УДК 519.681

О. И. ДУНАЕВСКАЯ, ассистент НТУ «ХПИ»;
Н. И. ЯЩУК, научный сотрудник НТУ «ХПИ»

РЕКУРРЕНТНАЯ ОЦЕНКА ПРОДОЛЖИТЕЛЬНОСТИ ОЖИДАНИЯ ПРОГРЕССОРА В ГЕНЕТИЧЕСКОМ АЛГОРИТМЕ

Розглянуто одне з актуальних питань вдосконалення генетичних алгоритмів. Запропоновано процедуру рекурентної оцінки тривалості очікування прогресора – особини, що забезпечує краще рішення у порівнянні з отриманими раніше. Задача розв'язана стосовно до NP-повних комбінаторних задач.

Рассмотрен один из актуальных вопросов совершенствования генетических алгоритмов. Предложена процедура рекуррентной оценки продолжительности ожидания прогрессора – особи, обеспечивающей лучшее решение по сравнению с полученными ранее. Задача решена применительно к NP-полным комбинаторным задачам.

One of the topical questions of genetic algorithm improvement was considered. There was suggested a recurrent estimation procedure of the waiting period for progressor – the individual, providing a better solution compared with those obtained previously. The problem is solved as applied to the NP-complete combinatorial task.

Введение. Хорошо известна высокая эффективность генетических алгоритмов (ГА) при решении многоразмерных комбинаторных задач [1, 2]. Это связано с конструктивно заложенной в ГА возможностью быстрого формирования большого числа потенциальных решений задачи в разных областях пространства поиска. ГА – адаптивный алгоритм. Как экспериментально показано в [3], рациональные значения основных параметров ГА (число особей в популяции, характер процедуры скрещивания, вероятность мутации и т.д.) существенно зависят от класса решаемой задачи (безусловная или условная оптимизация, одно- или многоэкстремальная задача, непрерывная или дискретная оптимизация). Там же предложена двухуровневая иерархическая алгоритмическая процедура настройки параметров ГА с учетом характера решаемой задачи. Однако, ни в этой работе, ни во многих других, посвященных проблеме эффективного использования ГА, не рассматривался принципиальный вопрос о числе непрогрессирующих поколений до останова. Проблема состоит в том, что априорное задание числа эпох ожидания или уровня приспособленности наилучшей особи, при достижении которых целесообразен останов решения задачи, не могут быть корректно обоснованы из каких-либо общих соображений или рекомендаций, привязанных к типу решаемой задачи. С другой стороны, важность этой проблемы определяется тем, что никак не обоснованная продолжительность ожидания появления прогрессивной особи может привести к ошибке в решении задачи (неконтролируемая погрешность недостижения оптимума).

В связи с этим поставим задачу построения процедуры оценки продолжительности ожидания прогрессора, учитывающей на каждом шаге достигнутый к этому моменту уровень качества решения.

Основные результаты.

Введем

T – временной интервал, прошедший от момента начала решения задачи,

$\varphi(R)$ – плотность распределения численного значения критерия на множестве возможных решений задачи,

$R(T)$ – лучшее решение, полученное к моменту T,

$P(T) = \int\limits_{R_{\min}}^{R(T)} \varphi(R)dR$ – вероятность того, что решение, соответствующее произвольно выбранной особи из очередной популяции, будет лучше (то есть обеспечивать меньшее значение критерия), чем $R(T)$.

При этом $P_n = 1 - (1 - P(T))^n$ – вероятность того, что хотя бы одно из n решений в очередной популяции окажется лучше, чем $R(T)$.

Пусть $N(T)$ – число эпох, в течение которых алгоритм ожидает появления решения, лучшего, чем $R(T)$. Тогда $P(N) = (1 - P(T))^{nN(T)}$ – вероятность того, что в течение $N(T)$ эпох не будет получено решение лучшее, чем $R(T)$. Если неполучение лучшего решения в течении $N(T)$ популяций трактовать как признак того, что лучшего решения не существует, то $P(N)$ – вероятность ошибки. Рассчитаем число эпох ожидания (значение $N(T)$), обеспечивающее заданную вероятность ошибки.

$$(1 - P(T))^{nN(T)} = P_з,$$

откуда

$$N(T) = \frac{\ln P_з(N(T))}{n \cdot \ln(1 - P(T))}.$$

Пусть $nN(T)$ – общее число особей, использованных для поиска лучшего решения в течении $N(T)$ эпох. Легко получить оценку среднего числа проб до получения прогрессивной особи.

Пусть p – вероятность получения решения, лучшего чем предыдущее в одиночном опыте.

Тогда

$P_1 = p$ – вероятность получения лучшего решения в первом опыте,

$P_2 = qp$ – вероятность получения лучшего решения во втором опыте, $q = 1 - p$,

$P_k = q^{k-1} p$ – вероятность получения лучшего решения в k-м опыте.

При этом среднее число опытов до получения лучшего решения будет равно

$$\overline{m} = p + 2qp + 3q^2 p + \ldots + kq^{k-1} p + \ldots = p\left(1 + 2q + 3q^2 + \ldots + kq^{k-1} + \ldots\right) =$$

$$= p\frac{d}{dq}\left(q + q^2 + q^3 + \ldots + q^k + \ldots\right) = p\frac{d}{dq}\left(\frac{q}{1-q}\right) = p\left(\frac{1-q+q}{(1-q)^2}\right) = \frac{p}{p^2} = \frac{1}{p}.$$

Трудности теоретической оценки значения $P(T)$ преодолеваются с использованием экспериментального исследования. Пусть для конкретной задачи после первой популяции, в которой отсутствует решение, лучшее, чем все предыдущие, фиксируется продолжительность ожидания (число эпох, в которых также не оказалось лучшего решения) и при этом получена последовательность – $M_1, M_2, \ldots, M_k, \ldots, M_{s+1}$. Введем рекуррентную зависимость продолжительности ожидания на очередном этапе улучшения решения от значения этой продолжительности на предыдущем этапе

$$M_{k+1} = a_1 M_k + a_2 M_k^2. \tag{1}$$

Найдем неизвестные параметры a_1 и a_2.

Введем

$$H = \begin{pmatrix} M_1 & M_1^2 \\ M_2 & M_2^2 \\ \ldots & \ldots \\ M_s & M_s^2 \end{pmatrix}; \; A = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}; \; Y = \begin{pmatrix} M_2 \\ M_3 \\ \ldots \\ M_{s+1} \end{pmatrix}.$$

Тогда

$$A = \left(H^T H\right)^{-1} H^T Y. \tag{2}$$

Пример. В качестве тестовой выберем задачу коммивояжера для по порядку перенумерованных $n = 120$ пунктов, расположенных на кольцевой дороге, представляющей собой окружность радиуса $R = 1$. Понятно, что оптимальное решение – маршрут вдоль этой дороги, и его длина равна $S = 2\pi \approx 2 \cdot 3.1416 = 6.2832$.

Исследование искомой зависимости было проведено, начиная от момента, когда ситуация ожидания лучшего решения возникла в течение нескольких (например, пятидесяти) циклов работы ГА подряд. При этом была получена следующая последовательность продолжительностей ожидания: 192, 263, 382, 563, 918, 1836, 4964, 24362.

Теперь, используя (2), получим

$$A_n^T = \begin{pmatrix} a_1^n & a_2^n \end{pmatrix} = \begin{pmatrix} 1{,}354 & 7{,}16 \cdot 10^{-4} \end{pmatrix}.$$

Полученные значения (a_1, a_2) могут быть использованы в (1) для оценки продолжительности ожидания на очередном этапе работы ГА.

Исследуем зависимость значений коэффициентов (a_1, a_2) в (1) от размерности задачи. С этой целью эта же задача решалась для $n_1 = 30$, $n_2 = 60$, $n_3 = 90$ пунктов, расположенных на этой же дороге. В результате были получены

$$A_{n_1}^T = \left(a_1^{n_1}, a_2^{n_1}\right) = \begin{pmatrix} 1{,}07 & 0{,}92 \cdot 10^{-4} \end{pmatrix};$$

$$A_{n_2}^T = \left(a_1^{n_2}, a_2^{n_2}\right) = \begin{pmatrix} 1{,}11 & 2{,}31 \cdot 10^{-4} \end{pmatrix};$$

$$A_{n_3}^T = \left(a_1^{n_3}, a_2^{n_3}\right) = \begin{pmatrix} 1{,}18 & 3{,}62 \cdot 10^{-4} \end{pmatrix};$$

Введем параболическую зависимость значений коэффициентов модели (1) от числа пунктов n:

$$a_k^{(n)} = c_{k0} + c_{k1} n + c_{k2} n^2, \; k = 0;1. \qquad (3)$$

Искомые коэффициенты получим, обрабатывая полученные результаты с использованием МНК.

Введем

$$H = \begin{pmatrix} 1 & n_1 & n_1^2 \\ 1 & n_2 & n_2^2 \\ 1 & n_3 & n_3^2 \\ 1 & n_4 & n_4^2 \end{pmatrix}; \; C_1 = \begin{pmatrix} c_{10} \\ c_{11} \\ c_{12} \end{pmatrix}; \; C_2 = \begin{pmatrix} c_{20} \\ c_{21} \\ c_{22} \end{pmatrix}; \; A_1 = \begin{pmatrix} a_1^{(n_1)} \\ a_1^{(n_2)} \\ a_1^{(n_3)} \\ a_1^{(n_4)} \end{pmatrix}; \; A_2 = \begin{pmatrix} a_2^{(n_1)} \\ a_2^{(n_2)} \\ a_2^{(n_3)} \\ a_2^{(n_4)} \end{pmatrix};$$

$$C_1 = \left(H^T H\right)^{-1} H^T A_1, \; C_2 = \left(H^T H\right)^{-1} H^T A_2. \qquad (4)$$

В результате расчетов по формулам (4), получим

$$C_1 = \begin{pmatrix} 1.116 \\ -2.51 \cdot 10^{-3} \\ 3.722 \cdot 10^{-5} \end{pmatrix}, \; C_2 = \begin{pmatrix} 1.183 \cdot 10^{-4} \\ -2.282 \cdot 10^{-6} \\ 5.972 \cdot 10^{-8} \end{pmatrix}.$$

При этом соотношение (1) с учетом (3) имеет вид

$$M_{k+1}(n) = \exp\left(\left(c_{10} + c_{11}n + c_{12}n^2\right) + \left(c_{20} + c_{21}n + c_{22}n^2\right)M_k\right)M_k .$$

Полученное соотношение позволяет оценить продолжительность ожидания появления прогрессивной особи в задаче коммивояжера для заданного числа пунктов n на каждом этапе улучшения уже достигнутого решения.

Выводы. Рассмотрена технология оценки продолжительности ожидания прогрессивной особи при решении задачи коммивояжера с помощью генетического алгоритма. Эта оценка формируется рекуррентно и учитывает результаты работы ГА на предыдущем шаге улучшения решения. Исследована зависимость параметров полученного оценочного соотношения от размерности задачи.

Список литературы: **1.** *Goldberg D.* Genetic Algorithms / *D. Goldberg*. – MA: Addison Wesley, 1989. – 210 p. **2.** *Лысенко Ю. Г.* Нейронные сети и генетические алгоритмы / *Ю. Г. Лысенко, Н. Н. Иванов, А. Ю. Минц.* – Донецк: ООО «Юго-Восток, Лтд», 2003. – 265 с. **3.** *Серая О. В.* Многоиндексные модели логистики в условиях неопределенности / *О. В. Серая.* – Х. : ФОП Стеценко И. И., 2010. – 512 с.

Надійшла до редколегії 07.02.2012