УДК 519:68

Р.Х. Ахмадов, Н.И. Ящук

Национальный технический университет «ХПИ», Харьков

ОТЫСКАНИЕ ГАМИЛЬТОНОВА ПУТИ НА НЕПОЛНОДОСТУПНОМ СТОХАСТИЧЕСКОМ ГРАФЕ

Классическая задача отыскания гамильтонова пути на графе рассмотрена для случая, когда длины ребер графа заданы неточно. Предложено решение задачи для следующих вариантов учета неопределенности: элементы матрицы – случайные величины с известной плотностью распределения; элементы матрицы заданы значениями математического ожидания и дисперсии; элементы матрицы определяются в предположении о наихудшей плотности распределения случайных значений полезности, для которой вероятность попадания случайной полезности в недопустимый диапазон максимальна.

Ключевые слова: *неполнодоступный граф, гамильтонов путь, длины ребер заданы неточно, генетический алгоритм.*

Введение

Пусть граф задан множеством n вершин и соединяющих их дуг. Задача состоит в построении замкнутого маршрута обхода всех вершин, не проходящего ни через одну из них дважды. Такой маршрут называется гамильтоновым [1].

Эффективность маршрута определяется значением некоторой выбранной характеристики (например, суммарная стоимость обхода или его продолжительность и т.п.).

В качестве такой характеристики удобно использовать широко трактуемое понятие «полезность» (в смысле Дж. Неймана и О. Моргенштерна [2]. В соответствии с этим будем считать, что задана числовая матрица размера $n \times n$, каждый элемент которой, например, лежащий в i-й строке и j-м столбце, задает значение полезности s_{ij} использования в маршруте участка (i, j), начинающего в вершинеі и заканчивающегося в вершине j. Граф неполнодоступен, то есть некоторая вершина j может быть недостижима непосредственно из некоторой вершины i.

В этом случае значение полезности s_{ij}, соответствующее элементу (i, j) матрицы будет равно нулю.

Задача построения замкнутого маршрута без петель, обходящего все вершины, называется задачей коммивояжера [3].

Формальная постановка задачи имеет вид: найти набор $X = (x_{ij})$, доставляющий минимум линейной форме

$$L(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij}\, x_{ij},$$

компоненты которой удовлетворяют следующим ограничениям

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j = 1,2,...,n,$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = 1,2,...,n,$$

$$u_i - u_j + n x_{ij} \le n - 1,$$

$$i = 1,2,...,n, \; j = 1,2,...,n.$$

Здесь

n – число пунктов,

s_{ij} – полезность участка (i, j),

x_{ij} – индикатор, равный единице, если в маршруте имеется звено, соединяющее непосредственно пункты (i, j), и равный нулю в противном случае.

Выполнение приведенных выше ограничений обеспечивает получение замкнутого маршрута без петель.

Эта задача принадлежит к классу трудных комбинаторных, так называемых NP - полных задач. Эффективность известных алгоритмов ее решения невелика – успешное решение может быть получено только для задач сравнительно небольшой размерности $(n < 20)$.

На практике возникает необходимость решения этой задачи гораздо более высокой размерности. Определенный оптимизм вызывает возможность использования для решения этой задачи генетического алгоритма (ГА) [4].

Задан неполнодоступный неориентированный граф с n вершинами.

Рассмотрим ситуацию, когда матрица, элементы которой определяют значения полезности соответствующих ребер графа, задана неточно.

Для решения задачи построения маршрута в этом случае используем генетический алгоритм. При этом технология расчета фитнесс-функции должна конструироваться с учетом того, каким образом описана неопределенность исходных данных.

Поставим задачу отыскания маршрута с максимальной полезностью для следующих вариантов неопределённости: элементы матрицы – случайные величины с известной плотностью распределения; элементы матрицы заданы значениями математического ожидания и дисперсии; элементы матрицы определяются в предположении о наихудшей плотности распределения случайных значений полезности, для которой вероятность попадания случайной полезности в недопустимый диапазон максимальна.

Основной материал

Простейший вариант задачи возникает, если случайные значения полезности определены своими плотностями распределения.

Будем считать, что случайное значение полезности S_{ij} использования участка (i, j) в маршруте описывается своей плотностью распределения $\phi(s_{ij})$. Простейший метод отыскания наилучшего маршрута в этом случае – оптимизация в среднем. При этом для каждого участка (i, j) рассчитывается среднее значение полезности

$$m_{ij} = \int_0^{\infty} s_{ij}\phi\left(s_{ij}\right) ds_{ij}, \ i = 1,2,...,n, \ j = 1,2,...,n \ . \quad (1)$$

Теперь каждому гамильтонову пути, заданному порядком обхода вершин графа $J = \{i_1, i_2,.., i_n\}$, i_k - номер вершины графа, занимающей k-ю, позицию в маршруте, ставится в соответствие значение фитнесс-функции

$$L(J) = \sum_{k=1}^{n-1} m_{i_k,\, i_{k+1}}. \quad (2)$$

При этом решением задачи будем считать маршрут, полученный генетическим алгоритмом, с максимальной суммарной средней полезностью.

Оптимизация в среднем обеспечивает получение приемлемых маршрутов, когда дисперсии полезности участков невелики. Если это не так, то реальный маршрут, полученный в результате оптимизации в среднем, может оказаться недопустимо плохим. В этом случае более надежный маршрут может быть определен следующим образом. Для каждого участка (i, j) рассчитывается вероятность того, что случайное значение полезности использования этого участка s_{ij} превзойдет допустимый порог $s_{ij}^{(0)}$.

$$P_{ij} = \text{Вер}\left\{s_{ij} \ge s_{ij}^{(0)}\right\} = \int_{s_{ij}^{(0)}}^{\infty} \phi(s_{ij})ds_{ij}. \quad (3)$$

При этом количественной характеристикой маршрута будет вероятность того, что полезность всех его участков превысит пороговую

$$P = \prod_{k=1}^{n-1} P_{k,k+1} \ . \quad (4)$$

Тогда решением задачи будет маршрут, соответствующий максимальному найденному ГА значению фитнесс-функции (4).

Во многих практических ситуациях в силу недостаточности исходных статистических данных отсутствует возможность получения достаточно адекватного описания плотностей распределения случайных значений полезности участков. Однако доступная для обработки выборка позволяет рассчитать оценки основных статистических характеристик этих случайных величин: средние значения и дисперсии.

В этом случае целесообразно отыскивать маршрут $S = \left\{s_1, s_{2,} ..., s_n\right\}$, который максимизирует комплексный критерий

$$I = \frac{L(S)}{D(S)}, \quad (5)$$

$$L(S) = \sum_{k=1}^{n-1} m_{s_k, s_{k+1}}, \quad D(S) = \sum_{k=1}^{n-1} D_{s_k, s_{k+1}}, \quad (6)$$

$$D_{s_{k,k+1}} = \int_0^{\infty} s_{k,k+1}^2 \phi(s_{k,k+1}) ds_{k,k+1} - m_{k,k+1}^2. \quad (7)$$

Смысл фитнесс-функции (5)–(7) понятен: наилучшим будет маршрут, максимизирующий суммарную среднюю полезность при минимизации суммарной ее дисперсии.

Рассмотрим, наконец, метод построения фитнесс-функции в условиях малой выборки исходных данных. Суть метода состоит в отыскании для каждого участка (i, j) наихудшей плотности распределения случайной величины его полезности при заданных значениях ее математического ожидания m_{ij} и дисперсии σ_{ij}^2.

Соответствующая формулируется следующим образом: найти плотность распределения $f(s_{ij})$, максимизирующую вероятность недостижения случайной величиной полезности S_{ij} порогового значения $s_{ij}^{(0)}$.

Формальная модель задачи:

целевая функция

$$P_{ij}^{(s)} = \int_0^{s_{ij}^{(0)}} f(s_{ij})ds_{ij} \Rightarrow \max,$$

ограничения

$$\int_0^{\infty} f(s_{ij})ds_{ij} = 1,$$

$$\int_0^{\infty} s_{ij}f(s_{ij})ds_{ij} = m_{ij},$$

$$\int_0^{\infty} (s_{ij} - m_{ij})^2 f(s_{ij})ds_{ij} = \sigma_{ij}^2.$$

Полученная задача является задачей континуального линейного программирования [5]. При этом, как показано в [5], искомая вероятность попадания в недопустимый интервал $[0, s_{ij}^{(0)}]$ определяется соотношением

$$P_{ij}^{(S)} = \frac{\sigma_{ij}^2}{m_{ij}^2 + \sigma_{ij}^2 - 2m_{ij}s_{ij}^{(0)} + (s_{ij}^{(0)})^2}.$$

Тогда количественной характеристикой маршрута является вероятность того, что полезность всех его участков будет ниже пороговой

$$P^{(S)} = \prod_{k=1}^{n-1} P_{k,k+1}^{(S)}. \quad (8)$$

Решение задачи – маршрут, соответствующий минимальному найденному ГА значению фитнесс-функции (8).

Выводы

Таким образом, рассмотрены методы решения задачи отыскания гамильтонова пути на графе с использованием генетического алгоритма. Показано, что конструктивные особенности и характер предлагаемого метода формирования фитнесс-функции зависят от уровня неопределенности имеющихся данных. Во всех рассмотренных ситуациях описанные методы в максимальной степени учитывают реальный объем информационного обеспечения технологии статистической обработки исходных данных.

Список литературы

1. Берж К. Теория графов и её применение / К. Берж; пер. с франц. – М.: ИИЛ, 1962. – 320 с.

2. Дж. фон Нейман. Теория игр и экономическое поведение / Дж. фон Нейман., О. Моргенштерн; пер. с англ. – М.: Наука, 1970. – 707 с.

3. Flood M. The Traveling Salesman Problem / M. Flood // Operations Research, 1958. - N6. – P. 791 – 814.

4. Goldberg D. Genethic Algorithms / D. Goldberg. – MA: Addison Wesley, 1989. – 210 p.

5. Раскин Л.Г. Прикладное континуальное линейное программирование / Л.Г. Раскин, И.О. Кириченко, О.В. Серая. – Х.: АВВ МВС Украины, 2004. – 292 с.

Поступила в редколлегию 10.08.2015

Рецензент: д-р техн. наук, проф. Л.Г. Раскин, Национальный технический университет «ХПИ». Харьков.

ВІДШУКАННЯ ГАМІЛЬТОНОВА ШЛЯХУ НА НЕПОВНОДОСТУПНОМУ СТОХАСТИЧНОМУ ГРАФІ

Р.Х. Ахмадов, Н.І. Ящук

Класична задача відшукання гамільтонова шляху на графі розглянута для випадку, коли довжини ребер графа задані неточно. Запропоновано вирішення задачі для наступних варіантів обліку невизначеності: елементи матриці - випадкові величини з відомою щільністю розподілу; елементи матриці задані значеннями математичного очікування й дисперсії; елементи матриці визначаються в припущенні про найгіршу щільності розподілу випадкових значень корисності, для якої ймовірність попадання випадкової корисності в неприпустимий діапазон максимальна.

***Ключові слова**: неповнодоступний граф, гамільтонів шлях, довжини ребер задані неточно, генетичний алгоритм.*

DETERMINATION OF HAMILTONIAN PATH ON THE INCOMPLETE STOCHASTIC COLUMN

R.H. Akhmadov, N.I. Yashchuk

TThe classical problem of finding a Hamiltonian path in a graph is considered for the case when the length of the edges are set inaccurately. The solution of the problem is proposed for the following account of uncertainty: the elements of the matrix are random variables with known density distribution; the value of the elements are defined by the expectation and variance; elements of the matrix are determined on the assumption of worst-density distribution of the random utility values for which the probability of hitting the random utility in the invalid range reaches the maximum.

***Keywords**: incomplete count, the Hamilton's path lengths of the edges are set inaccurately, genetic algorithm.*