УДК 519.854

Д.А. ГОБОВ

О СХОДИМОСТИ МОДИФИЦИРОВАННОГО АЛГОРИТМА УСКОРЕННОГО ВЕРОЯТНОСТНОГО МОДЕЛИРОВАНИЯ

Ключевые слова: *комбинаторная оптимизация, сходимость, алгоритмы вероятностного моделирования, G-алгоритм.*

ВВЕДЕНИЕ

Задачи комбинаторной оптимизации встречаются во многих областях практической деятельности (например, в задачах прогнозирования, составления планов, управления ресурсами и транспортными средствами, составления бюджета, сетевой маршрутизации и многое другое). В настоящее время перспективным направлением в создании методов дискретной оптимизации является использование тех или иных вероятностных механизмов. Рассмотрим один из известнейших примеров алгоритмов такого класса — методы стохастического локального поиска, в частности метод имитационного отжига [1]. Кроме оценки эффективности данного метода путем проведения вычислительных экспериментов, важное значение имеют теоретические исследования его сходимости. Так, в работах [2–4] определены достаточные условия для последовательности температур, при выполнении которых в алгоритмах отжига последовательность решений, получаемых на каждой итерации, сходится к оптимальному вектору по вероятности. В [5, 6] предложены температурные расписания, которые удовлетворяют необходимым и достаточным условиям для сходимости по вероятности к множеству глобальных решений. Другой подход предложен в [7], где исследована сходимость алгоритма имитационного отжига, который гарантирует нахождения именно глобального решения. Близкими по схеме к методу имитационного отжига являются алгоритмы ускоренного вероятностного моделирования (G-алгоритмы) (см. [8] и литературу к ней). В [9] рассмотрен вопрос сходимости классического G-алгоритма. В данной работе предлагается новая модификация алгоритма ускоренного вероятностного моделирования и исследуется его сходимость.

1. МОДИФИКАЦИЯ АЛГОРИТМА УСКОРЕННОГО ВЕРОЯТНОСТНОГО МОДЕЛИРОВАНИЯ И ЕГО СХОДИМОСТЬ

Задача комбинаторной оптимизации характеризуется тремя основными аспектами [10]:

- пространство поиска Ω, состоящее из конечного числа точек (изредка счетного);
- целевая функция $f(x)$, которая определена на пространстве поиска Ω;
- ограничения на пространство поиска.

Более формально она заключается в следующем: необходимо найти такую точку $x^* \in \Omega$, что $x^* = \arg\min_{x \in \Omega} f(x)$. Не ограничивая общности, будем полагать, что $f(x) > 0$, а Ω — конечное метрическое пространство с метрикой d.

Определение 1. Окрестностью точки $x \in \Omega$ произвольного радиуса $\rho > 0$ называется множество $N(x) = \{y \in \Omega, d(x, y) \le \rho\}$, где $d(x, y)$ — определенная на пространстве Ω метрика.

Приведем общую схему G-алгоритма [8]:

procedure G-search (x);
 begin
  $\mu_0 := 0$; $h := 0$;
  while (не выполнен критерий останова) **do**
   begin
    while (не выполнено условие равновесия) **do**
     begin
      $y :=$ очередная точка окрестности x;
      $\Delta := f(y) - f(x)$;
      $p(x, y) := \min\left\{1, 1 - \dfrac{\Delta}{z \cdot f(x)}\right\}$;
      $\xi := \mu_h + \text{random}[0,1] \cdot (1 - \mu_h)$;
      if $(p \ge \xi)$ **then**
       begin
        $x := y$;
       end
    end;
   $\mu_{h+1} = G(\mu_h)$;
   $h := h + 1$;
  end;
 return x;
end.

Здесь $G : [0,1) \to [0,1)$ — некоторая строго монотонная функция, а random[0,1] — датчик случайных чисел из отрезка [0,1]. Для такой формы G-алгоритма нельзя теоретически доказать сходимость к глобальному оптимуму, например, в отличие от алгоритма имитационного отжига. Рассмотрим следующий случай. Пусть $\mu_h = w > 0$, а x^* — локальный оптимум, для которого выполняется: $\forall\, y \in N(x^*)\ f(y) > f(x^*)[1 + (1 - w)z]$. Отсюда

$$\forall\, y \in N(x^*)\ p(x^*, y) < 1 - \frac{f(x^*)[1 + (1 - w)z] - f(x^*)}{zf(x^*)} = 1 - 1 + w = w.$$

Таким образом, $p< \xi \ \ \forall \ y \in N\ (x^*)$. Это означает, что в приведенном выше случае алгоритм не сможет выйти из локального оптимума. Поэтому дальше предлагается модификация указанного G-алгоритма, которая позволяет теоретически обосновать скорость сходимости к глобальному оптимуму. Суть модификации заключается в новом вероятностном механизме, который применяется для отсева точек в окрестности текущего решения. Предлагается рассчитывать вероятность перехода от точки x к произвольной точке y окрестности $N\ (x)$ по формуле

$$p\ (x,\ y) = \min \left\{ 1, 1 - \frac{\Delta}{g}\ (1-\mu) - \frac{\mu}{2}\ (1 + \operatorname{sgn}\ (\Delta)) \right\},$$

где $\operatorname{sgn}\ (x) = \begin{cases} 1, & x > 0, \\ -1, & x \le 0, \end{cases}$ $g = z \cdot \bar{f}$, $\bar{f} \ge f\ (x)\ \ \forall\ x \in \Omega$.

Будем считать, что известна верхняя граница $\bar{f}$ целевой функции $f\ (x)$: поскольку пространство Ω конечно, то она может быть рассчитана с помощью методов оценки для конкретного вида задачи (например, для квадратичной задачи о назначениях [11, 12]). Рассмотрим значение вероятности при $\Delta \le 0$ и $\Delta > 0$.

1. Пусть $\Delta > 0$, тогда

$$p\ (x,\ y) = \min \left\{ 1, 1 - \frac{\Delta}{g}\ (1-\mu) - \frac{\mu}{2}\ (1 + \operatorname{sgn}\ (\Delta)) \right\} =$$

$$= \min \left\{ 1, 1 - \frac{\Delta}{g} + \frac{\Delta}{g}\mu - \mu \right\} = \min \left\{ 1, \left(1 - \frac{\Delta}{g} \right) (1-\mu) \right\}.$$

2. Если $\Delta \le 0$, то имеем

$$p\ (x,\ y) = \min \left\{ 1, 1 - \frac{\Delta}{g}\ (1-\mu) - \frac{\mu}{2}\ (1 + \operatorname{sgn}\ (\Delta)) \right\} = \min \left\{ 1, 1 - \frac{\Delta}{g}\ (1-\mu) \right\}.$$

Из того, что $\Delta \le 0$ и $1 > \mu \ge 0$, следует $1 - \frac{\Delta}{g}\ (1-\mu) > 1$. Таким образом, $p\ (x,\ y) = \min \left\{ 1, 1 - \frac{\Delta}{g}\ (1-\mu) \right\} = 1$. Важно отметить, что в модифицированном алгоритме случайная величина $\xi \in [0, 1]$.

Вычислительную схему с использованием указанного вероятностного механизма назовем G^*-алгоритмом и представим в следующем виде:

procedure G^*-search (x);
  begin
    $\mu_0 := 0;\ h := 0;$
      while (не выполнен критерий останова) **do**
        begin
          while(не выполнено условие равновесия) **do**
            begin
              $y :=$ случайная точка окрестности x;
              $\Delta :\equiv f\ (y) - f\ (x)$;
              $p\ (x,\ y) := \min \left\{ 1, 1 - \frac{\Delta}{g}\ (1-\nu) - \frac{\mu}{2}\ (1 + \operatorname{sgn}\ (\Delta)) \right\}$
              $\xi :=$ random $[0, 1]$;
              if $(p \ge \xi)$ **then**
                begin
                  $x := y$;
                end
            end;
          $\mu_{h+1} = G\ (\mu_h)$;
          $h := h + 1$;
        end;
    return x;
  end.

Пусть $\Gamma(V, E)$ — ориентировочный граф, который соотносится с задачей оптимизации следующим образом: вершина i в Γ — это точка $x_i \in \Omega$, а ребра соединяют вершину i с точками, которые входят в заданную окрестность x_i. Тогда можно считать, что G^*-алгоритм строит случайный маршрут на этом графе. Переход из одной точки пространства Ω в другую зависит только от текущего состояния. Это позволяет рассматривать траекторию G^*-алгоритма как цепь Маркова, а при фиксированном значении μ — как однородную цепь Маркова.

Рассмотрим случай, когда граф Γ сильно связанный, т.е. для двух произвольных вершин i и j в V существует ориентированная цепь из i до j. Такой случай, в частности, моделирует роботу на пространстве перестановок. Для любой вершины i в Γ обозначим $M(i)$ множество инцидентных ей вершин. Если считать, что когда цепь Маркова находится в состоянии i, каждое состояние из его $M(i)$ может быть сгенерировано G^*-алгоритмом с одинаковой вероятностью, которая равняется $1/|M(i)|$, то вероятность перехода из i в j при заданном значении μ рассчитывается следующим образом:

$$P_{ij}(\mu) = \begin{cases} 0, & \text{если } j \notin M(i),\ j \neq i, \\ \dfrac{1}{|M(i)|} \min\left\{1, 1 - \dfrac{\Delta}{g}(1-\mu) - \dfrac{\mu}{2}(1 + \operatorname{sgn}(\Delta))\right\}, & \text{если } j \in M(i),\ j \neq i, \end{cases}$$

$$P_{ii}(\mu) = 1 - \sum_{j \in M(i)} P_{ij}(\mu).$$

Используем эту модель для доказательства сходимости G^*-алгоритма.

Определение 2 [7]. Алгоритм оптимизации сходится, если цепь Маркова содержит глобально оптимальное состояние. Есть и другие определения сходимости, например сходимость по вероятности вектора состояния к оптимальному вектору [2], но для наших исследований наиболее адекватно именно приведенное определение.

Пусть имеем целевую функцию $f(x)$, граф $\Gamma(V, E)$, построенный согласно приведенным выше правилам. Пусть μ^* — максимальное значение параметра μ, $\theta = \max\limits_{i \in V,\, j \in M(i)} \{f(x_i) - f(x_j)\}$. Из того, что Ω — конечное множество, а функция f ограничена величиной $\bar{f}$, имеем $\bar{\theta} < \infty$, $\underline{\theta} > 0$. Обозначим степень и диаметр $\Gamma(V, E)$ как d и D соответственно. Очевидно, что для любого $i \in V$ величина p_{ij} при любом значении μ не меньше величины $(1-\mu^*)\left(1 - \dfrac{\theta}{g}\right)$, если $j \in M(i)$.

Согласно [7] для алгоритма, который строит случайный маршрут на сильно связанном графе $\Gamma(V, E)$, построенном согласно приведенным выше правилам, траекторию которого рассматривать как цепь Маркова и для любых $i, j \in V$ вероятность перехода из вершини i в вершину j больше 0 при $j \in M(i)$, то математическое ожидание количества шагов, необходимых для того, чтобы посетить состояние, отвечающее глобальному оптимуму (opt) при произвольном начальном состоянии X, не превышает величины $\left(\dfrac{1}{d} p'\right)^{-D}$, где $p' = \min(p_{ij})$, $i, j \in V$. G^*-алгоритм удовлетворяет этим условиям. Таким образом получаем, что математическое ожидание количества шагов, необходимых для того, чтобы посетить состояние opt для G^*-алгоритма, не превышает величины $\left[d\left(\dfrac{g}{(1-\mu^*)(g-\theta)}\right)\right]^D$.

Теорема 1. G^*-алгоритм сходится за количество шагов, которое не превышает величины $2k\left[d\,\dfrac{g}{(1-\mu^*)(g-\theta)}\right]^D$, с вероятностью не меньше $(1-2^{-k})$ при любом начальном состоянии, где k — натуральное число.

Доказательство. Пусть $E = 2\left\lceil d\,\dfrac{g}{(1-\mu^*)(g-\theta)} \right\rceil^D$.

Доказательство того, что вероятность не посещения состояния opt на протяжении kE шагов меньше или равна величине 2^{-k}, проведем с помощью математической индукции.

Шаг 1. $k=1$. Пусть B — количество шагов алгоритма до того момента, когда состояние opt будет посещено.

Согласно неравенству Маркова получаем

$$P\left[B \le 2\left\lceil d\,\frac{g}{(1-\mu^*)(g-\theta)} \right\rceil^D\right] \ge 1 - \frac{\left\lceil d\,\dfrac{g}{(1-\mu^*)(g-\theta)} \right\rceil^D}{2\left\lceil d\,\dfrac{g}{(1-\mu^*)(g-\theta)} \right\rceil^D} = 1 - \frac{1}{2} = \frac{1}{2}.$$

Шаг 2. Пусть теорема выполняется для всех $k \le (r-1)$.

Шаг 3. Докажем теорему для $k = r$.

Пусть $X_E, X_{2E}, \ldots, X_{(r-1)E}$ — множества состояний цепи Маркова, которые были посещены на протяжении шагов $E, 2E, \ldots, (r-1)E$ соответственно. Рассмотрим два события.

Событие *A*: состояние opt не было посещено на протяжении первых E шагов.

Событие *B*: состояние opt не было посещено на протяжении следующих $(r-1)E$ шагов.

Тогда вероятность P того, что opt не будет посещен за rE шагов, рассчитывается следующим образом: $P = P[B/A] * P[A]$, где $P[A]$ — вероятность наступления события A, $P[B/A]$ — вероятность наступления события B при условии, что настало событие A.

Величина $P[B/A]$ зависит только от состояния цепи Маркова на шаге E и количества шагов $(r-1)E$. Отсюда $P = P[A]\sum\limits_{i \in V} P[B/X_E = i]\,P[X_E = i]$, но $P[A] \le 1/2$ и $P[B/X_E = i] \le 2^{-(r-1)}$ для каждого $i \in V$ (согласно гипотезе индукции).

Таким образом, имеем $\dfrac{1}{2}2^{-(r-1)} = 2^{-r}$.

Теорема доказана.

Рассмотрим задачи оптимизации, определенные на пространстве Ω перестановок размерности n.

Определение 3. Расстоянием $d(a, b)$ между двумя произвольными перестановками a и b, $a, b \in \Omega$, будем называть минимальное количество транспозиций, которые превращают перестановку a в b.

Напомним, что под транспозицией понимается обмен местами двух элементов перестановки [13].

Рассмотрим случай окрестности с радиусом 1:

$$d = C_n^2 = \frac{n!}{(n-2)!\,2!} = \frac{n(n-1)}{2}, \quad D = n-1.$$

С учетом этого имеем очередной результат.

Следствие 1. G^*-алгоритм на пространстве Ω перестановок размерности n с радиусом окрестности 1 сходится за количество шагов, которое не превышает величины $2k\left[\frac{n\,(n-1)}{2}\,\frac{g}{(1-\mu^{*})(g-\theta)}\right]^{n-1}$ с вероятностью, не меньшей $(1-2^{-k})$ при произвольном начальном состоянии, где k — натуральное число.

2. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Для сравнения эффективности G^{*}- и G-алгоритма решались квадратичные задачи о назначениях (КЗН), для которых известны или точные решения, или лучшие из известных. Реальные и тестовые задачи взяты из библиотеки [14].

Вычисления выполнялись на рабочей станции со следующими характеристиками: процессор AMD Turion 2,19 Ghz, RAM 512 Мбайт. Для получения статически обоснованных результатов по каждой задаче проводилась серия экспериментов с одинаковыми параметрами алгоритма и разными начальными приближениями целевой функции. Некоторые результаты ряда известных задач приведены в табл. 1, где n — размерность задачи, q — среднее улучшение значений целевой функции относительно начального варианта решения, которое рассчитывает по формуле

$$q=\frac{1}{C}\sum_{i=1}^{C}\frac{f_{i\,\text{на÷}}-f_i^{*}}{f_{i\,\text{на÷}}}100\%,$$

где $f_{i\,\text{на÷}}$ — начальное приближение целевой функции в эксперименте i, полученное методом Монте-Карло, f_i^{*} — значение целевой функции в эксперименте i, которое отвечает найденному данным алгоритмом решению; t — время работы соответствующего алгоритма (в с), C — объем серии экспериментов по решению задачи (для всех задач выбиралось $C=50$).

Результаты вычислительного эксперимента свидетельствуют, что для задач практически всех размерностей в среднем эффективность G^{*}-алгоритма была не хуже эффективности G-алгоритма.

В табл. 2 приведены значения среднеквадратичного отклонения (σ) и величина доверительного интервала (δ) для величины улучшения значения целевой функции.

Приведенные данные свидетельствуют, что доверительные интервалы достаточно малы, поэтому оценки эффективности алгоритмов можно считать достоверными.

Таблица 1. Результаты решения КЗН

Название задачи	n	**G--алгоритм**		**G^{*}-алгоритм**	
		q	t	q	t
Chr20	20	70,78	0,5	70,85	0,51
Chr20	20	71,60	0,54	71,80	0,51
Chr20	20	76,44	0,54	76,14	0,56
Lip40a	40	2,41	3,9	2,41	3,8
Lipa50	50	1,98	7,36	1,98	7,47
Esp64a	64	58,07	9,1	58,14	9,17
Sco72	72	14,80	26,4	14,93	27,1
Sco81	81	14,72	47,03	14,86	48,08
Tai90a	90	1,12	8,16	1,12	22,31
Tai100b	100	30,39	35,92	30,40	40,30

Таблица 2. Среднеквадрптичное отклонение и величина доверительного интервала

Название задачи	n	**G--алгоритм**		**G^{*}-алгоритм**	
		α	δ	α	δ
Chr20	20	5,39	5,42	5,07	5,1
Chr20	20	5,32	5,35	4,45	4,47
Chr20	20	6,92	6,96	5,68	5,71
Lip40a	40	0,33	0,33	0,23	0,23
Lipa50	50	0,19	0,19	0,20	0,20
Esp64a	64	2,79	2,8	2,83	2,84
Sco72	72	0,9	0,9	0,77	0,78
Sco81	81	0,71	0,71	0,73	0,74
Tai90a	90	0,069	0,07	0,08	0,08
Tai100b	100	1,46	1,47	1,46	1,47

ЗАКЛЮЧЕНИЕ

Применение алгоритмов ускоренного вероятностного моделирования для решения широкого круга задач комбинаторной оптимизации как отдельных алгоритмов или в составе метаэвристик вызвало необходимость исследования сходимости алгоритмов этого класса.

Для предложенной модификации *G*-алгоритма получена теоретически обоснованная оценка скорости сходимости, которая не зависит от начального приближения. Проведен вычислительный эксперимент для сравнения эффективности классического и модифицированного *G*-алгоритмов.

Полученные результаты позволяют утверждать, что изменения, внесенные в вероятностный механизм, не ухудшают эффективность алгоритма, таким образом, предложенный алгоритм может использоваться для решения широкого круга прикладных задач.

Дальнейшие разработки будут направлены на исследование влияния вида функции *G*, значений параметров алгоритма и характеристик задачи на эффективность алгоритма, разработка механизма автоматического подбора изменяемых параметров алгоритма.

СПИСОК ЛИТЕРАТУРИ

1. Kirkpatrick S., Gelatt C.D., Vecchi M.P. Optimization by Simulated Annealing // Science. — 1983. — N 4598. — P. 671–680.
2. Geman S., Geman D. Stochastic relaxation, Gibbs distribution, Bayesian restoration of image // IEEE Trans. PAMI. — 1984. — N 6. — P. 721–741.
3. Anily S., Federgruen A. Simulated annealing methods with general acceptance probabilities // J. of Applied Probability. — 1987. — N 35. — P. 657–667.
4. Johnson A.W. and Jacobson S.H. On the convergence of generalized hill climbing algorithms // Discrete Appl. Mathemat. — 2002. — N 119. — P. 37–57.
5. Gidas B. Nonstationary markov chains and convergence of the annealing algorithm // J. of Statistic. Physics. — 1985. — N 39 — P. 73–131
6. Hajek B. Cooling Shedules for Optimal Annealing // Math. Oper. Res. — 1988. — N 13. — P. 311–329.
7. Rajasekaran S. On Simulated Annealing and Nested Annealing // J. of Global Optimization. — January 2000. — N 1. — P. 43–56.
8. Гуляницкий Л.Ф. Сергиенко И.В. Метаэвристический метод деформируемого многогранника в комбинаторной оптимизации. // Кибернетика и системный анализ. — 2007. — № 6. — С. 70–79.
9. Турчин О.Я. Дослідження збіжності алгоритму прискореного ймовірнісного моделювання, що розв'язує певний клас задач комбінаторноВ оптимізаціВ // Вісн. КиВв. ун-ту. — 2003. — № 1. — С. 216–221.
10. Пападимитриу Х., Стайглиц К. Комбинаторная оптимизация. — М.: Мир, 1985. — 510 с.
11. Gilmore P.C. Optimal and suboptimal algorithms for the quadratic assignment problem // SIAM J. on Appl. Mathemat. — 1962. — N 10. — P. 305–310.
12. Hadley S.W., Rendl F., and Wolkowicz H. A new lower bound via projection for the quadratic assignment problem // Mathemat. of Oper. Res. — 1992. — N 17. — P. 727—739.
13. Сергиенко И.В., Каспшицкая М.Ф. Модели и методы решения на ЭВМ комбинаторных задач оптимизации. — Киев: Наук. думка, 1981. — 288 с.
14. QAPLIB // http://www.opt.math.tu-graz.ac.at/qaplib

Поступила 19.04.2007