УДК 004.657

© 2021 г. **Ю.А. Григорьев**, д-р техн. наук,
О.Ю. Плужникова

(Московский государственный технический университет им. Н.Э. Баумана)

ОЦЕНКА МОЩНОСТЕЙ АТТРИБУТОВ В СОЕДИНЕНИИ ТАБЛИЦ ПРИ ОПТИМИЗАЦИИ ЗАПРОСОВ К РЕЛЯЦИОННОЙ БАЗЕ ДАННЫХ

Проанализирована проблема оценки кардинальности промежуточных таблиц соединений в процессе вычисления стоимости плана выполнения запроса к реляционной базе данных. Предложен способ вычисления мощностей атрибутов, позволяющий уменьшить ошибку оценки кардинальности. Доказаны теоремы, подтверждающие непротиворечивость предложенного способа.

Ключевые слова: стоимостная модель, оптимизатор запросов, мощность атрибута, кардинальность таблиц.

DOI: 10.22250/isu.2021.67.3-18

Введение

Выбор оптимального плана выполнения запроса в реляционной СУБД был и остается серьезной проблемой. Она в основном связана с тем, что для оценки стоимости плана требуется знать значения селективности таблиц. Селективность – это доля записей таблицы, удовлетворяющих условию запроса к базе данных. Но на этапе оценки и выбора плана они не известны. Большая ошибка в расчете селективности приводит к большой ошибке в оценке кардинальности. Кардинальность – это число записей таблицы, удовлетворяющих условию запроса. Ошибка в их оценке приводит к выбору далеко не оптимального плана. Время выполнения запроса по такому плану может быть большим.

Вот мнение ведущего специалиста по вопросам оптимизации Гая Ломана (Guy Lohman), работающего в исследовательском отделе научного цен-

тра Almaden Research Center компании IBM: «Корень всех зол, ахиллесова пята оптимизации запросов – это оценка размера промежуточных результатов, известных как кардинальность. Модель кардинальности может легко внести ошибки на многие порядки! С такими ошибками вопрос не в том, почему оптимизатор выбрал плохой план. Скорее, возникает вопрос: зачем оптимизатору выбирать достойный план (при таких исходных данных)?» [1].

Были предприняты попытки или уменьшить ошибки в расчете селективности, или повысить устойчивость к таким ошибкам при выборе плана, или вообще отказаться от расчета селективности и менять план в процессе выполнения запроса. Для каждого этапа поиска оптимального плана и его реализации научным сообществом были предложены важные решения. Некоторые из них рассмотрены ниже.

Классическая схема выбора оптимального плана выглядит так:

- 1) генерируется множество планов;
- 2) рассчитывается стоимость каждого плана;
- 3) реализуется план с наименьшей стоимостью.

Планы отличаются друг от друга:

- 1) методом чтения записей из исходных таблиц;
- 2) порядком (планом) соединений исходных таблиц;
- 3) методом соединения таблиц.

Рассмотрим методы в следующей последовательности: методы чтения записей, методы соединения таблиц, методы планирования соединений таблиц.

Существуют два метода чтения исходных таблиц: чтение записей с помощью индекса (Index Scan) и последовательное чтение таблицы (Sequential Scan). Ошибка в селективности может привести к неправильному выбору метода чтения. Например, расчетная селективность получилась высокой, а фактическая – низкой. Оптимизатор выберет последовательное чтение всей таблицы, а на самом деле надо было назначить индексное сканирование. В [2] был разработан метод Smooth scan, который автоматически реализует плавный переход от одного метода чтения записей к другому при получении данных о селективности во время выполнения запроса.

Базы данных поддерживают различные методы соединения таблиц [3, 4]: NLJ (Nested Loops Joins), SMJ (Sorting and Merge Joins), HJ (Hash Joins). Они стали классическими. Ошибка в оценке селективности может привести к неправильному выбору метода соединения и увеличению времени выполнения запроса. В [4] предложен гибридный метод соединения G-join (Generalized Join). Если соединяемые таблицы R и S отсортированы, то

используется метод SMJ. Если нет, то значения атрибутов соединения хешируются как в HJ, сортируются в каждой странице хеш-раздела и сохраняются на диске. Требуемые страницы хеш-разделов R и S читаются в ОП при выполнении соединения.

Рассмотрим теперь методы планирования соединений таблиц. В [5] предлагается оценивать стоимость плана с помощью гистограмм, построенных самой СУБД (метод LEC). Идея методов CostGreedy [6] и SEER [7] заключается в уменьшении пространства оптимального набора планов (POSP) до небольшого числа качественных планов запросов. Это позволяет уменьшить число кэшируемых планов и повысить устойчивость к ошибкам оценки селективности.

Но проблема остается. Методы [5 – 7] используют модель расчета стоимости плана. Но точность таких расчетов существенно зависит от точности оценки селективности и кардинальности таблиц, используемых в запросе. Большая ошибка в их оценке может привести к выходу за границы значений показателей, в рамках которых метод выбора плана сохраняет устойчивость [8].

Были предприняты попытки избежать указанной проблемы соединений таблиц. В [9] получена общая оценка отклонения стоимости сгенерированного плана от оптимального значения (правило q^4). Появились методы, где селективность и кардинальность вообще не рассчитываются [10, 11]. Например, в [10] строится и сохраняется параметрический оптимальный набор планов (POSP) по диапазону селективности некоторого атрибута («букет» планов). Запрос начинает выполняться по плану из POSP с наименьшей стоимостью. При превышении этой стоимости выполнение запроса автоматически переключается на план со следующей по порядку стоимостью. И так далее, пока запрос не будет выполнен. Но сложность реализации этого метода велика, так как исполнители запросов существующих СУБД работают как конвейеры по выбранному плану. Прервать работу конвейера и переключиться на другой – непростая задача. Более того, POSP можно построить только для ограниченного набора запросов для заданных диапазонов изменения селективности некоторых параметров.

В следующих разделах статьи рассматривается проблема оценки кардинальности таблиц в процессе вычисления стоимости плана. Делается вывод о необходимости дополнительной оценки мощностей атрибутов таблиц в соединении. Далее предлагается способ их оценки, обеспечивающий непротиворечивость такой оценки.

Проблема оценки кардинальности промежуточных таблиц соединений в процессе вычисления стоимости плана

Стоимостная оптимизация запросов использовалась в первой реляционной СУБД System R [12]. Она и сейчас применяется в современных СУБД [13]. Во введении было показано, что важной остается проблема корректной оценки селективности и кардинальности таблиц при вычислении стоимости плана выполнения запроса к базе данных. Рассмотрим процесс вычисления в оптимизаторе кардинальности таблиц на примере соединений трех таблиц:

```
SELECT *  
FROM R1,R2,R3  
WHERE R1.a=R2.a and R2.b=R3.b and Pr1 and Pr2 and Pr3;
```

(1)

где Pr_i – это предикат таблицы R_i , $i = 1, 2, 3$. Предикат может включать комбинацию элементарных условий типа « $R_i.Atr_i$ оператор $const_i$ », связанных логическими операциями AND, OR, NOT. Оператор – это оператор сравнения ($=$, $>$, $<$, LIKE и др.). Предикату Pr_i соответствует подзапрос Q_i , после выполнения которого получается таблица T_i . Подзапрос реализуется или с помощью индекса, или используется последовательное чтение таблицы, или применяется и то, и другое (метод Smooth scan). Далее выполняется соединение таблиц T_i .

Порядок соединений таблиц (оптимальный план) определяет оптимизатор СУБД. Он рассматривает два плана:

- 1) $T_1 \bowtie_a T_2 \bowtie_b T_3$,
- 2) $T_2 \bowtie_b T_3 \bowtie_a T_1$.

План $T_1 \bowtie T_3 \bowtie_b T_2$ не анализируется, если таблицы R_1 и R_3 не имеют общих атрибутов [14]. Для каждого плана рассчитывается стоимость и выбирается план с наименьшей стоимостью. Стоимость оценивается на основе значений кардинальности таблиц. Рассмотрим, как оценивается кардинальность, на примере плана 1. Сначала рассчитывается кардинальность промежуточной таблицы $T_{12} = T_1 \bowtie_a T_2$:

$$|T_{12}| = |T_1| \cdot |T_2| / \max(IaR_1, IaR_2) = |T_1| \cdot |T_2| \cdot p_{12},$$
(2)

где IaR_1 , IaR_2 – мощности атрибута соединения «a» в исходных таблицах R_1 и R_2 (мощность атрибута – это число разных значений атрибута), величина $p_{12} = 1 / \max(IaR_1, IaR_2)$ интерпретируется как вероятность, что кортеж декартова произведения $T_1 \times T_2$ удовлетворяет предикату соединения $T_1.a = T_2.a$.

Далее рассчитывается кардинальность результирующей таблицы

$$T_{123} = T_1 \bowtie_a T_2 \bowtie_b T_3:$$

$$|T_{123}| = |T_{12}| \cdot |T_3| / \max(IbR_2, IbR_3) = |T_1| \cdot |T_2| \cdot |T_3| \cdot p_{12} \cdot p_{23}, \quad (3)$$

где IbR_2 , IbR_3 – мощности атрибута соединения «b» в исходных таблицах R_2 и R_3 ; величина $p_{23} = 1/\max(IbR_2, IbR_3)$ интерпретируется как вероятность, что кортеж декартова произведения $T_{12} \times T_3$ удовлетворяет предикату соединения $T_{2.b} = T_{3.b}$. Предполагается, что предикаты соединений $T_{1.a} = T_{2.a}$, $T_{2.b} = T_{3.b}$ независимы. Процесс вычисления кардинальности промежуточных таблиц можно продолжить по формуле (3), если число соединений $n > 2$:

$$T_1 \bowtie_a T_2 \bowtie_b T_3 \bowtie_c T_4 \dots$$

В работе [13] выполнен анализ качества работы оптимизатора при выборе оптимального плана. В частности, выполнена оценка точности расчета оптимизатором кардинальности таблиц, получаемых в процессе соединений (см. формулы (2), (3)).

Для анализа авторы [13] выбрали базу данных Интернет-фильмов Internet Movie Data Base (IMDB) [15]. Как и в большинстве реальных наборах данных, IMDB полна корреляций и неравномерных распределений данных, и, следовательно, гораздо сложнее, чем большинство синтетических наборов данных. Набор данных IMDB был загружен [13] в 5 реляционных систем баз данных: PostgreSQL, HyPer и три коммерческих СУБД (DBMS A, DBMS B, DBMS C). На основе базы данных IMDB в [13] были реализованы аналитические SQL-запросы. В работе [13] рассматриваются упорядоченные соединения (это самая важная задача оптимизации запросов). Запросы включают от 3 до 16 соединений таблиц, в среднем по 8 соединений для каждого запроса. Набор запросов состоит из 33 структур запросов, каждый из которых имеет от 2 до 6 вариантов, которые отличаются только селективностью. В результате в общей сложности использовались 113 запросов. В зависимости от селективности предикатов базовой таблицы варианты одной и той же структуры запроса имеют разные оптимальные планы запросов, которые дают значительно различающееся время выполнения (иногда на порядок). Кроме того, некоторые запросы имеют более сложные предикаты выбора, – например, дизъюнкции или поиск подстроки с использованием LIKE. Все эксперименты в [13] по оценке производительности были выполнены на сервере с двумя процессорами Intel Xeon X5570 (2,9 ГГц) и в общей сложности с 8 ядрами. Система имеет 64 ГБ ОП, поэтому вся база данных IMDB полностью кэшируется в ОП.

На рис. 1 приведена q-ошибка оценки оптимизатором кардинальности промежуточных таблиц (числа записей, удовлетворяющих условию в каждом соединении) по сравнению с истинной кардинальностью для запросов с $n=6$ соединениями. Например, если истинная кардинальность соединения

равна 100, то оценки 10 или 1000 обе имеют q-ошибку, равную 10. Но в первом случае значение q откладывается ниже горизонтальной оси «1», а во втором – выше. Приводится q-ошибка в зависимости от номера соединения в запросе. Видно, что оптимизатор недооценивает (underestimation) кардинальность с увеличением номера соединения. Это справедливо для всех протестированных СУБД. В DBMS C ошибка достигала почти 8 порядков.

Рис. 1. q-ошибка оценки кардинальности для запросов с 6 соединениями по сравнению с истинной кардинальностью (для реального наполнения БД) [13].

В PostgreSQL мощности атрибутов соединений (IaR_1 , IaR_2 , IbR_2 , IbR_3 в формулах (2), (3)) оцениваются на основе выборки фиксированного размера. Чтобы исключить влияние случайности при оценке мощностей, авторы [13] вычислили эти значения точно и заменили оценочные значения истинными. После этого они повторили серию экспериментов для PostgreSQL. Картина почти не изменилась (см. PostgreSQL на рис. 1).

Чтобы измерить влияние неверной оценки кардинальности промежуточных таблиц на производительность запросов, авторы [13] внедрили оценку этой кардинальности различными СУБД в PostgreSQL, а затем выполнили полученные оптимизатором планы (планы 1). Использование одного и того же механизма (PostgreSQL) планирования и выполнения запросов позволяет абстрагироваться от различных механизмов обработки запросов в других СУБД. Кроме того, они ввели в оптимизатор истинные кардинальности промежуточных таблиц и выполнили полученные планы (планы 2). Далее сравнили время выполнения запросов, реализованных с использованием планов 1 и 2. Доля запросов, которые выполнялись более чем в 2 раза медленнее по планам 1 по сравнению с планами 2, составила от 22% до 47% для разных СУБД (PostgreSQL – 36%), а медленнее в 1.1 раза – от 43% до 65% запросов. В [13] делается вывод о низком качестве оценок кардинальности промежуточных таблиц в соединении, выполняемых оптимизаторами СУБД.

На наш взгляд, есть две важные причины плохой работы оптимизатора:

1) необходимость использовать мощности атрибутов соединения в таблицах T_i , которые получились после выполнения подзапросов, а не в исходных таблицах R_i , как в (2) и (3). Например, для запроса (1) и атрибута «а» может случиться, что $Ia_1 \leq |T_1| < IaR_1 \leq |R_1|$, $Ia_2 \leq |T_2| < IaR_2 \leq |R_2|$ и $\max(Ia_1, Ia_2) < \max(IaR_1, IaR_2)$. Здесь Ia_1, Ia_2 – мощности атрибута «а» в таблицах T_1, T_2 , что приведет к росту p_{12} и $|T_{12}|$ в (2). Эти рассуждения справедливы и для атрибута соединения «b»;

2) в некоторых случаях нельзя считать независимыми предикаты соединений. Например, в (3) необходимо использовать мощность атрибута соединения «b» в промежуточной таблице T_{12} (Ib_{12}), а не в T_2 (Ib_2) и не в исходной таблице R_2 (IbR_2). Может случиться, что $Ib_{12} \leq |T_{12}| < Ib_2 \leq |T_2| \leq IbR_2 \leq |R_2|$ и $\max(Ib_{12}, Ib_3) < \max(IbR_2, IbR_3)$. Здесь Ib_2, Ib_3 – мощности атрибута «b» в таблицах T_2, T_3 , что приведет к росту p_{23} и $|T_{123}|$ в (3).

Можно предложить простые рекомендации, позволяющие устранить перечисленные выше причины плохой работы оптимизатора.

I. Использовать следующие мощности атрибутов соединения в подзапросах:

$$Ix_i = \min(|T_i|, IxR_i), i=1...m, \quad (4)$$

где x – атрибут соединения; i – номер таблицы подзапроса; m – число подзапросов. Равенство (4) следует из того факта, что мощность Ix_i атрибута таблицы не может быть больше числа записей $|T_i|$ в этой таблице.

II. Использовать указанные ниже значения мощностей атрибутов в промежуточных таблицах. Промежуточная таблица (ПТ) – это таблица, которая получается при k -м соединении, $k < n$. Например, T_{12} – это ПТ в плане $T_1 \bowtie_a T_2 \bowtie_b T_3$.

1) «х» – атрибут соединения в T_i и T_j :

$$Ix(T_i \bowtie_x T_j) = \min(Ix_i, Ix_j). \quad (5)$$

Равенство (5) следует из того факта, что мощность атрибута соединения «х» в таблице $T_i \bowtie_x T_j$ не может превышать мощности этого атрибута в таблицах-аргументах; выражение (5) можно использовать рекурсивно;

2) атрибут «у» не является атрибутом соединения в T_i или T_j . Он может быть атрибутом соединения в следующих соединениях этого плана.

Можно предложить способ оценки мощности атрибута «у» в таблице

$T_i \bowtie_x T_j$. Назовем его «способ 1», чтобы было проще на него ссылаться.

Способ 1.

а) $y \in T_i$,

$$\begin{aligned} \text{если } |T_i \bowtie_x T_j| < |T_i|, \text{ то } I_y(T_i \bowtie_x T_j) &= |T_i \bowtie_x T_j| \cdot I_{y_i} / |T_i|, \\ \text{иначе } I_y(T_i \bowtie_x T_j) &= I_{y_i}; \end{aligned} \quad (6)$$

б) $y \in T_j$,

$$\begin{aligned} \text{если } |T_i \bowtie_x T_j| < |T_j|, \text{ то } I_y(T_i \bowtie_x T_j) &= |T_i \bowtie_x T_j| \cdot I_{y_j} / |T_j|, \\ \text{иначе } I_y(T_i \bowtie_x T_j) &= I_{y_j}. \end{aligned} \quad (7)$$

Выражения (6), (7) можно трактовать так: если число записей в соединении становится меньше числа записей в таблице-аргументе, то мощность атрибута «у» в соединении пропорционально уменьшается.

В следующих разделах анализируются на непротиворечивость полученные способом 1 оценки мощности атрибута «у» в промежуточной таблице. Будем говорить, что способ 1 дает непротиворечивую оценку, если при изменении плана выполнения соединений не изменяются оценки кардинальности промежуточных таблиц и мощностей атрибутов в этих таблицах. Это важно, так как хотя порядок соединений таблиц влияет на время выполнения запросов, указанные величины должны быть инвариантными относительно изменения плана. Способ 1 будем называть способом на основе равенства отношений кардинальности и мощности.

Способ 1 дает непротиворечивую оценку для промежуточной таблицы (ПТ), полученной при одном соединении двух таблиц (ПТ-1). При изменении плана (перестановка исходных таблиц местами) не изменяется кардинальность ПТ-1 (см. (2)). Легко показать, что мощности атрибутов в ПТ-1 также не изменяются при изменении плана (см. (5) – (7)).

Для ПТ с двумя и более соединениями это не очевидно. Дело в том, что для ПТ есть еще вложенные промежуточные таблицы. При изменении плана вложенные ПТ могут быть другими и давать противоречивые результаты.

Сначала мы рассмотрим ПТ с двумя соединениями (ПТ-2), а затем распространим результаты анализа на ПТ с $n > 2$ соединениями (ПТ-n). В каждой из перечисленных ниже двух групп планов вложенные ПТ (показаны в скобках) не меняются: 1-я группа – $T_{(12)3}$, $T_{(21)3}$, $T_{3(12)}$, $T_{3(21)}$; 2-я группа – $T_{(23)1}$, $T_{(32)1}$, $T_{1(23)}$, $T_{1(32)}$. Поэтому достаточно рассмотреть планы $T_{(12)3} = T_{123}$ и $T_{(23)1} = T_{231}$. В дальнейшем будем использовать в выражениях (2) и (3) мощности атрибутов для таблиц T_i вместо R_i . С учетом сказанного перепишем

выражения (2) и (3) для разных планов.

Для плана

$$T_{123}=T_1 \bowtie_a T_2 \bowtie_b T_3 \quad (8)$$

имеем

$$|T_{12}| = |T_1| \cdot |T_2| / \max(Ia_1, Ia_2), \quad (9)$$

$$|T_{123}| = |T_{12}| \cdot |T_3| / \max(Ib_{12}, Ib_3). \quad (10)$$

Для плана

$$T_{231}=T_2 \bowtie_b T_3 \bowtie_a T_1 \quad (11)$$

по аналогии с (9) и (10) получим

$$|T_{23}| = |T_2| \cdot |T_3| / \max(Ib_2, Ib_3), \quad (12)$$

$$|T_{231}| = |T_{23}| \cdot |T_1| / \max(Ia_{23}, Ia_1). \quad (13)$$

Анализ способа оценки мощностей атрибутов в таблице на основе равенства отношений кардинальности и мощности (способ 1)

Лемма 1. Пусть $IaR_1=IaR_2$ и $IbR_2=IbR_3$. Тогда, если $Ix_i < Ix_j$, то $|T_i|=Ix_i$, где $x=a|b$, $(i, j)=(1,2)|(2,1)$.

Доказательство. Докажем лемму для варианта $x=a$ и $(i, j)=(1,2)$. Для остальных вариантов лемма доказывается аналогично.

$Ia_1 < Ia_2 \leq IaR_2$, так как T_2 – это подзапрос R_2 . Отсюда и из условия леммы следует, что $Ia_1 < IaR_1$. Из (4) получим $Ia_1 = \min(|T_1|, IaR_1) = |T_1|$.

Лемма доказана.

В дальнейшем операции умножения ($\cdot$) и деления ($/$), как принято, имеют одинаковый приоритет и выполняются последовательно, если нет круглых скобок.

Теорема 1. Способ 1 оценки мощностей атрибутов дает непротиворечивую оценку кардинальности соединения ПТ-2, если $IaR_1=IaR_2$ и $IbR_2=IbR_3$.

Доказательство. Рассмотрим план T_{123} (см. (8)).

Случай 1. $|T_{12}| < |T_2|$.

В этом случае $Ia_2 > Ia_1$, и из (9) получим

$$|T_{12}| = |T_1| \cdot |T_2| / \max(Ia_2, Ia_1) = |T_1| \cdot |T_2| / Ia_2. \quad (14)$$

Действительно, если $Ia_2 \leq Ia_1$, то $|T_{12}| = |T_1| \cdot |T_2| / Ia_1 \geq |T_2|$, так как $|T_1| / Ia_1 \geq 1$ (мощность атрибута не может превышать числа записей в таблице). А это противоречит рассматриваемому случаю. Получим

$$|T_{123}| = |T_{12}| \cdot |T_3| / \max(Ib_{12}, Ib_3), Ib_{12} = |T_{12}| \cdot Ib_2 / |T_2|. \quad (15)$$

Выражение для Ib_{12} следует из (7) для $x=a, y=b, i=1, j=2$.

Вариант 1. Пусть $Ib_{12} \geq Ib_3$.

Тогда $Ib_2 \geq |T_2|/|T_{12}| \cdot Ib_3 > Ib_3$ (для случая 1) и $Ib_3 = |T_3|$ (см. лемму 1). Из (14) и (15) следует $Ib_{12} = |T_1| \cdot |T_2| / Ia_2 \cdot Ib_2 / |T_2| \geq (|T_3| = Ib_3)$. Отсюда имеем

$$|T_1| \cdot Ib_2 \geq |T_3| \cdot Ia_2. \quad (16)$$

Для случая 1 и варианта 1 ($Ib_{12} \geq Ib_3$) получим из (15)

$$|T_{123}| = |T_2| \cdot |T_3| / Ib_2. \quad (17)$$

Теперь для случая 1 и варианта 1 оценим $|T_{231}|$ для плана T_{231} (см. (11)). С учетом, что для 1-го варианта $Ib_2 > Ib_3$, получим

$$|T_{23}| = |T_2| \cdot |T_3| / \max(Ib_2, Ib_3) = |T_2| \cdot |T_3| / Ib_2, \quad (18)$$

$|T_1| = Ia_1$ (по лемме 1, так как $Ia_1 < Ia_2$). Также $|T_{23}| < |T_2|$, поскольку $(|T_3| = Ib_3) / Ib_2 < 1$. Поэтому из (6) следует, что $Ia_{23} = |T_{23}| \cdot Ia_2 / |T_2|$. И в силу неравенства (16) имеем $Ia_{23} = |T_2| \cdot |T_3| / Ib_2 \cdot Ia_2 / |T_2| \leq (|T_1| = Ia_1)$. Отсюда с учетом (18) получим

$$|T_{231}| = |T_{23}| \cdot |T_1| / \max(Ia_{23}, Ia_1) = |T_2| \cdot |T_3| / Ib_2. \quad (19)$$

Из (17) и (19) следует, что для случая 1 и варианта 1 способ 1 оценки мощностей атрибутов дает непротиворечивую оценку кардинальности соединения ПТ-2.

Вариант 2. Пусть $Ib_{12} < Ib_3$.

Из (15) и (14) получим $Ib_{12} = |T_1| \cdot |T_2| / Ia_2 \cdot Ib_2 / |T_2| < Ib_3$. Следовательно, с учетом, что $|T_1| = Ia_1$, получим

$$Ia_1 \cdot Ib_2 < Ia_2 \cdot Ib_3. \quad (20)$$

Из (15) и (14) получим для 2-го варианта ($Ib_{12} < Ib_3$)

$$|T_{123}| = |T_1| \cdot |T_2| \cdot |T_3| / Ia_2 / Ib_3. \quad (21)$$

Теперь для случая 1 и варианта 2 оценим $|T_{231}|$ для плана T_{231} .

$$|T_{23}| = |T_2| \cdot |T_3| / \max(Ib_2, Ib_3). \quad (22)$$

1. Пусть $|T_{23}| < |T_2|$.

В этом случае $Ib_2 > Ib_3$ и, следовательно, $|T_{23}| = |T_2| \cdot |T_3| / Ib_2$ (из (22)) и $|T_3| = Ib_3$ (из леммы 1). Иначе при $Ib_2 \leq Ib_3$ получим противоречие: $|T_{23}| = |T_2| \cdot |T_3| / Ib_3 \geq |T_2|$, так как $|T_3| / Ib_3 \geq 1$ (мощность не может быть больше числа записей).

$$|T_{231}| = |T_{23}| \cdot |T_1| / \max(Ia_2, Ia_1), \quad (23)$$

$Ia_{23} = |T_{23}| \cdot Ia_2 / |T_2|$ (из (6)).

После подстановки $|T_{23}|$ получим $Ia_{23} = |T_2| \cdot |T_3| / Ib_2 \cdot Ia_2 / |T_2| > Ia_1$ (в силу (20) и с учетом $|T_3| = Ib_3$). Отсюда

$$|T_{231}| = |T_2| \cdot |T_1| / Ia_2 \cdot (|T_3| / Ib_3). \quad (24)$$

Дробь в круглых скобках равна 1.

2. $|T_{23}| \geq |T_2|$.

В этом случае $Ib_3 \geq Ib_2$. Иначе при $Ib_2 > Ib_3$ получим противоречие с указанным выше условием: $|T_{23}| = |T_2| \cdot |T_3| / Ib_2 < |T_2|$, $Ib_3 = |T_3|$, $Ib_3 / Ib_2 < 1$. Итак, $|T_{23}| = |T_2| \cdot |T_3| / Ib_3$ (см. (22)) и в силу (6) – мощность Ia_2 не меняется – получим

$$|T_{231}| = |T_{23}| \cdot |T_1| / \max(Ia_2, Ia_1) = |T_2| \cdot |T_3| \cdot |T_1| / Ib_3 / Ia_2. \quad (25)$$

Из (25), (24), (21) следует, что для случая 1 и варианта 2 способ 1 оценки мощностей атрибутов дает непротиворечивую оценку кардинальности соединения ПТ-2.

Случай 2. $|T_{12}| \geq |T_2|$. Тогда $Ib_{12} = Ib_2$ (см. (7)).

В этом случае $Ia_1 \geq Ia_2$. Иначе при $Ia_1 < Ia_2$ получим противоречие: $|T_{12}| = |T_1| \cdot |T_2| / Ia_2 < |T_2|$, если $(|T_1| = Ia_1) / Ia_2 < 1$. Из (9) и (10) получим для плана T_{123}

$$|T_{123}| = |T_1| \cdot |T_2| \cdot |T_3| / Ia_1 / \max(Ib_{12}, Ib_3), Ib_{12} = Ib_2. \quad (26)$$

Рассмотрим теперь план T_{231} (11) для этого случая. Согласно способу (6) оценки мощности атрибута, $Ia_{23} \leq Ia_2$. Но $Ia_2 \leq Ia_1$, поэтому из (12), (13) получим оценку $|T_{231}|$, совпадающую с (26).

Мы рассмотрели все возможные случаи и варианты. И везде способ 1 оценки мощностей атрибутов дает непротиворечивую оценку кардинальности соединения ПТ-2, т.е. $|T_{123}| = |T_{231}|$.

Теорема 1 доказана.

Теорема 2. Способ 1 оценки мощностей атрибутов дает непротиворечивую оценку мощности любого атрибута в соединении ПТ-2, если $IaR_1 = IaR_2$ и $IbR_2 = IbR_3$.

Мы не можем привести доказательство теоремы 2 в этой статье полностью. Продемонстрируем справедливость ее утверждения на одном примере: случай 1, вариант 1 (см. теорему 1). Пусть атрибут $u \in T_1$. Оценим мощность этого атрибута в таблицах T_{123} и T_{231} .

Сначала рассмотрим таблицу T_{123} . Из (14) следует, что $|T_{12}| \geq |T_1|$, так как $|T_2| \geq Ia_2$. Следовательно, из (6) имеем $Iu_{12} = Iu_1$. Из (17), (14) с учетом (16)

получим $|T_{123}| \leq |T_{12}|$. Тогда, учитывая (6) ($i=12, j=3$), выводим из (14), (17), что

$$I_{y123} = I_{y12} \cdot |T_{123}| / |T_{12}| = I_{y1} \cdot I_{a2} \cdot |T_3| / |T_1| / I_{b2}, \quad (27)$$

где $|T_3| = I_{b3}$, $|T_1| = I_{a1}$.

Теперь рассмотрим таблицу T_{231} . Так как $y \in T_1$, то $y \notin T_{23}$. Сравним число записей $|T_{231}| = |T_2| \cdot (|T_3| = I_{b3}) / I_{b2}$ (см. (19)) и $|T_1| = I_{a1}$. Видно, что $|T_{231}|$ явно зависит от числа записей $|T_2|$, а $|T_1|$ – нет. Но надо учитывать следующее: $I_{a23} \leq I_{a1}$, $|T_1| = I_{a1}$ и каждая из $|T_{231}| / I_{a23} = |T_2| / I_{a2}$ записей из T_{231} соединяется с одной записью из T_1 , т.е. каждая запись из T_1 дублируется в T_{231} $k = |T_2| / I_{a2}$ число раз. Поэтому надо сравнивать $|T'_{231}| = |T_{231}| / k = |T_3| \cdot I_{a2} / I_{b2}$ с $|T_1| = I_{a1}$. В силу (16) имеем $|T'_{231}| \leq |T_1|$. Следовательно, из (7) ($i=23, j=1$) получим

$$I_{y231} = I_{y1} \cdot |T'_{231}| / |T_1| = I_{y1} \cdot |T_3| \cdot I_{a2} / I_{b2} / |T_1|. \quad (28)$$

Величины (27) и (28) совпали. Аналогично можно доказать, что способ 1 оценки мощности атрибута дает непротиворечивую оценку мощности для других атрибутов из T_1 , T_2 и T_3 для разных случаев и вариантов, рассмотренных в теореме 1.

Теорема 3. Способ 1 оценки мощностей атрибутов дает непротиворечивую оценку кардинальности и мощностей атрибутов в соединении ПТ-п, если для любого соединения ПТ-2 в перестановке ПТ-п выполнены условия теоремы 1.

Доказательство. Теорему будем доказывать для левостороннего дерева соединений, которое наиболее часто используется на практике. Ниже знак $\bowtie$ соединения таблиц T_i заменен запятой. Приоритетность соединений будем обозначать круглыми скобками.

Пусть $ПТ-п = T_1, T_2, \dots, T_{n+1}$ и Q – некоторая допустимая перестановка указанных в ПТ-п таблиц (т.е. без декартовых произведений). Назовем перестановки эквивалентными, если они дают одинаковую оценку кардинальности и мощности любого атрибута в ПТ-п. Эквивалентность будем обозначать знаком $\equiv$. Представим Q в виде последовательности таблиц $Q = ((ПТ_1, T_1), P_1)$, где $ПТ_1$ – некоторая промежуточная таблица (м. б. пусто) и P_1 – некоторая перестановка. Из свойств ПТ-1 и ПТ-2 (см. доказанную теорему 1) имеем

$$\begin{aligned} Q &= ((T_1, ПТ_1), P_1) = ((ПТ_2, T_2), P_2) = (((T_1, ПТ_{12}), T_2), P_2) \equiv (((T_1, T_2), ПТ_{12}), P_2) = \\ &= ((T_{12}, ПТ_{12}), P_2) = ((ПТ_3, T_3), P_3) = (((T_{12}, ПТ_{123}), T_3), P_3) \equiv (((T_{12}, T_3), ПТ_{123}), P_3) = \\ &= ((T_{123}, ПТ_{123}), P_3) = \dots = ((T_{12\dots n}, T_{n+1}), P_{n+1}) = (T_1, T_2, \dots, T_{n+1}). \end{aligned} \quad (29)$$

Теорема 3 доказана.

Демонстрация способа 1 оценки мощностей атрибутов в таблице на примере хранилища данных типа ROLAP

Рассмотрим реляционное хранилище данных ROLAP типа «звезда» с таблицей фактов T_F и таблицами измерений $T_1, \dots, T_n$. Сравним оценки числа записей в промежуточных таблицах, выполненных с использованием способа 1 вычисления мощностей атрибутов (см. предыдущий раздел), и способа, используемого в реляционных СУБД (назовем его способ 0, см. мощности IaR_i и IbR_j в (2), (3)).

Определим мощности атрибутов соединения:

$$Ia_i = \min(|T_i|, IaR_i) = |T_i|, i=1\dots n, \text{ для таблиц измерений,} \quad (30)$$

$$Ib_i = \min(|T_F|, IaR_i) = IaR_i, i=1\dots n, \text{ для таблицы фактов.} \quad (31)$$

Будем рассматривать план ПТ- $n = (T_F \bowtie_{b1=a1} T_1 \bowtie \dots \bowtie_{bn=an} T_n)$ и полагать, что

$$k_i \geq k_{i-1}, k_i = |T_i|/IaR_i = |T_i|/Ib_i \leq 1, i=1\dots n. \quad (32)$$

Если это не так, то в плане можно сделать требуемую эквивалентную перестановку таблиц (см. теорему 3). Для 1-го промежуточного соединения с учетом (30), (31) получим

$$|T_{F1}| = |T_F| \cdot |T_1| / \max(Ib_1, Ia_1) = |T_F| \cdot |T_1| / Ib_1 = k_1 \cdot |T_F| \leq |T_F|. \quad (33)$$

Оценим мощности атрибутов таблицы F в промежуточной таблице T_{F1} с помощью способа 1 (см. (6)):

$$Ib_{F1i} = |T_{F1}| / |T_F| \cdot Ib_i = k_1 \cdot Ib_i, i=2\dots n. \quad (34)$$

Случай 1. Если $Ib_{F12} \geq Ia_2$, то для 2-го промежуточного соединения получим

$$|T_{F12}| = |T_{F1}| \cdot |T_2| / \max(Ib_{F12}, Ia_2) = (k_2/k_1) \cdot k_1 \cdot |T_F| \geq |T_{F1}|, \quad (35)$$

$$Ib_{F12i} = Ib_{F1i} = k_1 \cdot Ib_i, i=3\dots n, \text{ (см. (6) и (34)).} \quad (36)$$

Сравнивая (35) и (36) с (33) и (34), видим, что во 2-м соединении число записей возросло при использовании способа 1, а мощности атрибутов в T_{F12} не изменились.

Случай 2. Если $Ib_{F12} < Ia_2$, то с учетом (30) имеем для 2-го промежуточного соединения

$$|T_{F12}| = |T_{F1}| \cdot |T_2| / \max(Ib_{F12}, Ia_2) = |T_{F1}| \cdot |T_2| / Ia_2 = |T_{F1}|. \quad (37)$$

Таким образом, для этого случая число записей во 2-м соединении не

изменилось (см. (33)). Согласно (6), не изменились и мощности атрибутов в T_{F12} (см. (36)).

Продолжая выкладки с использованием *способа 1*, получим число записей для j -го промежуточного соединения

$$|T_{F1...j}| = k_1 \cdot |T_F| \cdot \prod_{i=1}^j q_i \geq |T_{F1...j-1}|, q_i = \{k_i / k_1 \mid 1\} \geq 1, j=1...n, \quad (38)$$

где k_i определено в (32).

Для используемого в СУБД *способа 0* оценки мощностей атрибутов получим из (2), (3)

$$|T_{F1...j}| = |T_F| \cdot \prod_{i=1}^j k_i, k_i \leq 1, j=1...n. \quad (39)$$

Наихудшим случаем является тот вариант, когда в силу неравномерности распределения ключей в таблице фактов T_F истинные кардинальности всех промежуточных таблиц близки к $|T_F|$. Из графиков на рис. 2а видно, что значение кардинальности промежуточной таблицы, полученное с помощью *способа 0*, убывает с ростом номера соединения j (см. формулу (39)), а с использованием *способа 1* возрастает (см. формулу (38)). При этом возрастание может быть нестрогим. Наличие горизонтальных участков зависит от значений коэффициентов k_i (32) и соотношений мощностей атрибутов соединений (см. выше случай 1 и случай 2).

Рис. 2. Сравнение способа 0 и способа 1 оценки мощностей атрибутов:
а) кардинальность промежуточных таблиц, б) q-ошибка.

На рис. 2б приведены q-ошибки оценки кардинальности промежуточных таблиц для способа 0 и способа 1 оценки мощности атрибута (по аналогии с рис. 1). Видно, что при использовании способа 0, который применяется в реляционных СУБД, q-ошибка оценки кардинальности существенно увеличивается с ростом номера соединения. Это согласуется с практическими

исследованиями, приведенными на рис. 1. И наоборот, предложенный способ 1 оценки мощностей атрибутов позволяет уменьшить значение q -ошибки.

Заключение

Рассмотрены современные методы реализации планов выполнения запросов к реляционным БД: методы чтения записей из исходных таблиц, методы соединения таблиц, методы планирования соединений таблиц. Показано, что при планировании соединений возникает проблема корректной оценки кардинальности промежуточных таблиц в процессе поиска оптимального плана выполнения запросов. Это связано с неправильной оценкой оптимизатором мощностей атрибутов соединений: в реляционных СУБД они определяются как мощности атрибутов в исходных таблицах (способ 0). При этом не учитывается, что после соединения эти мощности могут измениться. В статье предлагается новый способ оценки мощности атрибута на основе равенства отношений кардинальности и мощности в исходной и промежуточной таблицах (способ 1). На примере хранилища данных ROLAP показано, что точность оценки кардинальности таблиц повышается при использовании этого способа. Его можно реализовать в оптимизаторах СУБД с открытым кодом, – например, в MySQL, PostgreSQL.

ЛИТЕРАТУРА

1. *Guy Lohman*. Is query optimization a “solved” problem? <http://wp.sigmod.org/?p=1075>.
2. *Borovica-Gajic R. et al.* Smooth scan: Statistics-oblivious access paths // 2015 IEEE 31st International Conference on Data Engineering. IEEE. – 2015. – P. 315-326.
3. *Lewis J.* Cost-based Oracle fundamentals. – Apress, 2006.
4. *Graefe G.* New algorithms for join and grouping operations // Computer Science-Research and Development. – 2012. – Т. 27, №. 1. – P. 3-27.
5. *Chu F., Halpern J., Gehrke J.* Least expected cost query optimization: what can we expect? // Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems. – 2002. – P. 293-302.
6. *Darera P.N. et al.* On the production of anorexic plan diagrams // Proceedings of the 33rd international conference on Very large data bases. – VLDB Endowment, 2007. – P. 1081-1092.
7. *Harish D., Darera P.N., Haritsa J.R.* Identifying robust plans through plan diagram reduction // Proceedings of the VLDB Endowment. – 2008. – Т. 1, №. 1. – P. 1124-1140.
8. *Haritsa J.R.* Robust Query Processing: Mission Possible // 2019 IEEE 35th International Conference on Data Engineering (ICDE). IEEE – 2019. – P. 2072-2075.
9. *Moerkotte G., Neumann T., Steidl G.* Preventing bad plans by bounding the impact of cardinality estimation errors // Proceedings of the VLDB Endowment. – 2009. – Т. 2, №. 1. – P. 982-993.
10. *Dutt A., Haritsa J.R.* Plan bouquets: A fragrant approach to robust query processing // ACM

- Transactions on Database Systems (TODS). – 2016. – Т. 41, №. 2. – P. 1-37.
11. *Karthik S. et al.* Platform-independent robust query processing // IEEE Transactions on Knowledge and Data Engineering. – 2017. – Т. 31, №. 1. – P. 17-31.
 12. *Selinger P.G., Astrahan M.M., Chamberlin D.D., Lorie R.A., Price T.G.* Access path selection in a relational database management system. In SIGMOD. – 1979. – P. 23-34,
 13. *Leis V. et al.* How good are query optimizers, really? // Proceedings of the VLDB Endowment. – 2015. – Т. 9, №. 3. – P. 204-215.
 14. <https://postgrespro.ru/docs/postgrespro/9.5/planner-optimizer>.
 15. IMDB. <https://www.imdb.com/>.

E-mail:

Григорьев Юрий Александрович – grigorev@bmstu.ru;

Плужникова Ольга Юрьевна - pluzhnikova@bmstu.ru.