

КЕШУВАННЯ ВЕЛИКИХ ДАНИХ В БЕЗПРОВОДОВИХ МЕРЕЖАХ 5G

Малецький Д.В., Міночкін Д.А.

*Інститут телекомунікаційних систем КПІ ім. Ігоря Сікорського, Україна
E-mail: dmytro.maletskiy@gmail.com*

Big Data caching in wireless 5G networks

To cope with the data traffic growth new approaches to the networking should be implemented. The proactive caching architecture for 5G wireless networks is considered as a possible solution for improvements in user's quality-of-experience and the backhaul offloading problem.

Сучасні підходи такі, як розширення спектру, розгортання більшої кількості базових станцій (БС) та збільшення кількості вузлів в базовій мережі передачі даних вже не будуть актуальними з точки зору масштабованості, вартості та гнучкості для того, щоб впоратися з постійно зростаючими об'ємами даних в безпроводових мережах 5-го покоління. Такий безпрецедентний ріст інформаційного трафіку, що більшою мірою спричинений мобільним відео, соціальними медіа та різноманітними застосунками змушує мобільних операторів шукати інноваційні способи управління мережами, що постійно ускладнюються, та обмеженими ресурсами транспортних мереж. Необхідно розглянути перехід від нинішньої архітектури мобільних мереж до нової парадигми, що базуватиметься на збиранні та зберіганні інформації в дата центрах для її подальшого аналізу та прийняття рішень.

Рішенням, що може принести значні здобутки для мобільних операторів, є контекстно-усвідомлені мережі з хмарними або edge обчисленнями та використання аналізу великих даних.

На сьогоднішній день в [1] досліджено можливість використання великих даних в мобільних стільникових мережах з точки зору проактивного кешування. Враховуючи високу передбачуваність людської поведінки та великі кількості даних, що проходять через мережі операторів зв'язку, запропоновано архітектуру проактивного кешування даних. Ця архітектура оптимізує безпроводові мережі 5G, де велика кількість доступної інформації опрацьовується шляхом використання аналізу великих даних та засобів машинного навчання для оцінки популярності контенту. Нова архітектура може бути використана для edge кешування (на базовій станції), що призведе до підвищення рівня задоволення користувача та розвантаження транспортної мережі шляхом переносу контенту «ближче» до користувачів.

В прагненні забезпечити кращу масштабованість та гнучкість мереж, щодо обробки великих даних, вимоги до побудови мереж стають все більш залежними від програмного забезпечення. Майбутні великі мережі ставатимуть, ще складнішими, саме тому дата центрам та мережевій інфраструктурі мобільних операторів буде необхідно відслідковувати шаблони трафіку десятків мільйонів користувачів використовуючи різноманітні засоби збору статистичних даних користувачів (місцезнаходження, шаблони запиту на

трафік і т.п.) для більш детального аналізу.

Актуальним питанням постає управління мережами, керованими великими даними, в хмарному середовищі, так як зростання кількості даних досі залишається головним викликом для сучасної мобільної інфраструктури. В цьому світлі виникає технологія мобільних edge обчислень, у якій edge пристрої забезпечують «хмарні обчислення як можливість» в середині мережі радіо доступу, для виконання функцій зв'язку, зберігання та контролю. [2] Однак, для мереж 5G необхідно відмітити, що розгортання розподіленої хмарної інфраструктури біля кожної БС може призвести до зростання вартості в порівнянні з централізованим рішенням, враховуючи наявність сотень сайтів в мережі мобільного оператора. Більше того, для моделювання та передбачення просторово-часової поведінки користувача в мережах 5G, що орієнтовані на користувача, централізоване отримання мережевого трафіку має бути горизонтально масштабованим між серверами, що можливо лише в середині головного сайту для детального аналізу.

Очевидно, що єдиний шлях впоратися зі зростаючим мережевим трафіком даних, лежить через покращення управління даними та переносу даних з хмари на «край». За останні роки, вільна програмна платформа Nadoop успішно продемонструвала себе, як програмне рішення управління великими даними, що пропонує значне заощадження коштів в порівнянні з однорівневими інфраструктурами баз даних, можливість обробки різних форматів даних та паралельну обробку на кількох вузлах. Використання процесора обробки розподілених даних для аналізу користувацької поведінки та проактивного кешування контенту на «краю» може послабити трафік транспортної мережі та покращити якість користувацького досвіду зменшенням затримки.

Запропонована архітектура [1] збирає контекстуальну інформацію (наприклад, історію переглядів, інформацію про місцезнаходження) та передбачає просторо-часові вимоги для проактивного кешування продумано відібраного контенту на «краю» мережі. Представлена архітектура паралелізує обчислення та виконання алгоритму передбачення контенту на центральному сайті та розміщення кешу на БС. Таким чином, користувацькі потреби задовольняються значною мірою, за рахунок зменшення затримки та підвищення якості досвіду.

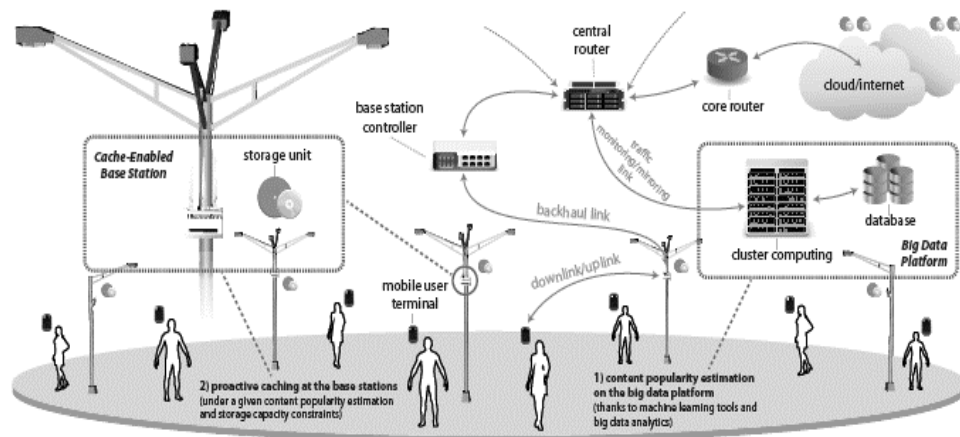


Рис. 1. Архітектура мережі.

На рис. 1 зображено архітектуру мережі, де платформа великих даних розгорнута на центральному сайті, для відслідковування/передбачення користувацьких запитів, в той час, як БС з можливістю кешування зберігають «стратегічний» контент передбачений платформою великих даних.

Зменшення навантаження транспортної мережі через edge кешування є значним викликом. В цьому сенсі, при розгляді розмірів контенту, вимог до швидкості передачі, транспортної мережі, тощо, необхідна об'єднана оптимізація матриці популярності контенту та розташування кешу контенту в певних малих стільниках. Крім того, під час прийняття рішення про кешування необхідно враховувати обмежену ємність сховищ маленьких базових станцій, безпроводових та транспортних з'єднань, великі розміри бібліотек та значну кількість користувачів з невідомими параметрами (наприклад, емпіричними значеннями популярності контенту). Виходячи з того, що розташування кешу може бути визначене у жорсткий чи наближений способи [3], маленькі базові станції навчаються та оцінюють незначну популярність контенту.

Основне призначення платформи великих даних, зберігати користувацький трафік даних та отримувати з нього корисну інформацію для прийняття рішень проактивного кешування.

З метою прийняття рішень платформа обробки даних в середині інфраструктури базової мережі має бути спроможна зчитувати і поєднувати дані з різних джерел та виносити зважені рішення швидко та надійно. Для цього, після відтворення інтерфейсу трансляції даних через засоби аналізування мережі, зібраний ряд даних експортується до платформи зберігання великих даних такої як Hadoop Distributed File System (HDFS) за допомогою методів для детального аналізу.

Після очищення даних, проводиться їх аналіз за допомогою високо рівневих мов запитів таких, як Hive Query language (HiveQL) та Pig Latin. Ціль даного етапу знайти взаємозв'язок між управляючими та інформаційними пакетами, наприклад місцезнаходження чи MSISDN користувача та запитуваний контент.

По закінченню машинного аналізу, щоб передбачити просторово-часову поведінку користувача для рішення проактивного кешування, результати аналізу зберігаються і використовуються повторно.

Розглянуто архітектуру проактивного кешування для безпроводових мереж 5G, що виконує обробку великої кількості даних на платформі великих даних використовуючи програмну платформу Hadoop та машинне навчання для оцінки популярності контенту, як одне з рішень подолання проблеми управління базової мережі та обмеженості транспортної мережі.

Література

1. Engin Zeydan, Ejder Bastug, Mehdi Bennis, Manhal Abdel Kader, Big Data Caching for Networking: Moving from Cloud to Edge, June 2016.
2. F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in Proceedings of the first edition of the MCC workshop on Mobile cloud computing, Helsinki, Finland, August 2012.
3. M. Ji, G. Caire, and A. F. Molisch, "Wireless device-to-device caching networks: Basic principles and system performance," IEEE Journal on Selected Areas in Communications, vol. 34, no. 1, pp. 176–189, Jan 2016.