

Развитие методов управления множеством транзакций и запросов пользователей на основе процедуры поиска точки входа во взвешенный граф в распределенных базах данных

Д. И. Фильгус

МИРЭА — Российский технологический университет
119571, Москва, пр-т Вернадского, 78

e-mail: dmif42@ya.ru

Аннотация. Разработан метод поиска кратчайшего гамильтонова пути в произвольном графе на основе рангового подхода, который обеспечивает высокую оперативность и существенное уменьшение погрешности решения задачи организации процесса управления множеством транзакций и запросов при их реализации в сетевых базах данных. Во многих случаях существующие решения не обеспечивают необходимых результатов по времени доступа и точности найденного решения. Использование разработанного метода позволяет минимизировать время простоя вычислительных устройств, сократить объемы и время передачи данных от одних исполнительных устройств другим, повысить общую масштабируемость, минимизировать время доступа к данным и пр. Важным достоинством предлагаемого метода является уменьшения числа элементарных операций, за счет использования процедуры нахождения оптимальной входной точки в граф, что приводит к существенному уменьшению времени на реализацию процедур формирования очереди выполнения операций в запросах. Предложенный алгоритм целесообразно использовать адаптивно, в комбинации с другими решениями, с целью минимизировать погрешностью решаемой задачи. В работе используются методы теории графов. Оценка эффективности решения задачи выполнена с использованием системного подхода, системного анализа и теории исследования операций. Обработка экспериментальных данных, полученных в ходе работы, проводилась в соответствии с положениями математической статистики.

Ключевые слова: граф, гамильтонов путь, запрос, ранговый подход, стянутое дерево путей, транзакция, ранг, распределенная база данных.

1. Введение

Анализ существующих методов решения задачи поиска кратчайшего гамильтонова пути, показал, что в этих методах, как правило, производится поиск решения по

исходному графу, без предварительного упорядочивания вершин графа либо с частичным упорядочиванием [1–3]. Такие подходы позволяют несколько уменьшить временную сложность алгоритмов поиска кратчайшего гамильтонова пути [4].

Исследования путей решения данной задачи показало, что неупорядоченность множества элементарных операций не позволяет определить исходную точку входа в структуру графа [5–7]. В связи с этим приходится осуществлять поиск либо от всех вершин, либо с применением некоторой процедуры отсечения неперспективных вариантов поиска [1, 8]. С целью выполнения операции определения точки входа в граф поиска, в ряде работ указывается возможность введения дополнительной фиктивной вершины. Данная вершина характеризуется тем, что весовые характеристики по путям доступа к вершинам следующего уровня являются равнозначными т. е. начало поиска кратчайшего пути выбирается по любому маршруту. На рис. 1 показан исходный граф поиска кратчайшего гамильтонова пути и на рис. 1 — граф с введением фиктивной вершины.

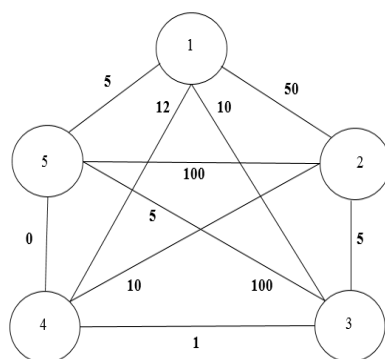


Рисунок 1. Исходный граф поиска кратчайших гамильтоновых пути

Рассмотрим процедуру определения точки входа в структуру графа, с целью сокращения вычислительной сложности алгоритма и сокращения времени нахождения кратчайшего гамильтонова пути при заданном уровне отклонения от оптимального решения. Для определения стратегии выбора оптимальной точки входа в граф введем эвристическое правило:

Оптимальной вершиной, для начала поиска кратчайшего гамильтонова пути, следует принимать ту вершину, для которой суммарное число максимальных весов, по направлениям движения к соседним вершинам, является наибольшим.

Данное правило основано на том, что существует большая вероятность того, что в кратчайшем гамильтоновом пути вершина S обязательно будет стоять на k -м

месте. Высказанное эвристическое правило подтверждается экспериментальными исследованиями.

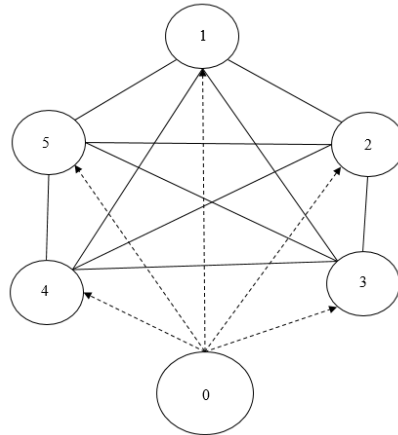


Рисунок 2. Граф поиска кратчайшего гамильтонова пути с введением фиктивной вершины

Исходя из исходных данных изложенных в [9] на первом этапе формируется матрица достижимости между элементарными операциями запросов и транзакций $A = [\alpha_{ij}]$:

$$\alpha_{ij} = \begin{cases} 1, & \text{если } i \Leftrightarrow j, \\ 0, & \text{в противном случае.} \end{cases} \quad (1)$$

Для иллюстрации данной процедуры используется граф представленный на рис. 1. Составленная матрица достижимости имеет вид представленный на рис. 3.

	S_1	S_2	S_3	S_4	S_5
S_1	–	50	∞	12	5
S_2	50	–	5	10	∞
S_3	∞	5	–	1	20
S_4	12	10	1	–	0
S_5	5	∞	20	0	–

Рисунок 3. Матрица достижимости между операциями запроса после процедуры декомпозиции

Уточнение. При выполнении процедуры формирования матрицы достижимости для неполносвязных графов производится их дополнение до полносвязных посредством присвоения весов отсутствующим связям равным ∞ . Выполнение данно-

го приведения обеспечивает возможность нахождения оптимального решения на последующих шагах реализации алгоритма поиска кратчайшего гамильтонова пути.

На основе анализа весовых характеристик матрицы достижимости формируется матрица условных индексов максимальных элементов по правилу:

$$i^* : \alpha_{i^*j} = \max [\alpha_{ij}], \quad (2)$$

где i^* — номер строки, в которой находится максимальный элемент, определенный при анализе столбцов, j — номер столбца, α_{ij} — числовое значение веса при переходе с S_i в S_j . Тогда матрица условных индексов формируется по следующему правилу:

$$\text{для } j = \overline{1, N}, \quad b_{ij} = \begin{cases} 1, & \text{если } i = i^* \\ 0, & \text{в противном случае} \end{cases} \quad (3)$$

Полученная в результате предыдущих вычислений матрица имеет вид, представленный на рис. 4.

	1	2	3	4	5	Σ
S_1	—	0	1	1	0	2
S_2	0	—	0	0	1	1
S_3	1	0	—	0	0	1
S_4	0	0	0	—	0	0
S_5	1	0	0	0	—	1

Рисунок 4. Матрица условных индексов

Пользуясь введенным эвристическим правилом выбора точки входа, на следующем этапе определим число единиц в каждой строке матрицы условных индексов:

$$S_{ij^*} = \sum_{j=1}^m b_{ij}, \quad \forall i \quad (4)$$

для приведенного примера максимальное число единиц равно 2 для первой строки. В результате проведенных вычислений на основе анализа результирующего столбца Σ , определим точку входа из соотношения:

$$S_{ij^*}^{\text{опт}} = \max S_{ij^*}, \quad \forall j^* \quad (5)$$

где $S_{ij^*}^{\text{опт}}$ — искомая точка входа в граф поиска кратчайшего гамильтонова пути. Найденное значение $S_{ij^*}^{\text{опт}}$ является исходным значением начала поиска для алгоритма [10]. В процессе поиска решения по соотношению (5) возможны 2 ситуации:

- среди полученных значений S_{ij}^* окажется одно значение являющееся максимальным;
- среди полученных значений S_{ij}^* окажется несколько значений, имеющих одинаковый вес и являющимися максимальными.

В первом случае последующий поиск производится от одной вершины $S_{ij}^{\text{опт}}$. Во втором случае, для избегания потери оптимального решения, необходимо производить поиск от всех вершин, удовлетворяющих соотношению (5). Схема алгоритма определения оптимальной точки входа приведена на рис. 5.

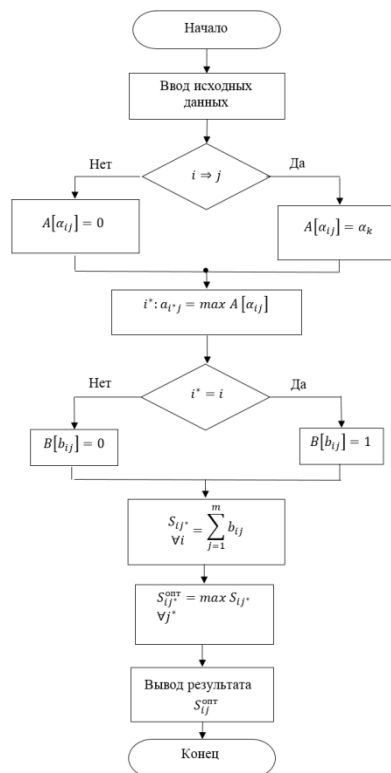


Рисунок 5. Схема процедуры поиска точки входа в граф с использованием анализа весовых характеристик

2. Алгоритм решения задачи поиска кратчайших гамильтоновых путей (ПКГП) в произвольных графах

Предлагаемый алгоритм ПКГП является результатом объединения алгоритма [10] и процедуры определения оптимальной точки входа в граф поиска кратчайшего гамильтонова пути (ПКГП ОТВ) (рис. 6).

В таком случае алгоритм ПКГП ОТВ будет состоять из следующей последовательности шагов:

- Шаг 1. Ввод исходных данных;
- Шаг 2. Формирование матрицы достижимости согласно (1);
- Шаг 3. Формирование матрицы условных индексов согласно правил (1)–(2);
- Шаг 4. Определение суммарного числа единиц в матрице условных индексов по соотношению (4);
- Шаг 5. Определение оптимальной точки входа в граф $G(V,E)$ (5).

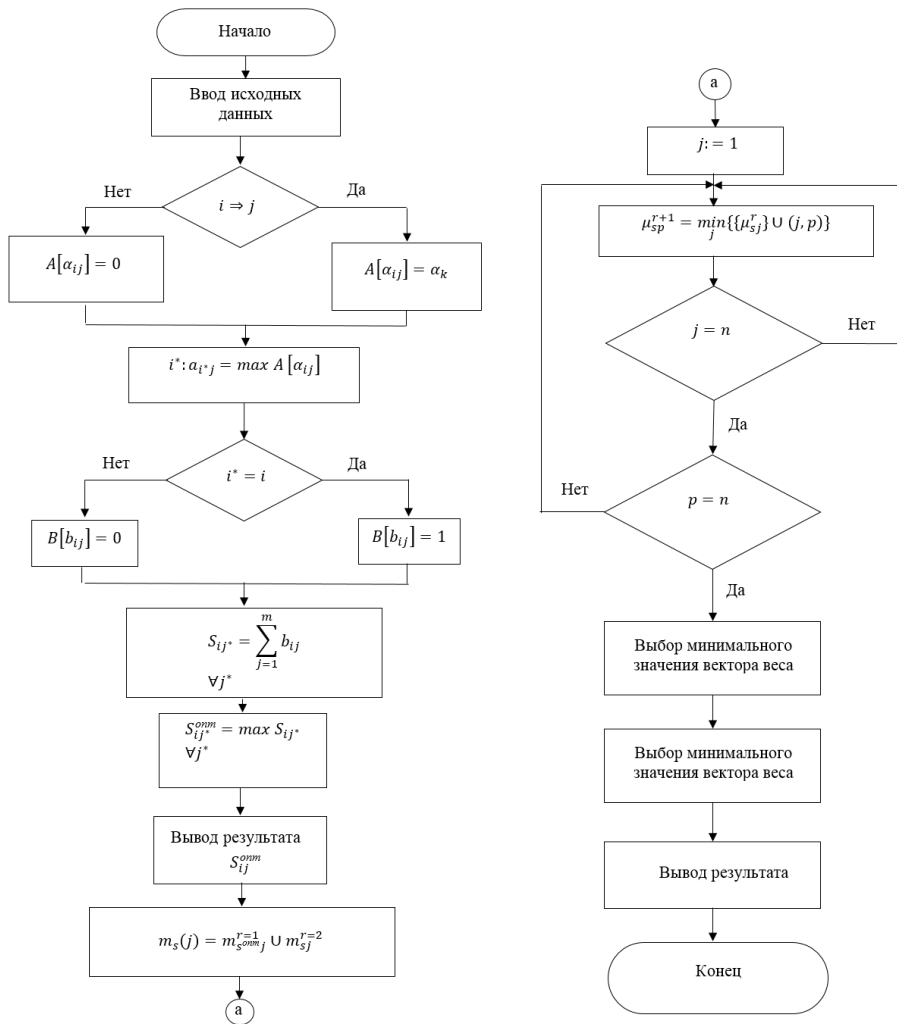


Рисунок 6. Схема работы алгоритма ПКГП ОТВ с использованием анализа весовых характеристик графа состояний

Шаг 6. В стянутом дереве всех путей используя процедуру А [10] определяем кратчайшие пути ранга $r = n - 1$ от вершины S ко всем остальным вершинам графа используя рекуррентное соотношение:

$$\mu_{sp}^{r+1} = \min\{\{\mu_{sp}^r\} \cup (j, p)\}; j = (\overline{1, n}); p = (\overline{1, n}); j \neq p. \quad (6)$$

Шаг 7. Среди всех построенных кратчайших путей ранга $r = n - 1$ на шагах 1 и 2 выбираем самый короткий в соответствии из соотношения (1) и алгоритм заканчивает работу.

Шаг 8. Вывод результата Π_{sj}^r .

Шаг 9. Конец работы алгоритма.

3. Результаты экспериментального исследования разработанного алгоритма

Для получения аналитических результатов экспериментального исследования алгоритма решения задачи поиска кратчайшего гамильтонова пути разработана программная модель, реализованная в прикладной тестирующей программе с условным названием Analizator.

Прикладная тестирующая программа Analizator предназначена для проведения тестов разработанного алгоритма и обработки результатов тестирования. Исходя из основного назначения прикладной программы предложено выделить три основных части программы, выполняющих функциональную нагрузку:

- область формирования исходных данных и настройки тестов;
- область анализа полученных результатов и их отображения на графиках;
- область отображения поранговых результатов тестирования и формирования записей в базе данных.

Использование указанных областей позволяет контролировать промежуточные значения реализации алгоритма, получать конечные результаты в виде графиков и таблиц базы данных. При этом дополнительно обеспечиваются следующие возможности:

- автоматический ввод расчетных данных;
- ручной ввод расчетных данных;
- отображение статистики;
- запись результатов в базу данных;
- простой и интуитивно доступный интерфейс.

При экспериментальном сравнении разработанных алгоритмов с известными весами ребер графа генерировались по равномерному закону распределения в диапазоне (0–50). Для получения среднего значения каждой точки графиков всех ана-

лизируемых характеристик решалось по 200 тестовых задач, все результаты статистического анализа получены с доверительной вероятностью 0,95.

Как видно из графиков, приведенных на рис. 7 при $n \geq 27$, алгоритмы Литтла, локального поиска и ПКГП имеют существенно более высокую временную сложность по сравнению с разработанным ПКГП ОТВ.

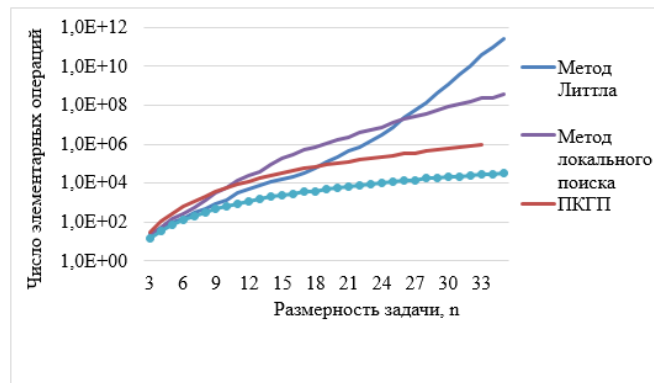


Рисунок 7. Зависимости числа элементарных операций от размерности решаемой задачи

Анализ производительности алгоритма показывает, что алгоритмы локального поиска имеют погрешность, лежащую в диапазоне от 7% до 29% и возрастающую с увеличением размерности решаемых задач. В разработанном алгоритме максимальное значение погрешности при $n > 15$ равняется $\Delta f = 2.24\%$. На больших размерностях наблюдается падение показателя погрешности и уже на размерности $28 < n < 100$ пиковое значение не превышает $\Delta f = 1.82\%$ (см. рис. 8).



Рисунок 8. Зависимость относительной погрешности от размерности решаемой задачи

Из графиков, приведенных на рис. 9, видно, что процент неточных решений для алгоритма ПКГП ОВТ принимает максимальное значение числа неточных решений, возникающих при решении задачи поиска кратчайшего гамильтонова пути при $p = 4.5\%$. Далее, в целом, наблюдается незначительный рост числа неточных решений.

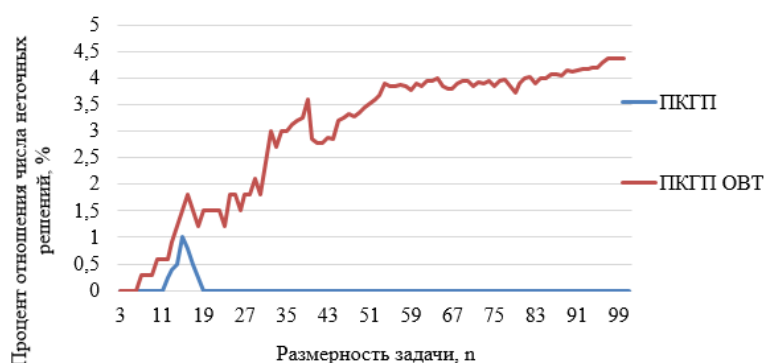


Рисунок 9. Зависимость числа неточных решений в % от размерности решаемой задачи

4. Заключение

Таким образом, предлагаемый способ решения задачи поиска кратчайшего гамильтонова пути может быть использован в сетевых базах данных для составления оптимальных расписаний выполнения запросов к базам данных, как на этапах реорганизации, так и на этапах сопровождения функционирования систем управления включающих базы данных. При этом предполагается существенное уменьшение времени на реализацию процедуры формирования очереди выполнения операций в запросах за счет уменьшения числа элементарных операций и числа обрабатываемых векторов в процедуре формирования очереди выполнения операций запроса.

Из экспериментального исследования разработанного алгоритма решения задач поиска кратчайшего гамильтонова пути следует, что в большинстве случаев удастся уменьшить временную сложность решения за счет незначительного увеличения вносимой погрешности. Так же из приведенного выше анализа видно, что предложенный алгоритм обладает большей производительностью чем [10] за счет использования процедуры поиска оптимальной точки входа в граф. Следует особо подчеркнуть, поскольку ПКГП ОВТ привносит дополнительную погрешность и % неточных решений, целесообразно использовать обозначенные алгоритмы адаптивно, реагируя на изменение размерности задачи для приведения погрешности решения задачи в диапазон $1\% < \delta f < 3\%$, при заданных временных ограничениях.

Тем самым возможно добиться существенного улучшения производительности совокупного решения по сравнению с известными методами.

Литература

- [1] Листровой С. В., Минухин С. В., Листровая Е. С. Разработка метода мониторинга распределенной вычислительной системы на основе определения кратчайших путей и кратчайших гамильтоновых циклов в графе // *Восточно-Европейский журнал передовых технологий*. 2015. Т. 6. № 4 (78). С. 32–45.
- [2] Дыйканов С. К. Задача о подсчете гамильтоновых путей и циклов // *Вестник современной науки*. 2016. Вып. 18. № 6.1. С. 8–9.
- [3] Ватутин Э. И., Дремов Е. Н., Мартынов И. А., Титов В. С. Метод взвешенного случайного перебора для решения задач дискретной комбинаторной оптимизации // *Известия Волгоградского государственного технического университета*. 2014. Вып. 137. № 10. С. 59–64.
- [4] Буй Д. Б., Скобелев В. Г. Модели, методы и алгоритмы оптимизации запросов в базах данных // *Компьютерные системы и информационные технологии*. 2014. № 2 (66). С. 43–58.
- [5] Частикова В. А., Власов К. А. Разработка и сравнительный анализ эвристических алгоритмов для поиска наименьшего гамильтонова цикла в полном графе // *Фундаментальные исследования*. 2013. № 10. С. 63–67.
- [6] Авезова Я. Э., Фомичев В. М. Условия примитивности и оценки экспонентов множеств ориентированных графов // *Прикладная дискретная математика*. 2017. № 35. С. 89–101.
- [7] Осиевский С. В., Тимченко А. А. Процедура определения точки входа на основе анализа векторов весовых характеристик исходного графа состояний // *Системы обработки информации*. 2002. Вып. 22. № 6. С. 281–284.
- [8] Федорин А. Н. Многокритериальные задачи ранцевого типа: разработка и сравнительный анализ алгоритмов: Дисс. ... канд. техн. наук: 05.13.18. Нижний Новгород, 2010. 132 с.
- [9] Листровой С. В., Голубничий Д. Ю., Листровая Е. С. Метод решения задач целочисленного линейного программирования с булевыми переменными на основе рангового подхода // *Электрон. моделирование*. 1998. № 6. С. 14–32.
- [10] Фильгус Д. И., Андрианова Е. Г., Раев В. К. Определение кратчайших гамильтоновых путей в произвольных графах распределенных баз данных // *Российский технологический журнал*. 2019. Т. 7. № 3.

Автор:

Дмитрий Игоревич Фильгус — аспирант кафедры инструментального и прикладного программного обеспечения Института информационных технологий, МИРЭА — Российский технологический университет

Development of transaction management methods and user requests based on the search for the entry point to the weighted graph in the distributed data-bases

D. I. Filgus

MIREA – Russian technological university
78, Prospect Vernadskogo, Moscow, Russia, 119571
e-mail: dmif42@ya.ru

Abstract. A method has been developed for finding the shortest Hamiltonian path in an arbitrary graph based on the rank approach, which provides high efficiency and a significant reduction in the error in solving the problem of organizing the process of managing multiple transactions and queries when they are implemented in network databases. In many cases, existing solutions do not provide the necessary results in terms of access time and accuracy of the found solution. Using the developed method allows minimizing idle time of computing devices, reducing the volume and time of data transfer from one device to another, increases overall scalability, minimizes data access time, etc. An important advantage of the proposed method is to reduce the number of elementary operations by using the optimal procedure. entry point in the graph, which leads to a significant reduction in time for the implementation of procedures for the formation of a queue of performing operations in queries. The proposed algorithm should be used adaptively, in combination with other solutions, in order to minimize the accuracy of the problem to be solved. Methods of graph theory used in this paper. The evaluation of the effectiveness of the task solution was performed using a systems approach, system analysis and the theory of operations research. Processing of experimental data obtained during the work was carried out in accordance with the provisions of mathematical statistics.

Keywords: graph, Hamiltonian path, query, rank approach, short tree of paths, transaction, rank, distributed database.

References

- [1] *Listrovoy S. V., Minukhin S. V., Listrovaya Ye. S. (2015) Vostochno-Yevropeyskiy zhurnal peredovykh tekhnologiy. 6-4(78):32–45.*
- [2] *Dyykanov S. K. (2016) Vestnik sovremennoy nauki. 18(6.1):8–9.*
- [3] *Vatutin E. I., Dremov E. N., Martynov I. A., Titov V. S. (2014) Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta. 137(10):59–64.*
- [4] *Buy D. B., Skobelev V. G. (2014) Komp'yuternyye sistemy i informatsionnyye tekhnologii. 2(66):43–58.*
- [5] *Chastikova V. A., Vlasov K. A. (2013) Fundamental'nyye issledovaniya. 10:63–67.*
- [6] *Avezova Ya. E., Fomichev V. M. (2017) Prikladnaya diskretnaya matematika. 35:89–101.*
- [7] *Osyevskyy S. V., Tymchenko A. A. (2002) Systemy obrobky informatsiyi. 22(6):281–284.*
- [8] *Fedorin A. N. Mnogokriterial'nyye zadachi rantsevogo tipa: razrabotka i sravnitel'nyy analiz algoritmov. Tesis. Nizhniy Novgorod, 2010.*
- [9] *Listrovoy S. V., Golubnichiy D. Yu., Listrovaya Ye. S. (1998) Elektron. modelirovaniye. 6:14–32.*
- [10] *Fil'gus D. I., Andrianova Ye. G., Rayev V. K. (2019) Rossiyskiy tekhnologicheskii zhurnal. 7(3)*