

УДК 004.853:629.4

ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ БОЛЬШИХ ДАННЫХ В ЭНЕРГЕТИКЕ ТЯГИ ПОЕЗДОВ



А. И. Давыдов
Доцент кафедры
«Информационная безопасность»
ФГБОУ ВО «Омский
государственный университет
путей сообщения», кандидат
технических наук, доцент
davydovai@bk.ru



С. О. Подгорная
Доцент кафедры
«Информатика и
компьютерная графика»
ФГБОУ ВО «Омский
государственный
университет путей
сообщения», кандидат
технических наук
ps.light@mail.ru



М. М. Соколов
Доцент кафедры
«Автоматика и
телемеханика»
ФГБОУ ВО «Омский
государственный
университет путей
сообщения», кандидат
технических наук, доцент
sokolovmm@mail.ru

А. И. Давыдов

Окончил Омский государственный университет путей сообщения. Область научных интересов связана с разработкой методов и алгоритмов построения систем поддержки принятия решений в области энергетики тяги поездов, а также методологии прогнозирования энергопотребления техническими системами.

М. М. Соколов

Окончил Омский государственный университет путей сообщения. Область научных интересов связана с моделированием электротехнических комплексов на железнодорожном транспорте с целью модернизации средств диагностирования устройств инфраструктуры и подвижного состава.

С. О. Подгорная

Окончила Омский государственный университет путей сообщения. Область научных интересов связана с исследованием методов анализа расхода топливно-энергетических ресурсов на тягу поездов.

Аннотация. Выполнен анализ различных подходов к использованию технологий анализа больших данных о расходе топливно-энергетических ресурсов (ТЭР) на тягу поездов для выявления скрытых закономерностей с целью выработки мероприятий по повышению энергетической эффективности одного из наиболее энергоемких сегментов энергетического баланса как железнодорожной отрасли, так и в целом промышленности страны.

Проанализированы основные методы и технологии анализа больших данных с целью их применения в рассматриваемой предметной области.

Ключевые слова: поддержка принятия решения, интеллектуальный анализ, расход энергии, анализ, прогнозирование.

Введение. Железнодорожный транспорт является одним из крупнейших потребителей топливно-энергетических ресурсов в России: на него приходится около 5 % всего вырабатываемого объема электрической энергии и около 11 % дизельного топлива. При этом значительная часть энергоресурсов (около 85 % электроэнергии и 70 %

дизельного топлива приходится на тягу поездов – работу по непосредственному выполнению перевозочного процесса).

В условиях необходимости поиска путей эффективного использования финансовых ресурсов, важнейшим направлением политики энергосбережения является качественный и всесторонний анализ данных о расходе энергии и факторах, оказывающих на него влияние с целью выявления элементов системы энергопотребления, на которые целесообразна выработка управляющих воздействий для сведения к минимуму случаев применения неоптимальных (с точки зрения энергетики) графиков движения поездов, а также использования неисправного подвижного состава. Упрощенная структура взаимосвязей в рассматриваемой предметной области приведена на рис. 1.

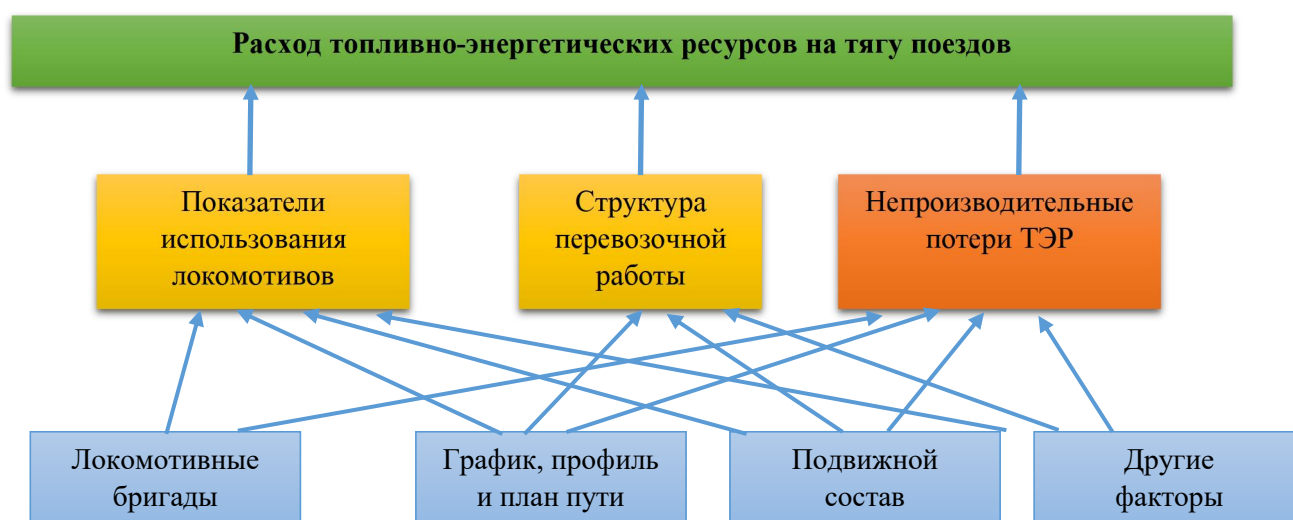


Рисунок 1. Структура взаимосвязей в системе управления энергопотреблением на тягу поездов

На рис. 1 к показателям использования локомотивов относится масса поезда, величина нагрузки на ось грузового вагона, техническая и участковая скорости, доля порожнего пробега грузовых вагонов; непроизводительные потери ТЭР – это простой подвижного состава в рабочем состоянии в ожидании работы, одиночное следование локомотивов, задержки поездов у запрещающих сигналов светофора, неграфиковые остановки, снижение скорости при проследовании мест с ограничением скорости и нагон опоздания поездов для вхождения в расписание. Под другими факторами подразумеваются прежде всего метеорологические факторы (температура атмосферного воздуха, ветер и т. д.), объем энергии рекуперации и применение различных энергосберегающих мероприятий и устройств.

Вся информация об указанных выше показателях (кроме метеофакторов) отражается в базе данных системы централизованной обработки маршрута машиниста (ЦОММ), автоматизированной системой управления «Нарушение безопасности движения», корпоративного информационного хранилища «Локомотивные парки» и др.

В настоящее время в информационных системах железнодорожного транспорта практически отсутствует аналитический аппарат [1-4]. Для решения указанных задач необходима зачастую трудоемкая ручная обработка в сторонних приложениях, что снижает оперативность принятия решений. Кроме того, в значительной части нормативных документов и методик применяется метод анализа от прошлого периода (месяца, квартала, года). При этом аналитику сложно сделать вывод – положительные или отрицательный эффект дали мероприятия этого года, простым снижением/увеличением

показателей здесь руководствоваться практически нереально ввиду специфики характера тягового энергопотребления.

Более 10 лет назад все железные дороги РФ перешли на единую систему ЦОММ, за этот период накоплен большой ретроспективный объем данных, которые при грамотном использовании инструментов интеллектуального анализа данных и технологий обработки больших данных могут дать возможность выявления новых закономерностей, что несомненно поможет сделать новый шаг в вопросах повышения энергетической эффективности тяги поездов [5-6].

Информация о расходе ТЭР на тягу поездов и влияющих на него факторах, обладает всеми необходимыми признаками bigdata. Понятие «большие данные» определяется через шесть основных свойств:

1 *Наличие большого объема* данных для того, чтобы обеспечить полноту и репрезентативность информации для реализации анализа энергопотребления на тягу.

2 *Исходные данные для анализа могут быть как структурированные, так и слабоструктурированные и неструктурированные* (именно эти данные имеют потенциал для поиска скрытых закономерностей между показателями энергетической эффективности).

3 *Данные должны быстро и регулярно накапливаться* (ежедневно на сети железных дорог совершается несколько десятков тысяч поездок, данные о которых ложатся в базу маршрутов машинистов, из которых в свою очередь формируется информация для вышележащих уровней детализации данных).

4 *Вариативность и комплексность данных по времени*, выражаемая в непостоянстве информационных потоков и гетерогенности их источников (поездки совершаются в различных условиях (погода, профиль и план пути, характеристики грузов и работы), разными локомотивными бригадами, на разном подвижном составе и т. д.).

5 *Достоверность* – все результаты поездок заносятся в базу данных на основании объективной информации – показаний приборов учета и других источников, в информационных системах для обеспечения достоверности данных используется набор специальных проверок и входного контроля.

6 *Ценность*: согласно теории и практики применения методов анализа данных для получения репрезентативных результатов исследования необходимы как можно большие объемы данных.

Среди методов и техник для анализа больших данных в области энергетики тяги поездов целесообразно применять следующие:

1 *Машинное обучение* – для формирования и исследования математических моделей для реализации инновационных технологий нормирования показателей энергопотребления на тягу поездов. Потребуется в первую очередь значительного увеличения ресурсов средств вычислительной техники, но это будет компенсировано оперативностью, достоверностью и точностью построенных решений и прогнозов.

2 *Построение частных нейросетевых моделей* для прогнозирования объемов энергопотребления и энергии рекуперации.

3 *Собственно технологии интеллектуального анализа данных* для добычи данных с целью их использования для прогнозирования, классификации, нейронных сетей, построения деревьев решений. Вызывают интерес такие задачи применения data mining как: анализ отклонений показателей, резко отличающиеся от генеральной совокупности (с целью выявления тяговых подвижных единиц для первоочередного направления на техническое обслуживание и ремонт, а также для определения локомотивных бригад, системно превышающих установленный нормативный расход ТЭР во всех поездках); кластеризация и классификация для распределения локомотивных бригад, локомотивов, подразделений по группам.

4 *Предиктивная аналитика* для решения задач планирования и прогнозирования энергопотребления и перевозочной работы на будущие периоды, выявления показателей, повлиявших на результат в уже прошедшем периоде.

5 *Построение имитационных моделей* для построения гипотез «что будет если изменить показатель X ?».

6 *Визуализация данных* – для возможности оперативного анализа полученных результатов.

Приведем пример первичной обработки данных о 3738 поездках в электрической тяге, в грузовом движении на одном из наиболее грузонапряженных участков Транссибирской магистрали. Была построена модель по алгоритму «случайный лес классификации». В качестве параметра разбиения была выбрана нагрузка на ось грузового вагона (именно этот показатель в настоящее время считается основным при оценке энергетического результата поездки), в качестве выходной константы – расход ТЭР.

Таблица 1. Результат работы алгоритма «случайный лес» в табличной форме

Узел 1	Узел 2	Количество точек, попавших в ветвь	Среднее значение нагрузки на ось, т	Среднее отклонение в узле по нагрузке на ось, т	Результат «Факт ТЭР» в ветви
2	3	786	17,89109	34,67570	3105
4	5	139	14,52374	49,84656	900
		7	1,60000	6,40000	
		132	15,20909	42,82355	
6	7	647	18,61453	28,45701	7835
8	9	619	18,49483	28,68136	7690
10	11	613	18,56166	28,08187	6205
		530	18,37698	27,83822	
12	13	83	19,74096	28,02917	6530
		18	22,42778	3,85201	
14	15	65	18,99692	32,17168	6760
		15	16,17333	36,57929	
		50	19,84400	27,74006	
		6	11,66667	42,84889	
		28	21,26071	16,17810	

На рис. 2 представлено дерево решений после применения алгоритма «случайный лес». Планируется проведение дальнейших исследований по классификации показателей энергопотребления на тягу поездов и влияющих на него факторов. Результат работы алгоритмов классификации может быть использован при разработке новых алгоритмов определения норм удельного расхода ТЭР на различных уровнях организационной структуры. В настоящее время установленные нормы достаточно субъективны, во многом зависят от «экспертного мнения» специалиста по теплотехнической работе. Реализация подобных алгоритмов позволит оперативно учитывать изменение показателей энергопотребления.

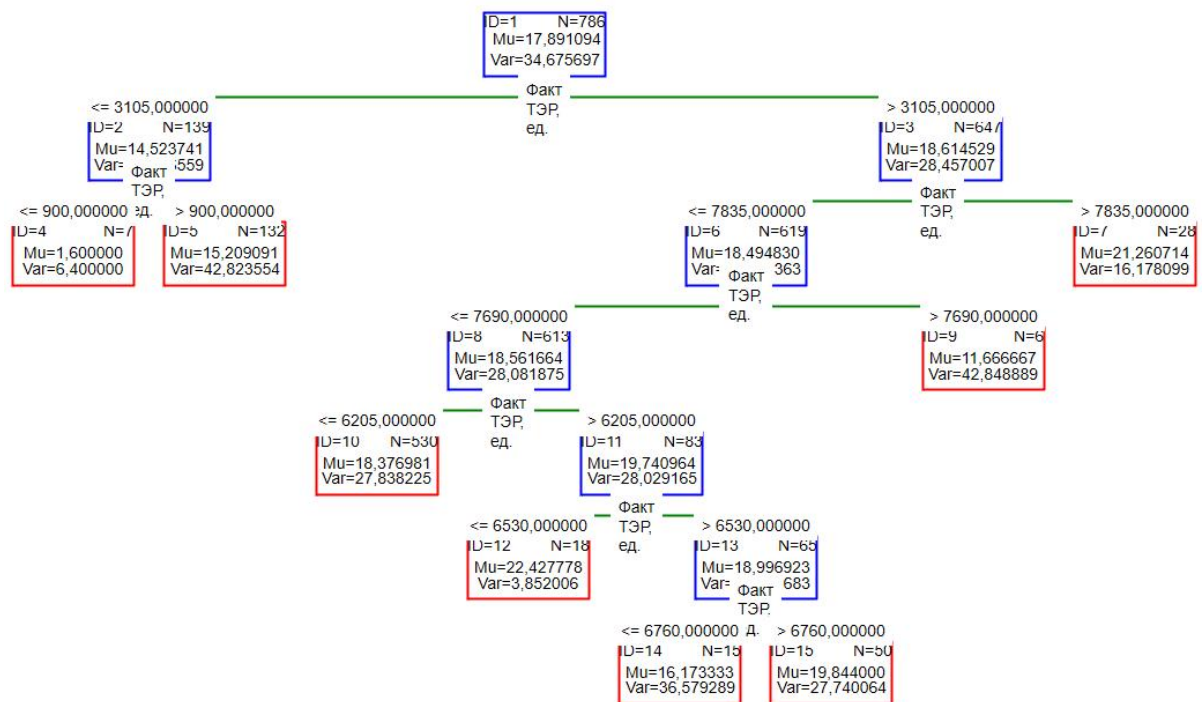


Рисунок 2. «Случайный лес» расхода ТЭР на тягу поездов по нагрузке на ось вагона

Исследование выполнено в рамках государственного задания № 109-03-2024-004 (научная тема ЕКТУ-2023-0003) по теме «Разработка концептуальной модели системы поддержки принятия решений в области анализа и прогнозирования расхода топливно-энергетических ресурсов на тягу поездов».

Список литературы

- [1] Davydov, A. Energy Efficiency Level Analysis of Train Traction in a Structural Unit / A. Davydov // Networked Control Systems for Connected and Automated Vehicles : Conference proceedings, St.Petersburg, 08–10 февраля 2022 года. Vol. 510-2. – Switzerland: Springer Nature Switzerland AG, 2023. – P. 437-446. – DOI 10.1007/978-3-031-11051-1_43. – EDN CNVGMY.
- [2] Сидорова, Е. А. Управление энергопотреблением на тягу поездов за счет повышения эффективности функционирования автоматизированной системы обработки информации об энергозатратах локомотивов / Е. А. Сидорова, С. П. Железняк, С. О. Подгорная. – Текст : непосредственный // Известия Транссиба. – 2018. – № 4 (36). – С. 143–154.
- [3] Подгорная, С. О. Прогнозирование показателя энергетической эффективности локомотивов / С. О. Подгорная, Е. А. Сидорова, А. И. Давыдов. – Текст : непосредственный // Известия Транссиба. – 2022. – № 2 (50). – С. 85–95.
- [4] Давыдов, А. И. Совершенствование информационного обеспечения анализа энергоэффективности тяги поездов / А. И. Давыдов // Научные революции: сущность и роль в развитии науки и техники : сборник статей по результатам Международной научно-практической конференции, Тюмень, 22 ноября 2022 года. – Sterlitaamak: Общество с ограниченной ответственностью "Агентство международных исследований", 2022. – С. 74-76. – EDN BQZNNQ.
- [5] Иванченко, В. И. Управление энергоэффективностью производственных процессов железнодорожных потребителей / В. И. Иванченко, А. А. Комяков // Известия Кыргызского государственного технического университета им. И. Раззакова. – 2023. – № 2(66). – С. 1139-1144. – DOI 10.56634/16948335.2023.2.1139-1144. – EDN XODVVD.
- [6] Комяков, А. А. Нормирование и прогнозирование расхода топливно-энергетических ресурсов в производственных процессах на железнодорожном транспорте с использованием информационных технологий / А. А. Комяков // Вестник Уральского государственного университета путей сообщения. – 2020. – № 2(46). – С. 35-43. – DOI 10.20291/2079-0392-2020-2-35-43. – EDN QUOQNG.

Авторский вклад

Давыдов Алексей Игоревич – руководство исследованием по поиску технологий применения методов интеллектуального анализа данных для их реализации в системах поддержки принятия решения в области энергетики тяги поездов, формирование структуры статьи. Исследование методов анализа больших данных для применения в области энергетики тяги поездов.

Подгорная Светлана Олеговна – описание структуры взаимосвязей в системе управления энергопотреблением на тягу поездов, исследование современной информационной базы для проведения анализа энергопотребления на тягу поездов.

Соколов Максим Михайлович – оценка соответствия элементов базы данных о расходе энергии на тягу поездов и факторах, оказывающих на него влияние, признакам больших данных. Проведение расчетов в программном обеспечении StatSoft Statistica.

BIG DATA MINING IN THE POWER ENGINEERING OF TRAIN TRACTION

A.I. Davydov

*Associate Professor of the
Department of Information
Security, Omsk State University of
Railway Engineering, Candidate of
Technical Sciences, Associate
Professor*

Podgornaya S. O.

*Associate Professor of the
Department of Computer Science
and Computer Graphics Omsk
State University of Railway
Engineering, Candidate of
Technical Sciences*

M. M. Sokolov

*Associate Professor of the
Department of Automation and
Telemechanics Omsk State
University of Railway Engineering,
Candidate of Technical Sciences,
Associate Professor*

Abstract. The analysis of various approaches to the use of big data analysis technologies on the consumption of fuel and energy resources for train traction has been carried out to identify hidden patterns in order to develop measures to improve the energy efficiency of one of the most energy-intensive segments of the energy balance of both the railway industry and the country's industry as a whole.

The main methods and technologies of big data analysis are analyzed in order to apply them in the subject area under consideration.

Keywords: decision support, intelligent analysis, energy consumption, analysis, forecasting.