

УДК 004.657, 004.658.3  
ББК 32.97

## **АЛГОРИТМ ПОИСКА АЛЬТЕРНАТИВНЫХ МАРШРУТОВ ДЛЯ ЗАПРОСОВ БАЗ ДАННЫХ С МАТЕРИАЛИЗОВАННЫМИ ПРЕДСТАВЛЕНИЯМИ<sup>1</sup>**

**Дятчина Д. В.<sup>2</sup>, Погодаев А. К.<sup>3</sup>**

*(Липецкий государственный технический университет,  
Липецк)*

*В данной статье даны основные определения: эквивалентные запросы, маршруты выполнения запроса, альтернативные маршруты, полная группа маршрутов; подробно описан алгоритм поиска альтернативных маршрутов для запроса в базах данных с материализованными представлениями.*

Ключевые слова: материализованные представления, база данных, альтернативные маршруты.

### **1. Введение**

В настоящее время в различных сферах человеческой деятельности важную роль играет оперативность принятия решений и скорость получения информации, влияющей на эти решения. Получить оперативно информацию из базы данных часто не представляется возможным из-за наличия большого объема данных и сложности запросов в использующейся информационной системе. Поэтому создание и применение методов, позволяющих повышать эффективность выполнения запросов, является активно развивающейся областью исследований. Под повышением эффективности выполнения запросов – оптимиза-

---

<sup>1</sup>Работа финансируется РФФИ в виде гранта 13-07-97519 р\_центр\_а

<sup>2</sup> Дарья Васильевна Дятчина, ассистент ( dd.8383@mail.ru).

<sup>3</sup> Анатолий Кириянович Погодаев, доктор технических наук, профессор (pak@stu.lipetsk.ru).

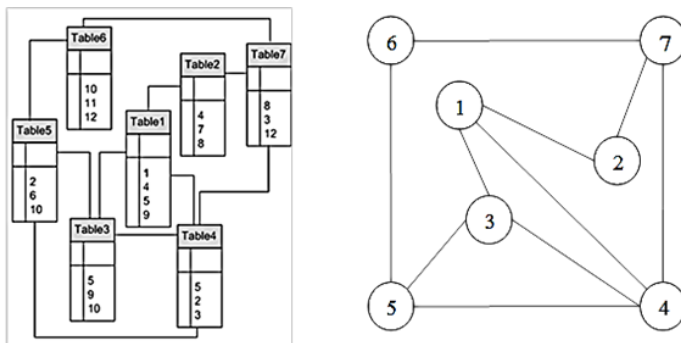
цией запросов – понимается сокращение времени их выполнения.

Существует множество подходов к минимизации времени получения требуемой информации из базы данных, основанных на оптимизации запросов: логической и семантической.[1] Эти подходы основаны на внутреннем преобразовании запросов, изменении последовательности выполнения операций реляционной алгебры. Мало исследованной является возможность оптимизации запросов в базах данных, содержащих дублируемую информацию. Содержание дубликатов информации позволяет строить различные запросы для получения одной и той же информации, причем время их выполнения может сильно варьироваться. Одним из способов контролируемого хранения и применения дублируемой информации является использование материализованных представлений. Материализованные представления, впервые появившиеся в СУБД Oracle, являются таблицами базы данных, хранящими результаты выполнения запросов. Целостность данных в таблицах материализованных представлений поддерживается периодической синхронизацией или использованием инструментов триггеров.

Таким образом, задаче поиска множества различных семантически эквивалентных вариантов выполнения запросов в базах данных с материализованными представлениями посвящена данная статья.

## **2. Поиск альтернативных маршрутов выполнения запроса**

Структура БД представляется в виде графа, где вершины соответствуют таблицам, а ребра – связям между ними. Схеме выполнения запроса соответствует множество вершин графа (таблиц БД, используемых в запросе) и ребер между ними (связей между таблицами, используемыми в условиях запроса) см.рис.1. Эти множества вершин и ребер образуют маршрут выполнения запроса на графе [2].



*Рис.1. Пример графоструктурной модели БД*

При использовании материализованных представлений в БД появляются другие маршруты соединения таблиц, применяемых в запросах.

Для корректного использования различных вариантов маршрутов в задаче оптимизации запросов дается ряд определений.

**Определение 1.** Два запроса называются эквивалентными, если при их выполнении получается семантически эквивалентный результат.

**Определение 2.** Маршрутом выполнения запроса называется множество таблиц, используемых в запросе.

**Определение 3.** Два эквивалентных запроса образуют альтернативные маршруты, если множество таблиц, входящих в первый запрос, не равно множеству таблиц, входящих во второй.

**Определение 4.** N маршрутов называются альтернативными, если любые два из них попарно являются альтернативными маршрутами.

**Определение 5.** N альтернативных маршрутов образуют полную группу маршрутов выполнения запроса, если не существует альтернативного к любому из маршрутов этой группы, не входящего в эту группу.

В рамках данных определений задача поиска всевозможных вариантов выполнения запроса сводится к определению полной группы маршрутов выполнения этого запроса.

Этот поиск производится на основе анализа исходного запроса с использованием материализованных представлений, созданных в БД.

Логически материализованные представления являются запросами на выборку данных из исходных таблиц баз данных и других материализованных представлений, в физическом смысле – представляют собой таблицы с данными.

Разработанный алгоритм поиска полной группы маршрутов извлечения информации заключается в последовательной замене части условий исходного запроса на материализованные представления, удовлетворяющие этим условиям.

Обозначим  $MView$  – множество материализованных представлений, элементами которого являются таблицы, представляющие собой запросы, результатами выполнения которых являются соединения от двух до нескольких таблиц в одну. Таким образом, каждое материализованное представление  $MV_i$  из  $MView$  можно описать набором  $(A^i, T^i, E^i)$ , где  $i$  – номер материализованного представления;  $A^i$  – множество выбираемых атрибутов, входящих в запрос, определяющий  $i$ -ое материализованное представление;  $T^i$  – множество таблиц, входящих в запрос, определяющий  $i$ -ое материализованное представление;  $E^i$  – множество условий, входящих в запрос, определяющий  $i$ -ое материализованное представление. Пусть общее число равно  $s$ .

Алгоритм поиска альтернативных маршрутов для запроса:  
Шаг 0. Дано  $A$  – множество выбираемых атрибутов в исходном запросе, элементы которого равны  $(a_1, a_2 \dots a_n)$ , где  $n$  – общее число атрибутов в запросе;  $T$  – множество таблиц в исходном запросе, элементы которого равны  $(t_1, t_2 \dots t_m)$ , где  $m$  – общее число таблиц;  $E$  – множество условий в исходном запросе, элементы которого равны  $(e_1, e_2 \dots e_k)$ , где  $k$  – общее число условий. В свою очередь каждая таблица  $t_l$  содержит множество атрибутов  $A(t_l) = \{a_{lr}\}$ , где  $r$  – номер атрибута для  $l$  таблицы.  $T^i$  – множество таблиц материализованных представлений, состоящее из элементов  $\{t_1^i, t_2^i, \dots, t_r^i\}$

Полагаем  $i=1$  – номер материализованного представления,  $j=1$  – номер альтернативного запроса на уровне  $\text{num\_ur}$ ,  $\text{num\_ur}=1$ , количество элементов на уровне  $\text{col\_el\_ur}[\text{num\_ur}]=0$ .

Шаг 1. Если  $E^i \subset E$ , то формируем альтернативный запрос

$R_j^i[\text{num\_ur}]$  с множеством таблиц  $T_j^i[\text{num\_ur}] = T_j^{\text{inp}} \cup MV_i$ , где

$$T_j^{\text{inp}} = T \setminus \left\{ t_1 \in T \mid \left( (A(t_1) \cap A) \subseteq A^i \right) \wedge \begin{cases} \text{не } \exists e_k \in E, t_r^i \in T^i \text{ и} \\ a_k \in A^w(t_1) = (A(t_1) \cap A^w) - \\ - (A(t_1) \cap A^s) \text{ т.ч.} \\ A(t_r^i) \in e_m, a_k \in e_m \end{cases} \right\}$$

и условий

$E_j^i[\text{num\_ur}] = E_j^{\text{inp}}[\text{num\_ur}] \cup (e \in E^i \mid \exists t \in T_j^{\text{inp}}[\text{num\_ur}], A(t) \subset e)$ ,

$E_j^{\text{inp}}[\text{num\_ur}] = E \setminus E^i$ , заносим его в список альтернативных маршрутов первого уровня, переходим на шаг 2. Иначе переходим на шаг 3.

Шаг 2.  $\text{col\_el\_ur}[\text{num\_ur}] = j$ ,  $i = i+1$ ,  $j = j+1$ . Если  $i > s$ , переходим на шаг 4, иначе переходим на шаг 1.

Шаг 3.  $i = i+1$ . Если  $i > s$ , переходим на шаг 4, иначе переходим на шаг 1.

Шаг 4.  $\text{num\_ur} = \text{num\_ur} + 1$ ,  $j = 1$ ,  $\text{col\_el\_ur}[\text{num\_ur}] = 0$ . Если  $\text{col\_el\_ur}[\text{num\_ur}-1] = 0$ , поиск закончен, иначе переходим на шаг 5.

Шаг 5. Выбираем  $R_j^i[\text{num\_ur}-1]$ ,  $i = i+1$ . Переходим на шаг 6.

Шаг 6. Если  $E^i \subset E_j^i[\text{num\_ur}-1]$ , то формируем альтернативный запрос  $R_j^i[\text{num\_ur}]$  с множеством таблиц  $T_j^i[\text{num\_ur}] = T_j^{\text{inp}} \cup MV_i$ , где

$$T_j^{inp} = T \setminus \left\{ t_l \in T \mid \left( (A(t_l) \cap A) \subseteq A^i \right) \wedge \begin{pmatrix} \text{не } \exists e_k \in E, t_r^i \in T^i \text{ и} \\ a_k \in A^w(t_l) = (A(t_l) \cap A^w) - \\ - (A(t_l) \cap A^s) \text{ т.ч.} \\ A(t_r^i) \in e_m, a_k \in e_m \end{pmatrix} \right\}$$

и условий

$$E_j^i[num\_ur] = E_j^{inp}[num\_ur] \cup (e \in E^i \mid \exists t \in T_j^{inp}[num\_ur], A(t) \subset e),$$

$E_j^{inp}[num\_ur] = E \setminus E^i$ ; заносим его в список альтернативных маршрутов уровня  $num\_ur$ , переходим на шаг 7, иначе переходим на шаг 8.

Шаг 7.  $col\_el\_ur[num\_ur]=j$ ,  $j=j+1$ ,  $i=i+1$ . Если  $i>s$ , переходим на шаг 8, иначе переходим на шаг 6.

Шаг 8. Если  $j>col\_el\_ur[num\_ur]$ , переходим на шаг 4, иначе на шаг 5.

Схема алгоритма представлена на рис. 2.

Результатом поиска является дерево, состоящее из  $s$  уровней, представленное на рис.3. Вершиной дерева является исходный маршрут (запрос). Вершинами первого уровня будут являться альтернативные маршруты, полученные путем использования одного материализованного представления. Вершинами  $i$ -го уровня являются альтернативные маршруты, полученные путем использования  $i$  материализованных представлений. Последний уровень данного дерева не может превышать числа материализованных представлений, имеющихся в базе данных.

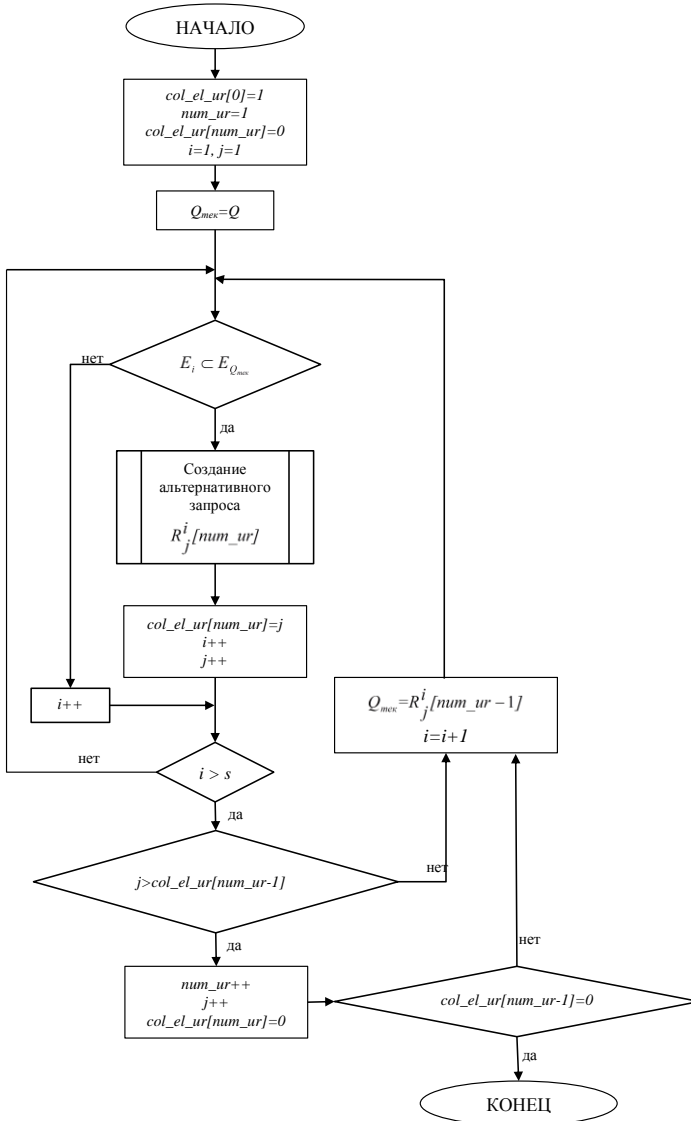
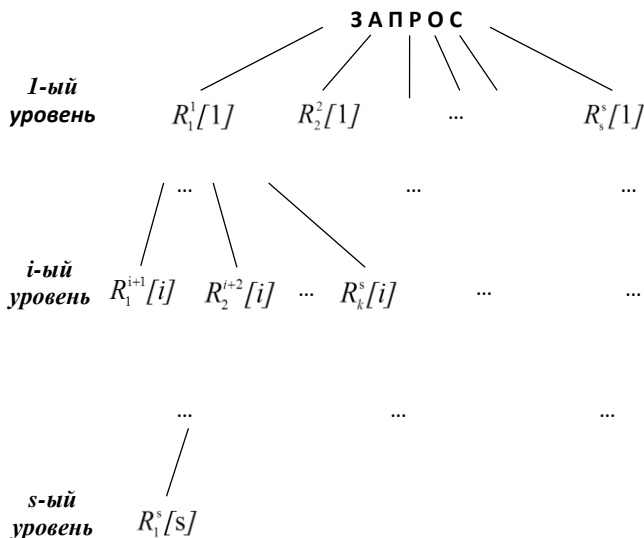


Рис.2. Алгоритм поиска полной группы маршрутов извлечения информации



*Рис.3. Дерево альтернативных маршрутов выполнения запроса*

### 3. Заключение

В данной статье подробно описан алгоритм поиска полной группы маршрутов выполнения запроса, основанный на использовании материализованных представлений, созданных в базе данных. Результаты работы данного алгоритма могут быть использованы для дальнейшей оптимизации запросов [3,4].

### Литература

1. ГАРСИА-МОЛИНА Г., УЛЬМАН Д., УИДОМ Д. *Системы баз данных* // Полный курс. Пер. с англ.- М.: Издательский дом «Вильямс», 2003. – С.1088.
2. ДЯТЧИНА Д.В. *Применение алгоритма оптимизация запросов на основе внесения контролируемой избыточно-*



- сти в базах данных // Вести ВУЗов Черноземья, №4 Липецк: 2012. – С. 48-51.
3. ПОГОДАЕВ А.К., ФЕДОРКОВА Г.О., ДЯТЧИНА Д.В. *Алгоритм оптимизации времени выполнения запросов в базах данных с альтернативными маршрутами соединения* // Вести ВУЗов Черноземья, №3 – Липецк: 2011. – С.51-54.
  4. DYATCHINA D.V. *Automated optimal query management system in databases with alternate table connection routes* // Proceedings of the Workshop on Computer Science and Information Technologies (CSIT'2013), Vienna-Budapest-Bratislava, September 15–21, 2013. Volume 3, Ufa State Aviation Technical University, 2013. – pp. 52-54.

## **FAULT FINDING ALTERNATIVE ROUTES FOR THE DATABASE QUERY MATERIALIZED VIEWS**

**Darya Dyatchina**, Lipetsk State Technical University, Assistant lecturer (dd.8383@ mail.ru).

**Anatoliy Pogodaev**, Lipetsk State Technical University, Doctor of Science, Professor (Lipetsk, pak@stu.lipetsk.ru).

*Abstract: in this article give the basic definitions: equivalent queries, the query routes, alternative routes, the full group of routes; detailed description of the algorithm to search for alternative routes for query databases with materialized views.*

**Keywords:** materialized views, the database, alternative routes.