

## 2.4. ВИЗУАЛИЗАЦИЯ РЕЛЯЦИОННЫХ БАЗ ДАННЫХ

Калинина Е.С.<sup>1</sup>, Манохина Т. В.<sup>2</sup>, Ступаков С. А.<sup>3</sup>  
<sup>123</sup>ФГБОУ ВО «Омский государственный университет путей сообщения»,  
Омск, Россия

*В статье рассмотрена проблематика изучения визуализации баз данных реляционно-го типа. В статье выделены преимущества и недостатки документо-ориентированной СУБД MongoDB на примере создания системы визуализации данных реального времени. Создана система визуализации данных реального времени, предназначенная для консолидации и обработки информации о событиях в высоконагруженных веб-системах. Система использует документо-ориентированную СУБД MongoDB для хранения полученных данных и JavaScript для обработки событий на стороне сервера, запускаемого с применением платформы Node.js.*

### Введение

Рост объема данных, ориентированных на пользователей, вызвал быстрое увеличение объема и типов данных, которые создаются, отображаются, анализируются и архивируются. Кроме того, изменяется количество новых наборов источников данных, включая датчики, системы глобального позиционирования (GPS), автоматизированные трекеры и системы мониторинга, создающие большие массивы данных. Эти большие объемы наборов данных, часто называемых «большими данными» (англ. Big Data), создают новые вызовы и возможности вокруг хранения, анализа и архивирования [1].

Параллельно с быстрым ростом объемов данных данные также становятся все более структурированными и рассеянными. Это означает, что традиционные методы управления данными, основанные на предварительном определении схемы данных и ссылок между данными, уже не отвечают современным требованиям обработки и анализа больших разнотипных слабо структурированных данных [2].

Документо-ориентированная система управления базами данных (ДОСУБД) – это система, которая сохраняет коллекцию текстовых данных и позволяет создавать запросы и обновления объектов. Обычно база данных включает в себя множество документов, соотносящихся по предмету, происхождению или применению [3]. Содержание каждого документа может быть свободным текстом, частично структурированным текстом, включающим несколько четко определенных атрибутов (например, название, автор, дата), или хорошо структурированным текстом, который, например, может быть закодирован с использованием языка XML. Иногда документы могут содержать мультимедийные компоненты [4].

ДОСУБД могут применяться для сохранения большого объема данных: изменений курсов валют, акций и ценных бумаг, трекинга использования сайта (анализа журналов Интернет-серверов с большим количеством пользователей), которые могут использоваться для разработки системы мониторинга (визуализации данных) реального времени. Таким образом, ДОСУБД предоставляет новые информационные технологии для издательств, цифровых библиотек, электронного правительства и электронного бизнеса в целом [5]. Поскольку новые технологии обработки больших массивов слабо структурированной информации еще не исследованы, выбранная тема данной статьи представляется актуальной.

### Методология исследования

Цель статьи – исследование визуализации баз данных реляционного типа.

### Основная часть

Существует много причин использовать нереляционные СУБД, но можно выделить две основные причины:

– Производительность разработки приложений. Много усилий в области разработки программного обеспечения уходит на отображение модели и структуры данных, находящихся в оперативной памяти компьютера, в реляционных базах данных. NoSQL хранилища данных могут обеспечить модель данных, которая будет лучше отвечать потребностям разработчиков приложений, тем самым упрощая взаимодействие с СУБД и, в результате чего, уменьшая количество написанного кода, время на отладку и внедрение.

– Большие масштабы данных. Организации считают целесообразным хранить большие объемы данных и обрабатывать их быстрее. Сделать это оказывается достаточно дорого, если вообще возможно, с использованием реляционных СУБД. Основной причиной является то, что реляционная база данных предназначена для запуска на одном компьютере, но, как правило, более экономно сохранять большие объемы данных и запускать вычисления на кластерах, состоящих из гораздо меньших и более дешевых машин. Многие нереляционные СУБД специально спроектированы и разработаны для работы на кластерах, поэтому они лучше подходят для сохранения больших объемов данных.

В литературе чаще всего выделяют следующие виды не реляционных СУБД [6]:

– Столбико-ориентированные СУБД. Структура данных в них контрастирует с кортежным (строчным) форматом реляционных СУБД. Это позволяет избегать использования места при хранении null-данных, просто не сохраняя столбцы, если нет данных для этого столбца. Столбцы не требуют никакого априорного определения. Кроме того, они способны хранить любые типы данных, поскольку данные могут быть сохранены в виде массива байтов. Такая структура данных, называемая Bigtable, используется, в частности, компанией Google.

– СУБД типа «ключ-значение». Они являются простой хэш-таблицей, поэтому весь доступ к данным осуществляется через первичный ключ. В качестве значений данных используются blob-объекты (англ. Binary Large Objects – большие объекты в бинарном формате). Поскольку СУБД типа «ключ-значение» всегда используют первичный ключ для доступа к данным, они, как правило, имеют высокую производительность и легко масштабируются.

– Базы данных на основе графов. Они позволяют сохранять объекты предметной области и связи между этими объектами. Объекты также называют узлами, обладающими свойствами. Узел представляет собой представление экземпляра объекта в применении. Связи между объектами известны как ребра, которые могут обладать свойствами. Ребра имеют направление; узлы организованы связями, позволяющими найти определенные закономерности между узлами. Организация графа позволяет сохранить данные один раз и затем интерпретировать по-разному на основе связей между объектами. Такая структура данных имеет много общего с представлением знаний на основе семантических сетей.

– Многомерные базы данных. Используются для интерактивной аналитической обработки агрегированных данных, при этом агрегация может быть по нескольким уровням иерархии, например год, квартал, месяц. Гиперкуб данных можно рассматривать как множество отношений реляционной базы данных по значениям каждого из измерений. Следовательно, носителем многомерной модели данных является отношение реляционной базы данных, изображенное как зафиксированное измерение. Следовательно, между реляционной и многомерной моделью существует взаимное отображение, а значит, можно данные, представленные в одной модели, превратить в другую модель. Поэтому построенные на многомерной модели технологии OLAP (On Line Analysis processing) [4] реализованы в распространенных СУБД реляционного типа MS SQL Server, Oracle и других. Многомерная модель расширяет реляционную: множество операций, кроме традиционных реляционных, содержит операции среза, свертки, детализации, вращения.

– Объектные СУБД. В них данные моделируются в виде объектов, их атрибутов, методов и классов. Одним из известных представителей этого типа систем хранения и обработки данных является СУБД CACHÉ [5]. Помимо объектной модели она поддерживает также реляционную и многомерную, при этом представление данных в одной из моделей автоматически преобразуется в другие модели.

– документо-ориентированные СУБД. Документы являются основной концепцией в документо-ориентированных СУБД.

Была опубликована и представлена серия работ, объясняющая некоторые из ключевых элементов инфраструктуры созданных систем. Наиболее важные из этих публикаций представлены в [6–9].

После одобрения NoSQL двумя ведущими веб-гигантами – Google и Amazon – появилось несколько новых продуктов. Многие разработчики стали использовать методы и идеи NoSQL в своих приложениях. Меньше чем за 5 лет NoSQL и связанные с ним понятия для управления большими объемами данных получили широкое распространение и использование во многих известных компаниях, включая Facebook, Netflix, Yahoo, eBay, Hulu, IBM и т.д.

Проблемы параллелизма существуют и в реляционных базах данных (потеря результатов обновления, зависимость от незафиксированных результатов, несогласованная обработка результатов) [10], но у неструктурированных БД они более сложны.

2. Представление данных в виде агрегатов. С этой точки зрения они ближе к объектным базам данных. Например, в объектной СУБД CACHÉ [5] такие агрегаты называются глобалами. Поскольку в агрегате хранится вся информация для определенной задачи, чтобы свести к минимуму операции join между объектами, то следствием этого может быть дублирование информации в других агрегатах.

Исходя из вышеприведенного, выполнен анализ существующих документо-ориентированных (К) СУБД и выбор базовой для системы визуализации данных. На сегодняшний день существует достаточно много ДОСУБД. Они выпускаются под разными лицензиями и реализованы для использования на разных платформах, хотя некоторые могут использоваться на нескольких платформах одновременно. Несмотря на некоторые отличия, следует отметить, что каждая реализация документной базы данных базируется на основном понятии – «документе», описывающем одну запись в базе данных. При этом допускаются разные способы организации хранения списков документов.

Была реализована визуализация данных реального времени, предназначенная для консолидации и обработки информации о событиях в высоконагруженных веб-системах – рис. 1.

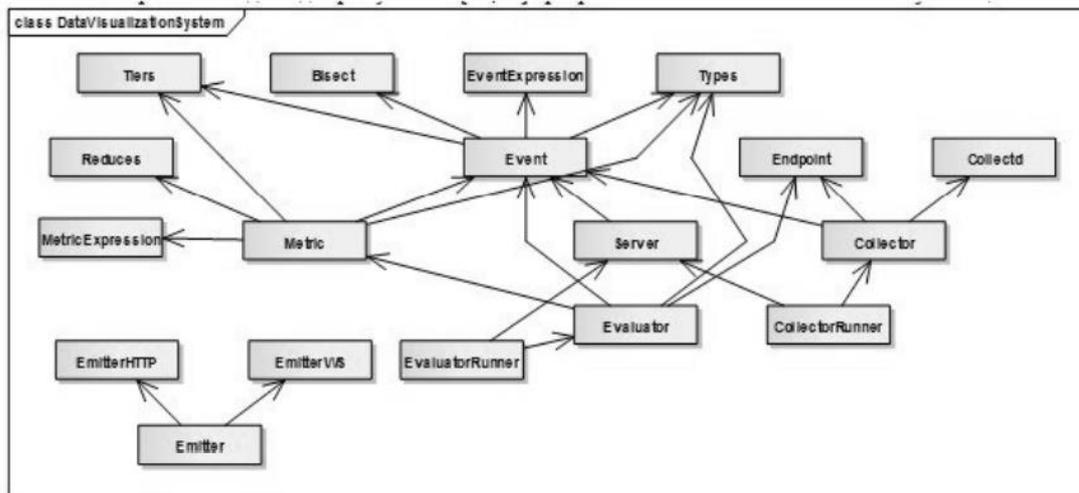


Рис. 1 – Диаграмма классов разработанной системы

Система использует документо-ориентированную СУБД MongoDB для хранения полученных данных и JavaScript для обработки событий на стороне сервера, запускаемого с применением платформы Node.js.

### Выводы

Практическим результатом проделанной работы является созданная система визуализации данных реального времени, предназначенная для консолидации и обработки информации о событиях в высоконагруженных веб-системах. Система использует документо-ориентированную СУБД MongoDB для хранения полученных данных и JavaScript для обработки событий на стороне сервера, запускаемого с применением платформы Node.js.

Преимущество системы состоит в формировании визуальной картины работы системы в режиме реального времени. К сожалению, черно-белый формат представления графического материала не позволяет наглядно представить результаты примера работы системы визуализации данных.

### Список литературы

1. Pedro Teixeira. Professional Node.js: Building Javascript Based Scalable Software. Indianapolis, IN. John Wiley & Sons, Inc. – 2012. – 412 p.
2. Абрамский М.М., Тимерханов Т.И. Сравнительный анализ использования реляционных и графовых баз данных в разработке цифровых образовательных систем // Вестник НГУ. Серия: Информационные технологии. 2018. №4. URL: <https://cyberleninka.ru/article/n/sravnitelnyy-analiz-ispolzovaniya-relyatsionnyh-i-grafovyh-baz-dannyh-v-razrabotke-tsifrovyyh-obrazovatelnyh-sistem> (дата обращения: 03.02.2023).
3. Васильева К. Н., Хусаинова Г. Я. Реляционные базы данных // Colloquium-journal. 2020. №2 (54). URL: <https://cyberleninka.ru/article/n/relyatsionnye-bazy-dannyh> (дата обращения: 03.02.2023).
4. Иванов В.И., Новиков С.П. Программное средство исследования различных алгоритмов кэширования в реляционных СУБД // Молодой исследователь Дона. 2018. №3 (12). URL: <https://cyberleninka.ru/article/n/programmnoe-sredstvo-issledovaniya-razlichnyh-algoritmov-keshirovaniya-v-relyatsionnyh-subd> (дата обращения: 03.02.2023).
5. Крэг Ларман. Применение UML 2.0 и шаблонов проектирования. – 3-е изд. – М.: Вильямс, 2006. – 736 с.
6. Применение CASE-средства VPwin 4.0 для информационного моделирования в системах обработки данных / С.В. Горин; А.Ю. Тандоев // СУБД. – М., 1995. – №3. – С. 125–127.
7. Реляционная база данных и её структура / Научная библиотека избранных естественно-научных изданий. — URL: [http://sernam.ru/book\\_cbd.php?id=2](http://sernam.ru/book_cbd.php?id=2) (дата обращения: 03.02.2023).
8. Савоськин И.В., Фирсов А.О. Исследование способов применения NoSQL и реляционных баз данных // E-Scio. 2019. №6 (33). URL: <https://cyberleninka.ru/article/n/issledovanie-sposobov-primeniya-nosql-i-relyatsionnyh-baz-dannyh> (дата обращения: 03.02.2023).
9. Свиридов А.А. Хранение топологии сети в реляционных базах данных // JSRP. 2013. №1. URL: <https://cyberleninka.ru/article/n/hranenie-topologii-seti-v-relyatsionnyh-bazah-dannyh> (дата обращения: 03.02.2023).
10. Федорова Г. Н. Разработка и администрирование баз данных / Г. Н. Федорова. – М.: Академия, 2015. – 58 с.

**References in Cyrillics**

1. Pedro Teixeira. Professional Node.js: Building Javascript Based Scalable Software. Indianapolis, IN. John Wiley & Sons, Inc. – 2012. – 412 p.
2. Abramskij M.M., Timerhanov T.I. Sravnitel'nyj analiz ispol'zovaniya relyacionnyh i grafovyyh baz dannyh v razrabotke cifrovyyh obrazovatel'nyh sistem // Vestnik NGU. Seriya: Informacionnye tekhnologii. 2018. №4. URL: <https://cyberleninka.ru/article/n/sravnitelnyy-analiz-ispolzovaniya-relyatsionnyh-i-grafovyh-baz-dannyh-v-razrabotke-tsifrovyyh-obrazovatelnyh-sistem> (data obrashcheniya: 03.02.2023).
3. Vasil'eva K. N., Husainova G. YA. Relyacionnye bazy dannyh // Colloquium-journal. 2020. №2 (54). URL: <https://cyberleninka.ru/article/n/relyatsionnye-bazy-dannyh> (data obrashcheniya: 03.02.2023).
4. Ivanov V.I., Novikov S.P. Programmnoe sredstvo issledovaniya razlichnyh algoritmov keshirovaniya v relyacionnyh SUBD // Molodoj issledovatel' Dona. 2018. №3 (12). URL: <https://cyberleninka.ru/article/n/programmnoe-sredstvo-issledovaniya-razlichnyh-algoritmov-keshirovaniya-v-relyatsionnyh-subd> (data obrashcheniya: 03.02.2023).
5. Kreg Larman. Primenenie UML 2.0 i shablonov proektirovaniya. – 3–e izd. – M.: Vil'yams, 2006. – 736 s.
6. Primenenie CASE–sredstva BPwin 4.0 dlya informacionnogo modelirovaniya v sistemah obrabotki dannyh / S.V. Gorin; A.YU. Tandoev // SUBD. – M., 1995. – №3. – S. 125–127.
7. Relyacionnaya baza dannyh i eyo struktura / Nauchnaya biblioteka izbrannyh estestvenno-nauchnyh izdanij. — URL.: [http://sernam.ru/book\\_cbd.php?id=2](http://sernam.ru/book_cbd.php?id=2) (data obrashcheniya: 03.02.2023).
8. Savos'kin I.V., Firsov A.O. Issledovanie sposobov primeneniya NoSQL i relyacionnyh baz dannyh // E-Scio. 2019. №6 (33). URL: <https://cyberleninka.ru/article/n/issledovanie-sposobov-primeneniya-nosql-i-relyatsionnyh-baz-dannyh> (data obrashcheniya: 03.02.2023).
9. Sviridov A.A. Hranenie topologii seti v relyacionnyh bazah dannyh // JSRP. 2013. №1. URL: <https://cyberleninka.ru/article/n/hranenie-topologii-seti-v-relyatsionnyh-bazah-dannyh> (data obrashcheniya: 03.02.2023).
10. Fedorova G. N. Razrabotka i administrirovanie baz dannyh / G. N. Fedorova. – M.: Akademiya, 2015. – 58 s.

**Ключевые слова**

Визуализация, базы данных, реляционные, особенности, изучение, применение

Калинина Екатерина Сергеевна, к. т. н, доцент кафедры «Информатика и компьютерная графика» ФГБОУ ВО «Омский государственный университет путей сообщения», Омск, Россия, [ekkalinina@mail.ru](mailto:ekkalinina@mail.ru)

Манохина Татьяна Витальевна, старший преподаватель кафедры «Информатика и компьютерная графика» ФГБОУ ВО «Омский государственный университет путей сообщения», [mtv-gups@mail.ru](mailto:mtv-gups@mail.ru)

Ступаков Сергей Анатольевич, к. т. н, доцент кафедры «Информатика и компьютерная графика» ФГБОУ ВО «Омский государственный университет путей сообщения», Омск, Россия, [stupakov1@yandex.ru](mailto:stupakov1@yandex.ru)

**Keywords**

Visualization, databases, relational, features, study, application

Kalinina Ekaterina Sergeevna, Candidate of Technical Sciences, Associate Professor of the Department of Informatics and Computer Graphics, Omsk State Transport University, Omsk, Russia, [ekkalinina@mail.ru](mailto:ekkalinina@mail.ru)

Manokhina Tatyana Vitalievna, Senior Lecturer of the Department of Informatics and Computer Graphics, Omsk State Transport University, [mtv-gups@mail.ru](mailto:mtv-gups@mail.ru)

Stupakov Sergey Anatolyevich, Candidate of Technical Sciences, Associate Professor of the Department of Informatics and Computer Graphics, Omsk State Transport University, Omsk, Russia, [stupakov1@yandex.ru](mailto:stupakov1@yandex.ru)

DOI: 10.34706/DE-2023-01-12

JEL classification C02 – Математические методы; M15 Управление информационными

The article deals with the problems of studying the visualization of databases of relational type. The article highlights the advantages and disadvantages of the MongoDB document-oriented DBMS using the example of creating a real-time data visualization system. A real-time data visualization system has been created, designed to consolidate and process information about events in highly loaded web systems. The system uses a document-oriented MongoDB DBMS to store the received data and JavaScript to process events on the server side, launched using the Node.js platform.