УДК 519.3

О.В. СЕРАЯ, канд. техн. наук, НТУ "ХПИ",
Л.В. БАЧКИР

СТОХАСТИЧЕСКАЯ ЗАДАЧА КОММИВОЯЖЕРА

Розглянуто проблему комівояжера високої розмірності для випадку, коли відстані між пунктами – випадкові величини. Запропоновано декомпозиційний алгоритм рішення задачі. Використано кластеризацію пунктів обходу та генетичний алгоритм пошуку найкоротших локальних шляхів.

The traveling salesman task of high dimension is considered in the case of when distances between points are random variable. The decomposition algorithm of task decision is offered. The cluster of bypass points and genetic algorithm of short local ways search is used.

Постановка проблемы и обзор литературы. Каноническая задача коммивояжера формулируется следующим образом: для заданной совокупности n пунктов при известной матрице расстояний между ними необходимо отыскать кратчайший маршрут без потерь обхода этих пунктов [1]. При этом под "расстоянием" между пунктами понимают не обязательно длину пути, это могут быть и некоторые другие величины (например, затраты на переезд, продолжительность переезда и т.п.). Задача коммивояжера это NP-полная задача – время ее решения экспоненциально растет с увеличением числа пунктов. Поэтому известные алгоритмы решения этой задачи эффективны только в случае невысокой размерности ($n \approx 20$) [2, 3]. Сравнительно недавно показана возможность использования для решения задачи коммивояжера генетических алгоритмов (ГА) [4, 5]. Как показывает практика, при этом размерность решаемых в приемлемое время задач существенно возрастает ($n \approx 100$). Однако, во многих случаях возникает необходимость решения этой задачи гораздо более высокой размерности ($n \approx 1000$). Для решения такой задачи может быть использован следующий декомпозиционный алгоритм. Сначала все множество пунктов с применением какой-либо процедуры кластеризации разбивается на некоторое число групп. Далее отыскивается рациональный порядок обхода этих групп и устанавливаются точки входа и выхода для смежных групп. Теперь для каждой отдельной группы определяется кратчайший маршрут, соединяющий точки входа и выхода, и, наконец, эти маршруты объединяются. Эта процедура просто реализуется и очень эффективна.

Задача существенно усложняется, если расстояния между пунктами – случайные величины. Понятно, что стандартный ГА в этой ситуации непосредственно применен быть не может. В связи с этим **целью статьи** является разработка технологии использования ГА для решения стохастической задачи коммивояжера.

Постановка задачи. Пусть в задаче коммивояжера заданы плотности распределения $\varphi_{ij}(R_{ij})$ случайных значений "расстояния" R_{ij} между пунктами i и j, например, в виде нормального распределения [6]

$$\varphi_{ij}(R_{ij}) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left\{-\frac{\left(R_{ij}-m_{ij}\right)^2}{2\sigma_{ij}^2}\right\},\ i=1,2,...,n,\ j=1,2,...,n. \qquad (1)$$

С учетом стохастической природы "расстояний" между пунктами введем понятие удаления одного пункта от другого, количественно оцениваемое как вероятность того, что случайное "расстояние" между пунктами окажется больше некоторого определенным образом выбранного порогового. При этом будем считать, что удаление одного пункта от другого тем больше, чем выше эта вероятность. В соответствии с (1) удаление пункта j от пункта i оценивается по формуле

$$u_{ij} = \int_{R_l}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left\{-\frac{\left(R_{ij}-m_{ij}\right)^2}{2\sigma_{ij}^2}\right\} dR_{ij}, \qquad (2)$$

где R_l – некоторый выбранный порог.

Понятно, что введенный показатель может быть использован для оценки удаления не обязательно смежных на маршруте пунктов. С учетом этого показателя сформулируем задачу отыскания кратчайшего маршрута.

Основные результаты. Решение задачи начинается с кластеризации [7, 8] пунктов обхода с учетом их удалений друг от друга. Процедуру кластеризации опишем вначале для случая, когда число кластеров p и центры группирования заданы (в качестве каждого из центров группирования выбирается какой-то из пунктов). Процедура кластеризации является пошаговой. Пусть к очередному $(l+1)$-му шагу l пунктов уже распределены по кластерам, образовав разбиение множества E всех подлежащих кластеризации пунктов на два подмножества: E_l^+ (распределенные) и E_l^- (не распределенные). Для каждого из оставшихся $(n-p-l)$ пунктов подмножества E_l^-, например, i-го, по формуле (2) рассчитывается удаление u_{ij} до каждого, например, j-го центра группирования. В этой формуле R_l – значение порога, вычисленное к l-му шагу. Теперь на множестве $\{u_{ij}\}$, $i \in E_l^+$, $j=1,2,...,p$, отыскиваем пару (i_0, j_0) по правилу

$$(i_0, j_0) = \arg\min_{\substack{i \in E_l^- \\ j}} \{u_{ij}\}. \qquad (3)$$

При этом пункт i_0 добавляется к кластеру j_0 и вычисляется новое значение порога $R_{l+1} = m_{i_0 j_0}$. Кроме того, корректируется разбиение множества E на подмножества распределенных и нераспределенных пунктов:

$$E_{l+1}^{+} = E_{l+1}^{+} \cup i_0\,,\; E_{l+1}^{-} = E_{l+1}^{-} \setminus i_0\,.$$

На следующем шаге процедура повторяется до полного распределения по кластерам. В случае, если центры группирования не заданы для выбранного числа p кластеров отыскиваются p пунктов так, чтобы минимальное удаление между ними было максимально. В результате решения задачи кластеризации получаем p групп пунктов. Пусть теперь E_k – множество номеров пунктов, вошедших в k-й кластер, $k = 1, 2, ..., p$. Далее в каждой группе рассчитаем положение центра тяжести, используя соотношения

$$X_{ц.т.}^{(k)} = \frac{1}{n_k}\sum_{i \in E_k} x_{ki}\,;\quad Y_{ц.т.}^{(k)} = \frac{1}{n_k}\sum_{i \in E_k} y_{ki}\,.$$

Здесь x_{ki} – абсцисса i-го пункта из k-й группы, y_{ki} – ордината i-го пункта из k-й группы, n_k – число пунктов в k-й группе.

Полученные координаты p точек – центров тяжести кластеров используем для отыскания с помощью обычного ГА маршрута рационального обхода этих кластеров. После этого для каждой пары смежных на полученном маршруте кластеров найдем кратчайшую перемычку. С этой целью составим множество пар пунктов, принадлежащих разным кластерам из выбранной пары смежных. Пусть n_1 – число пунктов первого из пары кластеров, а n_2 – число пунктов второго из них. По формуле (2) рассчитаем множество значений $u_{i_1 i_2}$, $i_1 = 1, 2, ..., n_1$, $i_2 = 1, 2, ..., n_2$, из которых найдем пару (i_{10}, i_{20}) наименее удаленных по правилу (3). При этом в соотношении (2) в качестве значения порога используем половину расстояния между центрами тяжести выбранной пары кластеров. Теперь пункт i_{10} полагаем пунктом выхода из первого кластера смежной пары, а пункт i_{20} – пунктом входа во второй. Процедура повторяется для всех смежных пар кластеров.

На очередном этапе общей процедуры в каждом кластере решается стохастическая задача отыскания кратчайшего маршрута для заданных пунктов входа и выхода. С этой целью построим стохастический аналог детерминированного генетического алгоритма. Принципиальным элементом соответствующей вычислительной процедуры является технология расчета критерия качества особей, вошедших в популяцию, сформированную на очередном шаге ГА. Пусть кластер содержит s пунктов. Введем индикатор

$$w_{ij} = \begin{cases} 1, \text{если соответствующий выбранной особи маршрут связывает} \\ \quad \text{пункты } i \text{ и } j \text{ непосредственно,} \\ 0, \text{в противном случае,} \quad i = 1, 2, ..., s-1, \;\; j = 2, 3, ..., s. \end{cases}$$

Тогда случайная длина маршрута, соединяющего найденные на предыдущем этапе пункты входа и выхода рассматриваемого кластера, определяется соотношением

$$R_\Sigma = \sum_{i=1}^{s-1}\sum_{j=2}^{s} R_{ij} w_{ij}\,.$$

С учетом (1) считаем, что плотность распределения случайной величины R_Σ имеет вид

$$\varphi(R_\Sigma) = \frac{1}{\sqrt{2\pi}\sigma_\Sigma}\exp\left\{-\frac{(R_\Sigma - m_\Sigma)^2}{2\sigma_\Sigma^2}\right\};\; m_\Sigma = \sum_{i=1}^{s-1}\sum_{j=2}^{s} m_{ij} w_{ij}\,;\; \sigma_\Sigma^2 = \sum_{i=1}^{s-1}\sum_{j=2}^{s} \sigma_{ij}^2 w_{ij}\,.$$

Тогда вероятность того, что случайная величина R_Σ превысит некоторый порог R_Π вычисляется по формуле

$$P(R_\Sigma \geq R_\Pi) = \int_{R_\Pi}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_\Sigma}\exp\left\{-\frac{(R_\Sigma - m_\Sigma)^2}{2\sigma_\Sigma^2}\right\} dR_\Sigma = \int_{\frac{R_\Pi - m_\Sigma}{\sigma_\Sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du\,. \quad (4)$$

Поставим теперь задачу отыскания маршрута $W = (w_{ij})$, минимизирующего (4). Ясно, что эта задача эквивалентна максимизации

$$L(x) = \frac{R_\Pi - m_\Sigma}{\sigma_\Sigma} = \frac{R_\Pi - \sum_{i=1}^{s-1}\sum_{j=2}^{s} m_{ij} w_{ij}}{\left(\sum_{i=1}^{s-1}\sum_{j=2}^{s} \sigma_{ij}^2 w_{ij}\right)^{\frac{1}{2}}}\,. \quad (5)$$

В качестве R_Π естественно выбрать длину кратчайшего маршрута, связывающего начальную и конечную точки кластера, рассчитываемую с использованием обычного ГА в предположении, что расстояния между пунктами детерминированы и определяются набором $\{m_{ij}\}$.

Преобразуем выражение (5). Число звеньев в любом связном маршруте, соединяющем s пунктов без петель, равно $s-1$. Действительно, так как необходимые условия, определяющие требования к маршруту, имеют вид

$$\sum_{i=1}^{s-1} w_{ij} = 1, \; j = 2, ..., s; \quad \sum_{j=2}^{s} w_{ij} = 1; \; i = 1, 2, ..., s-1,$$

то

$$\sum_{i=1}^{s-1} \sum_{j=2}^{s} w_{ij} = s-1.$$

Тогда

$$L(x) = \frac{\frac{R_{\Pi}}{s-1} \sum_{i=1}^{s-1} \sum_{j=2}^{s} w_{ij} - \sum_{i=1}^{s-1} \sum_{j=2}^{s} m_{ij} w_{ij}}{\left(\sum_{i=1}^{s-1} \sum_{j=2}^{s} \sigma_{ij}^2 w_{ij} \right)^{\frac{1}{2}}} = \frac{\sum_{i=1}^{s-1} \sum_{j=2}^{s} \left(\frac{R_{\Pi}}{s-1} - w_{ij} \right) w_{ij}}{\left(\sum_{i=1}^{s-1} \sum_{j=2}^{s} \sigma_{ij}^2 w_{ij} \right)^{\frac{1}{2}}} = \frac{\sum_{i=1}^{s-1} \sum_{j=2}^{s} r_{ij} w_{ij}}{\left(\sum_{i=1}^{s-1} \sum_{j=2}^{s} \sigma_{ij}^2 w_{ij} \right)^{\frac{1}{2}}}. \quad (6)$$

Выражение (6) для любого допустимого маршрута определяет значение критерия его качества тем большее, чем меньше вероятность (4) превышения длиной этого маршрута заданного порога.

В связи с этим ясно, что критерий (6) может быть использован при формировании новой популяции ГА в результате селекции после применения стандартных процедур скрещивания, мутации, рекомбинации. Описанная модификация ГА для каждого кластера обеспечивает отыскание наилучшего в выбранном смысле маршрута. Теперь, связывая индивидуальные маршруты кластеров "перемычками", найденными ранее, получим искомый маршрут, являющийся решением задачи.

Выводы. Таким образом, предложена вычислительная процедура решения задачи коммивояжера для случая, когда расстояния между пунктами – случайные величины с известной плотностью распределения. Для решения задачи использован декомпозиционный алгоритм, редуцирующий исходную сложную задачу к последовательности существенно более простых. Предложенная методика позволяет успешно решать стохастические задачи коммивояжера высокой размерности.

Список литературы: **1**. *Miller C.E., Tucker A.W., Zemlin R.A.* Integer Programming Formulation of Traveling Salesman Problems // Journal of the Association of Computing Machines. – 1960. – 7. – P. 326–329. **2.** *Литтл Дж., Мурти К., Суини Д., Кэрел К.* Алгоритм для решения задачи о коммивояжере // Экономические и математические методы. – 1965. – Е. 1. – № 1. – С. 34–41. **3.** *Корбут А.А., Финкельштейн Ю.Ю.* Дискретное программирование. – М.: Наука, 1969. – 328 с. **4.** *John Holland.* Genetic algorithms and the optimal allocation of trials. SIAM Journal on computation. – 1973. – 2. – P. 88–105. **5.** *Лысенко Ю.Г., Иванов Н.Н., Минц А.Ю.* Нейронные сети и генетические алгоритмы. – Донецк: Юго-восток, 2003. – 230 с. **6.** *Венцель Е.С.* Теория вероятностей. – М.: ГИФМЛ, 1962. – 564 с. **7.** *Дюран Б., Одел П.* Кластерный анализ. – М.: Статистика, 1977. – 126 с. **8.** Классификация и снижение размерности / *Айзазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.* / Под ред. Айвазяна С.А. – М.: Финансы и статистика, 1989. – 607 с.

Поступила в редакцию 25.09.2006

УДК 620.179.16:620. 79.17

Г.М. СУЧКОВ, д-р техн. наук, НТУ "ХПИ",
Е.А. АЛЕКСЕЕВ, канд. ф.-мат. наук, РИ НАНУ (г. Харьков),
В.В. ЗАХАРЕНКО, канд. техн. наук, РИ НАНУ (г. Харьков),
Р.А. МОТИЕНКО, РИ НАНУ (г. Харьков),
Е.Л. НОЗДРАЧЕВА, НТУ "ХПИ",
А.В. ДОНЧЕНКО, ООО "Квазар-микро" (г. Киев),
М.В. АНАНЬИНА, НТУ "ХПИ"

ЭФФЕКТИВНОСТЬ ОБРАБОТКИ ИНФОРМАЦИИ В СОВРЕМЕННЫХ ПРИБОРАХ НЕРАЗРУШАЮЩЕГО КОНТРОЛЯ

Наведено приклад реалізації в ЕМА дефектоскопі алгоритму обробки інформації з використанням програмного та апаратного методів. Експериментально встановлена можливість підвищення чутливості приладу приблизно в 10 разів.

The example of realization the algorithm of treatment information with use the programming and device methods in the EMA fault detector are given. The opportunity of increase of sensitivity of the device approximately in 10 times is experimentally established.

Постановка проблемы и анализ литературы. Основные способы и методы дефектоскопии металлоизделий хорошо известны и широко применяются в промышленности Украины и других стран [1]. Одним из перспективных методов ультразвукового контроля является электромагнитно-акустический (ЭМА) [2]. Однако, ему свойственна недостаточная чувствительность [3]. Устранить отмеченный недостаток возможно за счет применения современных методов обработки информации, получаемой при контроле изделий [4 − 7].

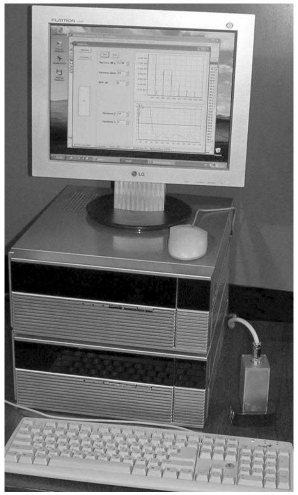

Рис. 1. Внешний вид дефектоскопа

Анализ известных литературных источников [8 − 9] позволил установить, что для достижения поставленной цели необходим комплексный подход − обработка информации на всех стадиях формирования полезного сигнала аппаратным и программным путем.

Цель статьи – повышение чувствительности приборов неразрушающего контроля.

ЭМА дефектоскоп. Рассмотрим современный прибор неразрушающего контроля на примере ЭМА дефектоскопа для контроля калиброванных прутков круглого и шестигранного сечения объемными сдвиговыми и поверхностными волнами. Внешний вид