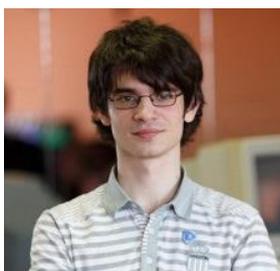


ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ BIG DATA В СФЕРЕ ТРАНСПОРТА



А.А. Александров
Аспирант кафедры информатики
БГУИР



И.И. Пилецкий
Научный руководитель совместной
лаборатории БГУИР-ИВА и АЦКТ
ИВМ, архитектор отделения по про-
граммному обеспечению ИВА IT Park,
доцент кафедры информатики
БГУИР, кандидат физико-матема-
тических наук

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
E-mail: zxybteyman@gmail.com.

Abstract. This paper describes some use cases of Big Data technologies for the sphere of transport, such as schedule planning, failure prediction of transport equipment based on diagnostic and vibration data, issues with transmitting of big amounts of data to remote datacenter using slow data transmitting channels.

Использование технологий Big Data в сфере транспорта позволяет решить различные не-тривиальные задачи в сфере транспорта.

Одной из проблем использования технологий Big Data на транспорте является проблема доставки больших данных в центры обработки данных. Если для городского транспорта, где покрытие мобильными сетями 3G приближается к 100%, проблем в отправке больших объёмов данных нет (данные можно накапливать на борту транспортного средства и передавать их в другое время, например, когда транспорт находится в парке), то для транспорта, такого как грузовые автомобили, поезда, грузовые самолёты, есть определенная проблема в доставке больших объёмов данных.

В местах недостаточного уровня сигнала для качественной передачи данных между базовой станцией оператора сотовой связи и установленным оборудованием, в блоке передачи предусмотрена «буферизация» (запись) данных на внутреннюю память. В момент появления устойчивой связи с качественным сигналом «буферизованные» данные передаются на сервер в автоматическом режиме.

Поскольку зачастую для достижения поставленных задач необходима высокая частота дискретизации снимаемых параметров (например, мировые координаты, моментальный расход топлива, состояние транспортного средства, состояние агрегатов транспортного средства и др.), существует проблема разработки сложных алгоритмов сжатия и регулярной доставки этой информации или, когда появляется стабильный канал связи.

Какую информацию можно собирать? В первую очередь это точные координаты с привязкой ко времени. На основании этой информации можно производить планирование и составление точных графиков и расписаний движения транспорта. В Минске в качестве пилотного проекта реализована система управления светофором на перекрестке улиц Козлова и Платонова на базе RFID-меток. Данную систему можно усовершенствовать, используя телеметрию с трамвая для упреждающего управления светофором. Система может подстроить ра-

боту светофора таким образом, чтобы минимизировать задержку трамвая на перекрестке и переходить к следующей фазе работы светофора сразу после прохода трамвая через перекресток.

В качестве передаваемых данных для локомотивов минимально должны быть: геокоординаты нахождения локомотива и точное время; серия, номер и депо приписки локомотива, на котором установлено оборудование; скорость перемещения локомотива.

Текущая парадигма ИТ систем, эксплуатируемых на железных дорогах – это *сопровождение процесса перевозок и эксплуатации железнодорожных объектов* [1].

Данные факты являются серьезной проблемой для создания среды интеллектуализации железных дорог, которая может быть выражена как: анализ спроса погрузки и доставки груза, анализ пропускной железнодорожной сети, определения маршрутов, составления расписания и регулирования графиков движения поездов, контроля состояния подвижного состава (локомотивов, вагонов) и диагностирование поломки подвижного состава, увеличения пропускной способности существующих линий (за счет точных знаний координат ПС, сокращения дистанции между ПС, увеличения скорости движения), решения задач транспортной логистики (в том числе и предоставление полной информации о грузе и его транспортировке грузовладельцам) и др.

Новая парадигма ИТ систем на железных дорогах – *интеллектуальные системы прогнозного типа, анализа, выбора и принятия оптимальных и упреждающих решений*, т.е. наряду с существующими технологиями, требуются интеллектуальные системы, предоставляющие полный цикл логистических услуг, обеспечивающие анализ и выбор оптимальных вариантов и решений опережающих события, безопасность транспортировки грузов и эксплуатации транспортных средств.

Основными факторами интеллектуализации железных дорог являются:

- технологическая оснащенность: идентификация подвижного состава, груза, точное знание геокоординат подвижного состава и груза на геополигоне дороги в режиме *real-time*;
- беспроводные системы мониторинга, позволяющие отслеживать актуальные координаты подвижного состава и грузов, находящихся в пути, и маршруты их следования;
- информация о техническом состоянии подвижного состава и ситуации сети железных дорог (на перегонах, на станциях) в режиме, близкому к *real-time*;
- интеграция взаимодействия компонент новых и уже существующих, обработка всех железнодорожных документов в центральных серверах, минимизируя обработку данных на линейном уровне;
- управленческие обоснованные решения на основе анализа больших и сверх больших объемов данных, поступающих со всей дороги – принятие оперативных решений в режиме *real-time*, а далее тактических и стратегических решений.

Так, сбор телеметрической информации текущего состояния двигателя позволяет принимать как оперативные, так и стратегические решения. Например, применительно к автобусу, такой информацией являются данные о моментальном расходе топлива, оборотах двигателя, текущая передача, нагрузка на ось и т.д. Обычно анализируют лишь данные ошибок двигателя, игнорируя походную информацию. Однако, при фиксации этих данных с высокой частотой дискретизации, можно провести анализ этих данных и выявить сложные режимы работы двигателя, оптимизировать конфигурацию двигателя для минимизации расхода топлива и обеспечения оптимального режима работы двигателя.

Также эти данные могут помочь для расчета текущего пассажиропотока на основании данных о текущей нагрузке на ось и выявлять высоконагруженные маршруты с дальнейшим принятием решений о корректировке графика движения и увеличения числа единиц подвижного состава на данном маршруте.

Малоизученной сферой применения Big Data и предиктивной аналитики на транспорте является предсказание неисправностей агрегатов и механизмов на основании данных объективного контроля и вибромониторинга. Собрав большую базу штампов работы агрегатов,

можно с высокой долей вероятности фиксировать аномалии работы двигателя и предугадывать выход из строя того или иного агрегата. Учитывая большие вычислительные мощности ЦОД и развитие технологий обработки такого рода данных можно получить информацию об отказе агрегата с точностью до движущейся детали. Объёмы сырых данных исчисляются гигабайтами в час, поэтому есть необходимость в разработке алгоритмов, позволяющих оптимизировать передаваемый поток сырых данных с датчиков таким образом, чтобы не потерять в достоверности передаваемых сэмплов и не потерять важную информацию при сжатии. Так, например, в 2015 только локомотивов на БелЖД было 920 единиц, при активном использовании только 70% прописного парка, передаче только минимальных данных о локомотиве (100 байт), каждые 10 секунд объем передаваемых данных будет не менее 556.041.600 байт в сутки. Здесь не учтен большой объём данных о работе агрегатов локомотива и расходе топлива, данные о составе поезда, данные о бригаде, а также обратно передаваемые данные на локомотив.

Новые технологии идентификации объектов, сбора и передачи информации позволяют уточнить традиционно решаемые задачи:

- *Уровня операционного управления*, которые требуют принятия правильных решений в режиме близкому к real-time.

- *Уровня текущего или тактического управления*, которые требуют поиска оптимальных вариантов решения часто возникающих типовых задач. Данные используются те же что были собраны ранее, на уровне операционного управления, но размещенные в базах данных и хранилищах железнодорожных систем.

- *Уровня стратегического управления*, которые требуют выработки планов деятельности на основе качественных прогнозов и различных видов анализа.

Проанализировав вышесказанное, можно выделить несколько направлений для исследований:

- Использование Big Data для решения проблемы составления расписаний и графиков движений транспорта (городского, дальнемагистрального, железнодорожного). Сюда же входит вопрос оптимизации расходов при эксплуатации транспорта с учетом рельефа местности (например, для грузовых составов поездов вес которых, может превышать 4000 тонн, весьма критичным вычисление текущей скорости состава с точностью до 0,1 км/ч, на поворотах, подъемах/спусках, линейных участках, перед светофором, что может позволить сэкономить топливо/электроэнергию на разгон и торможение до 10-15%).

- Исследование применения Big Data для предиктивной диагностики и мониторинга состояния транспорта и подвижного состава, разработка программно-аппаратных комплексов для предсказания выхода из строя тех или иных агрегатов, сбор и анализ больших массивов данных с физических датчиков вибрации, температуры и т.д., что позволит отказаться от плановых ремонтов ТО1, ТО2 и др., а перейти непосредственно к требуемому обслуживанию подвижного состава.

- Исследование методов передачи больших массивов данных в условиях узкого канала, методики сжатия с учетом специфики данных, минимизацию задержек от отправки данных до получения обратного ответа с рекомендациями о параметрах движения, техобслуживания и т.д.

В рамках программы исследования загруженности городских автобусов и перспективы замены их на гибридные электробусы было разработано устройство, устанавливаемое в моторный отсек автобуса и подключаемое к внутренней CAN-шине контроллера двигателя Mercedes ADM3. Устройство было предназначено для непрерывного сбора информации с высокой частотой о работе двигателя с привязкой к текущему времени. Информация записывалась на SD-карту, раз в неделю инженеры производили считывание информации и передачу её для дальнейшего анализа в лабораторию. Устройство «каталось» на протяжении нескольких недель на автобусном маршруте №18 г. Минска. Результатом анализа полученного массива данных явилось решение о целесообразности перехода на гибридные двигательные установки. Решение основывалось на анализе режимов работы двигателя в городских условиях и нагрузки

на двигатель, времени простоев на остановках, светофорах, динамики разгона-остановки и прочих факторов.

На данный момент ведется разработка прибора для сбора данных с транспортных средств, имеющий в себе канал связи на базе GSM-модема, привязку с текущим координатам посредством систем геопозиционирования GPS-ГЛОНАСС, при этом имеющим минимальные габариты и возможность работать в агрессивных условиях машинных отделений.

ЦОД лаборатории БГУИР-ИВА [2], может использоваться как база в пилот проектах для хранения и обработки данных с датчиков.

Наиболее перспективным направлением является применение сервисов платформы IBM Bluemix и когнитивного суперкомпьютера IBM Watson [3, 4], возможности которого позволяют совместить обработки больших объемов данных, получаемых с различных датчиков. IBM Bluemix предоставляет специальные модули для получения информации с IoT-устройств, к которым можно отнести разрабатываемый автономный прибор сбора данных.

Когнитивные системы, приложения и сервисы, аналитика (Watson) + IoT (Internet of Things) – Интернет вещей (например, автомобили, локомотивы, полигон дороги), позволяют с минимальными усилиями создать сеть взаимодействия M2M (Machine-to-Machine), что в будущем обеспечит переход к безлюдному производству.

Литература

[1]. Пилецкий И. И. «Один из методов построения и модернизации корпоративных приложений» Материалы конференции - “Software Engineering Conference (Russia) SEC(R) 2007”, Moscow, November 1-2, 2007.

[2]. И.И. Пилецкий и др. Виртуальная ИТ среда БГУИР для исследования Big Data и VCL, с. 21-32, BIG DATA and Predictive Analytics. Использование BIG DATA для оптимизации бизнеса и информационных технологий : сборник материалов междунар. науч.-практ. конф. / редкол. : М.П. Батура [и др.]. – Минск : БГУИР, 2015. – 220 с. ISBN 978-985-543-146-7. - С. 21-32.

[3]. What is Bluemix [Электронный ресурс] / IBM developerWorks. – 2015-2017. – Режим доступа: <https://www.ibm.com/developerworks/cloud/library/cl-bluemixfoundry/>. – Дата доступа: 20.03.2017.

[4]. IBM RCIS Watson Cloud Cognitive University [Электронный ресурс] / IBM Developer Works. – 2016-2017. – Режим доступа: <https://www.ibm.com/developerworks/community/groups/service/html/communitystart?communityUuid=bc004137-b64a-4378-ac02-2caf59c56c2a>. – Дата доступа: 20.03.2017.