

ТЕХНОЛОГІЇ BIG DATA В АВТОМОБІЛЬНОМУ СЕРВІСІ: ПРИНЦИПИ ТА ПРОБЛЕМИ

Валерій ГРИЦУК

*Харківський національний автомобільно-дорожній університет, Харків,
Україна, ORCID 0000-0002-3780-7815, e-mail: valeri.gritsuk@gmail.com*

Використання технологій Big Data у сфері автомобільного сервісу стало актуальною темою останніх років (сам термін Big Data був запроваджений у 2008 році редактором журналу «Nature» Кліффордом Лінчем [1]). Розробка інноваційних технологій в поєднанні з розширеними можливостями аналізу даних дозволяє автосервісним компаніям вирішувати проблеми ефективніше і забезпечувати високу якість обслуговування своїх клієнтів.

Big Data – це термін, що використовується для опису сукупності підходів, інструментів та методів обробки великих обсягів структурованих та неструктурованих даних, які не можуть бути ефективно оброблені за допомогою звичайних методів обробки. Основними характеристиками великих даних є постійне збільшення їх обсягу, що вимагає масштабованості, паралельних обчислень та розподіленості засобів роботи з Big Data. Для цього використовуються нереляційні бази даних (NoSQL), модель MapReduce для обробки інформації, компоненти кластерної екосистеми Hadoop [2], а також мови програмування R та Python, а також спеціалізовані продукти Apache [3] (Spark, AirFlow, Kafka, HBase та ін.) [4]. Згідно зі звітом MarketsandMarkets [5], прогнозований розмір ринку Big Data становитиме близько 138,9 мільярдів доларів США до 2023 року.

Big Data характеризується трьома вимірами, які відомі як 3V (концепція була розроблена компанією Meta Group у 2001 році): об'єм або обсяг даних (Volume), швидкість накопичення нових даних та їх обробки (Velocity) та різноманітність типів даних, які можуть оброблятися (Variety). Деякі дослідники до них стали відносити Veracity и Value – про достовірність та цінність отриманих, проаналізованих даних (таким чином концепцію можна вважати як 5V). Основні характеристики Big Data в автомобільній промисловості можна сформулювати наступним чином:

- Об'єми даних: в автомобільній промисловості генеруються великі об'єми даних, які потрібно збирати та обробляти.
- Різноманітність даних: дані можуть бути отримані з різних джерел, таких як датчики в автомобілях, системи GPS, соціальні медіа, платіжні системи та інше.

- Швидкість обробки даних: в автомобільній промисловості та сервісі важливо отримувати дані в режимі реального часу та швидко обробляти їх для прийняття рішень.

Грунтуючись на попередніх визначеннях, основні принципи роботи з великими даними у контексті автомобільного сервісу, можна сформулювати наступним чином. Ці принципи повинні включати горизонтальну масштабованість, відмовостійкість та локальність даних. Завдяки горизонтальній масштабованості можлива обробка великих обсягів даних без погіршення продуктивності. Відмовостійкість допомагає передбачити можливі випадки виходу з ладу обчислювальних вузлів та запобігти втраті даних. Локальність даних дозволяє зменшити витрати на передачу даних і забезпечити їх ефективну обробку на машині, на якій вони зберігаються. Усі ці принципи розробляються з метою ефективної обробки великих обсягів даних, що є важливим для автомобільного сервісу.

Автомобільний сервіс збирає різноманітні дані з автомобілів, що дозволяє покращити якість обслуговування та розвивати нові послуги для клієнтів. Основні типи даних, які можуть збиратися в автомобільному сервісі:

1. Дані про стан автомобіля: ці дані включають інформацію про технічний стан автомобіля, включаючи параметри двигуна, системи гальмування, системи безпеки і електронні системи управління. Вони збираються за допомогою датчиків, що встановлені на автомобілі.
2. Дані про рух автомобіля: ці дані включають інформацію про швидкість, пройдену відстань, час руху, маршрут і позицію автомобіля. Вони можуть бути зібрані за допомогою GPS-трекерів, які встановлюються в автомобілях.
3. Дані про водіїв: ці дані включають інформацію про водіїв, включаючи їх особисті дані, вік, стаж водіння, інформацію про аварії та правопорушення. Їх можна зібрати від клієнтів під час реєстрації в сервісі.
4. Дані про клієнтів: ці дані включають інформацію про клієнтів, включаючи їх контактні дані, інформацію про автомобілі, що були обслуговані, інформацію про оплату послуг. Ці дані збираються під час реєстрації клієнта в сервісі та під час обслуговування автомобіля.
5. Дані про сервіс: ці дані включають інформацію про виконані роботи, використані запчастин дати і типи проведених ремонтів та замінені запчастин, технічні характеристики автомобіля, та інші подібні дані.

З цими даними можна працювати, використовуючи Big Data і аналітику, щоб допомогти автомобільним компаніям розуміти поведінку своїх клієнтів, виявляти проблеми з автомобілями та забезпечувати більш ефективний сервіс та ремонт автомобілів.

Збирання та аналіз даних, що виникають під час експлуатації автомобілів, є дедалі більш важливими для автомобільної промисловості. Очікується, що в майбутньому велика кількість даних буде зібрана з сенсорів та інших пристроїв, що забезпечують зв'язок між автомобілем та зовнішнім середовищем, що забезпечить автомобільним компаніям ще більше можливостей для аналізу та покращення своїх послуг та продуктів.

Тенденції розвитку використання технологій Big Data в автомобільній промисловості включають:

1. Збільшення кількості сенсорів та інших пристроїв збору даних в автомобілях, що призводить до збільшення обсягу даних, які можна зібрати з кожного автомобіля.
2. Використання штучного інтелекту та машинного навчання для аналізу даних та виявлення закономірностей в автомобільній індустрії. Це може допомогти підвищити ефективність та надійність автомобілів, зменшити витрати на обслуговування та ремонт, а також покращити безпеку водіїв та пасажирів.
3. Розробка нових сервісів та продуктів, які базуються на даних, зібраних з автомобілів. Наприклад, системи навігації, які пропонують маршрути з оглядом на затори та інші перешкоди, або системи діагностики, які можуть попередити про можливі поломки до того, як вони стануть серйозними.
4. Розвиток екосистеми даних, яка дозволяє обмінюватись даними між автомобілями та іншими системами, такими як міські інфраструктури та інші автомобілі. Це може допомогти покращити безпеку на дорогах та зменшити затори, забезпечити кращу інтеграцію з громадським транспортом та іншими видами мобільності.

У цілому, використання технологій Big Data в автомобільній промисловості може допомогти підвищити ефективність, зменшити витрати та покращити безпеку на дорогах. Однак, разом з цим виникають нові проблеми збору та захисту даних, деякі з яких стають принциповими, а саме:

1. Проблема захисту конфіденційності даних: збір та зберігання великих даних може включати особисті та конфіденційні дані, які можуть бути використані для ідентифікації та дискримінації користувачів. Захист конфіденційності даних може включати шифрування даних та забезпечення доступу до них тільки з обмеженим колом осіб.
2. Проблема якості даних: великі дані можуть бути неповними, неточними та недостовірними. Це може бути спричинено помилками при зборі даних, недостатньою кількістю джерел даних та проблемами з їхньою обробкою. Розробка та використання методів та алгоритмів, які можуть виправляти та покращувати якість даних, можуть допомогти вирішити цю проблему.
3. Проблема безпеки даних: збереження великих даних може стати об'єктом кібератак, які можуть призвести до витоку даних та їхнього використання в злочинних

цілях. Для запобігання цим проблемам, необхідно використовувати різноманітні методи та інструменти забезпечення безпеки, включаючи аутентифікацію, авторизацію та шифрування даних.

Однак, крім проблем збору і захисту даних, існують і інші проблеми, пов'язані з великими даними. Нижче наведено кілька з них:

1. Проблема інтерпретації даних. При обробці великих обсягів даних, дуже важливо забезпечити правильну інтерпретацію даних, тому що навіть маленька помилка може призвести до невірних висновків і великих втрат.
2. Проблема завантаження та зберігання даних. Великі обсяги даних можуть займати значні обсяги пам'яті, що може призвести до зниження продуктивності системи. Також, зберігання даних може вимагати значних витрат на інфраструктуру і обслуговування.
3. Проблема безпеки даних. Якщо дані зберігаються в хмарі, то вони можуть бути викрадені зловмисниками. Тому важливо забезпечити надійний захист від несанкціонованого доступу до даних.
4. Проблема релевантності даних. Іноді великі обсяги даних можуть бути непотрібними, або нерелевантними, що може призвести до перевантаження системи.
5. Проблема етики. Іноді збір та використання великих даних може викликати питання етики, особливо, якщо вони містять конфіденційну інформацію про людей або їх поведінку.

Узагальнюючи, можна сказати, що збір та обробка великих даних в автомобільній промисловості є складним процесом, який потребує високих технічних знань і великих фінансових витрат.

У загальному, стаття описує важливість збору та аналізу великих даних в автомобільній промисловості, а також вказує на складнощі, пов'язані зі збором і захистом цих даних. Розглянуто основні види даних, які збираються в автомобільному сервісі, включаючи технічну та геолокаційну інформацію, дані про користувачів та їх поведінку. Описано основні принципи роботи з великими даними, такі як горизонтальна масштабованість, відмовостійкість та локальність даних. Захист великих даних від несанкціонованого доступу та зловживань є важливою проблемою, яку слід вирішувати в автомобільній промисловості. Для ефективної роботи з великими даними необхідні високопродуктивні технології та підходи. Зрештою, використання великих даних може допомогти автомобільній промисловості у покращенні процесів виробництва, обслуговування та розробки нових технологій.

Список використаних джерел

- [1]. С. Lynch, "How do your data grow? ". *Nature* 455, 28–29 (2008).
<https://doi.org/10.1038/455028a>

- [2]. Apache Hadoop. URL: <https://hadoop.apache.org/>_(дата звернення 29.04.2023)
- [3]. Apache Spark. URL: <https://spark.apache.org/>_(дата звернення 29.04.2023)
- [4]. Big Data. URL: <https://www.it.ua/knowledge-base/technology-innovation/big-data-bolshie-dannye> (дата звернення 29.04.2023)
- [5]. MarketsandMarkets. URL: <https://www.marketsandmarkets.com/>(дата звернення 29.04.2023)