

УДК 004.67

М.А.Максимчук

МОДУЛЬ АНАЛІТИЧНОЇ ОБРОБКИ ДАНИХ ПРОГРАМНИХ ЗАСОБІВ ПІДТРИМКИ ПРОЦЕСУ ОПТИМІЗАЦІЇ ПОКРИТТЯ ОПЕРАТОРА МОБІЛЬНОГО ЗВ'ЯЗКУ

**Національний
технічний університет Ук-
раїни
«Київський політехнічний
інститут»**

**Кафедра програмного
забезпечення
комп'ютерних систем**

**Науковий керівник –
Заболотня Т.М., к.т.н.**

Дана стаття присвячена розробці модуля динамічної аналітичної обробки даних для інформаційної системи підтримки процесу оптимізації покриття оператора мобільного зв'язку.

Для вирішення цієї задачі запропоновано використання технології OLAP, котра дозволяє створювати ефективні програмні засоби обробки великих обсягів даних. Визначена структурно-алгоритмічна організація модуля обробки даних щодо показників якості роботи мережі. Модуль складається з трьох основних частин: бази даних OLAP, інтерфейсу для роботи з цією базою даних та модулів її формування. Передбачена можливість динамічної зміни структури бази даних OLAP. У статті запропонований алгоритм функціонування модуля аналітичної обробки даних, а також наведені результати дослідження ефективності його роботи.

Данная статья посвящена разработке модуля динамической аналитической обработки данных для информационной системы поддержки процесса оптимизации покрытия оператора мобильной связи.

Для решения этой задачи предложено использование технологии OLAP, которая позволяет создавать эффективные программные средства обработки больших объемов данных. Определена структурно-алгоритмическая организация модуля обработки данных относительно показателей качества работы сети. Модуль состоит из трех основных частей: базы данных OLAP, интерфейса для работы с этой базой данных и модулей ее формирования. Предусмотрена возможность динамического изменения структуры базы данных OLAP. В статье предложен алгоритм функционирования модуля аналитической обработки данных, а также приведены результаты исследования эффективности его работы.

The given article is dedicated to the development of the dynamic analytical data processing module of the software for supporting the optimization process for GSM network coverage.

The technology OLAP, which allows to create effective software to process large volumes of data, was suggested to solve this problem. The structure algorithmic organization of the analytical data processing module in relation to quality indicators of the mobile network was defined. The module consists of three main parts: database OLAP, the interface for work with this database and modules of its formation. The possibility of dynamic changing the OLAP database structure was included. The algorithm of the analytical data processing module functioning and results of research on the effectiveness of its work were presented at the article.

Ключові слова: OLAP, оптимізація покриття оператора мобільного зв'язку, якість покриття оператора мобільного зв'язку.

Вступ

Сучасна ситуація на ринку операторів мобільного зв'язку України характеризується постійно зростаючим рівнем конкуренції між компаніями за утримання існуючих та залучення нових абонентів до своїх мереж. У зв'язку з цим оператори разом із розробкою різних маркетингових заходів щодо підвищення доступності своїх послуг для абонентів вимушені приділяти максимум уваги питанню покращення якості зв'язку, що досягається, в основному, за рахунок введення нових базових станцій та зміною параметрів функціонування існуючих станцій. Звісно, більш вигідним для операторів рішенням є

оптимізація роботи вже встановлених станцій, ніж закупка та запуск нових.

Таким чином, стає зрозумілим, що оптимізація мережі та налаштування її параметрів є стратегічною задачею сучасних операторів мобільного зв'язку, успішне розв'язання якої забезпечує підвищення рівня якості зв'язку та ефективності використання коштів як операторів, так і їх абонентів, і, як ніколи, актуальна на даний момент. Серед заходів щодо проведення оптимізації мережі оператора мобільного зв'язку особливе місце посідає створення модуля динамічної обробки даних, які надходять від телекомунікаційного обладнання. На практиці реагу-

вання на зміни показників роботи стільникової мережі має бути досить швидким. Тому організація та впровадження у програмні засоби модуля, який здійснював би оперативну аналітичну обробку одержуваних даних, є актуальним питанням, а його вирішення дозволить підняти на новий рівень ефективність процесу оптимізації покриття операторів мобільного зв'язку.

Огляд існуючих рішень

У результаті проведеного огляду існуючих рішень у сфері комп'ютерного аналізу вмісту баз даних, пов'язаних з показниками роботи телекомунікаційного обладнання, було знайдено декілька програмних продуктів: Nokia NetAct Planner, автоматизовану систему RPS-2, Wireless InSite.

Nokia NetAct Planner [1] є найважливішою частиною концепції *Connecting@Care*, котра включає в себе широкий спектр рішень, що дозволяють операторам підтримувати на високому рівні якість, надійність та ефективність існуючих GSM і GPRS мереж, а також мереж третього покоління. Програмний продукт розроблений для різноманітних мереж, має сучасний дизайн і надає можливість оптимізації мережі.

Система *RPS-2* є пакетом, що використовується при проектуванні стільникових мереж [2]. Вона призначена для автоматизованого проектування мереж різноманітної архітектури, що застосовують різні стандарти передачі даних.

Ще однією системою автоматизованого проектування радіомереж та аналітичної обробки даних є програмний пакет *Wireless InSite*, розроблений компанією Remcon [3]. Цей пакет дозволяє інженерам точно оцінювати вплив фізичного оточення на роботу бездротових систем зв'язку за допомогою чисельних методів фізичної математики.

Проте, усі ці програмні продукти мають суттєвий недолік – вони не надають можливості здійснення *динамічної* аналітичної обробки даних, отриманих від мобільного оператора, а, натомість, вимагають вносити в систему для аналізу дані, які попередньо обробляються користувачем за допомогою запитів до БД, що запускаються у консолі менеджера СУБД, чи за допомогою окремих комерційних програмних продуктів, які не є загальнодоступними. При обробці великих обсягів статистичних даних на запити до БД витрачається багато часу. Таким чином, використання множини додаткових службових запитів не сприяє підвищенню ефективності програмних рішень щодо підтримки оптимізації покриття оператора мобільного зв'язку, і введення до складу останніх модуля динамічної

аналітичної обробки даних є актуальною задачею.

Постановка задачі

Виходячи з наведених вище аргументів, *метою* даної роботи є підвищення швидкодії програмного забезпечення підтримки оптимізації покриття стільникової мережі за рахунок розробки модуля динамічної обробки телекомунікаційних даних. У відповідності до поставленої мети *задачами дослідження* є:

- визначення підходу до обробки даних щодо якості покриття оператора мобільного зв'язку;
- розробка структурно-алгоритмічної організації та програмна реалізація модуля динамічної аналітичної обробки показників роботи обладнання згідно з обраним підходом;
- аналіз результатів експериментального дослідження ефективності розробленого програмного модуля.

OLAP-підхід до обробки багатовимірних даних

Для вирішення питання розробки модуля динамічної обробки телекомунікаційних даних у статті запропоновано застосувати підхід до обробки масивів даних великого обсягу, який базується на використанні результатів вивчення останніх робіт спеціалістів у сфері OLAP (Online Analytical Processing). Дана технологія оперативної аналітичної обробки даних використовує методи і засоби збору, зберігання та аналізу багатовимірних даних у цілях підтримки процесів прийняття рішень [4].

Причина використання OLAP для обробки запитів – це швидкість. Реляційні бази даних зберігають сутності в окремих таблицях, котрі зазвичай добре нормалізовані. Такий спосіб структурування даних є зручним для відносно простих баз даних, але складні багатотабличні запити у ній виконуються відносно повільно.

OLAP робить миттєвий знімок реляційної БД і структурує її в просторову модель для запитів. Заявлений час обробки запитів в OLAP складає близько 0,1% від аналогічних запитів до реляційної БД [5].

У процесі аналізу вихідних даних будуються залежності між значною кількістю різноманітних параметрів. Для аналізу цих залежностей дані подаються у вигляді багатовимірної моделі – гіперкуба (OLAP-куба).

Куб створюється на основі з'єднання таблиць із застосуванням схеми «зірки» або схеми «сніжинки» (у нашому випадку – схеми «сніжинки»). У центрі схеми знаходиться таблиця фактів, яка містить ключові факти, за якими виконуються запити. Таблиці з вимірами приєднуються до таблиці фактів. Ці таблиці показують,

як можуть аналізуватися агреговані реляційні дані. Кількість можливих агрегацій визначається кількістю способів, якими первинні дані можуть бути ієрархічно відображені. OLAP-куб містить у собі базові дані та інформацію про виміри (агрегати). Куб потенційно містить всю інформацію, яка може знадобитися для відповідей на будь-які запити. Через величезну кількість агрегатів повний розрахунок проводиться тільки для деяких вимірювань, для інших вимірювань обчислення виконують «за вимогою». У нашому випадку осі куба є вимірами, на яких відкладаються параметри, що описують характеристики покриття оператора мобільного зв'язку. На перетині осей вимірів знаходяться вихідні дані. Під час роботи модуля над кубом виконуються операції, метою яких є статистичний аналіз цих даних.

Структура модуля аналітичної обробки даних

Структура розроблюваного модуля динамічної аналітичної обробки даних складається з трьох основних частин: бази даних OLAP, інтерфейсу для роботи з цією базою даних та модулів її формування (див. рис.1). Джерелом інформації для програмного забезпечення є база даних MSSQL, яка розміщується на сервері SQL Server 2008. Передбачена можливість динамічної зміни структури бази даних OLAP.

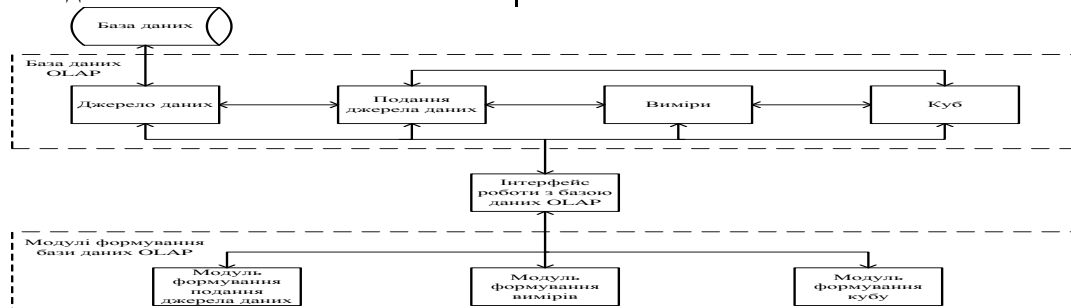


Рис. 1 Структурна схема модуля аналітичної обробки даних

Подання джерела даних отримується шляхом побудови схеми на основі потрібних таблиць, що знаходяться у базі даних, та віртуальним перенесенням потрібних даних з тих таблиць до подання джерела даних. Також створюються первинні ключі таблиць та зв'язки між таблицями. Після створення подання джерела даних створюються *виміри* й *OLAP-куб*. Осі куба є вимірами, на яких відкладаються параметри, що описують характеристики покриття оператора мобільного зв'язку. Виміри: дата, соти, контролери, регіони. На перетині осей вимірів знаходяться вихідні дані (з вхідної таблиці фактів) – міри та обчислювані міри (на основі вхідних даних). *Інтерфейс роботи з базою даних OLAP* виконує задачі ство-

Для роботи з базою даних MSSQL створено інтерфейс доступу до даних завдяки засобам технології ADO на платформі .Net.

Розглянемо основні частини модуля аналітичної обробки даних більш докладно.

База даних OLAP складається з джерела даних, подання джерела даних, вимірів і кубу.

Джерело даних отримується шляхом встановлення з'єднання з базою даних MSSQL.

Подання джерела даних – це логічна модель даних зв'язаних таблиць, уявлень і запитів з одного джерела даних, на основі якої розроблений проект оперативної аналітичної обробки (OLAP) інтелектуального аналізу даних на сервері Microsoft SQL Server. Ця модель є об'єктом, який містить метадані з вибраних об'єктів джерела даних, включаючи зв'язки між цими об'єктами, визначені в базі даних MSSQL. Подання джерела даних кешує метадані з джерел даних, на основі яких воно побудовано. Кешовані метадані дозволяють розробити проект служб Analysis Services без безперервного активного з'єднання з джерелом даних. Подання джерела даних дозволяє визначати підмножину даних, що заповнюють велике сховище даних. Крім цього воно дозволяє визначити однорідну схему на основі джерела даних.

рення або зміни бази даних OLAP-продукту Analysis Services. До *модулів формування бази даних OLAP* входять: модуль формування подання джерела даних, модуль формування вимірів та модуль формування кубу.

Алгоритм роботи модуля динамічної аналітичної обробки даних

Задачею розробленого програмного продукту є не тільки створення OLAP-кубу для підтримки оптимізації покриття оператора мобільного зв'язку, а й передбачення того, що параметри якості роботи мережі та параметри, що характеризують функціонування обладнання, можуть змінюватися з часом. Отже, програмний продукт повинен забезпечувати реагування на динамічну зміну структури вхідної бази, звідси

впливає, що OLAP-куб повинен динамічно змінюватись. Для цього розроблений узагальнений алгоритм роботи програми (див. рис. 2).

[Крок 1] – під'єднання до серверу;

[Крок 2] – перевірка, чи існує розроблювана база даних OLAP на сервері, якщо не існує, то

виконується процедура створення нової бази даних OLAP ([Крок 3]), якщо існує, то треба оновити дані, які могли оновитися у вхідній базі даних MSSQL, отже виконується [Крок 4];

[Крок 3] – створення нової бази даних OLAP;

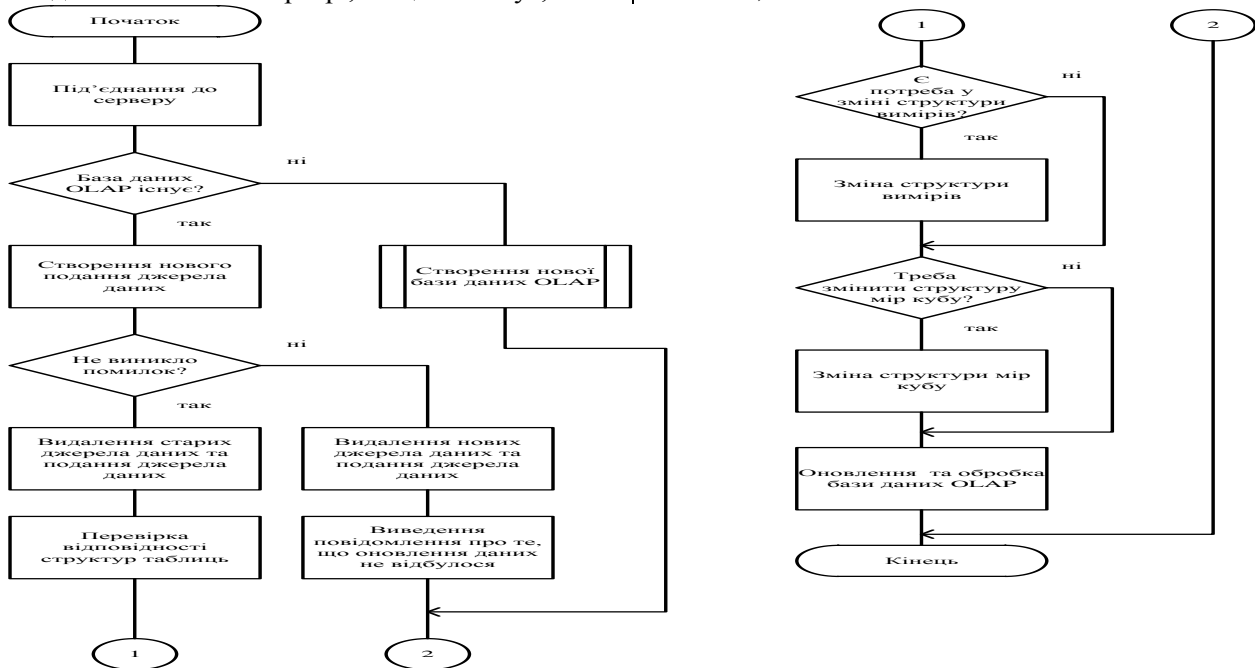


Рис. 2 Алгоритм роботи модуля аналітичної обробки даних

[Крок 4] – створення нового подання джерела даних на основі оновленої вхідної бази;

[Крок 5] – перевірка, чи є помилки на попередньому кроці (могла не створитися схема через невідповідність типів ключових стовпців або ж оновлені дані могли бути некоректними), якщо ні – старе джерело даних та подання можна видалити ([Крок 6]), якщо так – дані оновити неможливо, залишаємо у базі даних OLAP старі дані, при цьому нове подання джерела даних треба видалити ([Крок 7]);

[Крок 6] – видалення старих джерела даних та подання джерела даних;

[Крок 7] – видалення нових джерела даних та подання джерела даних;

[Крок 8] – виведення повідомлення про те, що дані не оновилися, а отже користувач може отримати результати обробки лише старих даних, виконання алгоритму закінчується;

[Крок 9] – дані оновилися, отже могла змінитися структура таблиць вхідної бази – виконується така перевірка;

[Крок 10] – перевірка, чи є необхідним змінити структуру вимірів, якщо треба (з'явилися нові чи зайві атрибути вимірів) – виконується [Крок 11], якщо ні – [Крок 12];

[Крок 11] – зміна структури вимірів;

[Крок 12] – дані оновилися, отже могла змінитися кількість параметрів якості роботи мережі покриття мобільного оператора, отже може бути необхідним змінити структуру мір OLAP-кубу, якщо треба – виконується [Крок 13], якщо ні – [Крок 14];

[Крок 13] – зміна структури мір кубу;

[Крок 14] – для збереження усіх змін у базі даних OLAP виконується оновлення та обробка цієї бази даних.

Результатом виконання даного алгоритму є сформований OLAP-куб, що містить усі параметри якості роботи мережі оператора мобільного зв'язку, усі характеристики обладнання мережі, а також є готовим для розгляду та виконання запитів за допомогою графічного інтерфейсу.

Отже, стислий алгоритм функціонування даного модуля виглядає так:

- Створення подання джерела даних, іменованих обчислень, ключів.
- Створення структури куба (мір, вимірів, атрибутів вимірів, зв'язків атрибутів, ієрархій атрибутів вимірів, правил агрегації та фільтрації).
- Виконання розгортки та обробки куба.
- Отримання за допомогою MDX-запитів проаналізованих даних.

Також в алгоритмі роботи даного модуля передбачено, що структура таблиць вхідних

даних та правила розрахунку показників можуть змінюватися. У результаті такої зміни в алгоритмі передбачена підпрограма, яка виконує зміну структури бази даних OLAP.

Аналіз ефективності роботи модуля

Програмна реалізація модуля аналітичної обробки даних щодо показників якості покриття оператора мобільного зв'язку виконана на базі Microsoft SQL Server 2008 (що включає в себе Microsoft Analysis Services). Для визначення ефективності роботи модуля був проведений порівняльний аналіз часу обробки даних за допомогою модуля та запитів у консолі менеджера СУБД. При тестуванні використовувався комп'ютер з такими характеристиками:

- тип центрального процесору Mobile DualCore Intel Core 2 Duo T5670, 1800 MHz (9 x 200);
- системна пам'ять 2039 МБ.

Обсяг таблиці фактів бази даних MSSQL – 2 мільйони записів. За кожного експерименту виконувався новий тип запиту. Результати експериментів наведені у табл. 1.

Таблиця 1 Порівняння часу обробки даних

Номер експерименту	Запити у менеджері СУБД	Модуль аналітичної обробки даних
1	150 с	10 с
2	347 с	13 с
3	294 с	9 с
4	503 с	20 с
5	228 с	11 с

У результаті проведених тестувань визначено, що даний модуль аналітичної обробки даних обробляє інформацію у десятки разів швидше. Отже, безперечно модуль є ефективним для вирішення задачі оперативної аналітичної обробки великих обсягів даних.

Висновки

У даній статті розглянуто питання підвищення швидкодії програмного забезпечення підтримки оптимізації покриття мобільного

оператора. Визначено основні недоліки існуючих програмних рішень цієї задачі.

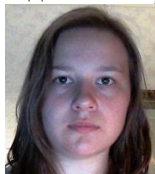
Запропоновано використання технології OLAP для забезпечення динамічної аналітичної обробки великих обсягів даних, які надходять від телекомунікаційного обладнання. Визначено структурно-алгоритмічну організацію програмного модуля аналітичної обробки таких даних. Дослідження на практиці швидкодії розробленого програмного забезпечення довели доцільність його впровадження у відповідні комплекси автоматизованої аналітичної обробки показників якості роботи стільникової мережі. Аналіз одних і тих самих баз даних з використанням OLAP-технології та без її використання показав у першому випадку зростання швидкості отримання статистичних даних.

Таким чином, використання розробленого модуля дозволяє відчутно підвищити швидкість роботи інженерів з вихідними даними, а отже, сприяє підвищенню рівня ефективності процесу оптимізації покриття операторів мобільного зв'язку. Отримані результати можуть бути використані у подальшому для створення засобів статистичної обробки даних у різноманітних предметних областях.

Список використаних джерел

1. Nokia NetAct™ Planner. For effective Network Planning [Електронний ресурс]. — Режим доступу : http://sysdoc.doors.ch/NOKIA/nokia_netact_planner.pdf
2. RPS2: Radio Planning System 2 [Електронний ресурс]. — Режим доступу : <http://www.rps2.ru/>
3. Wireless InSite® [Електронний ресурс]. — Режим доступу : <http://www.remcom.com/wireless-insite>
4. Бергер, А.Б. Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных [Текст] / А.Б. Бергер, И.В. Горбач. — СПб : БХВ-Петербург, 2007. — 529 с.
5. Федоров, А. Введение в OLAP-технологии Microsoft [Текст] / А. Федоров. — М. : Диалог-МИФИ, 2002. — 378 с.

Відомості про авторів



Максимчук Марина Анатоліївна – студентка факультету прикладної математики Національного технічного університету України «Київський політехнічний інститут», кафедра програмного забезпечення комп'ютерних систем, коло наукових інтересів: розроблення методів програмування, надійність та якість програмних систем.

E-mail: marinamaksimchuk@rambler.ru.



Заболотня Тетяна Миколаївна, к.т.н., старший викладач кафедри програмного забезпечення комп'ютерних систем Національного технічного університету України «Київський політехнічний інститут», коло наукових інтересів: об'єктно-орієнтоване програмування, розвиток теорії алгоритмів та обчислень, розроблення методів програмування для динамічних середовищ та агентно-орієнтованого програмування

E-mail: tatiana104@yandex.ru