

ШРАМЧЕНКО Б.Л., АХМАТОВ В.В.

ОБРОБКА ІЄРАРХІЧНИХ ДАНИХ У РЕЛЯЦІЙНІЙ БАЗІ ДАНИХ

SHRAMCHENKO B.L., AHMATOV V.V.

PROCESSING OF HIERARCHICAL DATA IN A RELATIONAL DATABASE

Storing and using hierarchical data structures in relational databases was not the main purpose of their creation. Therefore, SQL databases do not offer a standard way to work with such data. There are special databases for working with graphs. For example Gremlin, SPARAQL, Cypher. But if in the project there is no possibility to use them and it is necessary to work with hierarchical structures through SQL, certain methods which allow to solve this problem are applied.

The purpose of the study is to find an improved method of working with hierarchical structures, which will be both fast and intuitive. Investigate and compare different methods and algorithms, determine the speed of these methods and the influence of the depth of the hierarchical structure on the speed. Develop an application for testing methods of working with different tree structures in SQL.

Object and subject of the research is the features of methods and algorithms that allow to solve various problems of working with hierarchical data in SQL. The subject of research is to determine the pros and cons of each method, choose the best and test the performance of each method.

Scientific novelty and practical significance of the obtained results are included in working with hierarchical data in relational databases is common in many applications, so finding the best method will solve most of the problems faced by software developers.

Keywords: hierarchical data, relational database, management system, database requests, graphic interface.

Вступ

Зберігання та використання ієрархічних структур даних в реляційних базах даних не було головною метою їхнього створення. Тому SQL бази не пропонують стандартного способу роботи з такими даними. Для роботи з графами існують спеціальні бази даних. Наприклад Gremlin, SPARAQL, Cypher. Але якщо в проєкті немає можливості їх використовувати і потрібно працювати з ієрархічними структурами через SQL, то застосовують певні способи, які дозволяють вирішити цю проблему.

Мета дослідження - знайти удосконалений метод роботи з ієрархічними структурами, який буде водночас швидким та інтуїтивно зрозумілим. Дослідити та порівняти різні методи та алгоритми, визначити швидкодію цих методів та вплив глибини ієрархічної структури на швидкодію. Розробити додаток для тестування методів роботи з різними деревовидними структурами в SQL.

Постановка завдання

Об'єкт та предмет дослідження. Об'єктом дослідження є особливості методів та алгоритмів, які дозволяють вирішувати різні

проблеми роботи з ієрархічними даними в SQL. Предметом дослідження є визначення плюсів та мінусів кожного методу, вибір найкращого та тестування швидкодії кожного методу.

Методи та засоби дослідження. Для визначення оптимального методу потрібно порівнювати такі критерії: швидкість роботи, складність вставки нового вузла, переміщення вузла в структурі, підтримка цілісності даних, видалення вузла.

Наукова новизна та практичне значення отриманих результатів. Робота з ієрархічними даними в реляційних базах даних поширена в багатьох застосунках, тому знаходження найкращого методу вирішить велику частину проблем, з якими зустрічаються розробники програмного забезпечення.

Основна частина

Найпопулярнішим методом для зберігання ієрархічних структур є метод Adjacency list [2]. Цей метод дозволяє зручно та наочно зберігати дані, не має обмежень за висотою дерева, що представляє ієрархічну структуру, дозволяє оптимізовано вставляти, видаляти та змінювати позицію вершини дерева, але призводить до проблеми, що висловлюється у необхідності використовувати рекурсивні запити, які підтримують далеко не всі СУБД. Для роботи з такими даними краще за все використовувати СУБД з підтримкою рекурсивних запитів, наприклад PostgreSQL. Якщо використовується база даних без рекурсивних запитів, як MySQL, доведеться використовувати тимчасові таблиці та процедури.

Метод Nested sets використовує модель вкладених множин і являє собою техніку для представлення дерев в реляційних базах даних [3]. Ідея цього методу полягає в зберіганні маршруту обходу дерева у префіксному порядку. При цьому спочатку обходять корінь дерева, потім, вузли лівого піддерева у префіксному порядку, і потім, вузли правого піддерева. Порядок обходу зберігають в `left_key` та `right_key`. У поле `left_key` записують порядковий номер при вході в піддерево, а в `right_key` - порядковий номер при виході з піддерева. Завдяки цьому можна обійтись без рекурсії.

Метод вкладених множин працює набагато швидше попереднього і дозволяє без рекурсії робити вибірки. Але він має суттєвий недолік, який полягає у необхідності додаткового обходу всього дерева при додаванні нової вершини або при переміщенні існуючої. Наприклад, при додаванні нової вершини в нижній рівень дерева, необхідно оновлювати всі поля `left_key` та `right_key` у всіх вершинах, які знаходяться правіше та вище тієї, що додається.

Метод Closure table передбачає використання дерева, що представляє собою вкладені підмножини, кореневий вузол включає в себе всі підмножини першого рівня, які в свою чергу включають в себе підмножини другого рівня і так далі. Для зберігання даних таким способом потрібні дві таблиці. В першу таблицю треба записати всі підмножини, в другу - список входжень кожної підмножини у батьківську та рівні підмножин.

Переваги цього методу - це швидкі та прості запити, які виконуються без рекурсії.

Недоліки:

- дані зберігаються у вигляді, який не є інтуїтивно зрозумілим, його сприйняття ускладнене зв'язками між непрямыми потомками;
- необхідність використання тригерів для підтримки цілісності даних;
- операції переміщення та вставки вершин набагато складніші, ніж у методі Adjacency list.

Метод матеріалізованих шляхів, який називають ще Path enumeration або Materialized path. полягає в зберіганні повного шляху від вершини до вузла у прямому вигляді. Це найбільш інтуїтивно зрозумілий метод, який дозволяє без проблем проаналізувати структуру даних, створену з ним. Запити при використанні цього методу робити досить легко, але це впливає на швидкодію, оскільки пошук часто проходить за підрядком. Перевага вищеприписаного методу - інтуїтивно зрозумілий вигляд структури даних.

Недоліки методу:

- складна вставка та переміщення вершини;
- низька швидкодія;
- складність реалізації підтримки цілісності даних.

Висновки

Документо орієнтовані системи керування базами даних дозволяють за допомогою стандартної функціональності зручно та оптимізовано зберігати ієрархічні структури даних, але розробникам часто доводиться працювати з реляційними базами даних і не завжди є можливість використовувати документо орієнтовані. Тому були проаналізовані певні методи та алгоритми для роботи з такими даними в реляційних базах. Кожний метод має свої переваги та недоліки в певних ситуаціях. Тому вибір методу залежить від проблеми, яку вирішує програмний продукт. Якщо потрібно зберігати дані в інтуїтивно зрозумілому вигляді, а швидкість операцій не грає великої ролі, то правильним вибором буде

Materialized path. У випадку, коли потрібна робота з великою кількістю вкладених даних найкраще відповідає цій потребі метод підмножин Closure Table. Найбільш універсальним методом виявився список суміжних вершин, оскільки він не має необхідності в підтримці цілісності даних, вставка і переміщення не змінюють інші записи в таблиці, але для вибірки даних він потребує рекурсії.

Література

1. Joe Celko's - Trees and Hierarchies in SQL for Smartie 2012. – 257 с.
2. Python Data Structures and Algorithms - Benjamin Бака 2017. - 172 с.
3. Ієрархічна модель вкладених множин у реляційних базах даних - Б. Голуб 2010. - 107 с.

ШРАМЧЕНКО Б.Л., КОРОГОД Г.О., РАДЧУК А.Д.

МОДЕЛЮВАННЯ ПОВЕРХОНЬ ОДЯГУ МЕТОДАМИ КОНСТРУКТИВНОЇ ГЕОМЕТРІЇ ІЗ ЗАСТОСУВАННЯМ ГРАФІЧНИХ КЛЮЧІВ ПРОПОРЦІЙНОСТІ

SHRAMCHENKO B.L. KOROGOD G.O., RADCHUK A.D.

CLOTHS SURFACE SIMULATION BY CONSTRUCTIVE GEOMETRY METHODS WITH APPLICATION GRAPHICAL KEYS OF PROPORTIONALITY

Creation of software for the automated model building surface wear as the frame (of two mutually orthogonal lines belonging to the surface) is the subject of consideration in this work. Modern methods of obtaining initial data for the design of light industry and, in particular, clothes allow you to use not only the numerical values of dimensional attributes, but some form of curves on the surface of the human body. So the challenge presented by the transition curve to the surface of the future product, or building surface known curves belonging to the desired surface, while maintaining some smoothness of the surface.

To achieve the objective formulated such problems have been solved. The analysis graphic-plastic and surf-graphical keys proportionality building surface in terms of the existence of solution of the problem submitted to the original data. The necessary and sufficient conditions for the existence of surfaces that can be constructed using graphic-plastic and surf-graphical keys proportionality have been formulated. The software construction of the frame surface with graphic-plastic key proportionality has been developed. The software construction of the frame surface using surf-graphical key of proportionality has been developed. Means for input source and output boundary curves constructed frame on the monitor screen and on solid carrier of information. Using developed methods allows to reach high level of correspondence and results to teste of customer.

Keywords: surface graphic proportional key, graphic plastic proportional key, surface, frame surface, monitor screen, carrier of information.

Вступ

Графічні та графічно-аналітичні методи розрахунку елементів конструкцій, володіючи всіма перевагами геометричного методу в