

УДК 004.65

О. В. ПОЛОНЕВИЧ, канд. техн. наук;

О. М. ТКАЛЕНКО, канд. техн. наук, доцент;

В. Р. КОСЕНКО, канд. техн. наук;

К. В. ЖИВАГО, аспірант,

Державний університет телекомунікацій, Київ

ВИКОРИСТАННЯ ТЕХНОЛОГІЇ BIG DATA В ІТ-СИСТЕМАХ ОПЕРАТОРІВ МОБІЛЬНОГО ЗВ'ЯЗКУ

Розглянуто особливості технології Big Data. Виділено переваги та недоліки, які надає використання даної технології. Описано варіант реалізації Big Data на мережі оператора мобільного зв'язку.

Ключові слова: технологія Big Data; великі дані; обробка даних; оператор мобільного зв'язку.

ВСТУП

Майже щодня оператори мобільного зв'язку змушені відповідати на нові запити, які вимагає сучасність: упровадження нових технологій, розвиток нових цифрових послуг, підвищення швидкості та якості обслуговування. Окрім того, спостерігається стрімке падіння доходів операторів від традиційних послуг зв'язку, що за прогнозами не буде компенсоване зростанням доходів від передавання даних. При цьому вартість зберігання даних знижується, а монетизація інформації зростає. За таких умов новий тренд Big Data може стати важливою статтею доходів операторів мобільного зв'язку. Філософія Big Data передбачає новий підхід до збору, збереження, управління та аналізу масивів даних, оскільки їх очікуваний обсяг настільки великий, що обробити їх традиційними інструментами неможливо.

Аналіз даних дозволяє виявити певні і непомітні закономірності, які не може побачити людина. Необхідність упровадження Big Data зумовлюється також зміною вимог до швидкості аналізу даних та до отриманого результату. Якщо на старті розвитку data-аналітики було достатньо аналізу та звіту про події, які вже відбулися, то сьогодні, разом із оперативним аналізом, потрібен аналіз у реальному часі. Окрім того, сучасні реалії потребують прогнозування та оптимізації майбутніх подій, тобто їх передбачування, розрахунку поведінкових моделей та глобальних трендів на основі величезних масивів інформації, з яких і складається Big Data. Big Data дозволить оптимізувати всі сфери нашого життя — від державного управління до виробництва і телекомунікацій.

На вітчизняному ринку існує вже цілий пласт вирішень Big Data. Це швидше платформи, які складаються з безлічі модулів, що настроюються, індивідуальних для кожного конкретного бізнес-процесу. При цьому практично експоненціально збільшується сам склад оброблюваних операторами даних, з'являється і безліч додаткових цілей їх обробки, таких як маркетинг, надання окремих послуг, забезпечення безпеки сервісів клієнта.

ОСНОВНА ЧАСТИНА

Основи концепції Big Data

Великі дані припускають щось більше, ніж просто аналіз величезних обсягів інформації. Проблема не в тому, що організації створюють величезні обсяги даних, а в тому, що значна їх частина подана у форматі, який погано відповідає традиційному структурованому формату БД. Це веб-журнали, відеозаписи, текстові документи, машинний код або, наприклад, геопросторові дані. Все це зберігається в безлічі різноманітних сховищ, іноді навіть за межами організації. У результаті корпорації можуть мати доступ до величезного обсягу своїх даних і не мати необхідних інструментів, щоб встановити взаємозв'язок між цими даними і зробити на їх основі значущі висновки. Додайте сюди ті обставини, що дані зараз оновлюються все частіше, і ви отримаєте ситуацію, в якій традиційні методи аналізу інформації не можуть наздогнати величезні обсяги постійно оновлюваних даних, що врешті-решт і відкриває дорогу технологіям великих даних.

Отже, можна дати таке визначення: **Big Data**, або **великі дані** — це серія підходів, інструментів і методів обробки структурованих, слабо структурованих і неструктурованих даних величезних обсягів і значного різноманіття для отримання сприйнятливих для людини результатів, ефективних за умов безперервного приросту, розподілу по численних вузлах обчислювальної мережі, альтернативних традиційним системам управління базами даних і вирішень класу Business Intelligence.

Технології обробки Big Data вирішують три основні завдання: зберігання великих обсягів, структурізація розрізнених даних, швидкий і точний аналіз. У цьому і полягає специфіка їх відмінності від традиційних аналітичних систем (див. таблицю).

Як визначальні характеристики для великих даних виокремлюють «три V» — volume (обсяг), velocity (швидкість), variety (різноманіття). Під обсягом розуміють величини фізичного обсягу. Під швидкістю — як швидкості приросту, так

Порівняльна таблиця класичних баз даних і Big Data

Параметр	Класичні бази даних	Big Data
Обсяг даних	Десятки терабайт	Сотні петабайт
Тип даних	Структуровані	Будь-які довільні дані в довільній формі
Гнучкість	Стандартні зумовлені звіти обмеженої глибини	Довільні звіти будь-якого формату та глибини
Час обробки	Хвилини, часи, дні	Реальний час, секунди, хвилини
Дублювання даних	Допускається	Мінімізується всіма способами

і необхідність високошвидкісної обробки і отримання результатів. Під різноманіттям — можливістю одночасної обробки різних типів структурованих і неструктурованих даних.

Поміж основних факторів, що змінюються, у сферах можливого застосування Big Data можна виокремити такі:

- збільшення кількості різних джерел даних;
- розвиток програмних та апаратних технологій;
- покращення продуктивності обчислень;
- поява розподільної технології обробки даних.

Переваги використання Big Data операторами мобільного зв'язку

Одним із найбільш пріоритетних завдань для кожного телеком-оператора є забезпечення надійної і стабільної роботи його сервісів. В інформаційних системах операторів зв'язку вкрай велика частка неструктурованих і різноманітних даних, що надходять від телеком-обладнання, із служб моніторингу і підтримки, служб по роботі з клієнтами, різних соціальних мереж тощо. Отже, розширення Big Data і прискорення темпу зростання стає об'єктивною реальністю, а їх коректна обробка — глобальною потребою ринку в цілому, галузь телекомунікацій не виняток.

Технології обробки Big Data дають змогу в режимі реального часу здійснювати моніторинг і аналіз завантаженості мережі, виявляти нестабільну роботу устаткування і різного роду несправності. При цьому обробка тих даних, що надходять із систем моніторингу стану телеком-обладнання, дозволяє заздалегідь прогнозувати необхідне обслуговування, ремонт або заміну цього устаткування, забезпечуючи безперервність процесу.

Big Data уможливорює прогнозування навантаження на базові станції операторів у конкретний момент часу завдяки аналізу геолокаційних даних абонентів та активності використання ними сервісів. У результаті визначаються проблемні ділянки, що вимагають відповідної модернізації з метою поліпшення якості мережі.

Інша важлива перевага використання Big Data — це можливість економії. По-перше, на ремонті і заміні устаткування за рахунок прогнозування його виходу з ладу і можливості вжиття запланованих превентивних заходів. По-друге, завдяки можливості побудови власних індивідуальних графіків обслуговування і заміни, виходячи з фактичного зносу, а не з рекомендацій постачальників, які мають власний інтерес у цьому питанні.

Оператори мобільного зв'язку мають унікальний набір та охоплення даних:

- фінансовий стан абонента;
- коло спілкування;
- коло інтересів;
- інформація про переміщення;
- пристрої, що використовуються;
- сайти, що відвідуються.

Ці дані можна монетизувати, якщо навчитися їх обробляти, а також використовувати для покращення внутрішньої роботи компанії. Виокремимо три напрямки [4]:

◆ високоточний маркетинг (*precise marketing*) — адресна пропозиція продуктів та послуг тим абонентам, які найбільш готові до їх придбання (нові тарифні плани, додаткові сервіси, платіжні термінали, індивідуальні роумінгові пропозиції та ін.);

◆ управління якістю послуг для абонентів (*Customer Experience Management*) для підвищення задоволеності з метою запобігання відтоку абонентів;

◆ оптимізація внутрішньої роботи оператора і планування розвитку (*ROI-based Network Optimization and Planning*) на основі обліку всіх об'єктивних факторів і суджень абонентів для забезпечення максимальних гарантій повернення інвестицій у найкоротші терміни; тестування кращих варіантів пропозицій (A/B-тестування); аналітика і підстроювання контексту дослідження в реальному часі.

Окрім того, оператори мобільного зв'язку мають змогу монетизувати свої інформаційні активи, тобто продаж у тій чи іншій формі даних, які мають оператори, своїм партнерам, щоб вони могли з їх допомогою вирішувати свої завдання. Наприклад, збір геолокаційних даних про абонентів дає можливість:

◆ адаптувати рекламу, знаючи місцезнаходження людини чи її передбачуваного місця призначення;

◆ виявляти затори, ґрунтуючись на кількості та швидкості переміщення абонентських терміналів уздовж шосе. Інформація про місцезнаходження і швидкість переміщення можуть слугувати також для визначення вартості нерухомості або розцінок для рекламних щитів;

◆ отримувати інформацію про те, в яких районах міста «кипить» нічне життя або скільки протестувальників зібралось на демонстрації;

♦ прогнозувати поведінку людей, проаналізувавши їх переміщення і список дзвінків.

Але Big Data не просто нова послуга — її використання видозмінює більшість бізнес-процесів оператора, підіймаючи на новий рівень існуючі сервіси і дозволяючи вивести на ринок принципово нові послуги [1] (рис. 1).

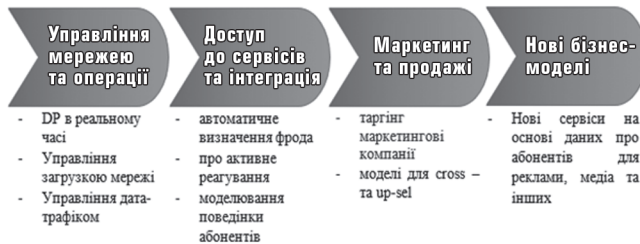


Рис. 1. Ланцюжок створення нових бізнес-моделей на основі Big Data

Реалізація концепції Big Data в мобільному зв'язку

Приклад реалізації можливостей Big Data на основі відкритого ПО Hadoop наведено на рис. 2



Рис. 2. Варіант реалізації Big Data в мережі оператора мобільного зв'язку [1]

Набір бібліотек та утиліт Hadoop включає в себе такі інструменти:

♦ CEM (Customer Experience Management) — управління клієнтським досвідом;

♦ RA (Revenue Assurance) — система «гарантованих доходів»;

♦ API (Application Programming Interface) — додаток з відкритим кодом для вбудовування в інтерфейси інших програм тощо.

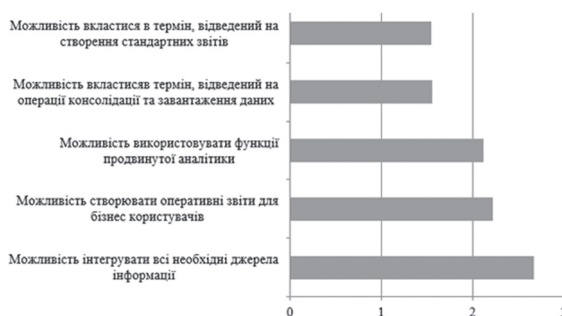


Рис. 3. Відповіді респондентів із числа телеком-операторів: 3 — найбільш важливо; 0 — найменш важливе джерело

Опитування телеком-компаній свідчать, що збираючи великі обсяги даних, телеком-компанії

далеко не повною мірою використовують ту цінність, яка в цих даних міститься. Так, дослідження компанії Heavy Reading (рис. 3) показало, що проблеми неможливості інтегрованого аналізу всіх наявних у телеком-операторів джерел інформації і недостатня оперативність аналізу цих даних — головні перепони, які постають перед компаніями при розвитку систем оперативної аналітики.

Основні проблеми Big Data

Попри величезні можливості цієї технології не слід забувати про проблеми, які з'являються з розповсюдженням використання Big Data.

Персональні дані та їхня недоторканість. Упровадження Big Data ставить під загрозу приватність користувачів, що може негативно позначитись на особистій безпеці. Паралельно необхідно додержуватись правильного балансу між бізнес-процесами та дотримуватись законодавства щодо обмеження і контролю доступу до особистих даних. Великі корпорації можуть маніпулювати цими даними, роблячи нас певною мірою своїми заручниками. Навіть якщо припустити, що Big Data буде використовуватися лише в шляхетних цілях, немає ніякої гарантії, що персональні дані зможуть бути надійно захищені від зловмисників та хакерів.

Дискримінація. Коли все відомо завдяки Big Data, окремих людей можуть обмежувати, наприклад, у доступі до банківських кредитів через можливу недостатню надійність чи переплачувати за медичне страхування через несприятливі умови проживання в певних місцевостях. Або навіть державні структури та приватні компанії захочуть обмежити нас у доступі до певних сервісів та ресурсів.

Складність інтеграції з існуючими системами. Питання про зберігання великих обсягів інформації пов'язане з необхідністю організації певних умов, тобто зі створенням простору і можливостей. Що стосується швидкості, то вона пов'язана не стільки з уповільненням і гальмуванням при використанні застарілих методів обробки, скільки з інтерактивністю — результат тим продуктивніший, чим швидше відбувається процес обробки інформації.

Проблема неструктурованості зумовлена **роздільністю джерел, їх форматом і якістю**. Для успішного об'єднання і обробки Big Data потрібна робота з їх підготовки, аналітичні інструменти або системи.

Великий вплив має і **межа «кількості» даних**. Визначити обсяг досить складно, а з огляду на це проблематично прорахувати, в яких фінансових вкладеннях і необхідних технологіях постане потреба. Проте для певних величин, наприклад терабайт, сьогодні успішно застосовуються постійно вдосконалювані нові методи обробки.

Відсутність загальноприйнятих принципів роботи з Big Data — ще одна проблема, яка ускладнюється згаданою раніше неоднорідністю потоків. Для її вирішення створюються нові методи аналізу Big Data.

Складнощі також викликають підбір даних для аналізу і алгоритм дій. Сьогодні немає чіткого розуміння того, які дані мають цінну інформацію і вимагають аналітики Big Data, а якими можна знехтувати. Також на ринку недостатньо професіоналів, які впораються з глибинним аналізом, сформулюють звіт про вирішення завдання і, відповідно, дадуть прибуток.

З огляду на це виокремимо основні проблеми в процесі впровадження проектів Big Data (рис. 4).

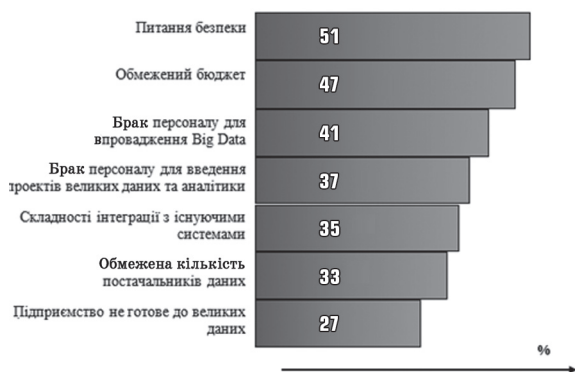


Рис. 4. Основні проблеми під час впровадження проектів Big Data

ВИСНОВКИ

На Big Data чекає велике майбутнє, і той, хто навчиться збирати та аналізувати величезні масиви даних так, щоб на їх основі передбачати появу тих

чи інших подій, відкриє додаткове джерело монетизації великих даних. Це, у свою чергу, може дати унікальні конкурентні переваги, надати нові можливості людям і, зрештою, зробить життя людей більш зручним. Тим не менш, слід зазначити, що розвиток і впровадження технологій Big Data має супроводжуватись удосконаленням підходів в області інформаційної безпеки і конфіденційності даних абонентів.

У статі систематизовано інформацію про концепцію Big Data, розглянуто основні можливості, які надає її реалізація на мережах операторам мобільного зв'язку. Сформульовано основні проблеми, які стоять перед системами зберігання масивних даних.

Запропонований варіант реалізації Big Data на мережі оператора мобільного зв'язку може бути використано як основа для реалізації аналогічних вирішень і проведення подальших досліджень.

Список використаної літератури

1. Степутин А. Н., Николаев А. Д. Мобильная связь на пути к 6G / Инфра-Инженерия. Москва–Вологда, 2017. С. 224–238.

2. Фрэнкс Б. Революция в аналитике. Как в эпоху Big Data улучшить ваш бизнес с помощью операционной аналитики. Москва, 2017. 320 с.

3. Сагинов В. Л. Big Data как один из ключевых драйверов развития существования ИКТ отрасли // Междунар. бизнес-форум «Big Data: технологии и бизнес» 15 октября 2015, Москва.

4. Колесов А. Что могут дать заказчику большие данные? URL:

<http://www.interface.ru/home.asp?artId=37711>

Рецензент: доктор техн. наук, доцент **К. П. Сторчак**, Державний університет телекомунікацій, Київ.

О. В. Полоневич, О. М. Ткаленко, В. Р. Косенко, К. В. Живаго

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ BIG DATA В ИТ-СИСТЕМАХ ОПЕРАТОРОВ МОБИЛЬНОЙ СВЯЗИ

Рассмотрены особенности технологии Big Data, выделены преимущества и недостатки, которые предоставляет использование данной технологии. Описан вариант реализации Big Data на сети оператора мобильной связи.

Ключевые слова: технология Big Data; большие данные; обработка данных; оператор мобильной связи.

O. V. Polonevych, O. M. Tkalenko, V. R. Kosenko, K. V. Zhyvaho

THE USING OF THE BIG DATA TECHNOLOGY IN IT-SYSTEMS OF THE MOBILE COMMUNICATION OPERATORS

In the article the peculiarities of the Big Data technology are examined, it is singled out the advantages and disadvantages that the using of this technology gives. The variant of the implementation of the Big Data on the mobile operator's link is described.

Keywords: Big Data technology; big data; data processing; mobile operator.