездную модель, которая хранит сведения по поездам, состояние, направление;
- 2) вагонную модель, которая хранит сведения по вагонам, их состоянию;
- 3) контейнерную и локомотивную модели;
- 4) отправочную (идентификатор отправки);
- 5) НСИ (нормативно-справочная информация);
- 6) заявки (номер ГУ-12, поступают из ИС «Этран»).

При этом каждая из отмеченных выше функциональных моделей может накапливать исторические сведения по движению и состоянию, а также содержать оперативные сведения по текущему положению. АСОУП является основной схемой по наполнению данными множества систем, таких как «Сириус», «Оскар-М» и другие [5, 6].

Автоматизированная система располагает сведениями огромного объема данных. Информационный комплекс систем, включающий АС, представляет собой хранилище данных, которое можно отнести к традиционным структурам big data, для чего необходимо подробнее рассмотреть структуру АСОУП.

Большой поток информации, которая наполняет ИС, содержит всю актуальную информацию об объектах на полигоне дорог РЖД, включая подход со смеж-



Рис. 1. Структурная схема АСОУП

ных дорог. Хронологически эта информация накапливается в исторических хранилищах. Каждая модель представлена множеством таблиц по ее деятельности. АСОУП представлена огромным количеством выходных справок и документов и отображена на рис. 1.

Объем накопленных данных в информационных источниках непрерывно увеличивается, архитектура взаимодействия постоянно расширяется и видоизменяется, что существенно усложняет ее структуру, ресурсоемкость и требует оптимизации уже функционирующих модульных структур для хранения и обработки больших данных [7].

Динамические процессы на полигонах железных дорог могут быть изучены и проанализированы по различным критериям. Большой объем информации позволяет моделировать процессы, нетипичные в рамках изучаемой модели, для поиска наилучшего решения, а также изучать накопленные данные посредством человеко-машинных систем для прогнозной аналитики [8].

Наиболее интересная информация для прогнозной аналитики грузопотоков представлена в вагонной модели АСОУП и выражена следующими сведениями: станция операции; нахождение на станции/в поезде; номер поезда и вагона в составе поезда; положение на станции; дислокация вагона; сведения по направлению, станции формирования и сортировки в пути следования; эксплуатационные данные и состояние, включая пробеги (текущие и плановые), вид плановых ремонтов, дату последнего ремонта, срок службы и т. д. Однако система оперативного управления представляет собой лишь объемное хранилище информации, данные из которого можно получать, запрашивая различные справочные сведения.

Для визуализации и анализа, понятного человеку-оператору (лицу, принимающему решение), необходимо структурировать огромный объем информации с целью получения сведений, пригодных для аналитики.

Рассмотрим перспективы применения методов big data для аналитики процессов грузоперевозки [9].

Ключевая сложность связана со скоростью получения и обработки данных, их большим количеством. Пользователям и технологам важно оперативно получать информацию, которая может храниться в различных источниках, аккумулируя сведения в единый ответ. Помимо оперативной обработки, существует проблема стремительного увеличения объема данных, связанного с ростом цифровизации и развитием технологий [10].

Программный комплекс ОАО «РЖД» включает в себя большой набор информационных систем (ИС), хранение и структура данных в которых отличаются. Обработка огромных массивов данных остается непрерывным процессом, что также накладывает ограничения на системы передачи данных. В связи с этим остается проблема задержек в ответах и превышающей норму нагрузки на сеть.

Разнообразие данных может привести к отсутствию прямой связи между системами по одинаковым объектам [3, 11].

Таким образом, необходимо иметь гибкую модель для передачи данных и унифицированную структуру для их хранения и обработки. На оперативность работы могут повлиять архитектура передачи данных, их объем и структура, а на минимизацию роста данных — структуризация информации, отсеивающая ложные и неструктурные сведения, что также может снизить вероятность дублирования.

## Обсуждение результатов

Возможным решением предлагается выбор модели для накопления сведений для последующей аналитики. Потребность в унификации и консолидации данных приводит к необходимости создания единого хранилища данных, которое сможет накапливать и обрабатывать информацию различного формата, описывая разного рода семантические правила, применение которых направлено на формализацию данных. При этом реализация единой формализации имеет ряд проблем и ограничений, которые обусловлены различными правилами хранения и наполнения данных в ИС, разнородными структурами и набором уникальных полей в системах, связь между которыми не всегда очевидно прослеживается, что может приводить к неоднозначному толкованию информации [12]. Помимо этого, огромные объемы проводимых операций при грузообороте лишь усложняют процесс обработки.

Важно отметить, что структура хранилища не должна быть абстрактна, но должна иметь модель, которая может быть описана графически, математически и формальным языком. При этом единое хранилище не должно являться результатом интеграции различных ИС и БД, а быть результатом их агрегации.

Единое хранилище данных реализуется исключительно с применением методов и технологий больших данных (big data) — это особый формат структуры данных, функционал которой обеспечивает параллельность многих вычислений для обработки различного типа данных с применением параллельного программирования. Данные, обрабатываемые методами big data, могут быть структурированы, частично структурированы по определенным критериям или не структурированы. Применение методов и средств больших данных позволяет увеличить техническую и экономическую эффективность, влиять на безопасность движения посредством поддержки в принятии технологических решений [13, 14].

Главные возможности big data выражаются тремя характеристиками модели 3V: volume (величина физического объема хранилища), velocity (скорость обработки, накопления и получения данных), variety (многообразие обработки различных структур данных) [9].

Для понимания перспективы применения методов больших данных необходимо выделить следующие сведения, интересующие для анализа события в рамках осуществления грузоперевозок: различные форс-мажоры, бросания (причины, сроки простоя), анализ причин несвоевременной доставки (срыв сроков доставки, величина срыва), дробление отправок, анализ потоков, стабильности перевозок по различным направлениям, исполнение плановых заявок, наличие окон и предупреждений, выполнение плана формирования, прогнозы грузопотоков по заявкам перевозок.

Модель обмена информацией между информационными источниками, описанными ранее, дана на рис. 2.

Модель, хранящая информацию по грузовым операциям, взаимодействует с поездной моделью, которая может обмениваться информацией с вагонной, контейнерной и локомотивной моделями. При этом ключом прямой связи между сущностями является номер вагона, представленный в каждой из них. Для наглядного понимания схемы передачи данных была построена даталогическая модель хранения данных, отражающая взаимосвязь между всеми сущностями и набором полей данных (рис. 3).

Описание сущностей таблиц представлено в табл. 1.

Модель заявок предназначена для хранения информации по полученным заявкам. Отправочная модель содержит информацию по накладным на осуществление грузоперевозок, вагонная модель хранит сведения по всем вагонам, модель картотеки содержит расширенную информацию по вагонам, их истории пробега и т. д. Остальные модели хранят информацию по доставленным грузам, способу доставки, колесным парам, ремонтам и т. д. [15].

Таким образом, все модели образуют единую систему хранения больших данных, состоящую из сущностей, каждая из которых должна включать номер вагона. В рамках данной статьи предлагается в качестве унифицированной модели принимать рейс вагона, включая все операции на пути следования, станции, время прибытия, отправления и т. д. Более подробно рейс может быть описан в отправочной

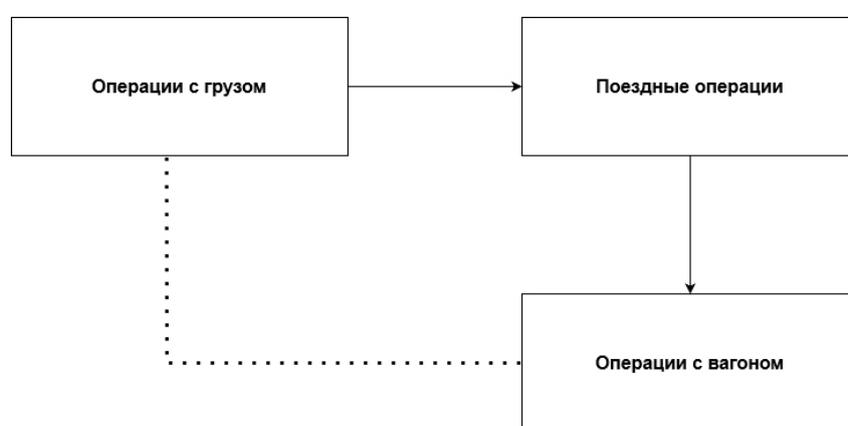


Рис. 2. Модель пооперационного информационного обмена

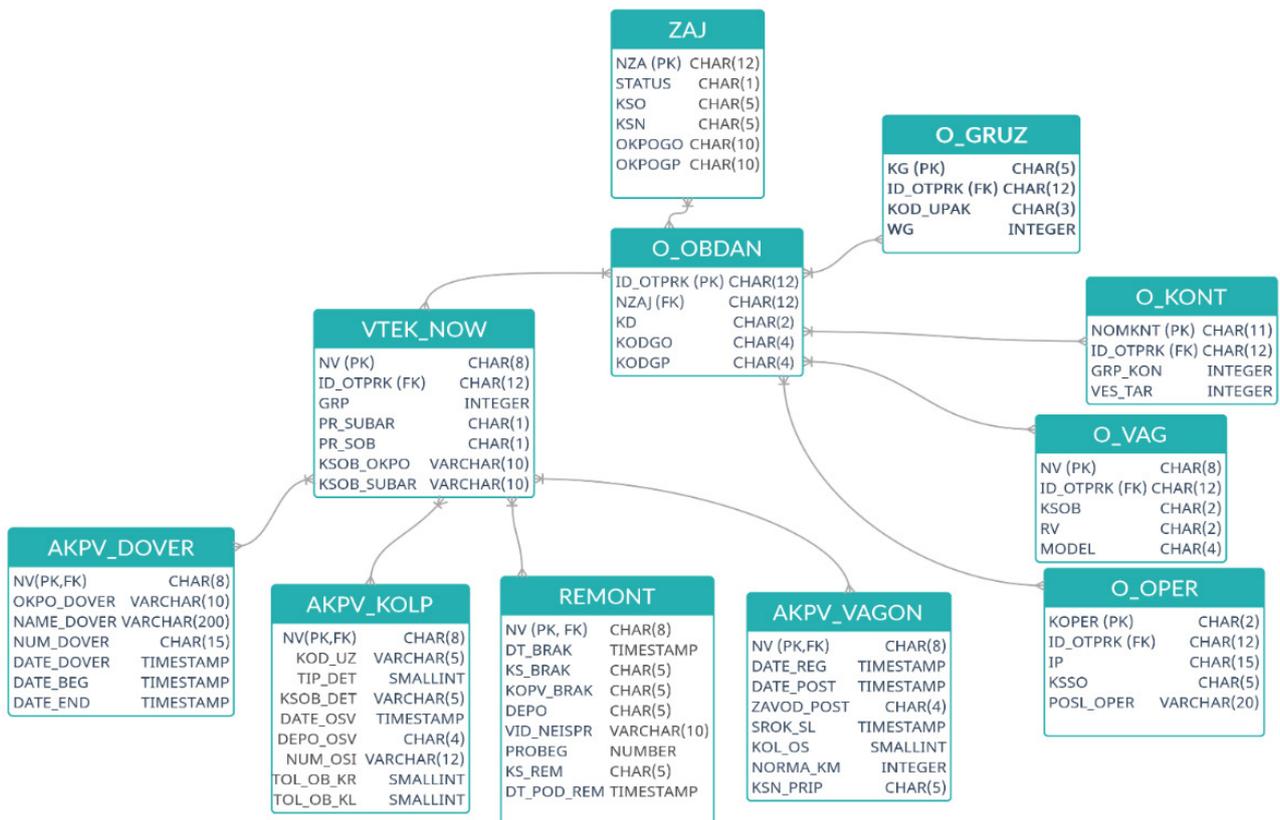


Рис. 3. Даталогическая модель сущностей

ТАБЛИЦА 1. Сущности хранения данных

Наименование сущности	Описание
Zaj	Заявка
O_obdan	Отправочная модель
Vtek_now	Вагонная модель
Akpv_vagon	Картотека
Akpv_dover	Доверенность
Akpv_kolp	Колесные пары
Remont	Ремонт
O_oper	Операции с отправкой
O_vag	Данные по вагонам
O_kont	Сведения по контейнеру
O_gruz	Грузы по отправкам

модели, включая станцию отправления, назначения, грузоотправителя, грузополучателя, вес груза, число вагонов, контейнеров, срок доставки и т. д.

В данном случае отправка подразумевается как некий элемент потока от станции отправления до станции назначения. Имея эти потоки за определенный период, можно производить анализ перевозок и предиктивную аналитику, выбирая критерий оценивания. Критериями могут быть нарушение сроков доставки, загруженность

станций, прогнозирование подхода или возможных ремонтов [16]. По каждому показателю можно вести предиктивную аналитику, при этом ключевым аспектом является архитектура построения сети, от которой будет зависеть возможность применения различных математических методов и нейросетевых моделей.

Данная модель может легко масштабироваться, добавляя критерии анализа, и быть кросс-платформенным решением при учете наличия единого стандарта хранения данных. Недостатком подхода к единой модели является сравнительно большой объем данных, который необходимо структурировать для последующей консолидации. Наличие большого количества данных подразумевает наличие дублирований и возможного отсутствия связей и неполноты информации. При этом очевидными преимуществами перехода к данной модели являются:

- хранение данных единого формата;
- возможность построения интуитивно понятной визуализации отношений между сущностями;
- оперативность доступа к информации и ее обработки в сравнении с аналогичными запросами;
- возможность использования современных технологий, применение которых позволит повысить эффективность поиска оптимального решения.

```

with ic as
(select kdn, dn, round(sum(cntvag)/30), kv,
round(sum(case when wg>0 then cntvag else 0 end)/30), kvg,
round(count(*)/30), ko,
round(sum(case when i.formras=60 then 1 else 0 end)/30), ds,
round(sum(case when i.formras=60 and kdo=kdn then 1 else 0 end)/30), ds01,
round(avg(i.raskr)), rs, round(avg(dt.vg-dto), 1), dd,
round(stddev(dt.vg-dto), 1), so,
round(sum(case when dt.vg>srdest+1 then cntvag else 0 end)/30), sr,
round(avg(srdest+1-dt.vg), 1), otkl,
round(stddev(srdest+1-dt.vg), 1), cotkl
from tsod.o_obdan i
where kdn in (select kd from nci.dor d where d.kstr=643 and kd not in (85,91)) and dt.vg>sysdate-30
group by kdn
)
select d.md Дорога
, kv вагонов, kvg груженых
, ko отправок , ds досылок, ds01 досыл_своих
, round(kv/(ko-ds), 1) ваг_в_отпр
, round(100*ds/ko) проц_дос
, rs расст
, dd ср_время
, so ср_откл
, round(rs/(dd*24)) скорость

```

ДОРОГА	ВАГОНОВ	ГРУЖЕННЫХ	ОТПРАВОК	ДОСЫЛОК	ДОСЫЛ_СВОИХ	ВАГ_В_ОТПР	ПРОЦ_ДОС	РАССТ	СР_ВРЕМЯ	СР_ОТКЛ	СКОРОСТЬ	ВАГ_СРВ	ПРОЦ_СРВ
1 ОКТ	14151	9720	3719	650	163	4,6	17	1830	6	5,1	13	1756	12
2 ДВОСТ	10522	6903	4656	852	615	2,8	18	3200	9,9	9,1	13	3292	31
3 С-КАВ	9018	6341	2426	320	137	4,3	13	1331	7,9	7,4	7	2043	23
4 Э-СИБ	16558	4564	8708	263	174	2	3	2800	7,8	9,6	15	1110	7
5 СВЕРД	9754	3603	3866	387	264	2,8	10	1676	7,1	6	10	1605	16
6 МОСК	6077	2818	4347	225	127	1,5	5	2851	8,6	6,5	14	1946	32
7 СЕВ	4851	2538	1782	106	53	2,9	6	959	5	4,6	8	564	12
8 Ю-УР	4563	2379	1776	199	65	2,9	11	840	5,2	5,6	7	988	22
9 В-СИБ	4934	2117	2039	88	54	2,5	4	2158	8,6	7,7	10	679	14
10 Ю-ВОС	5497	2107	1651	169	82	3,7	10	858	7,2	7	5	1255	23
11 КЕШ	5896	1862	2463	193	94	2,6	8	1160	7,4	6,9	7	1484	25
12 ГОРЬК	3146	1768	1362	128	58	2,5	9	1162	7,2	6,2	7	911	29

Рис. 4. Элемент скриншота рабочего окна

После выбора источника данных модели необходимо получить и обработать сведения по информационным источникам. Основным критерием в логистике в настоящее время считается доставка груза в срок.

Для примера потребности в хранении больших данных был проведен анализ грузовых отправок на полигоне «Октябрьская железная дорога — филиала ОАО «РЖД» за январь 2024 года.

Демонстрацией работы методов накопления больших данных является графическое отображение показателей. Графический вывод, как и методы запроса информации, может осуществляться различными способами в зависимости от ИС. В целях практической демонстрации была выбрана графическая среда разработки `sql Developer`, данные взяты из нескольких моделей хранения данных, представленных на рис. 3, и получены запросом на языке `sql`. Пример — иллюстрация запроса и вывода информации по выполненным доставкам дан на рис. 4.

Наблюдается отсутствие информации по некоторым направлениям доставки, что требует дополнительного анализа. В ИС должны закладываться модели и методы для поиска причин нарушений при осуществлении грузоперевозки.

В этом исследовании предлагаются технологии `big data`, которые должны быть направлены на сбор и хранение информации, наиболее полезной для анализа и отображения, доступ к которой будет оперативно получен. Целью дальнейших предиктивных изысканий может быть поиск нарушений или других критериев, результатом которого является процесс запоминания алгоритма поиска, то есть интеллектуальное решение, направленное на уточнение методов. Еще одним примером потребности в методах больших данных является задача прогноза отправления вагона в ремонт.

Заключительным этапом является выдача сведений клиенту, включающих:

- 1) объекты со срывом срока доставки;
- 2) объекты с прогнозом срыва срока доставки и достоверностью свершения этого факта;
- 3) по досылкам дать полные сведения, включающие исходную накладную с ее реквизитами по отправлению;
- 4) сводные сведения по дроблению отправок за период (прибывшие/совершенные на полигоне дороги);
- 5) сводные сведения по срыву доставок за период (прибывшие/совершенные на полигоне дороги).

Блок исходных данных для реализуемых задач предполагает применение данных, выгружаемых согласно элементам бизнес-операций из смежных систем, в частности:

- Автоматизированная система оперативного управления перевозками (АСОУП);
- Автоматизированная система планирования, учета и анализа проведения окон и выполнения хозяйственной работы (АС АПВО (АС АПВО-2) по сведени-

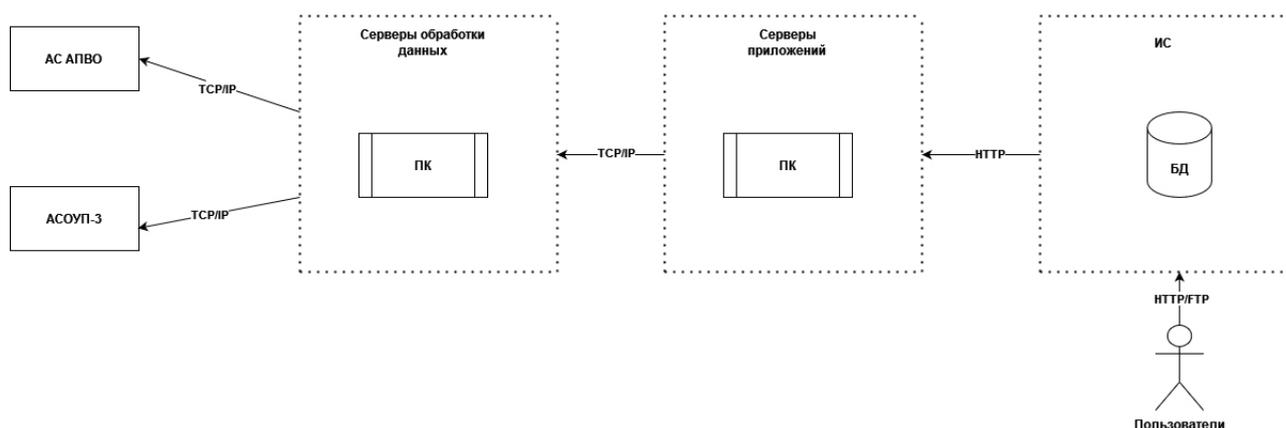


Рис. 5. Работа ИС

ям о числе ниток грузовых поездов по прогнозным графикам движения на выполнение ремонтно-путевых работ за год/месяц, а также ежесуточные оперативные изменения [17].

Пользовательские запросы в ИС обрабатываются и осуществляется поиск необходимых сведений в БД, которая накапливает информацию, поступающую с сервера приложения. Взаимодействие между ними осуществляется по http-протоколу. Сервер приложений получает обработанную информацию с сервера обработки данных по tcp/ip-протоколу. Сервер обработки, в свою очередь, запрашивает информацию из описанных АС также по tcp/ip-протоколу. Таким образом, клиент получает графическую информацию по запросу.

Дальнейшие расчеты по реализации системы доставки можно проводить при соответствующем программном обеспечении с использованием методик, изложенных в работах [18–24].

Блок-схема алгоритма получения необходимых сведений дана на рис. 5.

## Заключение

Выполненное исследование направлено на обоснование необходимости внедрения технологий big data для совершенствования перевозочного процесса по части хранения, накопления и обработки информации, а также к последующему анализу. Следующим этапом совершенствования транспортных технологических процессов должны быть выбор и проектирование имитационной модели архитектуры хранилища данных, которая в будущем может повлиять на дальнейшую интеграцию различных технологий и методов, которые в рамках данной статьи не рассматриваются.

Надо полагать, что интеграция и применение существующих технологий больших данных уже на этапе реорганизации взаимодействия позволят оптимизировать затраты на хранение данных посредством перехода к единой консолидированной структуре. Рассмотренный подход направлен на упрощение понимания структуры взаимодействия, а также на подготовку базы и моделей для последующей обработки.

## Библиографический список

1. Подхалюзина В. А. Анализ состояния железнодорожного транспорта в России // Экономика и социум. 2014. № 2–3 (11) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/analiz-sostoyaniya-zheleznodorozhnogo-transporta-v-rossii> (дата обращения: 17.01.2024).

2. Рогушина Ю. В. Разработка онтологической модели информационной потребности пользователя при семантическом поиске // Онтология проектирования. 2014. № 2 (12) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/razrabotka-ontologicheskoy-modeli-informatsionnoy-potrebnosti-polzovatelya-pri-semanticheskom-poiske> (дата обращения: 17.01.2024).

3. Власов А. И., Подорин А. А., Малеваный А. Ю. и др. Анализ визуальных моделей технологии больших данных при мониторинге перевозочного процесса на основе хранилища рейсов грузовых вагонов // Современные технологии. Системный анализ. Моделирование. 2020. № 3 (67) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/analiz-vizualnyh-modeley-tehnologii-bolshih-dannyh-pri-monitoringe-perevozochnogo-protsesa-na-osnove-hranilishcha-reysov-gruzovyh> (дата обращения: 17.01.2024).

4. Кравченко В. О., Крюкова А. А. Большие данные — практические аспекты и особенности // Academy. 2016. № 6 (9) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/bolshie-dannye-prakticheskie-aspekty-i-osobennosti> (дата обращения: 17.01.2024).

5. Москат Н. А. Методы повышения эффективности автоматизированной системы оперативного управления на железнодорожном транспорте // ИВД. 2018. № 1 (48) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/metody-povysheniya-effektivnosti-avtomatizirovannoy-sistemy-operativnogo-upravleniya-na-zheleznodorozhnom-transporte> (дата обращения: 15.01.2024).

6. Грошев Г. М., Климова Н. В., Суторовский А. В. и др. Автоматизация информационного обеспечения независимых участников мультимодальных перевозок контейнеров в морской порт в транспортном узле // Автоматика на транспорте. 2018. № 3 [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/avtomatizatsiya-informatsionnogo-obespecheniya-nezavisimyh-uchastnikov-multimodalnyh-perevozok-konteynerov-v-morskoj-port-v> (дата обращения: 17.01.2024).

7. Архипова Е. С. Роль обработки больших данных в управлении современным предприятием // Огарёв-Online. 2019. № 7 (128) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/rol-obrabotki-bolshih-dannyh-v-upravlenii-sovremennym-predpriyatiem> (дата обращения: 17.01.2024).

8. Маловецкая Е. В., Козловский А. П. Анализ моделей и принципов системного моделирования при построении прогнозных моделей погрузки грузов // International Journal of Open Information Technologies. 2020. № 12 [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/analiz-modeley-i-printsipov-sistemnogo-modelirovaniya-pri-postroenii-prognoznyh-modeley-pogruzki-gruzov> (дата обращения: 17.01.2024).

9. Назаренко Ю. Л. Обзор технологии «большие данные» (big data) и программно-аппаратных средств, применяемых для их анализа и обработки // European science. 2017. № 9 (31) [Электронный

ресурс]. URL: <https://cyberleninka.ru/article/n/obzor-tehnologii-bolshie-dannye-big-data-i-programmno-apparatnyh-sredstv-primenyaemyh-dlya-ih-analiza-i-obrabotki> (дата обращения: 17.01.2024).

10. Менщиков А. А., Перфильев В. Э., Федосенко М. Ю. и др. Основные проблемы использования больших данных в современных информационных системах // Столыпинский вестник. 2022. № 1 [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/osnovnyye-problemy-ispolzovaniya-bolshih-dannyh-v-sovremennyh-informatsionnyh-sistemah> (дата обращения: 17.01.2024).

11. Пугачев С. В., Хомоненко А. Д., Ярмолинский Ф. А. О разработке информационной системы грузоперевозок ОАО «РЖД» на основе безопасной интеграции приложений // Интеллектуальные технологии на транспорте. 2023. № 1 (33) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/o-razrabotke-informatsionnoy-sistemy-gruzoperevozok-oao-rzhd-na-osnove-bezopasnoy-integratsii-prilozheniy> (дата обращения: 17.01.2024).

12. Кравченко М. В., Никитин А. С., Спиридонов С. И. Об унификации обмена данными между разнородными средствами и системами в едином информационном пространстве // I-methods. 2020. № 2 [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/ob-unifikatsii-obmena-dannymi-mezhdu-raznorodnymi-sredstvami-i-sistemami-v-edinom-informatsionnom-prostranstve> (дата обращения: 17.01.2024).

13. Плясова С. В., Калинин А. Р., Зеленкина Е. В. Big data как объект оценки // Имущественные отношения в РФ. 2022. № 1 (244) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/bigdata-kak-obekt-otsenki> (дата обращения: 17.01.2024).

14. Акимов А. Е. Большие данные, искусственный интеллект и облачные технологии: цифровизация железных дорог // Инновации и инвестиции. 2023. № 3. [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/bolshie-dannye-iskusstvennyy-intellekt-i-oblachnye-tehnologii-tsifrovizatsiya-zheleznyh-dorog> (дата обращения: 17.01.2024).

15. Интерактивный автоинформатор для клиентов железнодорожного транспорта [Рукопись] : вып. квалиф. раб. ... степ. мгс. / Ф. А. Ярмолинский ; научный руководитель С. В. Пугачев ; рец. А. Н. Горкунов ; ФГБОУ ВО ПГУПС, кафедра ИВС. 2023. 147 с. 23 с.

16. Куренков П. В., Котляренко А. Ф. Внешнеторговые перевозки в смешанном сообщении: экономика, логистика, управление. Самара: СамГАПС, 2003. 636 с.

17. Куренков П. В. Материальные потоки в макрологистических системах: систематизация и классификация // Транспорт: наука, техника, управление. 2019. № 7. С. 21–26.

18. Куренков П. В., Давыдов С. В., Болгова Ю. С. Самарский центр консолидации грузопотоков в системе международных транспортных коридоров // Логистика сегодня. 2007. № 5. С. 312–322.

19. Куренков П. В., Багимов А. В. Взаимодействие отправителей и получателей каменного угля при экспортных перевозках в смешанном сообщении // Материалы международного научно-образовательного форума. Бургас: 2014. № 1 (5). С. 258–265.

20. Куренков П. В., Соловьева Л. В. Издержки взаимодействия компаний-операторов и ОАО «РЖД» // Материалы международного научно-образовательного форума. Бургас: 2014. № 1 (5). С. 266–275.

21. Куренков П. В., Соловьева Л. В. Логистические издержки взаимодействия компаний-операторов и ОАО «РЖД» // *Логистика*. 2014. № 4 (89). С. 24–27.

22. Куренков П. В., Солоп И. А., Чеботарева Е. А. и др. Оценка выполнения сроков доставки грузов на юге России // *Экономика железных дорог*. 2023. № 7. С. 13–25.

23. Солоп И. А., Чеботарева Е. А. Причинно-следственный анализ выполнения надежности доставки грузов железнодорожным транспортом в адрес потребителей Южного региона и портов Азово-Черноморского бассейна // *ИВД*. 2018. № 3 (50) [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/prichinno-sledstvennyy-analiz-vypolneniya-nadezhnosti-dostavki-gruzov-zheleznodorozhnym-transportom-v-adres-potrebiteley-yuzhnogo> (дата обращения: 17.01.2024).

24. Система управления окнами [Электронный ресурс]. URL: <https://niias.ru/products-and-services/products/asu/avtomatizirovannaya-sistema-planirovanie-uchet-i-analiz-provedeniya-okon-i-vypolneniya-khozyaystvenn/?ysclid=lrj4m2iauo361494077> (дата обращения: 17.01.2024).

Дата поступления: 29.01.2024

Решение о публикации: 01.03.2024

#### **Контактная информация:**

ЯРМОЛИНСКИЙ Федор Александрович — аспирант кафедры «Управление эксплуатационной работой» ФГБОУ ВО «Петербургский государственный университет путей сообщения Императора Александра I», Российская Федерация, 190031, Санкт-Петербург, Московский пр., 9, тел.: 8 (921) 766-06-53

ПОКРОВСКАЯ Оксана Дмитриевна — доктор технических наук, доцент, заведующий кафедрой «Управление эксплуатационной работой» ФГБОУ ВО «Петербургский государственный университет путей сообщения Императора Александра I», Россия, 190031, Санкт-Петербург, Московский пр., 9, тел.: 8 (965) 035-42-54

ПАСЕЧНИК Екатерина Дмитриевна — ассистент кафедры бизнес-информатики ФГБОУ ВО «Санкт-Петербургский государственный технологический институт (технический университет)», Россия, 190013, Россия, Санкт-Петербург, Московский пр., 24-26/49, литера А, тел.: 8 (911) 149-85-88

ПАКУЛИНА Елена Вячеславовна — ассистент кафедры «Наземные транспортно-технологические комплексы», ФГБОУ ВО «Петербургский государственный университет путей сообщения Императора Александра I», Россия, 190031, Санкт-Петербург, Московский пр., 9.

ТРАПЕЗНИКОВ Александр Александрович — преподаватель 1-й кафедры (организации повседневной деятельности и безопасности военной службы), Военный институт железнодорожных войск и военных сообщений ВА МТО им. генерала армии А. В. Хрулёва, [trapeznikova1979@mail.ru](mailto:trapeznikova1979@mail.ru)

## Features of the use of BIG DATA in the study of freight traffic on railway transport

F. A. Yarmolinsky<sup>1</sup>, O. D. Pokrovskaya<sup>1</sup>, E. D. Pasechnik<sup>1</sup>, E. V. Pakulina<sup>1</sup>,  
A. A. Trapeznikov<sup>2</sup>

<sup>1</sup> Emperor Alexander I St. Petersburg State Transport University, 9, Moskovsky ave., St. Petersburg, 190031, Russia

<sup>2</sup> Military Institute of Railway Troops and Military Communications of the VA MTO named after Army General A. V. Khrulev

**For citation:** Ermolinsky F.A., Pokrovskaya O. D., Pasechnik E. D. *i dr.* Features of the use of BIGDATA in the study of freight flows on railway transport // Bulletin of scientific research results. — 2024. — Iss. 1. — P. 107–122. (In Russian) DOI: 10.20295/2223-9987-2024-01-107-122

### Abstract

**Objective:** to characterize the features and prospects of using big data tools and technologies in the management of the transportation process on railways. **Methods:** neural network modeling, system analysis, forecasting, programming, big data, predictive analytics. **Results:** a datalogical model of entities for storing up-to-date data on cargo flows is proposed, and a structure for building a system for accumulating information is proposed. In addition, the paper examines the applied issues of solving the problems of storing, receiving and processing data using big data methods. **Practical significance:** Improving the management of railway transportation processes in the context of digital transformation in terms of obtaining more accurate forecasts.

**Keywords:** big data, cargo flows, predictive analytics, digitalization, forecast, neural network model.

### References

1. Podhalyuzina V. A. Analiz sostoyaniya zheleznodorozhnogo transporta v Rossii // *Ekonomika i socium*. 2014. No. 2–3 (11) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/analiz-sostoyaniya-zheleznodorozhnogo-transporta-v-rossii> (data obrashcheniya: 17.01.2024). (In Russian)
2. Rogushina Yu. V. Razrabotka ontologicheskoy modeli informacionnoj potrebnosti pol'zovatelya pri semanticheskom poiske // *Ontologiya proektirovaniya*. 2014. No. 2 (12) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/razrabotka-ontologicheskoy-modeli-informatsionnoj-potrebnosti-pol-zovatelya-pri-semanticheskom-poiske> (data obrashcheniya: 17.01.2024). (In Russian)
3. Vlasov A. I., Podorin A. A., Malevanyj A. Yu. i dr. Analiz vizual'nyh modelej tekhnologii bol'shih dannyh pri monitoringe perevoznogo processa na osnove hranilishcha rejsov gruzovyh vagonov // *Sovremennye tekhnologii. Sistemnyj analiz. Modelirovanie*. 2020. No. 3 (67) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/analiz-vizualnyh-modeley-tehnologii-bolshih-dannyh-pri-monitoringe-perevoznogo-protsesta-na-osnove-hranilishcha-reysov-gruzovyh> (data obrashcheniya: 17.01.2024). (In Russian)
4. Kravchenko V. O., Kryukova A. A. Bol'shie dannye — prakticheskie aspekty i osobennosti // *Academy*. 2016. No. 6 (9) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/bolshie-dannye-prakticheskie-aspekty-i-osobennosti> (data obrashcheniya: 17.01.2024). (In Russian)

5. Moskat N. A. Metody povysheniya effektivnosti avtomatizirovannoy sistemy operativnogo upravleniya na zheleznodorozhnom transporte // IVD. 2018. No. 1 (48) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/metody-povysheniya-effektivnosti-avtomatizirovannoy-sistemy-operativnogo-upravleniya-na-zheleznodorozhnom-transporte> (data obrashcheniya: 15.01.2024). (In Russian)
6. Groshev G. M., Klimova N. V., Sugorovskij A. V. i dr. Avtomatizatsiya informacionnogo obespecheniya nezavisimyh uchastnikov mul'timodal'nyh perevozok kontejnerov v morskoy port v transportnom uzle // Avtomatika na transporte. 2018. No. 3 [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/avtomatizatsiya-informatsionnogo-obespecheniya-nezavisimyh-uchastnikov-multimodalnyh-perevozok-konteynerov-v-morskoy-port-v> (data obrashcheniya: 17.01.2024). (In Russian)
7. Arhipova E. S. Rol' obrabotki bol'shikh dannyh v upravlenii sovremennym predpriyatiem // Ogaryov-Online. 2019. No. 7 (128) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/rol-obrabotki-bolshikh-dannyh-v-upravlenii-sovremennym-predpriyatiem> (data obrashcheniya: 17.01.2024). (In Russian)
8. Maloveckaya E. V., Kozlovskij A. P. Analiz modelej i principov sistemnogo modelirovaniya pri postroenii prognoznyh modelej pogruzki gruzov // International Journal of Open Information Technologies. 2020. No. 12 [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/analiz-modeley-i-printsipov-sistemnogo-modelirovaniya-pri-postroenii-prognoznyh-modeley-pogruzki-gruzov> (data obrashcheniya: 17.01.2024). (In Russian)
9. Nazarenko Yu. L. Obzor tekhnologii "bol'shie dannye" (big data) i programmno-apparatnyh sredstv, primenyaemyh dlya ih analiza i obrabotki // European science. 2017. No. 9 (31) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/obzor-tehnologii-bolshie-dannye-big-data-i-programmno-apparatnyh-sredstv-primenyaemyh-dlya-ih-analiza-i-obrabotki> (data obrashcheniya: 17.01.2024). (In Russian)
10. Menshchikov A. A., Perfil'ev V. E., Fedosenko M. Yu., et. al. Osnovnye problemy ispol'zovaniya bol'shikh dannyh v sovremennykh informacionnykh sistemah // Stolypinskij vestnik. 2022. No. 1 [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/osnovnye-problemy-ispolzovaniya-bolshikh-dannyh-v-sovremennykh-informatsionnykh-sistemah> (data obrashcheniya: 17.01.2024). (In Russian)
11. Pugachev S. V., Homonenko A. D., Yarmolinskij F. A. O razrabotke informacionnoj sistemy gruzoperevozok OAO "RZHD" na osnove bezopasnoj integracii prilozhenij // Intellektual'nye tekhnologii na transporte. 2023. No. 1 (33) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/o-razrabotke-informatsionnoy-sistemy-gruzoperevozok-oao-rzhd-na-osnove-bezopasnoj-integratsii-prilozhenij> (data obrashcheniya: 17.01.2024). (In Russian)
12. Kravchenko M. V., Nikitin A. S., Spiridonov S. I. Ob unifikacii obmena dannyimi mezhduraznorodnymi sredstvami i sistemami v edinom informacionnom prostranstve // I-methods. 2020. No. 2 [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/ob-unifikatsii-obmena-dannyimi-mezhduraznorodnymi-sredstvami-i-sistemami-v-edinom-informatsionnom-prostranstve> (data obrashcheniya: 17.01.2024). (In Russian)
13. Plyasova S. V., Kalinin A. R., Zelenkina E. V. Big data kak ob'ekt ocenki // Imushchestvennye otnosheniya v RF. 2022. No. 1 (244) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/bigdata-kak-obekt-otsenki> (data obrashcheniya: 17.01.2024). (In Russian)

14. Akimov A. E. Bol'shie dannye, iskusstvennyj intellekt i oblachnye tekhnologii: cifrovizatsiya zheleznyh dorog // Innovatsii i investitsii. 2023. No. 3. [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/bolshie-dannye-iskusstvennyy-intellekt-i-oblachnye-tehnologii-tsifrovizatsiya-zheleznyh-dorog> (data obrashcheniya: 17.01.2024). (In Russian)
15. Interaktivnyj avtoinformator dlya klientov zheleznodorozhnogo transporta [Rukopis'] : vyp. kvalif. rab. ... step. mgs. / F. A. Yarmolinskij ; nauchnyj rukovoditel' S. V. Pugachev ; rec. A.N. Gorkunov ; FGBOU VO PGUPS, kafedra IVS. 2023. 147 p. 23 p. (In Russian)
16. Kurenkov P. V., Kotlyarenko A. F. Vneshnetorgovye perevozki v smeshannom soobshchenii: ekonomika, logistika, upravlenie. Samara: SamGAPS, 2003. 636 p. (In Russian)
17. Kurenkov P. V. Material'nye potoki v makrologisticheskikh sistemah: sistematizatsiya i klassifikatsiya // Transport: nauka, tekhnika, upravlenie. 2019. No. 7. P. 21–26. (In Russian)
18. Kurenkov P. V., Davydov S. V., Bolgova YU. S. Camarskij centr konsolidatsii gruzopotokov v sisteme mezhdunarodnyh transportnyh koridorov // Logistika segodnya. 2007. No. 5. P. 312–322. (In Russian)
19. Kurenkov P. V., Bagimov A. V. Vzaimodejstvie otpravitelej i poluchatelej kamennogo uglya pri eksportnyh perevozkah v smeshannom soobshchenii // Materialy mezhdunarodnogo nauchno-obrazovatel'nogo foruma. Burgas: 2014. No. 1 (5). P. 258–265. (In Russian)
20. Kurenkov P. V., Solov'eva L. V. Izderzhki vzaimodejstviya kompanij-operatorov i OAO "RZHD" // Materialy mezhdunarodnogo nauchno-obrazovatel'nogo foruma. Burgas: 2014. No. 1 (5). P. 266–275. (In Russian)
21. Kurenkov P. V., Solov'eva L. V. Logisticheskie izderzhki vzaimodejstviya kompanij-operatorov i OAO "RZHD" // Logistika. 2014. No. 4 (89). P. 24–27. (In Russian)
22. Kurenkov P. V., Solop I. A., Chebotareva E. A., et. al. Ocenka vypolneniya srokov dostavki gruzov na yuge Rossii // Ekonomika zheleznyh dorog. 2023. No. 7. P. 13–25. (In Russian)
23. Solop I. A., Chebotareva E. A. Prichinno-sledstvennyj analiz vypolneniya nadezhnosti dostavki gruzov zheleznodorozhnym transportom v adres potrebitelej YUzhnogo regiona i portov Azovo-CHernomorskogo bassejna // IVD. 2018. No. 3 (50) [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/prichinno-sledstvennyy-analiz-vypolneniya-nadezhnosti-dostavki-gruzov-zheleznodorozhnym-transportom-v-adres-potrebiteley-yuzhnogo> (data obrashcheniya: 17.01.2024). (In Russian)
24. Sistema upravleniya oknami [Elektronnyj resurs]. URL: <https://niias.ru/products-and-services/products/asu/avtomatizirovannaya-sistema-planirovanie-uchet-i-analiz-provedeniya-okon-i-vypolneniya-khozyaystvenn/?ysclid=lrj4m2iauo361494077> (data obrashcheniya: 17.01.2024). (In Russian)

Received: 29.01.2024

Accepted: 01.03.2024

#### **Author's information:**

Fyodor A. YARMOLINSKY — Postgraduate student of the Department "Operational Work Management" of the Emperor Alexander I St. Petersburg State Transport University, 9, Moskovsky ave., St. Petersburg, 190031, Russia, tel.: 8 (921) 766-06-53

Oksana D. POKROVSKAYA — Doctor of Technical Sciences, Associate Professor, Head of the Department “Operational Work Management” of the Emperor Alexander I St. Petersburg State Transport University, 9, Moskovsky ave., St. Petersburg, 190031, Russia, tel.: 8-965-035-42-54

Ekaterina D. PASECHNIK — Assistant of the Department of Business Informatics, St. Petersburg State Institute of Technology (Technical University), litera A, 24-26/49, Moskovsky ave., St. Petersburg, 190013, Russia, tel.: 8911-149-85-88

Elena V. PAKULINA — Assistant of the Department of Land Transport and Technological Complexes of the Emperor Alexander I St. Petersburg State Transport University, 9, Moskovsky ave., St. Petersburg, 190031, Russia

Alexander A. TRAPEZNIKOV — lecturer of the 1st department (organization of daily activities and security of military service), Military Institute (Railway troops and military communications) of the of the Military Academy of Logistics named after General of the Army A. V. Khrulyov of the Ministry of Defense of the Russian Federation; trapeznikovaa1979@mail.ru