ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ BIG DATA В ТРАНСПОРТНОЙ СФЕРЕ

Александров А.А., Пилецкий И.И.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Пилецкий И.И. – канд. физ.-мат. наук, доцент

Данная работа содержит сведения о некоторых аспектах использования технологий Big Data в транспортной сфере, таких как предиктивный анализ поломок транспортных средств, вопросы доставки больших массивов данных с подвижного состава. Рассмотрены некоторые варианты направлений для дальнейшего исследования применения технологий Big Data.

В настоящее время происходит повсеместное внедрение и использование технологий Big Data в различных сферах жизни, связанных с появлением технологических возможностей анализировать огромные массивы данных в некоторых проблемных областях. Эти технологии не обошли стороной и транспортную сферу.

Одной из проблем использования технологий Big Data в сфере транспорта является проблема доставки больших массивов информации в центры обработки данных. Если для городского транспорта, где покрытие сетями мобильных сотовых операторов 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 такого транспорта, как грузовые автомобили, поезда, грузовые самолёты, есть определенная проблема в доставке больших объёмов данных. Одним из вариантов решения проблемы передачи данных между базовой станцией оператора сотовой связи и установленным оборудованием в условиях плохого покрытия, в блоке передачи можно предусмотреть «буферизацию» (запись) данных во внутреннюю память. В момент появления устойчивой связи с качественным сигналом «буферизованные» данные передаются на сервер в автоматическом режиме.

Поскольку для достижения заданной точности необходима высокая частота дискретизации снимаемых параметров (таких, как 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. Как пример, в Минске в качестве пилотного проекта была 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 работу светофора таким образом, чтобы минимизировать задержку трамвая на перекрестке и переходить к следующей фазе работы светофора сразу после прохода трамвая через перекресток.

Для локомотивов минимально необходимый набор передаваемых данных следующий: гео-координаты нахождения локомотива и точное время; серия, номер и депо приписки локомотива, на котором установлено оборудование; скорость перемещения локомотива.

Проблемы, которые могут быть решены при внедрении технологий Big Data для системы железнодорожного транспорта, следующие: анализ спроса погрузки и доставки груза, анализ пропускной железнодорожной сети, определение маршрутов, составление расписания и регулирование графиков движения поездов, контроль состояния подвижного состава (ПС): локомотивов, вагонов, и диагностирование поломок ПС, увеличение пропускной способности существующих линий за счет точных знаний координат ПС, сокращение дистанции между ПС, увеличение скорости движения, решение задач транспортной логистики (в том числе и предоставление полной информации о грузе и его транспортировке грузовладельцам) и др. [1].

Сбор телеметрической информации текущего состояния двигателя позволяет принимать как оперативные, так и стратегические решения. Применительно к автобусному транспорту, 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 выборок 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 обьем 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.

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 0,1км/ч, на поворотах, подьемах/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 гео-позиционирования GPS-ГЛОНАСС, при этом имеющим минимальные габариты и возможность работать в агрессивных условиях машинных отделений. ЦОД лаборатории БГУИР-IBA [2], может использоваться как база в пилот проектах для хранения и обработки данных с датчиков.

Наиболее перспективным направлением является применение сервисов платформы IBM Cloud Platform и когнитивного суперкомпьютера IBM Watson [3, 4], возможности которого позволяют совместить обработки больших объёмов данных, получаемых с различных датчиков. IBM Cloud Platform 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 (например, автомобили, локомотивы, полигон дороги), позволят с минимальными усилиями создать сеть взаимодействия M2M (Machine-to-Machine), что в будущем обеспечит переход к безлюдному производству.

Список использованных источников:

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 the IBM Cloud platform? [Электронный ресурс] / IBM developerWorks. – 2017-2018. – Режим доступа: https://console.bluemix.net/docs/overview/ibm-cloud-platform.html#whatis. – Дата доступа: 19.03.2019.
4. IBM RCIS Watson Cloud Cognitive University [Электронный ресурс] / IBM Developer Works. – 2016-2019. – Режим доступа: https://www.ibm.com/developerworks/community/groups/service/html/communitystart?communityUuid=bc004137-b64a-4378-ac02-2caf59c56c2a. – Дата доступа: 19.03.2019.

ПРОБЛЕМЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ОС WINDOWS

Андреюк Д. С., Кадушко А. А., Малиновская Е. Д.

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Стройникова Е. Д. – ассистент кафедры информатики

В процессе изучения предметной области была собрана и систематизирована информация о вредоносном программном обеспечении и информационных атаках. В результате была смоделирована и протестирована вредоносная программа, сокрытая в обыденном для обычного пользователя приложении.

В работе были изучены уязвимости ОС Windows относительно сохранности конфиденциальных и личных данных. Основной акцент был сделан на деятельности вредоносного программного обеспечения на примере вирусов, червей и троянов. Была приведена статистическая информация о деятельности вредоносного программного обеспечения с момента их возникновения. Продемонстрированы прецеденты наиболее масштабных и результативных вирусных атак за XX – XXI вв., таких как WannaCrypt0r, Jerusalem, CIH и подобных. Выведены критерии сравнения и классификации разного рода атак, изучаются их предпосылки, реализация, результаты и специфика. Рассмотрены особые беспрецедентные случаи осуществления атак (например, с использованием «умной» бытовой техники). Изучена также возможность осуществления DDoS-атак в социально-приемлемых целях. В процессе рассмотрения деятельности вредоносного программного обеспечения сделан упор на механизмы шифрования и распространения в прогностических интересах. Изучены способы сокрытия признаков деятельности вредоносного программного обеспечения (встраивание в код полезной программы, маскировка под неё и др.) с целью выявления специфики вредоносного программного обеспечения и планирования контрмер в виде разработки антивирусного программного обеспечения. В целом подход к изучению вредоносного ПО в данной работе использует концепцию «Знай врага в лицо».

В практической части продемонстрирована работа искусственно созданной программы, деятельность которой в не тестовом случае однозначно нежелательна. Пример представляет собой игру, запуск которой приводит к шифрованию файлов на жестком диске шифром Цезаря. Продемонстрирован механизм сокрытия, проникновения и шифрования самодельного трояна. Рассмотрены различные шифры (Цезаря, Виженера, многоалфавитная замена и др.), методы поиска ключей и дешифрования (на основе шифротекста, открытого текста и др.). Произведена классификация шифров, их сравнение, рассмотрено их использование в предметной деятельности. Изучены возможные исходы деятельности вредоносного программного обеспечения, способы восстановить утраченные или поврежденные данные. Предложены возможные меры профилактики утраты конфиденциальных данных. Спрогнозирована дальнейшая деятельность вредоносного программного обеспечения. Осуществлена рефлексия последующих цифровых эпидемий и их возможного исхода.

Список использованных источников:

1. WannaCry 2.0: наглядное подтверждение того, что вам обязательно нужно правильное решение для надежного бэкапа [Электронный ресурс]. – 2017. – Режим доступа: https://habr.com/ru/company/acronis/blog/328796/.
2. Холодильник атакует: как киберпреступники используют бытовую технику [Электронный ресурс]. – 2016. – Режим доступа: https://www.rbc.ru/technology_and_media/13/11/2016/5825cf889a79475b671ff971/.
3. Вредоносное ПО, вошедшее в историю. Часть II [Электронный ресурс]. – 2017. – Режим доступа: https://habr.com/ru/company/ua-hosting/blog/407621/.
4. Азбука безопасности [Электронный ресурс]. – 1998. – Режим доступа: https://kaspersky.antivirus.lv/rus/threats/safetyabc/.
5. Вирусы XXI века [Электронный ресурс]. – 2016. – Режим доступа: https://geekbrains.ru/posts/xxi_viruses/.
6. Самые разрушительные компьютерные вирусы начала XXI века [Электронный ресурс]. – 2018. – Режим доступа: https://www.iguides.ru/main/other/samye_razrushitelnye_kompyuternye_virusy_nachala_xxi_veka/.