

ИСПОЛЬЗОВАНИЕ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОГО АНАЛИЗА ДЛЯ РЕШЕНИЯ ЗАДАЧИ КОММИВОЯЖЕРА

Василий Шуть ¹⁾, Игорь Прожерин ²⁾

1) Василий Шуть к.т.н., доцент, Беларусь Брест Пушкинская-78-105

2) Игорь Прожерин преподаватель, Беларусь Брест Ленинградская-7,81, igor_prozherin@tyt.by

Аннотация: Рассматривается задача о коммивояжере и ее решение методом корреляционно-регрессионного анализа. Приводится математическая модель данного метода, который позволяет решить задачу при использовании минимального объёма памяти $n^2 + 3n$ и имеет малую трудоёмкость $\sim o(n)$.

Ключевые слова: задача коммивояжера, корреляционно-регрессионный анализ

1. ОБЗОР РАЗЛИЧНЫХ МЕТОДОВ

Задача коммивояжера занимает центральное место среди труднорешаемых задач комбинаторной (дискретной) оптимизации. Все существенные идеи решения таких задач или были первоначально предложены для решения задачи коммивояжера, или, как правило, прошли проверку на этой задаче [1].

Задача коммивояжера имеет множество вариантов и в основном классифицируются типом матрицы расстояний (симметричная, несимметричная, метрическая, неметрическая) и по методу решения (точные и приближённые алгоритмы). К точным методам относятся: полного перебора, динамического программирования, ветвей и границ, множителей Лагранжа, методы отсекающих плоскостей, композитные алгоритмы [2]. К приближённым методам относятся: эвристические, асимптотические и генетические алгоритмы [3].

В развитии вычислительных алгоритмов решения трудных задач комбинаторной оптимизации и, конечно, в первую очередь задачи коммивояжера в последние годы четко обозначились две тенденции. Первая состоит в разработке алгоритмов, содержащих большое число вычислительных процедур, реализующих вычисления различных нижних оценок, набор высокоэффективных эвристик, использование множителей Лагранжа и отсекающих плоскостей, различные правила разбиения и ветвления, процедуры анализа и упрощения информации. Программные комплексы, реализующие такие алгоритмы, часто содержат программы управления вычислительным процессом и программы решения задач с участием человека в диалоговом режиме. Такие алгоритмы ориентированы на скрупулезный учёт специфики

исходных данных задачи и призваны доставлять решение за минимальное время вычислительной системы. Конечно, в этом случае достаточно большими становятся расходы на алгоритмическое и программное обеспечение алгоритма [2].

Другая тенденция состоит в снижении расходов на программную реализацию алгоритмов. Исходная задача формулируется таким образом, что для ее решения используются только коммерческие пакеты или пакеты программ стандартного математического обеспечения ЭВМ [2].

Эвристические алгоритмы решения задачи коммивояжера являются основным инструментом решения практических задач. Появление эвристик обусловлено, в первую очередь, излишней чувствительностью точных алгоритмов по отношению к специфике задачи и наличию дополнительных условий. Эвристики представляют собой попытку учесть специфику задачи простыми средствами, создать прием, эффективный для решения задач с определённой особенностью [2].

Коммивояжер периодически посещает n городов. Расстояние между городами определяется числами c_{ij} из города i в j . Если прямого пути из города i в j нет, то c_{ij} принимает значение ∞ . Из этих расстояний составляется матрица расстояний C диагональ, которой заполняется ∞ .

Требуется составить маршрут передвижения коммивояжера, в котором:

- посещаются все города;
- ни один город, кроме начального, не посещается дважды;
- этот маршрут должен обладать минимальной суммарной протяжённостью.

Задача о коммивояжере сводится к поиску гамильтонова контура минимальной длины. [4]

Будем рассматривать дискретную, нецелочисленную, симметричную задачу коммивояжера.

Одним из самых простых алгоритмов решения задачи коммивояжера является алгоритм полного перебора. Этот алгоритм обладает очень существенным недостатком – время решения экспоненциально зависит от количества точек. Это означает, что реально для решения практических задач этот метод не применим. [4]

Алгоритм ветвей и границ более сложный, чем метод полного перебора. Его положительной стороной является точность решения. При реализации этого алгоритма на ЭВМ возникают проблемы использования ресурсов памяти и времени, т.к. требуется хранить большие объёмы информации. Данный алгоритм требует временных затрат $o(\sim 7n^2)$ и памяти $\sim (2n^3)$ ячеек [2,5].

Генетические алгоритмы были предложены в работах J. Holland, а затем развиты D. Goldberg и многими другими авторами. Эти алгоритмы основаны на моделировании размножения в природе и случайных изменениях, которые происходят во время мутации. Данный алгоритм не даёт оптимальности, но гарантирует какое-либо решение задачи за конечное заданное время t . Генетический алгоритм требует n^2 ячеек памяти, а трудоёмкость зависит не только от количества вершин графа, но и от ограничивающего параметра $o(k)$. Качество решения такого алгоритма будет зависеть от двух параметров k и n , причём при $k \rightarrow \infty$ качество алгоритма стремится к 100%. [1,6]

В работах Гимади Э. Х. [7,8] рассматривается приближенный алгоритм решения задачи коммивояжера. Плоскость разбивается параллельными линиями на конечный ряд уровней. В каждом уровне точки соединяются частичными путями, а затем уровни соединяются между собой перемычками. Этот алгоритм не отвечает на следующие вопросы: каким образом выбирать линии для разбиения, и каким должно быть число уровней? Данные вопросы являются наиболее важными, так как от их решения зависит точность полученных результатов (оптимального контура).

2. АЛГОРИТМ КОРРЕЛЯЦИОННО-РЕГРЕССИОННОГО АНАЛИЗА

В настоящей работе решается задача определения вектора, указывающего направление, по которому необходимо производить деление плоскости параллельными линиями. Таким вектором является линия регрессии множества точек (городов). В работе рассматривается только линейная регрессия, хотя может быть и любая другая, наиболее адекватно отражающая геометрию расположения городов.

Второй задачей, решаемой настоящей статьёй является определение оптимального числа уровней деления.

Рассмотрим процесс построения гамильтонова контура по точкам расположенным на плоскости с использованием линии регрессии, показывающей тенденцию расположения совокупности точек:

1. определяем коэффициенты линии регрессии по координатам точек;
2. преобразуем координаты точек в систему, повернутую на угол наклона линии регрессии, причём чтобы значения ординаты были положительными;
3. разбиваем плоскость линиями параллельными оси абсцисс на уровни;
4. формируем частичные пути в каждом уровне;
5. объединяем частичные пути между собой таким образом, что конец пути полученного на первом уровне соединяется с концом пути полученного на втором уровне, а начало пути полученного на втором уровне соединяется с началом пути полученного на третьем уровне и т.д., а затем крайние точки полученного пути замыкаются между собой.

В результате выполнения алгоритма получим путь движения коммивояжера. Рассчитываем длину данного пути по матрице расстояний.

Решение задачи зависит от выбора количества уровней. Понятно, что данное число зависит от количества точек. Поэтому необходимо определить каким образом проявляется эта зависимость. Можно предложить различные методы определения количества уровней, например, в зависимости от количества точек с использованием коэффициента, в зависимости от плотности точек.

Рассмотрим метод определения количества уровней в зависимости от количества точек с использованием коэффициента. Коэффициент можно определить экспериментальным путём следующим образом:

1. решить задачу при различных значениях числа уровней $k = 1 \dots n$, где n - количество точек;
2. определить минимальный полученный путь и количество уровней при этом;
3. определить коэффициент отношения количества точек к числу уровней, полученному на шаге 2;
4. выполнить шаги 1,2 и 3 при различном количестве точек $\{10, 50, 100, 500, 1000\}$;
5. усреднить полученные коэффициенты, т.о. получим число определяющее целесообразное количество уровней.

Трудоёмкость такого алгоритма будет зависеть от количества точек $\sim o(n)$, а количество памяти необходимое для решения задачи будет зависеть от количества точек, а именно $n^2 + 3n$ ячеек.

Построим математическую модель корреляционно-регрессионного метода решения задачи коммивояжера.

В общем случае любая линия на плоскости может быть представлена уравнением:

$$y = a_0 + a_1 x \quad (1)$$

Коэффициенты уравнения (1) или коэффициенты линии регрессии определяются следующими выражениями [9]:

$$a_1 = \left(\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i * \sum_{i=1}^n y_i \right) / \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right); \quad (2)$$

$$a_0 = \frac{1}{n} \left(\sum_{i=1}^n y_i - a_0 * \sum_{i=1}^n x_i \right). \quad (3)$$

Преобразуем координаты точек в систему, повернутую на угол наклона линии регрессии, причём чтобы значения ординат были положительными, используем преобразование:

$$x_i' = (x_i - x) * \cos \alpha + (y_i - y) * \sin \alpha; \quad (4)$$

$$y_i' = -(x_i - x) * \sin \alpha + (y_i - y) * \cos \alpha \quad (5)$$

где:

$$\sin \alpha = \frac{a_1}{\sqrt{1 + a_1^2}}; \quad \cos \alpha = \frac{1}{\sqrt{1 + a_1^2}}.$$

Разбиваем плоскость линиями параллельными оси абсцисс на уровни следующим образом:

$$pos = \left\lceil \frac{x_i'}{m} \right\rceil, \quad (6)$$

где pos - номер уровня, к которому отно-

сится точка, m - количество необходимых уровней.

При формировании частичных путей в каждом уровне создаются списки из координат точек в зависимости от уровня. Для получения гамильтонова контура частичные пути объединяются между собой так: конец пути полученного на первом уровне соединяется с концом пути полученного на втором уровне, а начало пути полученного на втором уровне соединяется с началом пути полученного на третьем уровне и т.д. Пример объединения точек показан на рис. 1.

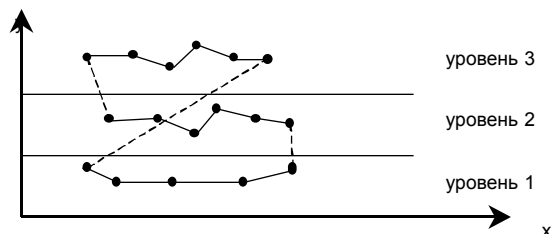


Рис.1 Объединение частичных путей в гамильтонов контур.

3. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ

Теоретическое сравнение рассмотренных алгоритмов по вычислительной трудоёмкости и количеству памяти представлено в табл. 1.

Таблица 1. Теоретическое сравнение затрачиваемых ресурсов

| Тип алгоритма | Трудоёмкость | Память | Качество |
|--------------------------------------|----------------|---------------|----------------|
| Прямого перебора | $n!$ | n^2 | Точный |
| Ветвей и границ | $o(\sim 7n^2)$ | $\sim (2n^3)$ | Точный |
| Генетический | $o(k * n)$ | $\sim n^2$ | Зависит от k |
| Приближённый | $o(n \ln n)$ | $\sim n^2$ | Зависит от n |
| Корреляционно-регрессионного анализа | $\sim o(n)$ | $n^2 + 3n$ | Приближённый |

Информация из табл. 1 представлена на рис. 2 в форме графика трудоёмкости вычислений алгоритмов.

Из рис. 2 видно, что метод корреляционно-регрессионного анализа является наименее трудоёмким. При этом объём используемой памяти не больше, чем в других алгоритмах.

3. ВЫВОДЫ

В заключении следует отметить, что разработка алгоритмов поиска гамильтонова контура является важной задачей, которую человек решает уже многие годы. Известные алгоритмы, предложенные ранее, либо требуют использования большого объёма памяти, либо имеют большую трудоёмкость для решения задачи коммивояжера.

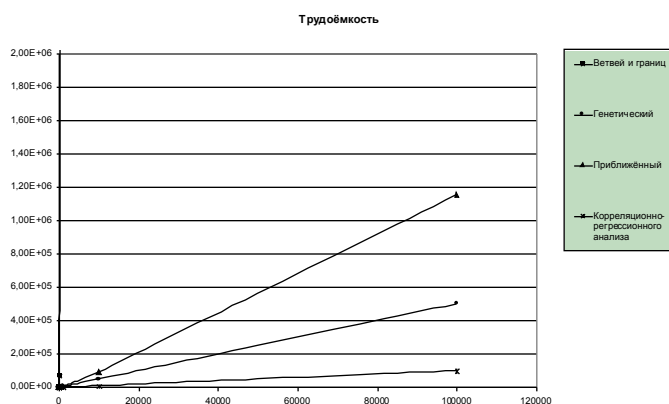


Рис. 2 График трудоемкости вычислений алгоритмов

В данной статье был предложен алгоритм быстрого поиска гамильтонова контура с помощью корреляционно-регрессионного анализа, который вобрал в себя достоинства ранее предложенных методов, а именно минимальные требования к памяти при достаточно быстром приемлемом решении задачи. Данный метод относится к приближенным методам решения задачи коммивояжера и ни в коем случае не претендует на окончательное решение данной задачи.

6. ЛИТЕРАТУРА

[1] Меламед И. И., Сергеев С. И., Сигал И. Х., *Задача коммивояжера. Вопросы теории // Автоматика и телемеханика.* – М.: Наука, 1989. №9. с.3-33.

[2] Меламед И. И., Сергеев С. И., Сигал И. Х., *Задача коммивояжера. Точные методы // Автоматика и телемеханика.* – М.: Наука, 1989. №10. с.3-29.

[3] Меламед И. И., Сергеев С. И., Сигал И. Х., *Задача коммивояжера. Приближенные алгоритмы // Автоматика и телемеханика.* – М.: Наука, 1989. №11. с.3-26.

[4] Новиков Ф. А., *Дискретная математика для программистов,* СПб.: Питер, 2000.–304 с.:ил.

[5] Перепелица В. А., Гимади Э. Х., *К задаче нахождения минимального гамильтонова контура на графе со взвешенными дугами, Сб. «Дискретный анализ», Новосибирск, вып. 15, 1969, 57-65.*

[6] Kaluga V. V., Muravjev S. A., Siridonov S. V., Telyatnikov R. V., *Application of genetic algorithms for solutions of the task is frequent – territorial plannings group radio electronic equipment, International Conference of Neural Networks and Artificial Intelligence ICNNAI'99| Proceedings. Edited by Vladimir Golovko, - Brest: BPI, 1999, 224p.*

[7] Гимади Э. Х., Перепелица В. А., *Асимптотический подход к решению задачи коммивояжера, Сб. «Управляемые системы», Новосибирск, вып. 12, 1974, 35-45.*

[8] Гимади Э. Х., Перепелица В. А., *Статистически эффективный алгоритм выделения гамильтонова контура (цикла), Сб. «Дискретный анализ», Новосибирск, вып. 22, 1973, 15-28.*

[9] Кузнецов А. В., Сакович В. А., Холод Н. И., *Высшая математика. Математическое программирование, Мн.: Выш. шк., 1994.*



Игорь Прожерин закончил Брестский государственный университет по специальности «Автоматизированные системы обработки информации» с отличием. В данный момент являюсь аспирантом кафедры «Информатики и прикладной математики» в Брестском государственном техникуме.

Занимаюсь задачами оптимизации, а именно труднорешаемыми (NP-полными) задачами. Одной из таких задач является задача коммивояжера. Если найти метод решения хотя бы одной NP-полной задачи, то можно решать любые задачи, принадлежащие этому классу. Поэтому данная статья посвящена именно задаче комми-

вожера. На брестском узле железной дороги возникла задача распределения работы трех козловых кранов на погрузочно-разгрузочной площадке (используется для перегрузки контейнеров с западной на восточную колею). Центральное место в данной задаче занимает алгоритм получения гамильтонова контура. Условия заказчика таковы, что задача должна решаться не более чем за 10 мин, поэтому основным является создание быстрого алгоритма решения задачи коммивояжера.