УЧЕНЫЕ ЗАПИСКИ КАЗАНСКОГО УНИВЕРСИТЕТА.
СЕРИЯ ФИЗИКО-МАТЕМАТИЧЕСКИЕ НАУКИ

2024, Т. 166, кн. 4 ISSN 2541–7746 (Print)
С. 639–650 ISSN 2500–2198 (Online)

ОРИГИНАЛЬНАЯ СТАТЬЯ

УДК 519.854.2: 519.245: 519.217.2 doi: 10.26907/2541-7746.2024.4.639-650

РЕШЕНИЕ ЗАДАЧИ КОММИВОЯЖЁРА МЕТОДОМ СТАТИСТИЧЕСКИХ ИСПЫТАНИЙ ПРИ ИСПОЛЬЗОВАНИИ СЛОЖНЫХ ЦЕПЕЙ МАРКОВА

С.В. Шалагин

Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ, г. Казань, 420111, Россия

Аннотация

Предложен метод решения задачи коммивояжёра (ЗК) с применением аппарата N-сложных цепей Маркова (N-ЦМ), последовательность состояний которых имитирует путь через N пунктов, каждый из пунктов посещается только один раз. Для каждого из пунктов задана вероятность перехода в один из l следующих пунктов, $l < N$. Выполнен анализ сложности реализации каждого из этапов предложенного метода в зависимости от заданных значений N и l. Получены оценки сложности генератора N-ЦМ на основе композиции конечного детерминированного автомата и вероятностного автомата без памяти. Сложность генератора N-ЦМ характеризуется объёмом множеств входов, внутренних состояний и выходов, а также объёмом памяти, требуемой для реализации функций переходов и выходов указанных автоматов. Даны оценки времени задержки функционирования генератора N-ЦМ. Рассчитаны вероятность генерирования допустимых путей, т. е. удовлетворяющих решению ЗК, и объём памяти, требуемой для хранения количества допустимых путей.

Ключевые слова: задача коммивояжёра, метод решения, сложная цепь Маркова, статистическое испытание, оценка сложности

Введение

Одной из важных задач комбинаторики является задача коммивояжёра или задача о странствующем торговце (англ. «Travelling Salesman Problem», TSP) [1]. Она сводится к поиску пути, который коммивояжёр проходит через промежуточные пункты по одному разу, возвращаясь в исходную точку, и является оптимальным по заданному критерию. Критерием оптимальности маршрута могут выступать различные параметры: минимальные время поездки, расходы на дорогу или длина пути. Возможно применение интегрального критерия, учитывающего все из указанных параметров. Существуют различные методы решения TSP [2]: полного перебора (грубой силы), ветвей и границ, случайного перебора, жадного алгоритма и др. Методы решения TSP делятся на точные [3] и приближённые [4].

Представляют интерес приближённые методы решения TSP [4, 5], основанные на статистических испытаниях. Общая идея метода статистических испытаний заключается в том, что вместо аналитического описания некоторого процесса производится его многократное вероятностное моделирование. В итоге получим множество реализаций, которые в совокупности можно обработать методами математической статистики. Мощность множества реализаций позволяет определить точность решения, получаемого на основе метода статистических испытаний. Приемлемую точность решения можно получить только при использовании средств вычислительной техники как общего, так и специального назначения [6, 7].

Метод статистических испытаний позволяет применить для решения широкого круга задач аппарат теории цепей Маркова (ЦМ) или марковских последовательностей [8]. Процессы, моделируемые при использовании ЦМ, характеризуются тем, что состояние ЦМ в следующий момент времени определяется только её состоянием в текущий момент времени. Существуют также сложные ЦМ, для которых состояние в следующий момент времени определяется множеством её состояний, в которых она находилась в последние N моментов времени, N-сложные ЦМ, $N > 1$ [9]. Известны различные реализации схемы генератора N-сложной ЦМ (далее — N-ЦМ), представленные на абстрактном и структурном уровнях [10–18]. В частности, в [10] показано, что схема генератора N-ЦМ, заданной стохастической матрицей (СМ), может быть представлена в виде определенной композиции двух конечных автоматов — детерминированного автомата Мура (КДА) и вероятностного автомата без памяти (ВА).

Ниже предложен подход к решению задачи TSP с применением N-сложной цепи Маркова при использовании метода статистических испытаний как метода, обладающего рядом преимуществ: малой связностью и устойчивостью к случайным сбоям. Это позволяет распараллелить процесс генерирования путей с последующим выбором среди найденных решений приближённого к оптимальному по заданному критерию.

1. Постановка TSP на основе N-ЦМ

Решим TSP для множества из N пунктов, один из которых выбран в качестве начального. Затем по определённому правилу произведем переход в следующий пункт, и так до тех пор, пока не будут пройдены все пункты, причём каждый из них посещается только один раз. Последовательность пройденных пунктов определим как путь V или, согласно определению из [10], как траекторию N-ЦМ. Последний пункт пути определим как финальный.

В TSP для каждого пункта заданы возможная прибыль коммивояжёра от перемещения в пункт j и затраты (материальные и/или временные) на перемещение из пункта i в пункт j. Первую величину обозначим как d_j, а вторую — как h_{ij}, $i, j = \overline{1, N}$. Абстрагируемся от конкретных параметров, определяющих значения выгоды и затрат. Для каждой пары пунктов введём функционал, характеризующий выгоду коммивояжёра от перемещения в пункт j из пункта i:

$$f_{ij} = d_j - h_{ij}. \tag{1}$$

Для каждого пункта на основании (1) рассчитаем вероятностный закон распределения перехода в следующий пункт. Особенностью этого закона является то, что вероятность перехода в уже пройденный пункт должна быть нулевой. Пусть из начального пункта в текущий пункт ведёт путь длины k, $k < N$. Тогда количество допустимых переходов для текущего пункта не будет превышать $N - k$.

Справедлива

Теорема 1. *Закон распределения вероятности перехода в следующий пункт j из текущего пункта i при прохождении пути V, включающего k заданных пунктов, определённый на основе* (1), *имеет вид*

$$\left\{\frac{q_{ij}}{Q_i}\right\}, \quad q_{ij} = \begin{cases} 0, & j \in V, \\ f_{ij}, & \text{иначе}, \end{cases} \quad \text{где} \quad Q_i = \sum_{j=1}^{N} q_{ij}, \quad i, j = \overline{1, N}. \tag{2}$$

Определим количество всевозможных путей движения коммивояжёра. При условии, что начальный пункт задан, верхняя его оценка равна величине $(N-1)!$ [1,2].

Для каждого пункта i переходы в другие пункты, ещё не включенные в путь V длины k, ранжируем по убыванию значений q_{ij}. Примем следующее ограничение: из заданного пункта может быть выполнен переход только в l пунктов, $l < N$, с максимальными значениями q_{ij}. Согласно этому ограничению некоторые переходы из одного пункта в другой окажутся запрещёнными. Тогда количество всевозможных допустимых путей будет ограничено сверху величиной $l^{N-l-1}\, l!$. При этом как минимум $(N-1)! - l^{N-l-1}\, l!$ путей будут недопустимыми ввиду того, что из каждого пункта разрешены переходы только в l других пунктов. Кроме того, если пройденный путь V включает в себя l пунктов, в которые разрешён переход из текущего пункта, то переход ни в один пункт из текущего пункта не разрешён. Следовательно, данный путь также является недопустимым.

Замечание. Для $l = 1$ алгоритм решения TSP сводится к жадному алгоритму поиска квазиоптимального решения. Тогда у N-ЦМ будет не более одного допустимого пути.

На основе вышеизложенного имеет место

Теорема 2. *Верхняя оценка количества всевозможных вариантов закона распределения вида* (2) *для текущего пункта i при заданном пути V длины k и f_{ij}, $k < N$, $i, j = \overline{1, N}$, определена как 2^l.*

Согласно теореме 2 и [10] закон N-ЦМ будет определён стохастической матрицей, содержащей $2^l N$ строк и N столбцов.

Теорема 3. *Путь V длины N является допустимым, если для любого из пунктов, включённых в V, в законе вида* (2) *существует хотя бы один ненулевой элемент: $\forall k = \overline{1, N} \;\; \exists\, q_{ij} = f_{ij}$.*

В соответствии с теоремами 1–3 произведем многократное генерирование N-ЦМ. Допустимый путь, сгенерированный наибольшее количество раз, является решением TSP. Метод статистических испытаний по определению является приближённым, поэтому описанное решение TSP является приближённым к оптимальному.

2. Генератор N-сложной цепи Маркова

Пусть генератор N-ЦМ состоит из композиции двух блоков согласно [10]: КДА и ВА. Рассмотрим задание КДА и ВА.

КДА определён в виде

$$(Z, S, Y, s(t+1) = \delta(z(t), s(t)), y(t) = \lambda(z(t), s(t))), \tag{3}$$

где $z \in Z$ — номер пункта, который посещает коммивояжёр в текущий момент времени, обозначим его $z(t)$; в Z включен отдельный элемент, характеризующий финальный пункт пути, $|Z| = N+1$; $s \in S$, S — множество пунктов, которые уже включены в путь длины k, $k < N$, $|S| = 2^N$; $y \in Y$ — номер варианта закона распределения вида (2) для пункта $z(t)$, согласно которому выбирается новый пункт $z(t+1)$, $z(t), z(t+1) \in Z$, $|Y| = w \leq 2^l N$, $l < N$. Значения $s(t+1) \in S$ и $y(t) \in Y$ определим на основе детерминированных функций перехода $s(t+1) = \delta(z(t), s(t))$ и выхода $y(t) = \lambda(z(t), s(t))$ соответственно.

ВА задан множеством

$$(Y, X, Z, z(t+1) = \mu(y(t), x(t))) , \tag{4}$$

где $y \in Y$ и $z \in Z$ определены как для КДА (3), $x(t) \in X$ — двоичная α-разрядная (псевдо)случайная величина, равномерно распределённая на интервале $[0, 2^\alpha - 1]$. Новый номер пункта — $z(t+1) \in Z$, $|Z| = N+1$, выберем при использовании стохастической функции μ, которая определена в соответствии с законом распределения вида (2) и значениями $y(t)$ и $x(t)$; функция μ в общем случае задана N двоичными значениями разрядности α, $(b_1, \ldots, b_N)$ в зависимости от значения $y(t)$, $0 \leq b_1 \leq \ldots \leq b_N \leq 2^\alpha - 1$; затем, когда определен вектор $(b_1, \ldots, b_N)$, согласно $x(t)$ выберем пункт k, $k = \overline{1, N}$, или финальный пункт с номером $(N+1)$, при условии $b_{k-1} < x(t) \leq b_k$, приняв $b_0 = 0$ и $b_{N+1} = 2^\alpha$. Формирование $z(t+1)$ как дискретной случайной величины с заданным законом распределения произведем на основе известных методов как параллельно, так и последовательно [17–19]. Согласно [16, 17] детерминированные и стохастические автоматные функции могут быть вычислены на распределённых вычислительных системах как общего, так и специального назначения при использовании нелинейных полиномиальных функций, определённых над конечным полем [16].

В дополнение к системе, включающей КДА и ВА, требуется определить память из $(N-1)$ ячеек разрядности $]\log_2 N[$ каждая. В эти ячейки запишем номера пунктов пути V, $z(2), \ldots, z(N)$, который N-ЦМ проходит из начального пункта $z(1)$ в финальный. Номер $z(1)$ в память не записываем, т. к. он известен. Финальный пункт пути с номером $(N+1)$ в память также не записываем, он служит признаком окончания генерирования пути $z(1), z(2), \ldots, z(t+1)$, $t \leq N$. Определим счётчик по модулю N для подсчёта длины пути t, пройденного от начального пункта до финального. Если длина пути t при выходе на финальный пункт равна N, то такой путь является допустимым. Если же выход в финальный пункт происходит при длине пути $t < N$, то такой путь считается недопустимым. Признак допустимости пути снимем с выхода счётчика по модулю N.

3. Метод решения TSP

Предложим метод определения решения TSP (далее — Метод), который включает подготовительный этап и два следующих этапа.

Подготовительный этап: определение генератора N-ЦМ на основе параметров N, (1) и l.

Этап 1. Формирование $f \leq F$ путей, соответствующих условию их допустимости в соответствии с теоремой 3; сохранение в памяти количества реализаций каждого из допустимых путей $V : z(1), \ldots, z(N)$.

Этап 2. Определение пути V, реализованного генератором N-ЦМ с максимальной частотой, как решения TSP.

При реализации Метода представляет интерес задача хранения количества различных вариантов (частот) реализации путей, допустимых согласно теореме 3,

$V : z(1), \ldots, z(N)$ (далее — ЗХЧ). Варианты V реализуются на этапе 1 Метода при использовании генератора N-ЦМ, определённого на подготовительном этапе Метода. Из полученных вариантов на этапе 2 Метода выбираем тот, который был сгенерирован с наибольшей частотой.

Самое очевидное решение ЗХЧ — применение памяти, имеющей организацию $2^{(N-1)]\log_2 N[}]\log_2 F[$, где F — максимально допустимое количество сгенерированных N-ЦМ. Значение $z(1)$ не участвует в определении адресов ячеек памяти, т. к. задано изначально. Перед генерированием N-ЦМ ячейки обнуляем. При каждом генерировании $V : z(1), \ldots, z(N)$ число в ячейке, которая находится по адресу, определяемому значениями $z(2), \ldots, z(N)$, увеличиваем на единицу. После генерирования f, $f \leq F$, N-ЦМ, выбираем ячейку (или ячейки) памяти по адресу $r(2), \ldots, r(N)$, где находится максимальное число. Значения одного или нескольких адресов, $r(2), \ldots, r(N)$, при заданном значении $z(1)$ соответствуют номерам пунктов пути $V : z(1), \ldots, z(N)$ — искомого решения TSP.

Второй вариант решения ЗХЧ — применение $(N-2)$-х блоков памяти (БП). Для начального участка пути в моменты времени 1, 2 и 3 достаточно хранить в первом БП количество пар всевозможных номеров $z(2)$ и $z(3)$, т. к. пункт $z(1)$ задан. Каждый из остальных $(N-3)$-х БП соответствует участку пути в три момента времени $(t-1)$, t и $(t+1)$, $t = \overline{3, (N-1)}$, и хранит количество троек всевозможных номеров пунктов $(t-1)$, t и $(t+1)$. БП под номерами $(t-1)$ и t хранят одинаковые значения $z(t)$ и $z(t+1)$, $t = \overline{2, (N-2)}$, с целью объединения отдельных участков в один путь V на этапе 2 Метода. БП под номерами $2, \ldots, (N-2)$ имеют организацию $2^{3]\log_2 N[}$ на $]\log_2 F[$, где F — максимально допустимое количество сгенерированных N-ЦМ. Первый БП имеет организацию $2^{2]\log_2 N[}$ на $]\log_2 F[$. Изначально ячейки $(N-2)$-х БП обнуляем. При каждом генерировании $V : z(1), \ldots, z(N)$ число в ячейке БП 1, которая находится по адресу, определяемому $z(2)$ и $z(3)$, увеличиваем на единицу. Аналогичную процедуру повторим для каждого БП t для ячейки по адресу, определяемому $z(t)$, $z(t+1)$ и $z(t+2)$.

Обозначим через S_t множество пунктов, которые принадлежат начальному участку пути V из t элементов, $t \leq N$. После генерирования f, N-ЦМ, $f \leq F$, выберем ячейку (или ячейки) БП 1 по адресу $r(2)$, $r(3)$, где находится максимальное число. Значениям адресов $r(2)$, $r(3)$ при заданном значении $z(1)$ соответствуют участки искомого пути $z(1)$, $z(2)$, $z(3)$. Затем выберем ячейки БП 2 с адресами $r(2)$, $r(3)$, $r(4)$, где находится максимальное число, $r(4) \in (Z\breve{\ } S_3)$. Значению адреса $r(2)$, $r(3)$, $r(4)$ соответствует участок искомого пути $z(1), \ldots, z(4)$. Затем последовательно выберем ячейки БП t, $t = \overline{3, (N-2)}$, с адресами $z(t)$, $z(t+1)$, $z(t+2)$, где находится максимальное число, $r(t+2) \in (Z\breve{\ } S_{t+1})$; $z(t)$, $z(t+1)$ определены на основе значения адреса ячейки БП $(t-1)$. Значения адресов $z(t)$, $z(t+1)$, $z(t+2)$ соответствуют участку искомого пути $z(1), \ldots, z(t+1)$ при заданном значении $z(1)$. В итоге для $t = (N-2)$ получим номера пунктов пути $z(1), \ldots, z(N)$ — искомого решения TSP.

Пример. Пусть задана N-ЦМ при $N = 7$. На основе метода статистических испытаний было сформировано 8 траекторий: 1-2-5-4-3-6-7, 1-3-5-2-4-7-6, 1-3-5-2-4-7-6, 1-2-7-5-4-3-6, 1-2-5-4-3-6-7, 1-2-5-6-7-3-4, 1-3-5-2-4-7-6 и 1-2-5-4-3-6-7.

Требуется один БП с организацией $2^{2]\log_2 7[} = 64$ на три и четыре БП с организацией $2^{3]\log_2 7[} = 512$ на 3. В БП 1 в ячейках, определяемых $z(2)$, $z(3)$ и соответствующих участкам путей (далее — участкам) 2-5, 3-5 и 2-7, запишем значения 4, 3 и 1, в остальных ячейках — нули. В БП 2 в ячейках, определяемых $z(2)$, $z(3)$, $z(4)$ и соответствующих участкам 2-5-4, 3-5-2, 2-7-5, 2-5-6, запишем значения 3, 3, 1 и 1 соответственно, в остальных ячейках — нули. В БП 3 в ячейках, определяемых $z(3)$, $z(4)$, $z(5)$ и соответствующих участкам путей 5-4-3, 5-2-4, 7-5-4 и 5-6-7, запишем значения 3, 3, 1 и 1 соответственно, в остальных ячейках — нули. В БП 4 в ячейках, определяемых $z(4)$, $z(5)$, $z(6)$ и соответствующих участкам 4-3-6, 2-4-7, 5-4-3 и 6-7-3, запишем значения 3, 3, 1 и 1 соответственно, в остальных ячейках — нули. В БП 5 в ячейках, определяемых $z(5)$, $z(6)$, $z(7)$ и соответствующих участкам 3-6-7, 4-7-6, 4-3-6 и 7-3-4, запишем значения 3, 3, 1 и 1 соответственно, в остальных ячейках — нули.

Найдем теперь решение TSP. Для БП 1 наиболее часто встречается начальный участок 1-2-5. Для БП 2 рассмотрим участки, которые начинаются с пунктов 2-5: 2-5-4 и 2-5-6. Частота появления участка 2-5-4 максимальная. Выберем участок пути 1-2-5-4. Для БП 3 рассмотрим участки, которые начинаются с пунктов 5-4: только участок 5-4-3 соответствует данному условию. Выберем участок пути 1-2-5-4-3. Для БП 4 рассмотрим участки, которые начинаются с пунктов 4-3: только участок 4-3-6 соответствует данному условию. Выберем участок пути 1-2-5-4-3-6. Для БП 5 рассмотрим участки, которые начинаются с пунктов 3-6: только участок 3-6-7 соответствует данному условию. Выберем в итоге путь 1-2-5-4-3-6-7 как решение TSP, полученное на основе метода статистических испытаний.

Первый вариант решения ЗХЧ требует один блок памяти, общий объём которого равен $2^{(N-1)]\log_2 N[}]\log_2 F[$ бит, а второй — множество из $(N-2)$-х блоков памяти общим объёмом $(2^{3]\log_2 N[(N-3)} + 2^{2]\log_2 N[})]\log_2 F[)$ бит. В итоге общая ёмкость $(N-2)$-х блоков памяти для реализации второго варианта будет в $\frac{2^{(N-3)]\log_2 N[}}{(2N-5)}$ раз меньше, чем для первого.

4. Анализ сложности предложенного Метода

Оценим сложность реализации этапов Метода: подготовительного и второго — на основе требуемой ёмкости памяти в битах, а первого — по времени задержки вычисления функций КДА вида (3) и ВА вида (4).

Подготовительный этап Метода. Оценим сложность реализации системы, включающей КДА вида (3) и ВА вида (4). Для КДА объёмы множеств входов (Z), внутренних состояний (S) и выходов (Y) будут определены как $|Z| = N+1$, $|S| = 2^N$ и $|Y| = w \le 2^l$. Функции переходов $s(t+1) = \delta(z(t), s(t))$ и выходов $y(t) = \lambda(z(t), s(t))$ в (3) требуют реализации преобразований вида $Z \times S \to S$ и $Z \times S \to Y$. Для ВА вида (4) объёмы множеств для входов S, X и выходов Z определены частично как для КДА вида (3), объём множества $|X| = 2^\alpha$, где α — разрядность равномерно распределённого (псевдо)случайного числа. Вероятностная функция выходов μ определена как множество из $|Y||X| = w2^\alpha$ законов (2) распределения дискретной случайной величины $z(t+1) \in Z$.

Для реализации функций перехода и выхода КДА вида (3), $\delta(z(t), s(t))$ и $\lambda(z(t), s(t))$ требуется память с организацией $2^{]\log_2(N+1)[+N} \times N$ для первой и $2^{]\log_2(N+1)[+N} \times]\log_2 w[$ для второй функции. Определим организацию памяти для задания стохастической функции μ ВА.

Если требуется извлечь значения $(b_1, \ldots, b_N)$ последовательно, то для задания μ нужна память с организацией $2^{l]\log_2(N+1)[+\alpha+]\log_2 N[} \times \alpha$; в случае параллельного извлечения N значений $(b_1, \ldots, b_N)$ память для задания μ должна иметь организацию $2^{l\cdot]\log_2(N+1)[+\alpha} \times (\alpha N)$, где α — разрядность равномерно распределённого (псевдо)случайного числа $x(t) \in X$. Если требуется извлекать из памяти значения $z(t+1)$ в соответствии с функцией $\mu(y(t), x(t))$, то такая память должна иметь организацию $2^{]\log_2 w[+\alpha} \times]\log_2(N+1)[$.

Этап 1 Метода. Проведем генерирование N-ЦМ. Время генерирования одного элемента N-ЦМ определено как время вычисления $z(t+1)$ на основе значений $z(t)$ и $s(t)$. Оценка этого времени определена формулой

$$T_{z(t+1)} = \max\left(\max\left(T(x), T(\lambda)\right) + T(\mu), T(\delta)\right), \tag{5}$$

где $T(x)$, $T(\lambda)$, $T(\mu)$ и $T(\delta)$ — времена вычисления равномерно распределённого (псевдо)случайного числа $x(t) \in X$, функции выхода КДА вида (3), функции выхода ВА вида (4) и функции перехода КДА вида (3) соответственно.

Согласно (5) время генерирования N-ЦМ будет определено как $N\,T_{z(t+1)}$. Перепишем условие, определённое в теореме 3, следующим образом:

$$\forall k = 1, \ldots, t, \quad t \le N, \quad \exists q_{ij} = f_{ij}. \tag{6}$$

Элементы N-ЦМ в моменты времени $t = 2, \ldots, (l-1)$, $l < N$, образуют последовательность, удовлетворяющую условию (6), с вероятностью 1, тогда как в моменты времени $l \le t < N$ существует вероятность сгенерировать путь, для которого условие (6) не выполнится (т. е. получим недопустимый путь). Поиск указанной вероятности сводится к решению известной задачи поиска значения вероятности гипергеометрического распределения при заданных параметрах [20], или схеме выбора t шаров, среди которых l являются красными, $l \le t$, из ящика (урны), в которой находится N шаров, из которых l красные, а $(N-l)$ — белые [21]. Искомая вероятность при заданных параметрах N, l и t равна $p(N,l,t) = C_{N-l}^{t-l}/C_N^t$. Справедлива

Теорема 4. *Вероятность генерирования N-ЦМ при заданном значении $l \le N$ и в каждый момент времени $t = 1, \ldots, N-1$, удовлетворяющая условию* (6)*, равна*

$$Pr(N,l) = \prod_{t=l}^{N-1} (1 - p(N,l,t)) = \prod_{t=l}^{N-1} (1 - C_{N-l}^{t-l}/C_N^t).$$

Из теоремы 4 вытекает

Следствие. *Вероятность того, что в момент времени t, $t \in \{l, \ldots, N-1\}$, при заданном значении $l \le N$, N-ЦМ удовлетворяет условию* (6)*, равна*

$$\ddot{P}r(N,l,t) = \prod_{s=l}^{t}(1 - p(N,l,s)) = \prod_{s=l}^{t}(1 - C_{N-l}^{t-l}/C_N^t),$$

а вероятность противоположного события равна $1 - \ddot{P}r(N,l,t)$.

Теорема 4 определяет вероятность того, что N-ЦМ, удовлетворяющая условию (6), будет сгенерирована за время $N\,T_{z(t+1)}$ в соответствии с (5). Следствие из теоремы 4 позволяет вычислить вероятность того, что при генерировании N-ЦМ вероятность генерирования недопустимого пути в момент времени t, $t \in l, \ldots, N$ составляет $1 - \ddot{P}r(N,l,t)$.

Этап 2 Метода. Выполним подсчёт частот реализации каждой из N-ЦМ, удовлетворяющих условию (6) для $t \le N$. N-ЦМ вида $z(1), \ldots, z(N)$, сгенерированная наибольшее количество раз, является приближённым решением TSP. Для первого варианта решения ЗХЧ требуется $2^{(N-1)]\log_2 N[}$ обращений к ячейкам памяти с выбором среди них наибольшего и с сохранением адреса ячейки, содержащей максимальное значение. Для второго варианта получим $2^{2]\log_2 N[} + (N-3)2^{3]\log_2 N[}$ обращений к ячейкам БП под номерами $1, \ldots, (N-2)$. Для каждого БП сохраняется адрес (адреса) ячейки, содержащей максимальное значение.

Заключение

Представлен подход к решению задачи коммивояжёра на основе предложенного Метода, который основан на генерировании N-сложной цепи Маркова при использовании генератора N-ЦМ в виде композиции КДА (3) и ВА (4). Получение решения сводится к реализации метода статистических испытаний. На подготовительном этапе Метода строится генератор N-ЦМ. Объёмы множеств КДА вида (3) и ВА вида (4) зависят от N и l следующим образом: объём Z линейно зависит от N, объём S зависит от N экспоненциально, объём Y экспоненциально зависит от l, объём X зависит от N и l лишь косвенным образом, определяя точность задания закона распределения вида (2). Количество ячеек памяти для хранения значений функций перехода и выхода КДА вида (3) экспоненциально зависит от N, тогда как размер ячеек памяти для значений функции перехода зависит от N линейно, а для значений функции выхода — логарифмически зависит от l. Количество ячеек памяти для хранения значений, требуемых для реализации стохастической функции μ ВА вида (4), последовательной и параллельной, зависит экспоненциально от величин α и l, а от величины N зависит линейно. Величина ячейки памяти, требуемой для последовательной реализации μ, линейно зависит от α — разрядности $x(t) \in X$, а для параллельной реализации — от произведения αN. Функции КДА и ВА могут быть реализованы также при использовании полиномиальных функций над полем Галуа $GF(2^2)$ на ПЛИС/FPGA, сложность их реализации в базисе ПЛИС представлена в [22].

На этапе 1 Метода производится генерирование N-ЦМ. Согласно (5) на время генерирования каждого из элементов N-ЦМ влияние оказывает задержка вычисления вероятностной функции ВА вида (4), а также максимум от задержек генерирования равномерно распределённой случайной величины и вычисления функций выхода КДА вида (3). Сумма значений $\max(T(x), T(\lambda)) + T(\mu)$ в (5) определяет время генерирования $z(t+1)$, если она больше времени вычисления значения $T(\delta)$.

Для хранения количества реализаций каждого из вариантов N-ЦМ длины N требуется память с организацией $2^{(N-1)]\log_2 N[}]\log_2 F[$, для хранения участков N-ЦМ длины три — $(N-3)$ БП с организацией $2^{3]\log_2 N[}$ на $]\log_2 F[$, а для начального участка длины два — один БП с организацией $2^{2]\log_2 N[}$ на $]\log_2 F[$, где F — максимально допустимое количество сгенерированных N-ЦМ. В обоих случаях размер ячейки логарифмически зависит от F, количество ячеек при росте N в первом случае растёт экспоненциально, а во втором — линейно.

Метод позволяет получить приближённое решение задачи коммивояжёра для N пунктов при следующих условиях: переход из текущего пункта возможен не более чем в $l \leq N$ других пунктов, которые ранее не посещались, и начальный пункт пути задан. На этапе 1 Метода реализуемо многократное параллельное генерирование различных допустимых путей при использовании вычислительных систем как общего, так и специального назначения. За счет статистической обработки множества полученных результатов предложенный Метод устойчив к случайным отклонениям в отличие от известных приближённых детерминированных методов. Данное обстоятельство открывает перспективы применения предложенного Метода для решения широкого круга задач в таких областях, как проектирование компьютерных систем на микро- и макроуровне, обработка изображений и видео, секвентирование ДНК, логистика.

Благодарности. Автор выражает благодарность профессору кафедры компьютерных систем, д. т. н. В.М. Захарову за ценные замечания по работе.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Литература

1. *Little J.D.C., Murty K.G., Sweeney D.W., Karel C.* An algorithm for the traveling salesman problem // Oper. Res. 1965. V. 11, No 6. P. 94–107. https://doi.org/10.1287/opre.11.6.972.
2. *Меламед И.И., Сергеев С.И., Сигал И.Х.* Задача коммивояжера. Вопросы теории // Автомат. телемех. 1989. № 9. С. 3–33.
3. *Меламед И.И., Сергеев С.И., Сигал И.Х.* Задача коммивояжера. Точные методы // Автомат. телемех. 1989. № 10. С. 3–29.
4. *Меламед И.И., Сергеев С.И., Сигал И.Х.* Задача коммивояжера. приближённые алгоритмы // Автомат. телемех. 1989. № 11. С. 3–26.
5. *Luu Q.T., Aibin M.* Traveling Salesman Problem: Exact solutions vs. heuristic vs. approximation algorithms // Baeldung. 2024. URL: https://www.baeldung.com/cs/tsp-exact-solutions-vs-heuristic-vs-approximation-algorithms.
6. *Бусленко Н.П., Шрейдер Ю.А.* Метод статистических испытаний (Монте-Карло) и его реализация на цифровых вычислительных машинах. М.: Гос. изд-во физ.-матем. лит., 1961. 228 с.
7. *Соболь И.М.* Численные методы Монте-Карло. М.: Наука, 1973. 312 с.
8. *Kemeny J., Snell J.* Finite Markov Chains. Ser.: The University Series in Undergraduate Mathematics / Kelley J.L., Halmos P.R. (Eds.). Princeton, NJ: D. Van Nostrand Co., 1960. 210 p. URL: https://archive.org/details/finitemarkovchai0000unse.
9. *Doob J.L.* Stochastic Processes. Ser.: Wiley Classics Library. V. 24. Hoboken, NJ: Wiley, 1990. 654 p. URL: https://books.google.ru/books?id=0GUGAQAAIAAJ.
10. *Альпин Ю.А., Захаров В.М.* Теоретико-автоматный метод описания и моделирования случайных процессов // Вероятн. мет. кибернет. 1983. Т. 19. С. 10–16.
11. *Гремальский А.А., Андроник С.М.* Генератор случайного марковского процесса. 1987. Номер патента: 1453403. URL: https://patents.su/5-1453403-generator-sluchajjnogo-markovskogo-processa.html.
12. *Гремальский А.А., Андроник С.М.* Генератор случайного марковского процесса. 1987. Номер патента: 1481755. URL: https://patents.su/5-1481755-generator-sluchajjnogo-markovskogo-processa.html.
13. *Гремальский А.А.* Генератор случайного марковского процесса. 1988. Номер патента: 1531093. URL: https://patents.su/6-1531093-generator-sluchajjnogo-markovskogo-processa.html.
14. *Юминов О.Б., Ирисов М.В., Дзюин С.В.* Генератор n-связной марковской последовательности. 1988. Номер патента: 1550501. URL: https://patents.su/2-1550501-generator-n-svyaznojj-markovskojj-posledovatelnosti.html.
15. *Добрыдень В.А.* Устройство для моделирования марковских процессов. 1975. Номер патента: 526909. URL: https://patents.su/6-526909-ustrojjstvo-dlya-modelirovaniya-markovskikh-processov.html.
16. *Захаров В.М., Нурутдинов Ш.Р., Шалагин С.В.* Полиномиальное представление цепей Маркова над полем Галуа // Вестн. КГТУ им. А.Н. Туполева. 2001. № 3. С. 27–31.
17. *Захаров В.М., Шалагин С.В., Гумиров А.И.* Аппаратно-программный модуль генератора марковских последовательностей на основе программируемых логических интегральных схем // Вестн. КГТУ им. А.Н. Туполева. 2022. Т. 78, № 4. С. 164–172.
18. *Захаров В.М., Шалагин С.В., Эминов Б.Ф.* Автоматные марковские модели над конечным полем. Казань: Спец. фонд управл. цел. капит. для разв. Казан. нац. исслед. техн. ун-та им. А.Н. Туполева, 2022. 328 с. URL: https://disk.yandex.ru/i/PxXZClzIRdDnzQ.

19. *Захаров В.М., Шалагин С.В., Гумиров А.И.* Генератор дискретной случайной величины с заданным законом распределения в архитектуре ПЛИС/FPGA // Вестн. ДГУ. Сер. 1: Ест. науки. 2023. Т. 38, № 3. С. 28–33. http://dx.doi.org/10.21779/2542-0321-2023-38-3-28-33.

20. *Rice J.A.* Mathematical Statistics and Data Analysis. Belmont, CA: Thomson Brooks/Cole, 2007. 698 p. URL: https://archive.org/details/mathematicalstat0000rice_c1o8_3ed.

21. *Feller V.* An Introduction to Probability Theory and Its Applications. V. 1. Ser.: Wiley Series in Probability and Statistics. New York, NY: Wiley, 1957. 669 p. URL: https://archive.org/details/introductiontopr0001will/mode/2up.

22. *Шалагин С.В.* Сложность вычисления нелинейных полиномиальных функций над полем $GF(2^2)$ на ПЛИС/FPGA // Поиск эффективных решений в процессе создания и реализации научных разработок в российской авиационной и ракетно-космической промышленности: Междунар. науч.-практ. конф. Т. II. Казань, Изд-во Казан. гос. техн. ун-та, 2014. С. 661–664.

Поступила в редакцию 15.08.2024
Принята к публикации 5.10.2024

Шалагин Сергей Викторович, доктор технических наук, доцент, профессор кафедры компьютерных систем

Казанский национальный исследовательский технический университет им. А.Н. Туполева

ул. Карла Маркса, д. 10, г. Казань, 420111, Россия

E-mail: *sshalagin@mail.ru*

ISSN 2541–7746 (Print)
ISSN 2500–2198 (Online)

UCHENYE ZAPISKI KAZANSKOGO UNIVERSITETA.
SERIYA FIZIKO-MATEMATICHESKIE NAUKI
(Proceedings of Kazan University. Physics and Mathematics Series)

2024, vol. 166, no. 4, pp. 639–650

ORIGINAL ARTICLE
doi: 10.26907/2541-7746.2024.4.639-650

Solving the Traveling Salesman Problem by Statistical Testing Using Complex Markov Chains

S.V. Shalagin

Kazan National Research Technical University named after A.N. Tupolev – KAI, Kazan, 420111 Russia

E-mail: *sshalagin@mail.ru*

Received August 15, 2024; Accepted October 5, 2024

Abstract

A method was proposed for solving the traveling salesman problem (TSP) using N-complex Markov chains (N-MC), the state sequence of which simulates a path through N points, each visited only once. Transition probabilities from each point to one of the next l points were set, where $l < N$. The complexity of implementing all stages of the method, depending on the values of N and l, was analyzed. Estimates of the complexity of the N-MC generator were obtained based on the composition of a finite deterministic automaton and a probabilistic memoryless automaton. The complexity of the N-MC generator is characterized by the volume of input sets, internal states, and outputs, as well as the amount of memory required to implement the transition and output functions of the automata. The delay time of the N-MC generator operation was also estimated. The probability of generating valid paths, i.e., those resolving TSP, and the memory requirements for storing valid paths were calculated.

Keywords: traveling salesman problem, method of solving, complex Markov chain, statistical test, complexity estimate

Acknowledgments. V.M. Zakharov (Professor of the Computer Systems Department at Kazan National Research Technical University named after A.N. Tupolev – KAI, Doctor of Technical Sciences) is thanked for his helpful comments on this study.

Conflicts of Interest. The authors declare no conflicts of interest.

References

1. Little J.D.C., Murty K.G., Sweeney D.W., Karel C. An algorithm for the traveling salesman problem. *Oper. Res.*, 1965., vol. 11, no. 6, pp. 94–107. https://doi.org/10.1287/opre.11.6.972.
2. Melamed I.I., Sergeev S.I., Sigal I.Kh. The traveling salesman problem. Issues in the theory. *Autom. Remote Control*, 1989, vol. 50, no. 9, pp. 1147–1173.
3. Melamed I.I., Sergeev S.I., Sigal I.Kh. The traveling salesman problem. Exact methods. *Autom. Remote Control*, 1989, vol. 50, no. 10, pp. 1303–1324.
4. Melamed I.I., Sergeev S.I., Sigal I.Kh. The traveling salesman problem. Approximate algorithms. *Autom. Remote Control*, 1989, vol. 50, no. 11, pp. 1459–1479.
5. Luu Q.T., Aibin M. Traveling Salesman Problem: Exact solutions vs. heuristic vs. approximation algorithms. *Baeldung*, 2024. URL: https://www.baeldung.com/cs/tsp-exact-solutions-vs-heuristic-vs-approximation-algorithms.
6. Buslenko N.P., Shreider Ju.A. *Metod statisticheskikh ispytanii (Monte-Karlo) i ego realizatsiya na tsifrovykh mashinakh* [The Statistical Test Method (Monte Carlo) and Its Use on Digital Computers]. Moscow, Gos. Izd. Fiz.-Mat. Lit., 1961. 228 p. (In Russian)
7. Sobol' I.M. *Chislennye metody Monte-Karlo* [Numerical Monte Carlo Methods]. Moscow, Nauka, 1973. 312 p. (In Russian)
8. Kemeny J.G., Snell J.L. *Finite Markov Chains.* Ser.: The University Series in Undergraduate Mathematics. Kelley J.L., Halmos P.R. (Eds.). Princeton, NJ, D. Van Nostrand Co., 1960. 210 p. URL: https://archive.org/details/finitemarkovchai0000unse.
9. Doob J.L. *Stochastic Processes.* Ser.: Wiley Classics Library. Vol. 24. Hoboken, NJ, Wiley, 1990. 654 p. URL: https://books.google.ru/books?id=0GUGAQAAIAAJ.
10. Al'pin Yu.A., Zakharov V.M. An automata-theoretic approach to the description and modeling of stochastic processes. *Veroyatn. Metody Kibern.*, 1983, vol. 19, pp. 10–16. (In Russian)
11. Gremal'skii A.A., Andronik S.M. Stochastic Markov process generator. Patent USSR no. 1453403, 1987. URL: https://patents.su/5-1453403-generator-sluchajjnogo-markovskogo-processa.html. (In Russian)
12. Gremal'skii A.A., Andronik S.M. Stochastic Markov process generator. Patent USSR no. 1481755, 1987. URL: https://patents.su/5-1481755-generator-sluchajjnogo-markovskogo-processa.html. (In Russian)
13. Gremal'skii A.A. Stochastic Markov process generator. Patent USSR no. 1531093, 1988. URL: https://patents.su/6-1531093-generator-sluchajjnogo-markovskogo-processa.html. (In Russian)

14. Yuminov O.B., Irisov M.V., Dzyuin S.V. N-state Markov chain generator. Patent USSR no. 1550501, 1988. URL: https://patents.su/2-1550501-generator-n-svyaznojj-markovskojj-posledovatelnosti.html. (In Russian)

15. Dobryden' V.A. A device for modeling Markov processes. Patent USSR no. 526909, 1975. URL: https://patents.su/6-526909-ustrojjstvo-dlya-modelirovaniya-markovskikh-processov.html. (In Russian)

16. Zakharov V.M., Nurutdinov Sh.R., Shalagin S.V. Polynomial representation of Markov chains over a Galois field. *Vestn. KGTU im. A.N. Tupoleva*, 2001, no. 3, pp. 27–31. (In Russian)

17. Zakharov V.M., Shalagin S.V., Gumirov A.I. Hardware and software module of the Markov sequence generator based on programmable logic integrated circuits. *Vestn. Kazan. Gos. Tekh. Univ. im. A.N. Tupoleva*, 2022, vol. 78, no. 4, pp. 164–172. (In Russian)

18. Zakharov V.M., Shalagin S.V., Eminov B.F. *Avtomatnye markovskie modeli nad konechnym polem* [Automata Markov Models over a Finite Field]. Kazan, Spets. Fond Upr. Tselevym Kap. Razvit. Kazan. Nats. Issled. Tekh. Univ. im. A.N. Tupoleva, 2022. 328 p. URL: https://disk.yandex.ru/i/PxXZClzIRdDnzQ. (In Russian)

19. Zakharov V.M., Shalagin S.V., Gumirov A.I. Discrete random variable generator with the given distribution law in FPGA-architecture. *Vestn. DGU. Ser. 1: Estestv. Nauki*, 2023, vol. 38, no. 3, pp. 28–33. http://dx.doi.org/10.21779/2542-0321-2023-38-3-28-33. (In Russian)

20. Rice J.A. *Mathematical Statistics and Data Analysis.* Belmont, CA, Thomson Brooks/ Cole, 2007. 698 p. URL: https://archive.org/details/mathematicalstat0000rice_c1o8_3ed.

21. Feller V. *An Introduction to Probability Theory and Its Applications.* Vol. 1. Ser.: Wiley Series in Probability and Statistics. New York, NY, Wiley, 1957. 669 p. URL: https://archive.org/details/introductiontopr0001will/mode/2up.

22. Shalagin S.V. The complexity of computing nonlinear polynomial functions over the $GF(2^2)$ field on PLD/FPGA. *Poisk effektivnykh reshenii v protsesse sozdaniya i realizatsii nauchnykh razrabotok v rossiiskoi aviatsionnoi i raketno-kosmicheskoi promyshlennosti: Mezhdunar. nauch.-prakt. konf.* [Exploring Effective Solutions for the Development and Implementation of Scientific Innovations in Russia's Aviation and Space Industry: Proc. Int. Sci. Pract. Conf.]. Vol. II. Kazan, Izd. Kazan. Gos. Tekh. Univ., 2014, pp. 661–664. (In Russian)

Для цитирования: *Шалагин С.В.* Решение задачи коммивояжёра методом статистических испытаний при использовании сложных цепей Маркова // Учен. зап. Казан. ун-та. Сер. Физ.-матем. науки. 2024. Т. 166, кн. 4. С. 639–650. https://doi.org/10.26907/2541-7746.2024.4.639-650.

For citation: Shalagin S.V. Solving the traveling salesman problem by statistical testing using complex Markov chains. *Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-Matematicheskie Nauki*, 2024, vol. 166, no. 4, pp. 639–650. https://doi.org/10.26907/2541-7746.2024.4.639-650. (In Russian)