Уральский государственный экономический университет

ФОРМАЛИЗАЦИЯ ИНФОРМАЦИИ И Big Data

02.03.03 -Математическое обеспечение и администрирование информационных систем, направленность (профиль) -разработка и администрирование информационных систем

http//vikchas.ru

Лекция 10. Тема **«Big Data традиционных информационных систем и квантовый компьютер»**

Часовских Виктор Петрович
доктор технических наук, профессор кафедры
ШИиКМ, ФГБОУ ВО «Уральский государственный
экономический университет

Екатеринбург 2026

1. Введение

В современном мире объём цифровой информации растёт с огромной скоростью. Каждый день создаются данные в социальных сетях, банковских системах, интернет-магазинах, медицинских учреждениях, промышленных датчиках, транспортных системах и государственных информационных ресурсах. Эти огромные массивы данных принято называть **Big Data**, или **большие данные**.

Одновременно с развитием традиционных информационных систем активно исследуются новые вычислительные технологии, среди которых особое место занимает **квантовый компьютер**. Он основан не на классических битах, а на квантовых состояниях — кубитах. Квантовые компьютеры потенциально способны ускорять решение отдельных сложных задач, включая оптимизацию, моделирование и обработку данных.

Цель лекции — рассмотреть, как Big Data используется в традиционных информационных системах, какие проблемы возникают при обработке больших данных и какую роль в будущем могут сыграть квантовые компьютеры.

2. Понятие Big Data

Big Data — это совокупность технологий, методов и инструментов для хранения, обработки и анализа данных очень большого объёма, высокой скорости поступления и разнообразного формата.

Часто Big Data описывают через концепцию **5V**:

1. **Volume — объём**
   Данные измеряются терабайтами, петабайтами и эксабайтами.
2. **Velocity — скорость**
   Данные поступают и должны обрабатываться в реальном времени или почти в реальном времени.
3. **Variety — разнообразие**
   Данные могут быть структурированными, полуструктурированными и неструктурированными.
4. **Veracity — достоверность**
   Не все данные точны: возможны ошибки, дубли, шумы и противоречия.

5. **Value — ценность**
   Главная задача Big Data — извлечь полезную информацию и получить практическую выгоду.

3. Источники больших данных

Большие данные формируются из множества источников:

- социальные сети и мессенджеры;
- поисковые системы;
- электронная коммерция;
- банковские операции;
- медицинские информационные системы;
- датчики Интернета вещей;
- системы видеонаблюдения;
- спутниковые и геоинформационные системы;
- производственные линии;
- мобильные устройства;
- государственные базы данных.

Например, интернет-магазин собирает данные о просмотрах товаров, покупках, поисковых запросах, отзывах и поведении пользователей. На основе этого можно строить рекомендации и прогнозировать спрос.

4. Традиционные информационные системы

Традиционная информационная система — это комплекс программных, аппаратных, организационных и информационных средств, предназначенных для сбора, хранения, обработки и передачи данных.

К ним относятся:

- банковские системы;
- бухгалтерские программы;
- CRM-системы;
- ERP-системы;

- системы управления базами данных;
- медицинские информационные системы;
- образовательные платформы;
- государственные информационные системы.

Главная особенность традиционных ИС — использование классической вычислительной архитектуры, основанной на битах, процессорах, оперативной памяти, базах данных и сетевых технологиях.

5. Big Data в традиционных информационных системах

В традиционных информационных системах Big Data применяется для решения следующих задач:

5.1. Аналитика и прогнозирование

Большие данные позволяют анализировать прошлые события и прогнозировать будущие. Например:

- прогнозирование спроса;
- оценка финансовых рисков;
- выявление мошенничества;
- прогнозирование отказов оборудования;
- анализ поведения клиентов.

5.2. Персонализация

Компании используют Big Data для индивидуального подхода к пользователям:

- рекомендательные системы;
- персональная реклама;
- подбор контента;
- индивидуальные предложения;
- анализ пользовательского опыта.

5.3. Автоматизация управления

Большие данные помогают автоматизировать принятие решений:

- управление запасами;

- маршрутизация транспорта;
- распределение ресурсов;
- управление производством;
- динамическое ценообразование.

5.4. Безопасность

Big Data используется для кибербезопасности:

- обнаружение подозрительной активности;
- анализ сетевого трафика;
- выявление атак;
- мониторинг событий безопасности;
- обнаружение аномалий.

6. Архитектура Big Data-систем

Традиционная архитектура Big Data обычно включает несколько уровней.

6.1. Сбор данных

Данные поступают из разных источников: приложений, датчиков, журналов событий, транзакций, веб-сервисов.

Инструменты:

- Apache Kafka;
- Flume;
- Logstash;
- API-интерфейсы;
- IoT-шлюзы.

6.2. Хранение данных

Для хранения больших объёмов информации используются:

- распределённые файловые системы;
- NoSQL-базы данных;
- облачные хранилища;
- хранилища данных;

- озёра данных.

Примеры технологий:

- Hadoop HDFS;
- Apache Cassandra;
- MongoDB;
- HBase;
- Amazon S3;
- Google Cloud Storage.

6.3. Обработка данных

Обработка может быть пакетной или потоковой.

Пакетная обработка:

- Hadoop MapReduce;
- Apache Spark.

Потоковая обработка:

- Apache Flink;
- Spark Streaming;
- Kafka Streams.

6.4. Аналитика и визуализация

Результаты анализа представляются в виде отчётов, графиков и панелей мониторинга.

Инструменты:

- Tableau;
- Power BI;
- Grafana;
- Kibana;
- Superset.

7. Основные проблемы Big Data

Несмотря на эффективность Big Data-технологий, существуют серьёзные проблемы.

7.1. Объём и масштабируемость

Чем больше данных, тем сложнее их хранить, передавать и обрабатывать. Требуются распределённые системы и мощная инфраструктура.

7.2. Качество данных

Ошибки, дублирование, неполнота и противоречивость данных снижают точность анализа.

7.3. Скорость обработки

Некоторые задачи требуют мгновенного результата, например обнаружение мошеннической операции в банке.

7.4. Безопасность и конфиденциальность

Большие данные часто содержат персональную, медицинскую или финансовую информацию. Поэтому важны защита, шифрование и контроль доступа.

7.5. Сложность анализа

Даже при наличии больших объёмов данных трудно найти действительно полезные закономерности.

7.6. Высокая стоимость

Системы хранения, вычислительные кластеры, облачные ресурсы и специалисты требуют значительных затрат.

8. Квантовый компьютер: основные понятия

Квантовый компьютер — это вычислительная система, использующая законы квантовой механики для обработки информации.

В классическом компьютере единицей информации является **бит**, который принимает значение 0 или 1.

В квантовом компьютере используется **кубит**. Кубит может находиться не только в состоянии 0 или 1, но и в их суперпозиции.

Основные свойства кубитов:

- **суперпозиция** — возможность находиться в нескольких состояниях одновременно;

- **запутанность** — связь между кубитами, при которой состояние одного зависит от состояния другого;
- **интерференция** — усиление правильных вариантов решения и ослабление неправильных.

Эти свойства позволяют квантовым алгоритмам эффективно решать некоторые задачи, которые для классических компьютеров являются очень сложными.

9. Отличие квантового компьютера от классического

Скопировать таблицу

Характеристика	Классический компьютер	Квантовый компьютер
Единица информации	Бит	Кубит
Состояния	0 или 1	Суперпозиция 0 и 1
Принцип работы	Логические операции	Квантовые операции
Устойчивость	Высокая	Чувствителен к шуму
Основная сфера	Универсальные вычисления	Специальные сложные задачи
Стадия развития	Зрелая технология	Развивающаяся технология

Важно отметить: квантовый компьютер не является просто «очень быстрым обычным компьютером». Он эффективен не для всех задач, а только для определённых классов задач.

10. Квантовые алгоритмы и их связь с Big Data

Квантовые компьютеры могут быть полезны для некоторых задач анализа больших данных.

10.1. Алгоритм Гровера

Алгоритм Гровера предназначен для поиска в неструктурированной базе данных. Он даёт квадратичное ускорение по сравнению с классическим перебором.

Если классическому алгоритму нужно в среднем проверить N элементов, то квантовому алгоритму требуется примерно √N шагов.

Применение:

- поиск записей;
- анализ больших массивов;
- криптоанализ;
- оптимизация запросов.

Однако для практического применения необходимо эффективно загрузить данные в квантовую систему, что пока является сложной задачей.

10.2. Алгоритм Шора

Алгоритм Шора позволяет эффективно разлагать большие числа на множители и находить дискретные логарифмы.

Это важно не столько для Big Data, сколько для информационной безопасности, так как многие современные криптосистемы основаны на сложности факторизации.

Применение:

- анализ угроз для RSA;
- пересмотр криптографической защиты данных;
- развитие постквантовой криптографии.

10.3. Квантовое машинное обучение

Квантовое машинное обучение — это направление, изучающее использование квантовых алгоритмов для задач искусственного интеллекта.

Возможные применения:

- классификация данных;
- кластеризация;
- оптимизация моделей;
- ускорение обучения;
- работа с высокоразмерными признаками.

Пока квантовое машинное обучение находится преимущественно на исследовательской стадии.

11. Потенциальная роль квантовых компьютеров в Big Data

Квантовые компьютеры могут быть полезны в следующих направлениях.

11.1. Оптимизация

Многие Big Data-задачи связаны с поиском оптимального решения:

- оптимальная маршрутизация транспорта;
- распределение ресурсов;
- управление логистикой;
- подбор инвестиционного портфеля;
- планирование производства.

Квантовые алгоритмы потенциально могут ускорить такие вычисления.

11.2. Быстрый поиск

Поиск по большим массивам данных может быть ускорен алгоритмом Гровера. Но реальная эффективность зависит от возможности квантового доступа к данным.

11.3. Моделирование сложных систем

Квантовые компьютеры особенно перспективны для моделирования квантовых систем:

- молекул;
- материалов;
- химических реакций;
- лекарственных соединений.

В таких задачах Big Data может сочетаться с квантовым моделированием и машинным обучением.

11.4. Анализ графов и сетей

Многие большие данные имеют форму графов:

- социальные связи;
- транспортные сети;

- финансовые транзакции;
- интернет-графы;
- связи между документами.

Квантовые методы могут использоваться для отдельных задач анализа графов и оптимизации.

12. Ограничения квантовых компьютеров для Big Data

Несмотря на большие перспективы, сегодня существуют серьёзные ограничения.

12.1. Проблема загрузки данных

Чтобы квантовый компьютер обработал данные, их нужно представить в квантовом виде. Загрузка классических Big Data в квантовую память может занять много времени и свести на нет преимущество квантовых вычислений.

12.2. Ограниченное число кубитов

Современные квантовые компьютеры имеют ограниченное количество кубитов и высокий уровень ошибок.

12.3. Декогеренция

Кубиты нестабильны и быстро теряют квантовые свойства под воздействием окружающей среды.

12.4. Ошибки вычислений

Квантовые операции подвержены шуму, поэтому требуется квантовая коррекция ошибок.

12.5. Специализированность

Квантовый компьютер не заменяет классические Big Data-платформы. Он может использоваться как специализированный ускоритель для отдельных задач.

13. Big Data и постквантовая безопасность

Одно из важнейших последствий развития квантовых компьютеров — угроза современным криптографическим системам.

Большие данные часто содержат:

- персональные данные;

- банковскую информацию;
- медицинские записи;
- коммерческие тайны;
- государственные данные.

Если мощный квантовый компьютер сможет взламывать распространённые криптосистемы, возникнет риск раскрытия огромных массивов информации.

Поэтому развивается **постквантовая криптография** — методы защиты, устойчивые к атакам квантовых компьютеров.

Основные направления:

- криптография на решётках;
- хэш-подписи;
- кодовые криптосистемы;
- многочленные криптосистемы.

Организации уже постепенно переходят к новым стандартам защиты данных.

14. Гибридная модель: классические Big Data + квантовые вычисления

Наиболее реалистичный сценарий будущего — не замена традиционных систем, а их интеграция с квантовыми вычислениями.

В такой модели:

- классические системы собирают и хранят данные;
- Big Data-платформы очищают и подготавливают данные;
- квантовый компьютер решает специализированную задачу;
- результат возвращается в классическую систему;
- пользователь получает аналитический отчёт через привычный интерфейс.

Пример:

1. Компания собирает данные о логистике.
2. Spark или Hadoop обрабатывают и агрегируют данные.
3. Квантовый алгоритм ищет оптимальные маршруты.
4. Результат передаётся в ERP-систему.

5. Менеджер видит рекомендации в панели управления.

15. Примеры возможного применения

15.1. Финансы

- оптимизация инвестиционного портфеля;
- оценка рисков;
- обнаружение мошенничества;
- моделирование рынков.

15.2. Медицина

- анализ медицинских данных;
- подбор лекарств;
- моделирование молекул;
- персонализированная медицина.

15.3. Логистика

- оптимизация маршрутов;
- управление складами;
- прогнозирование спроса;
- распределение транспорта.

15.4. Промышленность

- предиктивное обслуживание оборудования;
- оптимизация производства;
- контроль качества;
- цифровые двойники.

15.5. Кибербезопасность

- анализ больших потоков событий;
- обнаружение аномалий;
- защита данных постквантовыми алгоритмами;
- оценка криптографических рисков.

16. Современное состояние технологий

На сегодняшний день традиционные Big Data-технологии широко применяются в бизнесе, науке и государственном управлении. Они являются зрелыми и практически значимыми.

Квантовые компьютеры, напротив, пока находятся на этапе активных исследований и экспериментов. Уже существуют облачные платформы, через которые можно запускать квантовые программы:

- IBM Quantum;
- Google Quantum AI;
- Microsoft Azure Quantum;
- Amazon Braket;
- D-Wave Leap.

Однако современные квантовые устройства ещё не способны массово заменить классические вычислительные кластеры.

17. Перспективы развития

В будущем можно ожидать:

- роста числа и качества кубитов;
- развития квантовой коррекции ошибок;
- появления более мощных квантовых алгоритмов;
- интеграции квантовых сервисов в облачные платформы;
- развития гибридных квантово-классических систем;
- перехода информационных систем на постквантовую криптографию;
- появления новых подходов к анализу больших данных.

Вероятнее всего, квантовые компьютеры будут использоваться как специализированные ускорители для задач, где классические методы слишком медленны или затратны.

18. Заключение

Big Data является важной частью современных традиционных информационных систем. Большие данные позволяют организациям

принимать более точные решения, прогнозировать события, автоматизировать процессы и повышать эффективность работы.

Квантовый компьютер открывает новые перспективы для обработки сложных задач, связанных с оптимизацией, поиском, моделированием и безопасностью. Однако на современном этапе он не заменяет традиционные Big Data-системы, а рассматривается как их возможное дополнение.

Наиболее перспективным направлением является гибридный подход, при котором классические информационные системы выполняют сбор, хранение и предварительную обработку данных, а квантовые компьютеры применяются для решения отдельных вычислительно сложных задач.

Краткие выводы

1. Big Data — это технологии работы с большими, быстрыми и разнообразными данными.
2. Традиционные информационные системы используют Big Data для аналитики, прогнозирования, персонализации и безопасности.
3. Квантовый компьютер использует кубиты, суперпозицию и запутанность.
4. Квантовые компьютеры перспективны для оптимизации, поиска, моделирования и криптоанализа.
5. Прямая обработка Big Data на квантовом компьютере пока ограничена техническими проблемами.
6. Главная практическая модель будущего — гибридные квантово-классические системы.
7. Развитие квантовых технологий требует перехода к постквантовой защите данных.