

через протокол Bluetooth Low Energy с частотой дискретизации до 500 Гц, однако наиболее часто используемыми являются 100 и 250 Гц. На устройства PIQ наложены определённые ограничения по производительности, размеру аккумуляторов, низкому энергопотреблению, габаритам.

Так как пропускная способность протокола BLE довольно низкая, потеря данных при частоте даже 100 Гц во время их передачи на другие, более мощные устройства, достаточно вероятно, соответственно, качество обработки из-за потерь данных снижается. Поэтому необходимы алгоритмы, способные выполняться на самих сенсорных устройствах. Как следствие, устройство должно передавать уже обработанные данные, объём которых значительно меньше. Все алгоритмы разрабатываются с учётом данного ограничения.

Разработка методов и алгоритмов обработки данных инерциальных сенсоров проводится для устройства PIQ. Эти методы и алгоритмы должны преобразовывать и интерпретировать сенсорные данные в корректную и полезную для пользователей информацию. Разрабатываемые алгоритмы обработки данных включают в себя: определение промежутков данных, в период которых происходила какая-либо активность; вычисление кинематики движения и других параметров, интересных для пользователя-спортсмена; машинное обучение и распознавание движений на основе сенсорных данных.

Режим обработки данных при непосредственном их получении от датчиков необходим в связи с малым объёмом оперативной памяти и невозможностью хранить в ней большие массивы данных. Алгоритмы, обрабатывающие большие массивы данных и способные повторно использовать для вычислений уже полученные ранее сенсорные данные, проще и эффективнее. Однако эти алгоритмы не смогут быть применены на исследуемых сенсорных устройствах. Проблема потоковой обработки данных решается с помощью использования буфера данных минимальных размеров (например, для алгоритмов сглаживания достаточно буфера данных размером 5-11 точек), а также путём применения потоковых формул для подсчёта статистических данных (к примеру, таких, как среднее арифметическое, среднее квадратичное отклонение и дисперсия). Выявление активностей во время получения данных производится с использованием конечного автомата с состояниями, которые определяют начало, конец активности и некоторые другие состояния.

Тактовая частота микроконтроллеров сенсорных устройств довольно низкая. Поэтому необходимо разрабатывать достаточно простые по вычислениям методы и алгоритмы обработки данных, чтобы скорость работы алгоритма не уменьшалась из-за меньшей тактовой частоты микроконтроллера по сравнению с более мощными компьютерами. Однако при этом необходимо учитывать и качество обработки.

Фильтрация сенсорных данных, избавление от постоянной составляющей и внешних искажающих факторов в данных остаётся открытым вопросом и одним из важных пунктов в исследовании. На данный момент один этап фильтрации производится на сенсорах – применяется аппаратный фильтр низких частот. Также осуществляется калибровка датчиков, она основана на постоянной проверке состояния сенсора и в момент, когда сенсор не двигается, вычисляются постоянные составляющие данных и удаляются. Однако, как показывают результаты тестов, этого недостаточно для корректных вычислений кинематики движения и других пользовательских параметров.

Таким образом, при разработке алгоритмов обработки сенсорных данных необходимо предложить их простую реализацию, способную выполняться на сенсорных устройствах, с обеспечением режима потоковой обработки данных и достаточной точности предоставляемых потребителю результатов.

Список использованных источников:

1. PIQ – Homepage [Электронный ресурс]. – 2016. – Режим доступа: www.piq.com

АРХИТЕКТУРА ДЕЦЕНТРАЛИЗОВАННОЙ РАСПРЕДЕЛЁННОЙ ПОДСИСТЕМЫ ХРАНЕНИЯ ДАННЫХ В СИСТЕМАХ ПРИНЯТИЯ РЕШЕНИЙ

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

Скачихин А.А.

Бранцевич П.Ю. – к.т.н., доцент

Увеличение объёмов данных и сложности алгоритмов в системах принятия решений неизбежно ведут к распределённому хранению и обработке, в силу того, что ограниченные возможности одного компьютера не позволяют решать поставленные задачи за приемлемое время. Децентрализованная распределённая система хранения информации, учитывающая особенности предметной области, может дать ряд преимуществ по сравнению с реляционными базами данных, а также не специализированными распределёнными системами хранения, архитектурные решения которых направлены на максимальную универсальность.

Рассмотрим сценарий использования системы принятия решений на примере системы оценки технического состояния механизмов на основе вибрационных параметров. Данные с виброизмерительных датчиков падают в подсистему хранения и сбора информации, где, в зависимости от поставленных задач,

могут использоваться для принятия решений как в реальном времени, так и для обработки и статистического анализа в дальнейшем. Длительное хранение представляет наибольший интерес в рамках этой работы. На основе собранных данных подсистемы обработки данных выводят новые знания, которые в дальнейшем используются для интеллектуального анализа. При этом стоит отметить, что обработка данных возможна без их изменения. Второй важный аспект работы таких систем – отсутствие конкуренция на запись, ведь каждый датчик является отдельным источником данных. Третье, данные хранятся в неструктурированном виде, как последовательность измерений, что даёт возможность эффективно хранить их в бинарном виде.

Современные универсальные системы хранения информации, оказываются неоправданно сложными для задач хранения вибрационных сигналов. Неиспользуемыми остаются такие возможности как транзакционность, возможность задать строгую схему, механизмы блокировок при конкурентной записи. В тоже время универсальность существующих решений не позволяет применять эффективные методы сжатия, такие как дельта компрессия, а методы доступа к данным не могут быть оптимизированными учитывая особенности алгоритмов, которые в последствии будут обрабатывать полученные данные.

Вышеизложенные факты лежат в основе гипотезы, что, применив специализированную децентрализованную распределённую подсистему хранения данных, можно добиться повышения эффективности работы системы принятия решений. А именно, обеспечить:

- 1) уменьшение времени доступа благодаря эффективному двоичному хранению на диске блоками, соответствующими размеру окна алгоритмов;
- 2) уменьшение времени доступа благодаря отказу от транзакций и блокировок на операции чтения;
- 3) уменьшение объёмов занимаемого пространства на диске за счет применения методов сжатия однородных двоичных данных;
- 4) отказоустойчивость системы, гарантируемую путем репликации данных, на основе конфигурируемого признака (например, времени поступления данных);
- 5) децентрализацию, которая ведёт к тому, что запись в систему хранения может происходить полностью параллельно без участия управляющей стороны [1].

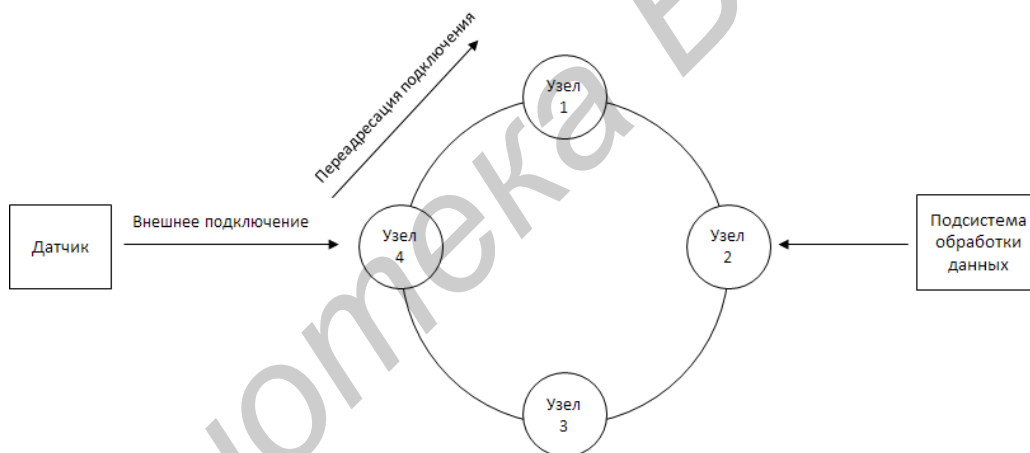


Рис. 1 – Упрощенная архитектура децентрализованной распределённой подсистемы хранения

Выделим следующие ключевые особенности архитектуры (рис.1):

- 1) количество узлов в подсистеме не является фиксированным и может быть изменено в соответствии с решаемой задачей;
- 2) так как подсистема является децентрализованной, датчики и подсистема обработки данных могут подключаться к любому узлу системы. Если необходимые данные не найдены на узле, то узел связывается с узлом, на котором данные располагаются. В случае с датчиками узел может принять решение переадресовать подключение на другой узел чтобы устранить необходимость передачи данных по сети. Подключение подсистемы обработки данных никогда не переадресовывается, так как в большинстве сценариев обращения на чтение требуют доступа к данным на нескольких узлах;
- 3) при старте подсистемы узлы обмениваются информацией о том, по какому принципу происходит репликация и сегментирование данных;
- 4) узлы, по возможности, кэшируют данные в оперативной памяти для ускорения доступа;
- 5) произвольный доступ к данным не возможен, это позволяет использовать дельта компрессию и существенно экономить дисковое пространство.

У такой реализации также есть ряд недостатков по сравнению с более универсальными аналогами:

- 1) узкая специализация может привести к тому, что придётся комбинировать данные из различных источников в промежуточное представление, что окажет существенное влияние на время обработки данных [2];
- 2) нет возможности вносить изменения в существующие данные в результате их обработки, что в ряде случаев будет предпочтительнее, чем хранение результатов отдельно, потребляя дополнительное дисковое пространство. Однако, для большинства сценариев использования изменение исходных данных представляется менее предпочтительным.

Таким образом, предлагается архитектура подсистемы хранения в системах принятия решений, позволяющая улучшить количественные характеристики системы путём проектирования с учётом особенностей предметной области.

Список использованных источников:

1. Zaharia, M. An Architecture for Fast and General Data Processing on Large Clusters / M. Zaharia // PHD Dissertation. – 3 February 2014 / Berkeley, USA. – 2014. P. 3 – 35
2. Pu, Q. FairRide: Near-Optimal, Fair Cache Sharing / Q. Pu, H. Li, M. Zaharia, A. Ghodsi, I. Stoica // NSDI. – 16 March 2016 / Santa Clara, USA. – 2016. P. 2 - 3

ПОСТРОЕНИЕ АНАЛИТИЧЕСКОГО СЕРВИСА ДЛЯ АНАЛИЗА СПЕКТРОВ

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

Усиков А. В.

Рудикова Л.В. – доцент, канд. физ-мат. наук

На сегодняшний момент накоплено довольно много материалов, которые представлены в определенных таблицах и атласах спектральных элементов. Непосредственная автоматизация процесса обработки спектрограмм, получаемых с помощью мобильного лазерного спектрометра, представляет собой программные инструменты, которые включают в себя достаточно узкий комплект возможностей. Недоступность инструмента, который дает возможность хранить и анализировать спектры в режиме реального времени с интеллектуальным поиском является актуальным направлением.

Установим необходимые требования к разрабатываемому сервису. В первую очередь в сервисе должны присутствовать возможности аналитической обработки данных, интеллектуального поиска данных (алгоритмы добычи данных), использования нейронных сетей для прогнозирования, корреляций, типовых образцов и исключений в больших объемах данных спектров. Необходимо отметить, что каждая из этих возможностей представляет собой самостоятельные части с точки зрения архитектуры.

Сервис должен равным образом предоставлять удаленный доступ к проанализированным и результирующим данным, поэтому необходимо предоставить публичную конечную точку пользователям для доступа в любое время к сервису и получения данных анализа.

Многоуровневая архитектура обеспечивает группировку связанной функциональности приложения в различных слоях, выстраиваемых вертикально, поверх друг друга. Слои слабо связаны, и между ними осуществляется явный обмен данными. Точное разбивание приложения на слои помогает поддерживать строгое распределение функциональности, что в свою очередь, обеспечивает гибкость, а также практичность и несложность сопровождения.

Функциональные области приложения разделяются на многослойные группы (уровни). Сервис состоит из шести взаимодействующих друг с другом слоев: уровень хранения данных, уровень надстроек, уровень доступа к данным, уровень бизнес-логики, уровень сервисов, уровень клиентов. Исходя из предъявляемых требований, для анализа спектров, полученных в результате лазерной экспрессной экспертизы будет приемлемым разрабатывать с применением многоуровневой архитектуры.

Для построения сервиса представлена многоуровневая архитектура. Подобная архитектура обеспечивает группировку связанной функциональности приложения в различных слоях, выстраиваемых вертикально, поверх друг друга. Слои слабо связаны, и между ними осуществляется явный обмен данными. Точное разбивание приложения на слои помогает поддерживать строгое распределение функциональности, что в свою очередь, обеспечивает гибкость, а также практичность и несложность сопровождения.

Каждый слой агрегирует ответственности и абстракции уровня, расположенного непосредственно под ним. При строгом разделении на слои составляющие одного слоя могут взаимодействовать только с составляющими такого же слоя или составляющими слоя, расположенного прямо под данным слоем. Более свободное разделение на слои позволяет составляющим взаимодействовать с составляющими того же и всех нижележащих слоев.

Внедрение подобного рода возможностей для обработки, анализа и хранения спектров, полученных в результате лазерной экспрессной экспертизы, дадут возможность внушительно усовершенствовать скорость аналитической обработки данных с интеллектуальным поиском, использованием нейронных сетей для прогнозирования сходств, корреляций, типовых образцов и исключений в больших объемах данных спектров. Следует отметить, что благодаря версионному хранению спектров, данные снимков каждый раз можно реконструировать до конкретного состояния в зависимости от требований..

Список использованных источников:

1. Кудрявцев, Ю. А. OLAP технологии: обзор решаемых задач и исследований / Ю.А. Кудрявцев // Бизнес-информатика. - 2008. - № 1. - С. 66-70.
2. Рудикова, Л. В. Разработка программного визуализатора спектров для поддержки лазерной экспрессной экспертизы //