

## ТЕХНИЧЕСКИЕ НАУКИ



УДК 004

### Анализ Big Data с использованием периферийной аналитики

*К.К. Казарян*

Донской государственный технический университет (г. Ростов-на-Дону, Российская Федерация)

**Аннотация.** Проанализирована теория совершенствования обработки данных в режиме реального времени. Рассмотрена постановка задачи построения аналитических данных и передачи оператору только полезных данных. Поставлены проблемы ограниченного пространства в хранении больших объемов данных и определены пути их возможного решения. В частности, предложено решение проблемы, связанной с обработкой данных в режиме реального времени, с применением периферийной аналитики.

**Ключевые слова:** IT, apple, samsung, oppo, huawei, смартфоны, техника, защита, Edge Analytics, Edge, microsoft, windows.

### Big Data Analysis with Edge Analytics

*Katorina K Kazaryan*

Don State Technical University (Rostov-on-Don, Russian Federation)

**Abstract.** The paper analyzes the theory of improving data processing in real time. The formulation of the problem of constructing analytical data and transmitting only useful data to the operator is considered. The problems of limited space in storing large amounts of data are posed and the ways of their possible solution are determined. A solution to the problem associated with real-time data processing using peripheral analytics is proposed.

**Keywords:** IT, apple, samsung, oppo, huawei, smartphones, appliances, protection, Edge Analytics, Edge, microsoft, windows.

**Введение.** В настоящее время все больше компаний сталкиваются с необходимостью регулярно проводить обработку поступающих и имеющихся у них данных и анализировать их в режиме реального времени. Для этой цели они используют «Интернет вещей», который является одним из главных факторов, вызывающих существенное изменение наиболее распространенных методов работы в данной области, поскольку потоки данных с датчиков требуют немедленной обработки и анализа для управления соответствующими системами. Цель данной работы — создание алгоритма анализа Big Data с использованием периферийной аналитики.

**Основная часть.** Для использования аналитики в реальном времени необходимо понимать аспекты ее проектирования в периферийных устройствах [1]. Причинами для использования инфраструктуры с пограничной аналитикой могут быть технические аспекты, затраты и необходимость соблюдения конфиденциальности данных.

Случаи надлежащего использования периферийной аналитики — это приложения, которые влияют на безопасность человека и требуют отказоустойчивости вычислительной архитектуры.

Приложения, требующие малой задержки между источниками данных, такие как датчики IoT и инфраструктура вычислительной аналитики, также требуют периферийной аналитики [2].

Еще одна причина для внедрения периферийной аналитики — соответствие требованиям и управление данными. Использование локальной инфраструктуры может помочь в соблюдении Общего регламента по защите данных (GDPR) и других правил за счет хранения и обработки ограниченных данных в странах, где они были получены.

Многие разработчики и специалисты по обработке и анализу данных в настоящее время пользуются преимуществами аналитических платформ более высокого уровня, доступных в облачных серверах.

Когда специалисты по данным работают в облачных платформах и могут получать больше вычислительных ресурсов на запрос по относительно низкой цене, они могут создавать сложные модели машинного обучения с множеством функций и параметров для полной оптимизации результатов.

Однако при развертывании моделей в инфраструктуре периферийных вычислений очень сложный алгоритм может значительно увеличить затраты на инфраструктуру, размер устройства и требования к питанию. К тому же развертывание высоконадежной и безопасной системы периферийной аналитики требует разработки и реализации отказоустойчивых архитектур, систем, сетей, программного обеспечения и моделей.

Несмотря на всю доступность оптимизации оборудования, профилактического обслуживания, контроля качества и критических предупреждений в среде граничных вычислений появляются другие проблемы, к примеру ограниченное пространство на оборудовании, физическая доступность, пропускная способность и более серьезные проблемы с безопасностью.

Это означает, что инфраструктура, к которой пользователь привык в своем центре обработки данных, может оказаться неработоспособным решением. Поэтому необходимо изучить новые технологии, разработанные с учетом архитектур периферийных вычислений.

В настоящее время наиболее распространенными в сфере периферийной аналитики являются функции обработки данных, в том числе фильтрация и агрегация данных [3].

Пограничные вычисления открывают следующие перспективы для интеллектуальных решений: цены на датчики падают, приложениям требуется больше аналитики в реальном времени, а сложность разработки оптимизированных финансовых и экономических алгоритмов для этой среды снижается.

С периферийной аналитикой связано несколько различных парадигм из области аналитики, машинного обучения и граничных вычислений [4]:

1. EdgeAnalytics. Относится к алгоритмам аналитики и машинного обучения, развернутым во внеоблачной инфраструктуре, известной как периферия. Эта инфраструктура географически распределена.

2. Непрерывная (потокковая) аналитика. Относится к анализу, выполняемому в режиме реального времени на основе данных, которые непрерывно обрабатываются.

3. Обработка событий. Способ обработки данных и принятия решений в режиме реального времени.

Рассмотрим создание алгоритма анализа Big Data с использованием периферийной аналитики в сфере образования. На сегодняшний день этот анализ имеет ряд недостатков: неполнота и сложность обработки данных, невозможность создания персонализированных образовательных программ без подробной информации о поведении студентов. Для устранения этих недостатков и был разработан алгоритм анализа Big Data с использованием периферийной аналитики, который состоит из следующих шагов:

1. Первый и самый важный шаг — определение конкретных целей анализа Big Data в образовании.

Например, можно поставить следующие цели:

- оптимизация процессов обучения для повышения эффективности образовательных программ;
- выявление проблемных областей в обучении и установление причин неуспеваемости студентов;
- повышение качества обратной связи между студентами и преподавателями;
- разработка персонализированных образовательных программ на основе индивидуальных потребностей студентов.

2. Разработка системы сбора данных. Для этого могут быть использованы различные периферийные устройства, такие как носимые устройства, датчики движения, камеры видеонаблюдения и т. д. Система должна быть разработана таким образом, чтобы собирать данные, которые необходимы для достижения поставленных целей.

3. После проектирования системы сбора данных нужно начать сбор и хранение данных. Для этого необходимо выбрать специализированные инструменты, такие как базы данных или хранилища данных. Важно убедиться, что система сбора и хранения данных обеспечивает их полную конфиденциальность и безопасность.

4. Далее необходимо обработать и проанализировать собранные данные. Для анализа больших объемов данных могут использоваться различные инструменты и технологии, такие как Hadoop, Spark и другие. Анализ может проводиться с помощью статистических методов, машинного обучения и др. Главная цель — выявление важных закономерностей и тенденций, которые могут быть использованы для улучшения процессов обучения.

5. После того как данные обработаны и проанализированы, можно приступать к использованию результатов анализа для улучшения процессов обучения. Например, можно определить наиболее эффективные методы обучения и разработать персонализированные образовательные программы для студентов.

6. Завершающий этап — оценка результатов внедрения системы анализа Big Data в образовании с помощью периферийной аналитики. Она позволит определить, насколько эффективно были решены поставленные задачи и какие изменения необходимо будет внести в будущем.

Таким образом, можно сделать вывод, что алгоритм внедрения анализа Big Data с помощью периферийной аналитики в образовательной сфере позволяет оптимизировать процесс обучения и создавать персонализированные образовательные программы на основе индивидуальных потребностей студентов.

**Заключение (выводы).** Разработан алгоритм анализа Big Data с использованием периферийной аналитики. Он позволяет проводить анализ данных на месте хранения данных без необходимости перемещения их на центральный сервер, что, в свою очередь, сократит время анализа и обработки результатов, а также снизит затраты на передачу данных. Результаты анализа могут быть визуализированы и использованы для принятия решений по оптимизации бизнес-процессов.

#### Библиографический список

1. Мокеев А.В. *Применение технологии СВИ для выполнения функций управления, защиты и автоматизации*. Цифровая подстанция. URL: <https://digitalsubstation.com/blog/2018/03/05/primenenie-tehnologii-sinhronizirovannyh-vektornyh-izmerenij-svi-dlya-vypolneniya-funksij-upravleniya-zashhity-i-nbsp-avtomatiki/> (дата обращения: 01.05.2023).
2. *Интеллектуальные информационные технологии*. Сайт факультета ПММ ВГУ. URL: <http://www.amm.vsu.ru/vmpit/ИТ.html> (дата обращения: 01.05.2023).
3. Шмаков Р. *Периферийные вычисления как неотъемлемая часть прогресса*. IT Week. URL: <https://www.itweek.ru/iot/article/detail.php?ID=222471> (дата обращения: 01.05.2023).
4. Джонсон Йохан. *AIoT и Edge-аналитика: мощная комбинация*. IT Week. URL: <https://www.itweek.ru/iot/article/detail.php?ID=221849> (дата обращения: 01.05.2023).

*Об авторе:*

**Казарян Каторина Камоевна**, студент кафедры «Вычислительные системы и информационная безопасность» Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), [kazaryan\\_katorina@mail.ru](mailto:kazaryan_katorina@mail.ru)

*About the Author:*

**Katorina K Kazaryan**, student of the Computer Systems and Information Security Department, Don State Technical University (1, Gagarin Sq., Rostov-on-Don, 344003, RF), [kazaryan\\_katorina@mail.ru](mailto:kazaryan_katorina@mail.ru)