

ВАСИЛЬЕВ Г. В.
МЕТОДИКА ОБРАБОТКИ БОЛЬШИХ ДАННЫХ ОТ МОБИЛЬНОГО ПРИЛОЖЕНИЯ В РЕАЛЬНОМ ВРЕМЕНИ

УДК 004.5:658, ВАК 2.3.1, ГРНТИ 50.41.21

Методика обработки больших данных от мобильного приложения в реальном времени

Methodology for processing big data from a mobile application in real-time

Г. В. Васильев

G. V. Vasilev

Комсомольске-на-Амуре
государственный университет, г.
Комсомольск-на-Амуре

Komsomolsk-na-Amure state
university, Komsomolsk-na-Amure

В статье автор рассказывает о своей методике обработки больших данных от мобильного приложения в реальном времени. Разработано новое современное мобильное приложение. Для хранения больших данных использовалась таблица NoSQL архитектуры. Для обработки больших данных разработано программное обеспечение на Python.

In the article, the author spoke about his method of processing big data from a mobile application in real-time. A new contemporary mobile application has been developed. To store big data, a NoSQL architecture table was used. New Python software has been developed for processing big data.

Ключевые слова: большие данные, NoSQL, Google Firebase, Google BigQuery, Java, Python, мобильное приложение

Keywords: big data, NoSQL, Google Firebase, Google BigQuery, Java, Python, mobile application

Введение

По мере распространения технологий, которые собирают и анализируют данные, расширяются и возможности бизнеса по анализу данных и извлечению из них новых идей. Сбор информации компании используют для целого ряда задач, включая лучшее понимание повседневных операций, принятие более обоснованных бизнес-решений и изучение своих клиентов.

Предприятия умеют извлекать все типы данных практически из каждого укромного уголка. Наиболее очевидными источниками являются активность потребителей на их веб-сайтах, страницах в социальных сетях, телефонные звонки клиентов и живые чаты. Многим компаниям потребительские данные помогают лучше понять и удовлетворить требования своих клиентов. Анализируя поведение клиентов, а также огромное количество отзывов,

компании могут быстро изменить свое цифровое присутствие, товары или услуги, чтобы они лучше соответствовали текущему рынку.

Автор провел анализ разных мобильных приложений в различных сферах деятельности [1-8]. Также автор проанализировал текущие решения по сбору данных и их хранение [9-16]. Автор предложит свою методику с помощью, которой можно будет повысить эффективность предприятия по анализу поведения клиентов.

Анализ данных

Данные можно разделить на несколько типов:

- **Персональные данные.** Эта категория включает личную информацию, такую как номера социального страхования и пол, а также неличную информацию, включая Ваш IP-адрес, файлы cookie веб браузера и идентификаторы устройств (которые есть на вашем электронном устройстве, например, ноутбук или смартфон).

- **Данные о взаимодействии.** Этот тип данных подробно описывает, как потребители взаимодействуют с веб-сайтом компании, мобильными приложениями, текстовыми сообщениями, страницами в социальных сетях, электронной почтой, платной рекламой и маршрутами обслуживания клиентов.

- **Поведенческие данные.** Эта категория включает сведения о транзакциях, такие как история покупок, информация об использовании продукта, например, повторяющиеся действия и качественные данные, например, информация о движении мыши.

- **Данные об отношении.** Этот тип данных включает показатели удовлетворенности потребителей, критерии покупки и многое другое.

Компании собирают данные различными способами из многих источников. Некоторые методы сбора носят высокотехнологичный характер, в то время как другие носят более дедуктивный характер. Хотя в этих процессах часто используется сложное программное обеспечение. Компании используют изобилие методов для сбора и обработки данных о клиентах по метрикам, проявляя интерес к типам данных, начиная от демографических данных и заканчивая поведенческими данными. Данные о клиентах можно собирать тремя способами:

- напрямую спрашивая клиентов;
- косвенно отслеживая клиентов;
- добавляя другие источники данных о клиентах к своим собственным данным.

Методика разработки программного обеспечения для работы с большими данными

Далее будет представлена разработка нового современного мобильного приложения для работы в финансовой сфере. Данное мобильное приложение собирает поведенческие данные пользователей. Для хранения данных использовалась NoSQL таблица. Также будет разработано программное

приложение для получения, анализа и обработки данных о пользователях из NoSQL таблицы.

Некоторый функционал в мобильном приложении использует сервис Firebase от Google. Google Firebase – это набор инструментов, который поддерживает разработку мобильных и веб-приложений.

Первым этапом для реализации процесса сбора данных по поведению пользователя в мобильном приложении, является аутентификацией. Аутентификация – процедура проверки подлинности, например:

- проверка подлинности пользователя путём сравнения введённого им пароля (для указанного логина) с паролем, сохранённым в базе данных пользовательских логинов;
- подтверждение подлинности электронного письма путём проверки цифровой подписи письма по открытому ключу отправителя;
- проверка контрольной суммы файла на соответствие сумме, заявленной автором этого файла.

Технология аутентификации обеспечивает контроль доступа к системам, проверяя, совпадают ли учетные данные пользователя с учетными данными в базе данных авторизованных пользователей или на сервере аутентификации данных. При этом аутентификация обеспечивает безопасность информационных систем.

Существует несколько типов аутентификации. В целях идентификации пользователей обычно используется идентификатор пользователя, а проверка подлинности происходит, когда пользователь предоставляет учетные данные, такие как пароль, который соответствует его идентификатору пользователя. Практика запроса идентификатора пользователя и пароля известна как однофакторная проверка подлинности. В последние годы компании усилили аутентификацию, запрашивая дополнительные факторы аутентификации, такие как уникальный код, который предоставляется пользователю через мобильное устройство при попытке входа в систему, или биометрическую подпись, такую как сканирование лица или отпечаток пальца. Это известно, как двухфакторная аутентификация.



Рисунок 1. Процесс регистрации

На Рисунке 1 представлен процесс регистрации в мобильном приложении.

На Рисунке 2 можно увидеть, что пользователь “t3g.ru” добавился в таблицу пользователей.



Identifier	Providers	Created ↓	Signed In	User UID
t3@g.ru	✉	Feb 2, 2022	Feb 2, 2022	tTYNlotiHjPzHNKq5hjOTNo3ZoW2
t4@g.ru	✉	Feb 2, 2022	Feb 2, 2022	b1NTxB82qwWkWFtj1XIDRDma..

Рисунок 2. Таблица пользователей

User UID будет служить для прохождения аутентификации пользователя во время совершения некоторых действий в мобильном приложении.

Далее, с помощью этих данных можно уже авторизоваться (Рисунок 3-4).



Рисунок 3. Окно авторизации

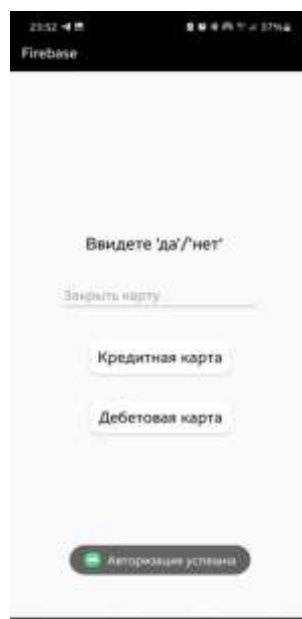


Рисунок 4. Успешная авторизация

После регистрации и авторизации пользователем в мобильном приложении можно отслеживать определенные действия пользователя. Например, действие открытие карты. Если ввести слово “да” и нажать на кнопку “Кредитная карта”, то в таблице записывается соответствующая запись. На рисунке 5 представлена часть кода, которая записывает значение в таблицу.

```
findViewById(R.id.credit_btn).setOnClickListener(view → {
    if(answer.getText().toString().equals("да"))
        myRef.child("users").child(user.getUid()).child("card_action").child("credit")
            .child("close").setValue(1);
}
```

Рисунок 5. Код записи значения в таблицу

На рисунке 6 представлена структура таблицы. Таблица является NoSQL архитектуры. В списке “users” хранятся значения привязанных UID пользователей. Далее для каждого пользователя есть два объекта “credit” и “debit”, то есть кредитная и дебетовая карта. Значение 0 в числовой переменной “close” означает, что карта открыта, а значение 1 означает, что карту нужно закрыть.

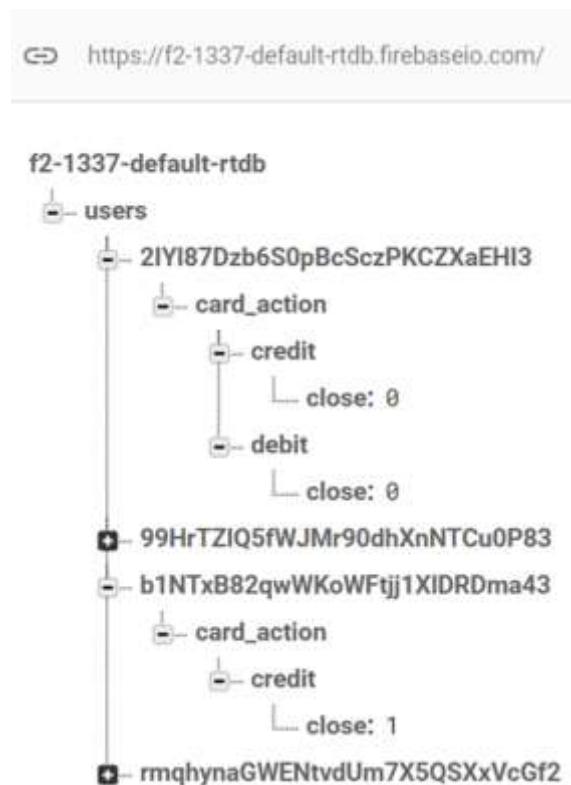


Рисунок 6. Структура таблицы

Преимущества такой таблицы является, что она работает в реальном времени. На рисунках 7-8 представлен процесс добавления данных в таблицу.



Рисунок 7. Ввод данных



Рисунок 8. Добавление данных в таблицу

На рисунках 9-10 представлен процесс изменения данных в таблице.



Рисунок 9. Изменение данных



Рисунок 10. Изменение данных в таблице

Результаты

После применения вышеописанной методики разработки программного обеспечения для работы с большими данными можно получать и обрабатывать данные. На рисунке 11 представлен результат выполнения программного обеспечения. Программное обеспечение считала данные о действиях пользователя из текущей таблицы.



Рисунок 11. Получение данных из таблицы

Заключение

В результате проведенного исследования была разработана методика, которая позволяет разработать новое современное программное обеспечение для пользователей. Программное обеспечение для пользователей является мобильным приложением, в котором отслеживаются важные события для бизнеса и в реальном времени записываются/обновляются/удаляются в NoSQL базе данных.

Также было разработано новое современное программное обеспечение, которое позволяет бизнесу формировать запросы к NoSQL базе данных и получать информацию о событиях клиентов.

Для разработки нового современного мобильного приложения для операционной системе Android применялись самые актуальные технологии и методики. Мобильное приложение было написано в IDE Android Studio на языке программирования Java. Для хранения большого объема данных использовалась NoSQL архитектура вида key:value. Для этой задачи была выбрана платформа Google BigQuery. Программное обеспечение для получения и обработки данных из NoSQL таблицы было разработано с помощью языка программирования Python.

Данное программное обеспечение легко масштабируется, позволяет обрабатывать большие объёмы данных, работает в real-time режиме. Таким образом предприятие легко сможет занять высокие позиции в рейтингах. Так как будет анализировать и понимать своих потребителей.

Компания Amazon одна из первых внедрила рекомендации товаров на основе интересов пользователей. А сегодня то же самое делают и десятки других компаний: Spotify, Pinterest, Netflix и многих других. Примерно 48% организаций используют big data, чтобы получить максимально полное представление о клиентском опыте.

Список использованных источников и литературы

1. Разумова О. И. Банковские мобильные приложения как средство цифровизации выдачи кредита в рамках проектного финансирования застройщиков // Заметки ученого. 2022. № 5. С. 101-109
2. Шадрин В. Г. Мобильные приложения банков как важнейший инструмент цифровых маркетинговых коммуникаций в финансовой сфере / В.Г. Шадрин, М.А. Кузин // Креативная экономика. 2022. Том 16. № 9. С. 3473-3486. DOI: 10.18334/ce.16.9.116302

3. Щетинин Н. С. Мобильное приложение для учета финансов // Информационные технологии в строительных, социальных и экономических системах. 2022. № 3(29). С. 104-110
4. Крамаренко Т. А. Разработка интеллектуальных мобильных приложений / Т.А. Крамаренко, Е.Л. Фешина, Т.В. Лукьяненко // Известия ЮФУ. технические науки. 2022. № 2(226). с. 249-259. DOI: 10.18522/2311-3103-2022-2-249-259
5. Широков С. А. Мобильное приложение "портфель инвестиций" / С.А. Широков, О.И. Белозеров // Научно-техническое и экономическое сотрудничество стран АТР в XXI веке. 2022. Том 2. С. 438-440
6. Дорохин М. А. Проектирование мобильного приложения для операционной системы android // Студенческий вестник. 2022. № 18-15 (210). С. 5-7
7. Зуева К. А. Сервисы для создания мобильных приложений / К.А. Зуева, В.Ю. Белаш // дневник науки. 2022. № 4(64)
8. Басшыкызы Д. Разработка мобильного приложения // Наука, техника и образование. 2022. № 3 (86). С. 39-40
9. Скрыпников А. В. Парсинг данных для мобильного приложения университета / А.В. Скрыпников, В.В. Денисенко, А.А. Берестовой // Вестник воронежского института ФСИН россии. 2022. № 2. С. 126-132
10. Васильев Г. В. Сбор данных о погоде через API Openweather / Г.В. Васильев, В.Д. Бердоносков, А.В. Васильев // Наука, инновации и технологии: от идей к внедрению: Материалы Международной научно-практической конференции. Комсомольск-на-Амуре, 2022. С. 86-88
11. Васильев, Г. В. Программный-аппаратный комплекс для сбора данных, которые применяются в обучении искусственного интеллекта в области энергетики / Г. В. Васильев, В. Д. Бердоносков // Молодежь и наука: актуальные проблемы фундаментальных и прикладных исследований: Материалы IV Всероссийской национальной научной конференции студентов, аспирантов и молодых ученых. В 4-х частях, Комсомольск-на-Амуре, 12-16 апреля 2021 года / Редколлегия: Э.А. Дмитриев (отв. ред.), А.В. Космынин (зам. отв. ред.). – Комсомольск-на-Амуре: Комсомольский-на-Амуре государственный университет, 2021. – Ч. 3. – С. 163-165. DOI: 10.17084/978-5-7765-1480-7-2021-163
12. Винников М. Д. Проектирование базы данных для мобильного приложения / М. Д. Винников, К. В. Гудков // Современные информационные технологии. 2022. № 35 (35). С. 117-119. DOI: 10.46548/SIT-2022-0035-0030
13. Васильев Г. В. Оптимизированное хранение больших данных с помощью Apache Hive / Г. В. Васильев, А. В. Васильев // Наука, инновации и технологии: от идей к внедрению: Материалы Международной научно-практической конференции. Комсомольск-на-Амуре, 2022. С. 89-91. DOI: 10.17084/978-5-7765-1502-6-2022-89
14. Васильев Г. В. Параллелизм и отказоустойчивость в обработки больших данных / Г. В. Васильев, А. В. Васильев // Актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности: Материалы I Международной научно-практической конференции молодых учёных.

Комсомольск-на-Амуре, 2021. С. 257-260. DOI: 10.17084/978-5-7765-1488-3-2021-257

15. Дронов Р. Д. Сравнительный анализ СУБД, используемых для мобильных приложений / Р. Д. Дронов, А. Н. Петрова // Наука, инновации и технологии: от идей к внедрению: Материалы Международной научно-практической конференции. Комсомольск-на-Амуре, 2022. С. 15-17

16. Васильев Г. В. Применение программной модели Map/Reduce в анализе больших данных в области медицины / Г. В. Васильев, А. В. Васильев // Актуальные проблемы информационно-телекоммуникационных технологий и математического моделирования в современной науке и промышленности: Материалы I Международной научно-практической конференции молодых учёных. Комсомольск-на-Амуре, 2021. С. 260-263. DOI: 10.17084/978-5-7765-1488-3-2021-260

List of references

1. Razumova O. I. Banking mobile applications as a means of digitalizing the issuance of a loan in the framework of project financing for developers // Notes of a scientist. 2022. No. 5. S. 101-109.

2. Shadrin V.G. Mobile applications of banks as the most important instrument of digital marketing communications in the financial sector / V.G. Shadrin, M.A. Kuzin // Creative Economy. 2022. Volume 16. No. 9. S. 3473-3486. DOI: 10.18334/ce.16.9.116302

3. Shchetinin N. S. Mobile application for financial accounting // Information technologies in construction, social and economic systems. 2022. No. 3(29). pp. 104-110

4. Kramarenko T. A. Development of intelligent mobile applications / T.A. Kramarenko, E.L. Feshina, T.V. Lukyanenko // Izvestiya SFedU. Technical science. 2022. No. 2(226). With. 249-259. DOI: 10.18522/2311-3103-2022-2-249-259

5. Shirokov S.A. Mobile application "investment portfolio" / S.A. Shirokov, O.I. Belozarov // Scientific, technical and economic cooperation of the Asia-Pacific countries in the XXI century. 2022. Volume 2. S. 438-440

6. Dorokhin M. A. Designing a mobile application for the android operating system // Student Bulletin. 2022. No. 18-15 (210). pp. 5-7

7. Zueva K.A. Services for creating mobile applications / K.A. Zueva, V.Yu. Belash // diary of science. 2022. No. 4(64)

8. Basshykyzy D. Development of a mobile application // Science, technology and education. 2022. No. 3 (86). pp. 39-40

9. Skrypnikov A.V. Data parsing for a university mobile application / A.V. Skrypnikov, V.V. Denisenko, A.A. Berestova // Bulletin of the Voronezh Institute of the Federal Penitentiary Service of Russia. 2022. No. 2. P. 126-132

10. Vasiliev G.V. Collection of weather data through API Openweather / G.V. Vasiliev, V.D. Berdonosov, A.V. Vasiliev // Science, innovations and technologies: from ideas to implementation: Proceedings of the International Scientific and Practical Conference. Komsomolsk-on-Amur, 2022, pp. 86-88

11. Vasiliev, G. V. Software and hardware complex for data collection, which are used in training artificial intelligence in the field of energy / G. V. Vasiliev, V. D. Berdonosov // Youth and science: actual problems of fundamental and applied research: Materials IV All-Russian National Scientific Conference of Students, Postgraduates and Young Scientists. In 4 parts, Komsomolsk-on-Amur, April 12-16, 2021 / Editorial Board: E.A. Dmitriev (responsible editor), A.V. Kosmynin (deputy editor-in-chief). - Komsomolsk-on-Amur: Komsomolsk-on-Amur State University, 2021. - Part 3. - P. 163-165. DOI: 10.17084/978-5-7765-1480-7-2021-163
12. Vinnikov M. D. Designing a database for a mobile application / M. D. Vinnikov, K. V. Gudkov // Modern information technologies. 2022. No. 35 (35). pp. 117-119. DOI: 10.46548/CIT-2022-0035-0030
13. Vasiliev G. V. Optimized storage of big data using Apache Hive / G. V. Vasiliev, A. V. Vasiliev // Science, innovation and technology: from ideas to implementation: Proceedings of the International Scientific and Practical Conference. Komsomolsk-on-Amur, 2022, pp. 89-91. DOI: 10.17084/978-5-7765-1502-6-2022-89
14. Vasiliev G. V. Parallelism and fault tolerance in big data processing / G. V. Vasiliev, A. V. Vasiliev // Actual problems of information and telecommunication technologies and mathematical modeling in modern science and industry: Proceedings of the I International Scientific and Practical Conference of Young Scientists. Komsomolsk-on-Amur, 2021. S. 257-260. DOI: 10.17084/978-5-7765-1488-3-2021-257
15. Dronov R. D. Comparative analysis of DBMS used for mobile applications / R. D. Dronov, A. N. Petrova // Science, innovations and technologies: from ideas to implementation: Proceedings of the International Scientific and Practical Conference. Komsomolsk-on-Amur, 2022, pp. 15-17
16. Vasiliev G. V. Application of the Map/Reduce software model in the analysis of big data in the field of medicine / G. V. Vasiliev, A. V. Vasiliev // Actual problems of information and telecommunication technologies and mathematical modeling in modern science and industry: Proceedings of the I International scientific-practical conference of young scientists. Komsomolsk-on-Amur, 2021, pp. 260-263. DOI: 10.17084/978-5-7765-1488-3-2021-260