

## О ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛАХ ДЛЯ ЧИСЛА ЛОКАЛЬНЫХ ОПТИМУМОВ

**А.В. Еремеев**<sup>1</sup>

доцент, д.ф.-м.н., e-mail: eremeev@ofim.oscsbras.ru

**К.Р. Ривс**<sup>2</sup>

профессор, e-mail: c.reeves@coventry.ac.uk

<sup>1</sup>Институт математики им. С. Л. Соболева Сибирского отделения РАН

<sup>2</sup>Университет Ковентри

**Аннотация.** Число локальных оптимумов является важным показателем сложности задач оптимизации для алгоритмов локального поиска. В настоящей работе обсуждаются некоторые методы нахождения доверительных интервалов для этого параметра в задачах, где мощность пространства поиска не позволяет выполнить полный перебор решений. Представлены результаты вычислительных экспериментов, полученные с использованием предложенных методов для модели  $\mathcal{NK}$  ландшафтов С. Кауффмана, задачи отыскания двоичной последовательности с минимальной автокорреляцией, задачи о вершинном покрытии и задачи распределения бункерных ёмкостей в производственных линиях.

**Ключевые слова:** локальный поиск, задача комбинаторной оптимизации, перепись Шнабеля, консервативный доверительный интервал.

### Введение

Значительное число эвристических методов решения задач комбинаторной оптимизации основано на поиске улучшающих решений в *окрестности* некоторой текущей точки. При этом могут быть использованы различные стратегии выбора новой текущей точки из окрестности, в которую будет сделан очередной переход. На практике такие методы быстро приходят к локальному оптимуму, после чего требуется либо начинать поиск с новой точки (см., например, [20, 30, 34]), либо прибегать к некоторому правилу постепенного выхода из окрестности, как это делается в алгоритмах поиска с запретами [14] и поиска с переменными окрестностями [16]. Как показывает практика, для успешного применения таких методов большое значение имеет удачный выбор системы окрестностей. Сложность индивидуальной задачи с заданной системой окрестностей характеризуется такими параметрами, как трудоёмкость выбора улучшающего решения в окрестности, среднее и максимальное число шагов до локального оптимума, гарантированная относительная погрешность локальных оптимумов и общее их число. Исследование генетических алгоритмов также

приводит к необходимости анализа структуры множества локальных оптимумов [26].

В настоящей работе предлагаются и обосновываются статистические методы построения доверительных интервалов для числа локальных оптимумов на основе результатов многократного выполнения локального поиска. Методы предназначены для исследования индивидуальных задач комбинаторной оптимизации, в которых большая размерность пространства решений не позволяет осуществить полный перебор всех его элементов за приемлемое время. Будем далее обозначать число локальных оптимумов задачи через  $\nu$ .

Из литературы известно несколько подходов к оцениванию величины  $\nu$  на основе результатов многократного выполнения локального поиска из случайно выбранных начальных точек. Один из подходов основывается на упрощающем предположении о том, что все локальные оптимумы имеют равные вероятности попадания в них (предположение *изотропности*) [27]. Другие подходы состоят в подборе параметров некоторого распределения вероятностей (экспоненциального, гамма, логнормального и т.д.) [12, 28] для описания вероятности попадания в различные локальные оптимумы или использовании таких непараметрических методов, как бутстреп-метод или метод складного ножа [28, 29]. Некоторые из этих подходов основываются на методах, разработанных в статистике и экологии для оценивания численности популяций (см., например, [24, 33]).

Настоящая работа представляет собой расширение краткой статьи [11] и содержит более детальный теоретический анализ и более широкое экспериментальное исследование. В п. 1 рассматриваются методы построения консервативных доверительных интервалов, основанные непосредственно на функции распределения наблюдаемой целочисленной случайной величины. Здесь же предложены методы их вычисления, основанные на первом повторении локального оптимума и на числе различных локальных оптимумов, полученных за время испытаний. В п. 2 приводятся описание тестовых примеров и результаты вычислительных экспериментов, проведённых с предложенными методами построения доверительных интервалов и с непараметрическим бутстреп-методом из [29]. Описание последнего содержится в приложении.

## 1. Методы построения доверительных интервалов

Рассмотрим сначала достаточно общую ситуацию, когда при любом значении представляющего интерес неизвестного параметра  $\nu$  функция распределения наблюдаемой целочисленной случайной величины  $T$  известна. Пусть  $\mathbf{F}(t; \nu) = \mathbf{P}\{T \leq t\}$  — функция распределения величины  $T$  при указанном значении  $\nu$ . Наблюдаемое значение величины  $T$ , полученное в фактически проведённом эксперименте, будем обозначать через  $\tau$ ,  $\tau \in \mathbb{N}$ .

Если упрощённо считать, что  $\nu$  может принимать любое вещественное значение и  $\mathbf{F}(t; \nu)$  непрерывно зависит от  $\nu$ , то при заданных вероятностях  $\alpha_1, \alpha_2$ , таких, что  $\alpha_1 + \alpha_2 < 1$ , доверительный интервал  $[\nu_1(\tau), \nu_2(\tau)]$  уровня

$(1 - \alpha_1 - \alpha_2) \cdot 100\%$  может быть построен решением уравнений

$$1 - \mathbf{F}(\tau; \nu_1) = \alpha_1, \quad \mathbf{F}(\tau; \nu_2) = \alpha_2 \quad (1)$$

относительно неизвестных  $\nu_1$  и  $\nu_2$  (см., например, [4, 22]).

Далее интервал  $\nu_1(\tau) \leq \nu \leq \nu_2(\tau)$  будем называть консервативным доверительным интервалом с доверительным уровнем  $(1 - \alpha) \cdot 100\%$ , если с вероятностью не менее  $1 - \alpha$  выполняются неравенства  $\nu_1(\tau) \leq \nu \leq \nu_2(\tau)$ . В случае  $\nu_2(\tau) = +\infty$  консервативный доверительный интервал будем называть консервативной нижней границей.

Поскольку в поставленной задаче параметр  $\nu$  принимает только натуральные значения, необходимо соответственно адаптировать (1).

**Утверждение 1.** Пусть

$$\nu_1(\tau) = \min\{\nu' \in \mathbb{N} : 1 - \mathbf{F}(\tau - 1; \nu') \geq \alpha_1\},$$

$$\nu_2(\tau) = \max\{\nu' \in \mathbb{N} : \mathbf{F}(\tau; \nu') \geq \alpha_2\},$$

тогда  $[\nu_1(\tau), \nu_2(\tau)]$  является консервативным доверительным интервалом целочисленного параметра  $\nu$  с уровнем значимости  $(1 - \alpha_1 - \alpha_2) \cdot 100\%$ , причём  $\mathbf{P}\{\nu < \nu_1(T)\} \leq \alpha_1$ ,  $\mathbf{P}\{\nu_2(T) < \nu\} \leq \alpha_2$ .

**Доказательство.** Пусть для некоторого значения  $\nu$  утверждение неверно. Тогда имеет место по крайней мере одно из двух неравенств:  $\alpha_1 < \mathbf{P}\{\nu < \nu_1(\tau)\}$ ,  $\alpha_2 < \mathbf{P}\{\nu_2(\tau) < \nu\}$ .

В первом случае

$$\begin{aligned} \alpha_1 < \mathbf{P}\{\nu < \min(\nu' : 1 - \mathbf{F}(T - 1; \nu') \geq \alpha_1)\} &\leq \mathbf{P}\{1 - \mathbf{F}(T - 1; \nu) < \alpha_1\} = \\ &= \mathbf{P}\{T - 1 \geq \min(\Theta : 1 - \mathbf{F}(\Theta; \nu) < \alpha_1)\} = \mathbf{P}\{T \geq \Theta_{\min}(\nu, \alpha_1)\}, \end{aligned}$$

где  $\Theta_{\min}(\nu, \alpha_1) = \min\{\Theta' : 1 - \mathbf{F}(\Theta' - 1; \nu) < \alpha_1\}$ . Однако

$$\mathbf{P}\{T \geq \Theta_{\min}(\nu, \alpha_1)\} = 1 - \mathbf{F}(\Theta_{\min}(\nu, \alpha_1) - 1; \nu) < \alpha_1 \quad (2)$$

по определению  $\Theta_{\min}$ , что приводит к противоречию.

Во втором случае

$$\begin{aligned} \alpha_2 < \mathbf{P}\{\max\{\nu' : \mathbf{F}(T; \nu') \geq \alpha_2\} < \nu\} &\leq \\ &\leq \mathbf{P}\{\mathbf{F}(T; \nu) < \alpha_2\} = \mathbf{P}\{T < \theta_{\min}(\nu, \alpha_2)\}, \end{aligned}$$

где  $\theta_{\min}(\nu, \alpha_2) = \min\{\theta : \mathbf{F}(\theta; \nu) \geq \alpha_2\}$ . Однако  $\mathbf{P}\{T < \theta_{\min}(\nu, \alpha_2)\} < \alpha_2$ .  $\square$

Модифицируем предыдущее утверждение с целью нахождения консервативной нижней границы заданного доверительного уровня  $\alpha$  в случае, когда распределение вероятностей случайной величины  $T$  зависит не только от  $\nu$ , но и от дополнительного вектора параметров  $\mathbf{y}$ , принадлежащего некоторому множеству  $\mathcal{Y}(\nu) \subseteq \mathbb{R}^r$ . Обозначим в таком случае функцию распределения для  $T$  через  $\mathbf{F}(t; \nu, \mathbf{y})$ .

**Утверждение 2.** Пусть

$$\nu_1(\tau) = \min \left\{ \nu' \in \mathbb{N} : 1 - \inf_{\mathbf{y} \in \mathcal{Y}(\nu')} \mathbf{F}(\tau - 1; \nu', \mathbf{y}) \geq \alpha \right\},$$

тогда  $\mathbf{P}\{\nu < \nu_1(T)\} \leq \alpha$ .

**Доказательство.** Достаточно в доказательстве предыдущего утверждения заменить определение  $\Theta_{\min}(\nu, \alpha)$  на

$$\Theta_{\min}(\nu, \alpha) = \min \left\{ \Theta' \in \mathbb{N} : 1 - \inf_{\mathbf{y} \in \mathcal{Y}(\nu)} \mathbf{F}(\Theta' - 1; \nu, \mathbf{y}) < \alpha \right\},$$

и соответствующим образом модифицировать неравенство (2):

$$\begin{aligned} \mathbf{P}\{T \geq \Theta_{\min}(\nu, \alpha)\} &= 1 - \mathbf{F}(\Theta_{\min}(\nu, \alpha) - 1; \nu, \mathbf{y}) \leq \\ &\leq 1 - \inf\{\mathbf{F}(\Theta_{\min}(\nu, \alpha) - 1; \nu, \mathbf{y}) : \mathbf{y} \in \mathcal{Y}(\nu)\} < \alpha. \end{aligned}$$

□

Если при любом  $t$  функция  $\mathbf{F}(t; \nu)$  (функция  $\mathbf{F}_{\mathcal{Y}}(t, \nu) = \inf_{\mathbf{y} \in \mathcal{Y}(\nu')} \mathbf{F}(\tau - 1; \nu', \mathbf{y})$ ) монотонна по  $\nu$ , то доверительный интервал из утверждения 1 (утверждения 2) может быть вычислен с помощью метода дихотомии. Ниже будет показано, что для рассматриваемых в настоящей работе оценок величины  $\nu$  монотонность по  $\nu$  имеет место.

### 1.1. Ожидание первого повторения

Пусть  $p_i$  — вероятность попадания локального поиска в локальный оптимум  $i$ ,  $i = 1, \dots, \nu$ , и пусть  $T$  — номер испытания локального поиска, при котором впервые получен ранее встречавшийся локальный оптимум. Функцию распределения случайной величины  $T$ , соответствующую вектору  $\mathbf{p} = (p_1, \dots, p_\nu)$ , будем обозначать через  $\mathbf{F}_{\mathbf{p}}(t; \nu)$ . В частном случае *изотропного распределения*, когда вероятности попадания во все локальные оптимумы одинаковы, введём обозначение  $\bar{\mathbf{p}} = (1/\nu, \dots, 1/\nu)$ .

Следующее утверждение показывает, что изотропное распределение является в некотором смысле экстремальным.

**Утверждение 3.** Для любых  $t, \nu \in \mathbb{N}$ ,  $t \geq 2$ , функция  $\mathbf{F}_{\mathbf{p}}(t; \nu)$  достигает своего единственного минимума при  $\mathbf{p} = \bar{\mathbf{p}}$ .

**Доказательство.** Пусть  $S$  — множество всех последовательностей длины  $t$  из элементов множества  $\{1, \dots, \nu\}$ , не содержащих повторений, и пусть  $\sigma(i)$  —  $i$ -й элемент последовательности  $\sigma$ . Тогда

$$\mathbf{F}_{\mathbf{p}}(t; \nu) = 1 - \mathbf{P}\{T > t\} = 1 - \sum_{\sigma \in S} \prod_{i=1}^t p_{\sigma(i)}.$$

$\mathbf{F}_{\mathbf{p}}(t; \nu)$  — непрерывная функция на компактном множестве

$$\Lambda = \left\{ (p_1, p_2, \dots, p_\nu) \geq 0 : \sum_{i=1}^{\nu} p_i = 1 \right\},$$

когда  $t$  фиксировано. Таким образом, существует распределение вероятностей попадания в локальные оптимумы, при котором достигается минимум функции  $\mathbf{F}_{\mathbf{p}}(t; \nu)$  на  $\Lambda$ . Покажем, что никакое распределение  $\mathbf{p}$ , кроме  $\bar{\mathbf{p}}$ , таким свойством не обладает.

Предположим противное: пусть минимум достигается на векторе  $\mathbf{p}$ , где  $p_i \neq p_j$  для некоторых  $i$  и  $j$ . Не теряя общности, предположим, что  $i = 1, j = 2, p_2 = \gamma p_1$  и  $\gamma < 1$ . Рассмотрим распределение вероятностей

$$p'_1 = p'_2 = \frac{p_1 + p_2}{2}, p'_3 = p_3, \dots, p'_\nu = p_\nu$$

и соответствующую ему функцию распределения  $\mathbf{F}_{\mathbf{p}'}(t; \nu)$ . Достаточно показать, что  $\mathbf{F}_{\mathbf{p}'}(t; \nu) < \mathbf{F}_{\mathbf{p}}(t; \nu)$ , т.к. это приведёт к противоречию с оптимальностью  $\mathbf{p}$ . Обозначим через  $S(j)$  подмножество  $S$ , состоящее из тех последовательностей, в которые входит элемент  $j, j = 1, 2$ . Тогда  $\mathbf{F}_{\mathbf{p}}(t; \nu) = 1 - a(\mathbf{p}) - b(\mathbf{p}) - c(\mathbf{p}) - d(\mathbf{p})$ , где

$$\begin{aligned} a(\mathbf{p}) &= \sum_{\sigma \in S(1) \cap S(2)} \prod_{i=1}^t p_{\sigma(i)}, & b(\mathbf{p}) &= \sum_{\sigma \in S(1) \setminus S(2)} \prod_{i=1}^t p_{\sigma(i)}, \\ c(\mathbf{p}) &= \sum_{\sigma \in S(2) \setminus S(1)} \prod_{i=1}^t p_{\sigma(i)}, & d(\mathbf{p}) &= \sum_{\sigma \in S \setminus (S(1) \cup S(2))} \prod_{i=1}^t p_{\sigma(i)}. \end{aligned}$$

Аналогично  $\mathbf{F}_{\mathbf{p}'}(t; \nu) = 1 - a(\mathbf{p}') - b(\mathbf{p}') - c(\mathbf{p}') - d(\mathbf{p}')$  и  $d(\mathbf{p}') = d(\mathbf{p})$ . Заметим, что если элемент 2 во всех последовательностях из  $S(2) \setminus S(1)$  заменить на 1, то будет получено множество последовательностей  $S(1) \setminus S(2)$ . Таким образом, с одной стороны,

$$\begin{aligned} b(\mathbf{p}) + c(\mathbf{p}) &= \sum_{\sigma \in S(1) \setminus S(2)} \prod_{i=1}^t p_{\sigma(i)} + \gamma \left( \sum_{\sigma \in S(1) \setminus S(2)} \prod_{i=1}^t p_{\sigma(i)} \right) \\ &= (1 + \gamma)p_1 \cdot \sum_{\sigma \in S(1) \setminus S(2)} \prod_{i: \sigma(i) \neq 1} p_{\sigma(i)}. \end{aligned}$$

С другой стороны,  $p'_1 = p'_2 = (p_1 + p_2)/2$ , поэтому

$$b(\mathbf{p}') + c(\mathbf{p}') = (p_1 + p_2) \sum_{\sigma \in S(1) \setminus S(2)} \prod_{i: \sigma(i) \neq 1} p_{\sigma(i)}$$

и  $b(\mathbf{p}) + c(\mathbf{p}) = b(\mathbf{p}') + c(\mathbf{p}')$ . Далее

$$a(\mathbf{p}) = \gamma p_1^2 \cdot \sum_{\sigma \in S(1) \cap S(2)} \prod_{i: \sigma(i) \neq 1, 2} p_{\sigma(i)},$$

$$a(\mathbf{p}') = \left( \frac{p_1(1+\gamma)}{2} \right)^2 \sum_{\sigma \in S(1) \cap S(2)} \prod_{i: \sigma(i) \neq 1,2}^t p_{\sigma(i)}.$$

Однако  $(1+\gamma)^2 > 4\gamma$ , поэтому  $a(\mathbf{p}') > a(\mathbf{p})$  и  $\mathbf{F}_{\mathbf{p}'}(t; \nu) < \mathbf{F}_{\mathbf{p}}(t; \nu)$ .  $\square$

Как следствие из доказанного утверждения вытекает свойство монотонности функции  $\mathbf{F}_{\bar{\mathbf{p}}}(t; \nu)$ .

**Следствие 1.**  $\mathbf{F}_{\bar{\mathbf{p}}}(t; \nu)$  является строго убывающей функцией от  $\nu$ .

**Доказательство.** Достаточно сравнить оптимальное по утверждению 3 значение  $\mathbf{F}_{\bar{\mathbf{p}}}(t; \nu)$  с величиной  $\mathbf{F}_{\mathbf{p}}(t; \nu)$ , где  $\mathbf{p} = (1/(\nu-1), 1/(\nu-1), \dots, 1/(\nu-1), 0)$ .  $\square$

Ввиду монотонности  $\mathbf{F}_{\bar{\mathbf{p}}}(t; \nu)$  по  $\nu$ , в изотропном случае доверительный интервал для  $\nu$  может быть вычислен методом дихотомии с использованием утверждения 1, при этом  $\mathbf{F}(t; \nu) = 1 - \binom{\nu}{t} t! / \nu^t$ . Консервативная нижняя граница  $\nu_1$  с заданным доверительным уровнем  $100 \cdot (1-\alpha)\%$ , справедливая при любом распределении  $\mathbf{p}$ , вытекает из утверждения 2. Далее будем обозначать эту границу через  $L(t, 1-\alpha)$ .

Если задача оптимизации имеет чрезвычайно большое число локальных оптимумов, а число испытаний локального поиска  $r$  оказалось недостаточным, чтобы встретить хотя бы один локальный оптимум более одного раза, то консервативная нижняя граница доверительного уровня  $100(1-\alpha)\%$  для  $\nu$  может быть вычислена в предположении о том, что следующее (гипотетическое) испытание приведёт к первому повторению. Оценим скорость роста оценки  $L(r+1, 1-\alpha)$  как функции от  $r$ . Заметим, что согласно утверждению 2,  $L(r+1, 1-\alpha)$  есть минимальное значение  $\mu$ , при котором  $\binom{\mu}{r} r! / \mu^r \geq \alpha$ . Однако

$$\begin{aligned} \binom{\mu}{r} \frac{r!}{\mu^r} &= \frac{\mu!}{(\mu-r)! \mu^r} = \exp \left\{ \sum_{j=1}^r \ln \left( 1 - \frac{r-j}{\mu} \right) \right\} \leq \\ &\leq \exp \left\{ - \sum_{j=1}^r \frac{r-j}{\mu} \right\} = \exp \left\{ - \frac{r(r+1)}{2\mu} \right\} \leq \exp \left\{ - \frac{r^2}{2\mu} \right\}, \end{aligned}$$

поэтому  $L(r+1, 1-\alpha) \geq 0.5r^2 / \ln(\alpha^{-1})$ , и, к примеру, при  $\alpha = 0.05$  имеем  $0.5r^2 / \ln(\alpha^{-1}) \approx 0.167r^2$ . При больших значениях  $r$  эта простая оценка хорошо аппроксимирует величину  $L(r+1, 1-\alpha)$ .

## 1.2. Метод переписи Шнабеля

Один из методов, используемых в экологии для оценки численности популяций, носит название *переписи Шнабеля* [32, 33]. Из популяции с неизменным составом производится отлов заранее заданного числа особей, каждая из которых имеет константную вероятность отлова. При отлове особь помечается (если она не была помечена ранее) и возвращается обратно в популяцию. Статистические оценки численности популяции вычисляются на основе данных о числе помеченных особей среди отловленных.

Для оценки числа локальных оптимумов метод переписи Шнабеля адаптируется следующим образом. Выполняются  $r$  испытаний локального поиска из начальных точек, выбранных случайным образом с одинаковым распределением (число  $r$  выбирается априори). Пусть  $K$  — случайная величина, равная числу различных локальных оптимумов, полученных за  $r$  испытаний. Обозначим через  $\mathbf{F}_{\mathbf{p}}(k; \nu, r) = \mathbf{P}\{K \leq k\}$  функцию распределения величины  $K$  в предположении о том, что  $\mathbf{p} = (p_1, \dots, p_\nu)$  — вектор вероятностей попадания в локальные оптимумы.

Для вычисления доверительных интервалов на параметр  $\nu$  известны приближенные методы, в которых используется аппроксимация распределения величины  $K$  нормальным распределением (см., например, [27, 33]), а также точные методы [13, 23], основанные на непосредственном вычислении функции распределения, подобно выражениям из утверждения 1.

Как было замечено в [7], в изотропном случае распределение вероятностей случайной величины  $K$  имеет вид:

$$\mathbf{P}\{K = k\} = \frac{\nu(\nu - 1)\dots(\nu - k + 1) S(r, k)}{\nu^r} = \frac{\nu! S(r, k)}{(\nu - k)! \nu^r}, \quad (3)$$

где  $S(r, k)$  — число Стирлинга второго рода, равное количеству способов разбить  $r$ -элементное множество на  $k$  непустых подмножеств без учета порядка подмножеств,  $S(r, k) = \frac{1}{k!} \sum_{j=0}^k (-1)^{k+j} \binom{k}{j} j^r$ .

Для вычисления оценки максимального правдоподобия  $\hat{\nu}(r, k)$  может быть рассмотрена величина

$$\mathcal{L}(\nu, r, k) = \ln(\mathbf{P}\{K = k\}) = \ln \nu! - \ln(\nu - k)! + \ln S(r, k) - r \ln \nu.$$

В [8] Дж. Даррохом предложено приближённо (с погрешностью не более единицы) вычислять  $\hat{\nu}(r, k)$ , как корень уравнения  $\mathcal{L}(\nu, r, k) = \mathcal{L}(\nu - 1, r, k)$ , т. е.

$$(\nu - 1)^r = \nu^{r-1}(\nu - k), \quad (4)$$

т. к.  $\ln \nu! - \ln(\nu - 1)! = \ln \nu$ . Корень уравнения (4) может быть вычислен с помощью метода Ньютона или методами оптимизации унимодальной функции на прямой. Заметим, что при  $k = r$  уравнение  $\frac{\partial}{\partial \nu} \mathcal{L}(\nu, r, k) = 0$  не имеет конечного решения, и оценка  $\hat{\nu}(r, k)$  лишена смысла.

Найденные в [8] асимптотические оценки дисперсии позволяют построить консервативный доверительный интервал с центром в точке  $\hat{\nu}(r, k)$ , используя упрощающее предположение о нормальности распределения данной оценки. Как показано в [13, 23], более точно вычислить консервативный доверительный интервал удаётся при использовании нормальной аппроксимации распределения  $K$ . Однако оба подхода могут давать существенные ошибки в ситуациях, когда предположение нормальности оказывается неуместным. Как видно из рис. 1, доверительный интервал из [8] даёт корректные оценки при достаточно больших значениях  $r$ . При малых значениях  $r$  возникают погрешности из-за нарушения гипотезы нормальности.

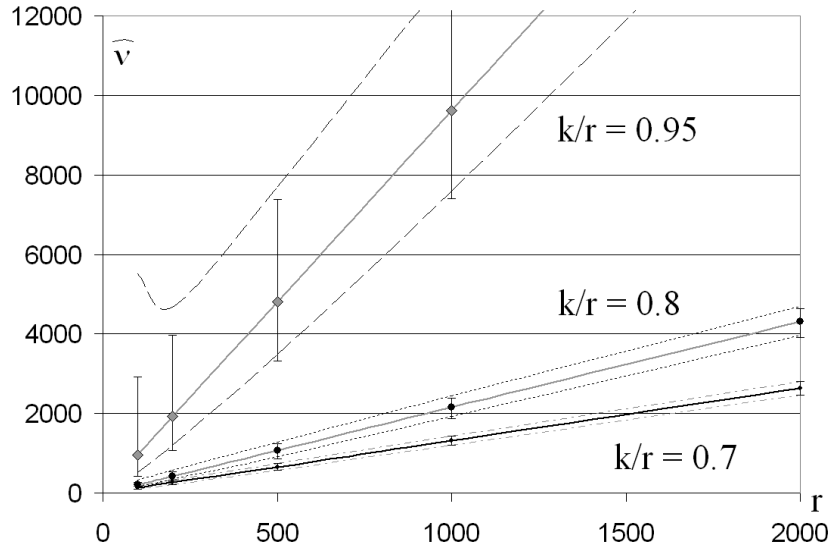


Рис. 1. Доверительные интервалы для  $\nu$ , построенные симметрично к оценке  $\hat{\nu}(r, k)$  по методу Дж. Дарроха [8] и по точным формулам из утверждения 1. Пунктирные линии показывают 95% доверительные интервалы, полученные в предположении нормальности  $K$ . Сплошные линии — оценки максимального правдоподобия с точными 95% доверительными интервалами.

Рассмотрим следствие основного результата А.М. Зубкова и Н.Н. Попова из [2], на основе которого будет получен аналог утверждения 3 для переписи Шнабеля. Пусть имеется  $r$  частиц, размещаемых случайным образом, независимо, с одинаковым распределением по ячейкам  $1, 2, \dots, \nu$ . Вероятность попадания в  $j$ -ю ячейку равна  $p_j$ ,  $j = 1, \dots, \nu$ ,  $\sum_{j=1}^{\nu} p_j = 1$ . Как показано в [2], в случае изотропного распределения вероятностей  $(p_1, \dots, p_{\nu}) = \bar{\mathbf{p}}$  число занятых ячеек  $\mu(r, \bar{\mathbf{p}})$  «стохастически больше» числа занятых ячеек  $\mu(r, \mathbf{p}')$  при любом другом распределении вероятностей  $\mathbf{p}'$ , т.е.  $\mathbf{P}\{\mu(r, \bar{\mathbf{p}}) \leq k\} \leq \mathbf{P}\{\mu(r, \mathbf{p}') \leq k\}$  для всех  $r \in \mathbb{N}$ ,  $k = 1, \dots, \nu$ .

Обозначая через  $\mathbf{F}_{\mathbf{p}}(k; r, \nu)$  функцию распределения случайной величины  $K$  при вероятностях попадания в локальные оптимумы  $p_1, \dots, p_{\nu}$ ,  $\mathbf{p} = (p_1, \dots, p_{\nu})$ , указанный выше результат из [2] записывается как

**Утверждение 4.** При любых фиксированных  $\nu \in \mathbb{N}$ ,  $r \in \mathbb{N}$ ,  $k = 0, \dots, r$ , функция  $\mathbf{F}_{\mathbf{p}}(k; r, \nu)$  достигает минимума в точке  $\mathbf{p} = \bar{\mathbf{p}}$ .

Утверждение 4 позволяет вычислять консервативную нижнюю границу  $L'(r, k, \mathbf{p})$  для  $\nu$  с заданным доверительным уровнем  $p$  ввиду утверждения 2. Корректность такой нижней границы при условии справедливости утверждения 4 была отмечена нами в [11]. Аналогичная нижняя граница предложена Ч. Мао в [21]. Заметим, что  $\mathbf{F}_{\mathbf{p}}(r; \nu, r) = 1$  при любом  $\nu$ , поэтому при  $k = r$  нижняя оценка  $L'(r, k, \mathbf{p})$  не определена. Подобно доказательству следствия 1, из утверждения 4 получаем

**Следствие 2.**  $\mathbf{F}_{\bar{\mathbf{p}}}(k; r, \nu)$  является невозрастающей функцией от  $\nu$  при любых  $r \in \mathbb{N}$ ,  $k = 0, \dots, r$ .



Последнее следствие было независимо доказано в [11,23] без использования результата [2]. Это следствие даёт основание применению метода дихотомии для построения нижней границы  $L'(r, k, \mathbf{p})$  в соответствии с утверждением 2, а в изотропном случае – консервативного доверительного интервала в соответствии с утверждением 1.

## 2. Экспериментальное исследование

### 2.1. Используемые задачи

Изложенные выше методы оценивания числа локальных оптимумов были опробованы на ряде комбинаторных задач оптимизации. Среди них задачи с целевыми функциями, построенными случайным образом по модели  $\mathcal{NK}$ -ландшафтов С. Кауффмана [19], задачи отыскания двоичной последовательности с минимальной автокорреляцией [15], задачи о вершинном покрытии и задачи распределения бункерных ёмкостей в производственных линиях [9]. Перечисленные задачи формулируются следующим образом.

**Модель  $\mathcal{NK}$ -ландшафтов** предложена в [19] для генерации индивидуальных задач безусловной оптимизации с настраиваемой «изрезанностью ландшафта» целевой функции  $F : \{0, 1\}^n \rightarrow \mathbb{R}^+$  при размерности пространства решений  $n = \mathcal{N}$ . Значение этой функции в точке  $\mathbf{x} \in \{0, 1\}^n$  определяется как среднее арифметическое значений вспомогательных функций  $F_i(x_i, x_{i_1}, \dots, x_{i_k})$ ,  $i = 1, \dots, n$ , каждая из которых зависит от соответствующей компоненты  $x_i$ , а также от других  $\mathcal{K}$  компонент  $x_{i_1}, \dots, x_{i_k}$ . Наборы номеров  $\{i_1, \dots, i_k\}$  на этапе генерации исходных данных выбираются случайным образом среди  $\mathcal{K}$ -элементных подмножеств множества  $\{1, \dots, i-1, i+1, \dots, n\}$  с равномерным распределением. Значения функции  $F_i$ ,  $i = 1, \dots, n$ , для каждого из  $2^{\mathcal{K}+1}$  возможных значений её аргумента назначаются датчиком случайных чисел, выбираемых из интервала  $[0, 1)$ . Эти значения также входят в исходные данные задачи.

**Задача отыскания двоичной последовательности с минимальной автокорреляцией** состоит в отыскании вектора  $\mathbf{y} \in \{-1, +1\}^n$ , минимизирующего сумму  $\sum_{j=1}^{n-1} c_j^2$ , где  $c_j$  — величина автокорреляции для последовательности  $\mathbf{y}$  при смещении  $j$ , которая в данном случае имеет вид  $c_j = \sum_{i=1}^{n-j} y_i y_{i+j}$ . Индивидуальная задача отыскания двоичной последовательности с минимальной автокорреляцией полностью описывается числом  $n$ .

**Невзвешенная задача о вершинном покрытии (ЗВП).** Данная задача является NP-трудной в сильном смысле (см., например, [1]) и имеет следующую постановку. Пусть дан граф  $G = (V, E)$  с множеством вершин  $V$  и множеством рёбер  $E$ . Подмножество  $C \subseteq V$  называется *вершинным покрытием* в  $G$ , если каждое ребро инцидентно хотя бы одной вершине из  $C$ . Требуется найти вершинное покрытие  $C^*$  минимальной мощности.

Пусть в графе  $G = (V, E)$  множество вершин имеет вид  $V = \{v_1, \dots, v_n\}$ . Для представления ЗВП в форме задачи безусловной оптимизации на множестве  $\{0, 1\}^n$  предположим, что элемент  $x_j$ ,  $j = 1, \dots, n$ , равен 1 тогда и только тогда, когда вершина  $v_j$  входит в искомое множество. Обозначим через  $C(\mathbf{x})$  множество вершин, соответствующее вектору  $\mathbf{x}$ . Пусть целевая функция  $f(\mathbf{x}) = \sum_{j=1}^n x_j$  в случае, если  $C(\mathbf{x})$  — вершинное покрытие; иначе вводится штраф за нарушение ограничений задачи и значение целевой функции полагается равным  $f(\mathbf{x}) = \sum_{j=1}^n x_j + u(\mathbf{x}) \cdot |V|$ , где  $u(\mathbf{x})$  — число рёбер в  $E$ , которым не инцидентна ни одна вершина из  $C(\mathbf{x})$ .

Для экспериментов с ЗВП далее используются тестовые примеры для задачи о наибольшей клике из коллекции DIMACS [18], преобразованные в индивидуальные задачи о вершинном покрытии с помощью известного преобразования (см., например, [1]).

### **Задача распределения бункерных ёмкостей в производственной линии.**

При управлении производственными линиями, в которых детали перемещаются от одной единицы оборудования (ЕО) к другой с помощью некоторого транспортного механизма, возникает следующая задача оптимизации объёмов бункеров. Предполагается, что вследствие отказов оборудования в процессе работы линии возникают остановки ЕО, случайные по моменту возникновения и длительности. Последствия отказов распространяются на смежные операции из-за невозможности передать деталь на следующую операцию или отсутствия деталей на входе ЕО. Наличие бункеров (ёмкостей) для складирования деталей между ЕО позволяет снизить влияние отказов на соседние операции и повысить среднюю производительность линии. Однако установка бункеров связана с дополнительными капитальными затратами и увеличением числа складированных деталей. Задача состоит в выборе объёмов бункеров с учётом средней производительности линии  $V(\mathbf{h})$  в стационарном режиме, капитальных затрат на установку бункеров  $J(\mathbf{h})$  и средней стоимости хранения деталей  $Q(\mathbf{h})$  в стационарном режиме.

Любая ЕО может находиться в состояниях «работа», «отказ» (идёт восстановление ЕО после её поломки), «блокировка» (невозможно передать обработанную деталь на следующую операцию) и «простой» (отсутствие деталей на входе ЕО). Пусть  $n$  — количество бункеров,  $m$  — число ЕО в линии. Через  $\mathbb{Z}_+$  обозначим множество неотрицательных целых чисел. Введём вектор переменных  $\mathbf{h} = (h_1, \dots, h_n) \in \mathbb{Z}_+^n$ , где  $h_i$  — объём  $i$ -го бункера,  $i = 1, \dots, n$ . Структура линии представляется ориентированным последовательно-параллельным графом  $G$ , у которого вершины  $b_0, \dots, b_{n+1}$  отвечают бункерам, а дуги  $a_1, \dots, a_m$  — единицам оборудования. Каждая дуга направлена от вершины входного бункера соответствующей ЕО к вершине его выходного бункера. Возможны кратные дуги, т.е. параллельно работающие ЕО с общими бункерами. Единственная вершина  $b_0$ , не имеющая входящих дуг, соответствует входному бункеру линии, а единственная вершина  $b_{n+1}$  без выходящих дуг — выходному. Каждой вершине  $b_j$  приписано максимальное число вмещаемых деталей  $h_j$  в бункер  $j$ , причём  $h_0 = h_{n+1} = \infty$ . Предполагается, что на входе линии всегда имеется до-

статочное число деталей и обработанные детали всегда могут быть помещены в выходной бункер, поскольку он имеет неограниченный объём.

Каждая дуга  $a_i$  графа  $G$  характеризуется тройкой параметров  $(T_i^O, T_i^B, U_i) \in \mathbb{R}_+^3$ . Находясь в состоянии «работа», ЕО с номером  $i$  имеет постоянное время обработки детали  $U_i$ ,  $i = 1, \dots, m$ . Отказы и восстановления различных ЕО происходят независимо, и время наработки на отказ, также как и время восстановления ЕО, имеет экспоненциальное распределение. Для каждой ЕО  $T_i^O$  — среднее время наработки на отказ, а  $T_i^B$  — среднее время восстановления. В состояниях «блокировка» и «простой» ЕО не отказывает.

Обозначим среднее число деталей в стационарном режиме через  $q_j(\mathbf{h})$ . Тогда критерий максимизации имеет вид

$$f(V(\mathbf{h}), B(\mathbf{h}), Q(\mathbf{h})) = T_{am}S(V(\mathbf{h})) - J(B(\mathbf{h})) - T_{am} \sum_{j=1}^{n-1} c_j q_j(\mathbf{h}),$$

где  $T_{am}$  — период амортизации линии;  $S(V(\mathbf{h}))$  — доход от продажи готовой продукции за единицу времени;  $c_j$  — затраты на хранение одной детали в бункере  $j$  в единицу времени;  $J(B(\mathbf{h}))$  — капитальные затраты на установку бункеров. В качестве метода приближенного вычисления средней производительности линии  $V(\mathbf{h})$  и среднего числа деталей  $q_j(\mathbf{h})$  в каждом бункере  $j$  используется метод из [6], основанный на замене участков линии эквивалентными ЕО с применением марковских моделей для пар последовательных и параллельных ЕО.

## 2.2. Результаты экспериментов

В экспериментах с описанными выше задачами используется система окрестностей радиуса 1, порождённая метрикой Хэмминга. Для задач, построенных по модели  $\mathcal{NK}$ -ландшафтов и задач отыскания двоичной последовательности с минимальной автокорреляцией и задач распределения ёмкостей бункерных устройств выполнялось по 1000 независимых испытаний алгоритма локального поиска. Для большинства задач вершинного покрытия этот параметр был увеличен до 10000.

В некоторых задачах малой размерности удавалось осуществить полный перебор элементов пространства решений — в таких случаях построенные доверительные интервалы сравниваются с истинным значением  $\nu$ . В случаях, когда полный перебор оказывался чрезмерно трудоёмким, доверительные интервалы сравнивались с бутстреп-оценкой из [10, 29], показавшей хорошие результаты в экспериментах [10, 17]. Описание бутстреп-метода из [29] приводится в приложении. При расчётах бутстреп-методом число независимых реализаций  $B$  всюду полагалось равным 1000.

**Модель  $\mathcal{NK}$ -ландшафтов.** Эксперименты по оценке числа локальных оптимумов в задачах, построенных по модели  $\mathcal{NK}$ -ландшафта, проводились при

различной «изрезанности ландшафта», определяемой величиной  $K$  в диапазоне от 2 до 12, в то время как размерность  $n = 15$  оставалась неизменной. Рис. 2 показывает истинное число локальных оптимумов и двухсторонние 95%-доверительные интервалы, вычисленные методом переписи Шнабеля в предположении изотропности  $\mathbf{p} = \bar{\mathbf{p}}$  с использованием утверждения 1. Доверительные интервалы вокруг значений максимума правдоподобия  $\hat{\nu}(r, k)$  рассчитаны в предположении изотропности. Рисунок также содержит нижние границы  $L(t, 0.95)$  и  $L'(t, 0.95)$  доверительного уровня 95%.

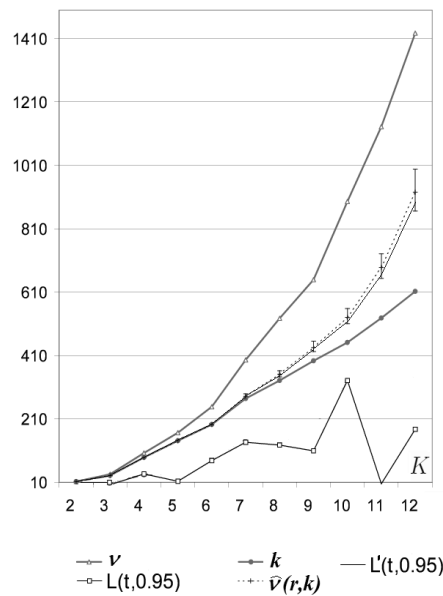


Рис. 2. Доверительные интервалы для числа локальных оптимумов в  $\mathcal{NK}$ -ландшафтах

График значений  $\nu$  на рис. 2 существенно отклоняется от оценки максимального правдоподобия по методу Шнабеля, что показывает неадекватность предположения об изотропности применительно к задачам данного класса. Нижние границы, основанные на ожидании первого повторения, оказываются даже ниже числа  $r$  обнаруженных оптимумов за  $k$  испытаний локального поиска.

**Задачи отыскания двоичной последовательности с минимальной автокорреляцией.** Рис. 3, подобно рис. 2, представляет результаты экспериментов с задачами отыскания двоичной последовательности с минимальной автокорреляцией. Длина последовательности  $n$  варьируется от 7 до 29. Для задач с  $n > 24$  полный перебор пространства решений оказался чрезмерно трудоёмким, поэтому точное значение величины  $\nu$  для задач большой размерности не изображено. Как видно из рисунка, бутстреп-метод даёт оценку наиболее близкую к истинному значению  $\nu$ .

Из приведённого графика видно, что при малой размерности задач ( $n < 12$ ) оценка максимального правдоподобия  $\hat{\nu}$ , основанная на методе переписи Шнабеля, даёт достаточно близкие к истинным значения. Далее с увеличением раз-

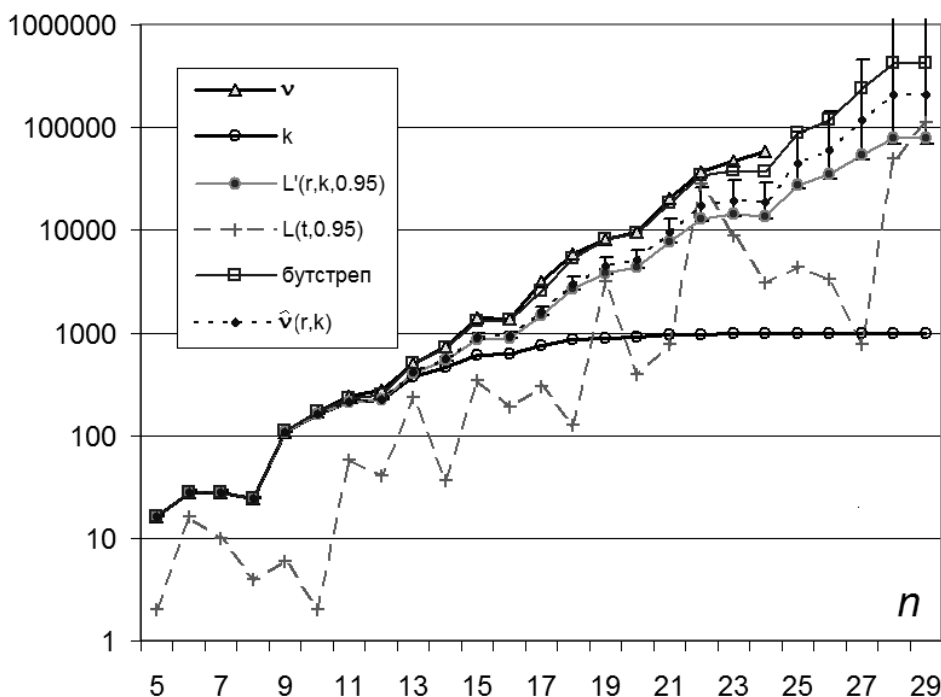


Рис. 3. Оценки числа локальных оптимумов в задачах отыскания двоичной последовательности с минимальной автокорреляцией

мерности задач предположение изотропности, по-видимому, становится неадекватным. Оценка  $L(t, 0.95)$ , в отличие от  $L'(r, k, 0.95)$ , ведёт себя очень нестабильно.

**Задачи о вершинном покрытии из коллекции DIMACS.** Результаты экспериментов по оцениванию числа локальных оптимумов для задач вершинного покрытия представлены в табл. 1. Здесь столбец «95% интервал» содержит 95% доверительные интервалы, построенные с использованием утверждения 1 в предположении изотропности.

Как выяснилось, задачи большой размерности в коллекции DIMACS характеризуются большими значениями величины  $\nu$ , в связи с этим для получения содержательных результатов оценивания числа локальных оптимумов рассматривались только графы с числом вершин не более 200. При  $r = 1000$  на всех тестовых примерах, кроме brock200-2 и brock200-4 не обнаруживалось ни одного повторения ранее найденных локальных оптимумов, а значит оценка, основанная на ожидании первого повторения, была не определена. Для других задач число испытаний  $r$  было увеличено до 10 000.

Задачи gn200p0.9-44, mapn a27, san200-0.9-1 и san200-0.9-3 не представлены в табл. 1 в связи с тем, что и при  $r = 10000$  вновь было получено  $k = r$ . В этих случаях нижняя граница может быть построена по формуле  $L(r + 1, 0.95) \approx 0.167r^2 = 1.67 \cdot 10^8$  (см. п. 1.1). Как видно из табл. 1, нижняя

Таблица 1. Оценки числа локальных оптимумов в ЗВП

Задача	$ V $	$r$	$k$	$L(t, 0.95)$	$L'(r, k, 0.95)$	бутстреп	95% интервал
john8-4-4	70	$10^4$	4509	572	5260	7145	5248-5398
brock200-2	200	$10^3$	972	346	13263	33761	12214-26212
mann a9	45	$10^4$	9406	7824	$7.6 \cdot 10^4$	$1.5 \cdot 10^5$	74808-87445
keller4	171	$10^4$	9805	$1.4 \cdot 10^5$	$2.3 \cdot 10^5$	$4.9 \cdot 10^5$	$2.2 \cdot 10^5$ - $2.9 \cdot 10^5$
brock200-4	200	$10^3$	999	$8.1 \cdot 10^5$	$1.1 \cdot 10^5$	$8.3 \cdot 10^5$	$9.0 \cdot 10^4$ - $2.0 \cdot 10^7$
john16-2-4	120	$10^4$	9939	$1.6 \cdot 10^5$	$6.6 \cdot 10^5$	$1.6 \cdot 10^6$	$6.4 \cdot 10^5$ - $1.1 \cdot 10^6$
san200-0.72	200	$10^4$	9945	$1.2 \cdot 10^6$	$7.2 \cdot 10^5$	$1.8 \cdot 10^6$	$7.0 \cdot 10^5$ - $1.2 \cdot 10^6$
san200-0.92	200	$10^4$	9995	$4.1 \cdot 10^6$	$4.8 \cdot 10^6$	$1.2 \cdot 10^7$	$4.3 \cdot 10^6$ - $3.1 \cdot 10^7$
c125.9	125	$10^4$	9999	$1.5 \cdot 10^7$	$1.1 \cdot 10^7$	$10^8$	$9.0 \cdot 10^6$ - $1.1 \cdot 10^9$
gn200p0.9-55	200	$10^4$	9999	$1.1 \cdot 10^6$	$1.1 \cdot 10^7$	$10^8$	$9.0 \cdot 10^6$ - $1.1 \cdot 10^9$

оценка, основанная на времени первого повторения, и в данной серии даёт неустойчивые результаты, в ряде случаев оказываясь ближе к оценке бутстрепа чем нижняя оценка, построенная по методу переписи Шнабеля (brock200-4, san200-0.72, c125.9).

Полученные данные показывают, что число локальных оптимумов в ЗВП близкой размерности может существенно различаться и зависит от структуры графа. Доверительный интервал, построенный по методу переписи Шнабеля в предположении изотропности, во многих случаях не содержит значения бутстреп-оценки. Таким образом, предположение изотропности в задачах о вершинном покрытии не оправдывается.

**Задача распределения ёмкостей бункеров.** Результаты оценивания числа локальных оптимумов в задачах распределения ёмкостей бункеров, созданных на основе реальных данных технологических линий автозавода Renault [6], показаны на рис. 4. Число промежуточных бункеров  $n$  и мощность множества допустимых решений в этих задачах приведена в табл. 2. Среди тестовых примеров 4 линии имеют последовательную структуру, а 2 линии AS7, AS8 имеют последовательно-параллельную структуру из 10 ЕО. Полный перебор пространства решений во всех задачах кроме AS3 оказался чрезмерно трудоёмким из-за значительного времени вычисления целевой функции. В связи с этим точное число локальных оптимумов для этих задач не вычислялось.

Как видно из рисунка, на этих задачах нижняя граница  $L'(r, k, 0.95)$ , соответствующая равновероятному попаданию в локальные оптимумы, оказывается близка к бутстреп-оценке. Можно предположить, что структура локальных оптимумов задачи распределения ёмкостей бункерных устройств близка к изотропной. Причиной тому может служить симметричность «ландшафта» целевой функции в этих задачах, отмеченная в работе [5].

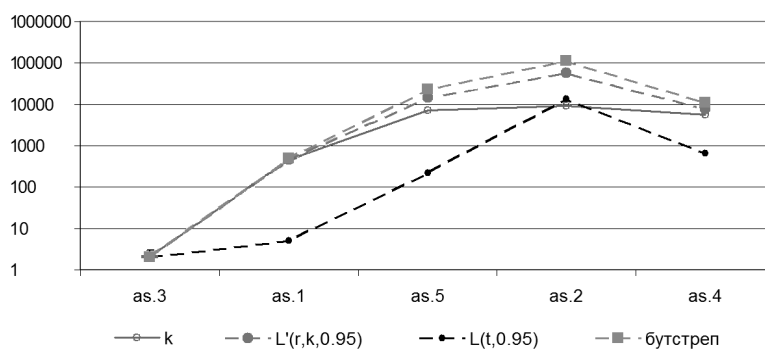


Рис. 4. Оценки числа локальных оптимумов в задачах распределения ёмкостей бункеров

Таблица 2. Тестовые задачи распределения ёмкостей бункеров

	$n$	допустимых решений
AS1	4	$4 \times 10^5$
AS2	9	$9 \times 10^{11}$
AS4	3	$4 \times 10^3$
AS6	13	$2 \times 10^{22}$
AS7	7	$9 \times 10^7$
AS8	7	$5 \times 10^8$

### 3. Заключение

Проведённые вычислительные эксперименты на известных тестовых примерах средней размерности показывают адекватность предложенных нижних оценок и бутстреп-метода для приближенного оценивания числа локальных оптимумов. В то же время построенные двусторонние доверительные интервалы зачастую не покрывают истинное значение оцениваемого параметра, что говорит о неприменимости к рассматриваемым задачам предположения о равновероятном попадании в локальные оптимумы алгоритма локального поиска.

Для оценивания числа локальных оптимумов снизу в большинстве случаев предпочтительнее использовать нижнюю границу, основанную на методе переписи Шнабеля, т.к. нижняя граница, основанная на времени первого повторения локальных оптимумов, даёт нестабильные результаты. Тем не менее последняя оценка может оказаться единственным приемлемым методом, когда локальные оптимумы повторяются чрезвычайно редко и априори сложно выбрать число испытаний, при котором будет получено хотя бы одно повторение ранее полученного локального оптимума.

Дальнейшие исследования могут быть направлены на использование предложенных методов в правилах остановки для эвристик локального поиска с

многократным перезапуском. В частности, необходимое условие для остановки алгоритма может формулироваться как условие превышения числа найденных локальных оптимумов над нижней границей заданного доверительного уровня. Также предложенные методы могут использоваться для классификации тестовых примеров по их сложности для алгоритмов, основанных на локальном поиске. Методы оценивания числа локальных оптимумов могут также быть использованы в биоинформатике (см., например, [31]).

## Благодарности

Работа выполнена при поддержке Программы фундаментальных научных исследований государственных академий наук на 2013-2020 годы, п. I.5.1.5. «Исследование и решение задач комбинаторной оптимизации с использованием целочисленного программирования». Авторы благодарны С.А. Клокову за помощь в доказательстве утверждения 3.

## Приложение

Настоящее приложение содержит описание бутстреп-метода для оценки числа локальных оптимумов, предложенного в [10, 29]. Бутстреп-метод предложен Б. Эфроном для решения целого ряда задач статистического анализа [3], и в частности, для непараметрической оценки смещения. Пусть  $X_1, \dots, X_n$  – независимые одинаково распределённые случайные величины с неизвестным распределением  $F$  на некотором пространстве  $\mathcal{X}$ . Предположим, что на основе наблюдаемых значений

$$X_1 = x_1, \dots, X_n = x_n$$

может быть вычислена оценка  $\hat{\theta}(x_1, \dots, x_n)$  некоторого интересующего нас функционального параметра распределения  $\theta(F) \in \mathbb{R}$  (например, математического ожидания, дисперсии или  $|\text{supp } F|$ , как в случае оценивания числа локальных оптимумов или особей в популяции). Интерес представляет распространённый случай, когда оценка имеет вид

$$\hat{\theta}(x_1, \dots, x_n) = \theta(\hat{F}),$$

где  $\hat{F}$  – эмпирическое распределение, приписывающее массу  $\frac{1}{n}$  каждой из точек имеющейся выборки  $x_1, \dots, x_n$ . Требуется оценить величину смещения оценки  $\hat{\theta} = \theta(\hat{F})$ :

$$\text{bias} \equiv \mathbf{E}_F[\theta(\hat{F}) - \theta(F)], \quad (5)$$

где  $\mathbf{E}_F$  обозначает математическое ожидание в предположении, что  $X_1, \dots, X_n$  – независимые одинаково распределённые случайные величины с распределением  $F$ .

В бутстреп-методе для оценки смещения  $\text{bias}$  при фиксированном  $\hat{F}$  рассматривается случайная выборка объема  $n$ :

$$X_1^* = x_1^*, \dots, X_n^* = x_n^*,$$



где  $X_1^*, \dots, X_n^*$  — независимые одинаково распределённые случайные величины с вероятностным распределением  $\hat{F}$ . Выборочное распределение интересующей нас случайной величины  $\hat{\theta}(\hat{F}) - \theta(F)$  аппроксимируется *бутстреп-распределением* случайной величины  $\hat{\theta}(X_1^*, \dots, X_n^*) - \theta(\hat{F})$ . Для получения бутстреп-оценки смещения  $\text{BIAS}_*$  в формуле (5) производится замена неизвестного распределения случайной величины  $\hat{\theta}(\hat{F}) - \theta(F)$  имеющимся бутстреп-распределением:

$$\text{BIAS}_* = \mathbf{E}_*[\hat{\theta}(X_1^*, \dots, X_n^*) - \theta(\hat{F})],$$

где  $\mathbf{E}_*$  обозначает математическое ожидание в предположении, что  $X_1^*, \dots, X_n^*$  — независимые одинаково распределённые случайные величины с вероятностным распределением  $\hat{F}$ , зафиксированным на его наблюдаемом значении.

Величину  $\text{BIAS}_*$  в некоторых случаях удаётся вычислить точно или приближённо теоретическими методами (см., например, гл. 1 в [3]), однако, наиболее часто для этого используется аппроксимация бутстреп-распределения методом Монте-Карло, где генерируется достаточно большое число  $B$  независимых реализаций  $x_{1b}^*, \dots, x_{nb}^*$ ,  $b = 1, \dots, B$ , случайной выборки  $X_1^*, \dots, X_n^*$  из распределения  $\hat{F}$ . В таком случае бутстреп-оценка смещения аппроксимируется величиной

$$\widehat{\text{BIAS}}_* = \frac{1}{B} \sum_{b=1}^B \hat{\theta}(x_{1b}^*, \dots, x_{nb}^*) - \theta(\hat{F}).$$

### Алгоритм 1. Бутстреп-оценка смещения

1. Построить эмпирическое распределение  $\hat{F}$ , приписывающее массу  $\frac{1}{n}$  каждой из точек имеющейся выборки  $x_1, \dots, x_n$ .

2. Для  $b = 1, \dots, B$  независимо выполнять:

2.1. Построить бутстреп-выборку из  $\hat{F}$ :

$$X_1^* = x_{1b}^*, \dots, X_n^* = x_{nb}^*.$$

2.2. Вычислить  $\hat{\theta}_b^* = \hat{\theta}(x_{1b}^*, \dots, x_{nb}^*)$ .

3. Вычислить бутстреп-оценку смещения

$$\widehat{\text{BIAS}}_* = \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta}).$$

4. Конец работы.

С использованием бутстреп-оценки смещения, найденной алгоритмом 1, может быть получена скорректированная бутстреп-оценка:

$$\hat{\theta}_* = \hat{\theta} - \widehat{\text{BIAS}}_*. \tag{6}$$

Применительно к задаче оценки числа локальных оптимумов описанный алгоритм позволяет оценивать смещение статистики  $\hat{\nu}(r, k)$ , вычисленной по

методу максимального правдоподобия. В данном случае  $\theta = \nu$ ,  $\mathcal{X} = \{1, \dots, \nu\}$  – множество всех локальных оптимумов, распределение  $F$  описывает случайную величину, равную номеру локального оптимума, получаемого алгоритмом локального спуска из случайно выбранной начальной точки. На шаге 1 алгоритма 1 по  $n = r$  испытаниям локального спуска строится эмпирическое распределение  $\hat{F} = (v_1/r, \dots, v_k/r)$ , где  $v_j$  – число испытаний, при которых был получен локальный оптимум с номером  $j$ ,  $j = 1, \dots, k$ .

На шаге 2.2. полагаем  $\hat{\nu}_b^* = \hat{\nu}(r, k_b^*)$ , где  $k_b^*$  – число различных значений в бутстреп-выборке  $(j_1^*, \dots, j_r^*)$  на итерации  $b$ . На шаге 3  $\hat{\nu}_b^* = \hat{\theta}_b^*$ ,  $b = 1, \dots, B$  и оценка смещения вычисляется по формуле

$$\widehat{\text{BIAS}}_* = \sum_{b=1}^B (\hat{\nu}_b^* - \nu(r, k)) / B.$$

При нарушении условия изотропности локальные оптимумы, куда вероятность попадания локального спуска мала, имеют мало шансов быть представленными в эмпирическом распределении  $\hat{F}$ . В таком случае  $\hat{F}$  оказывается ближе к равномерному распределению по сравнению с истинным распределением вероятностей  $F$ . В результате скорректированная бутстреп-оценка (6) отклоняется в сторону оценки максимального правдоподобия  $\hat{\nu}(r, k)$ . Приведённые в [10] результаты эксперимента подтверждают это.

## ЛИТЕРАТУРА

1. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. М. : Мир, 1982.
2. Зубков А.М., Попов Н.Н. Отношение частичного порядка, порождённое распределениями числа занятых ячеек // Матем. заметки. 1982. Т. 32, № 1. С. 97–102.
3. Эфрон Б., Тибширани Г.Г.Р. Нетрадиционные методы многомерного статистического анализа. М. : Финансы и статистика, 1988.
4. Севастьянов Б.А. Курс теории вероятностей и математической статистики. М. : Наука, 1982.
5. Сигаев В.С. О свойствах локальных оптимумов задачи размещения буферных накопителей // Вестник Омского университета. 2006. № 4. С. 1–3.
6. Ancelin B., Semery A. Calcul de la productivité d'une ligne intégrée de fabrication // RAIRO Automat., Productiq. Informatiq. Industrielle. 1987. V. 21. P. 209–238.
7. Craig C.C. Use of marked specimens in estimating populations // Biometrika. 1953. V. 40, N. 1–2. P. 170–176.
8. Darroch J.N. The multiple-recapture census. I: estimation of a closed population // Biometrika. 1958. V. 45, N. 3–4. P. 343–359.
9. Dolgui A., Eremeev A., Kolokolov A., Sigaev V. A genetic algorithm for the allocation of buffer storage capacities in a production line with unreliable machines // Journal of Mathematical Modelling and Algorithms. 2002. V. 1, N. 2. P. 89–104.
10. Eremeev A.V., Reeves C.R. Non-parametric estimation of properties of combinatorial landscapes // Applications of Evolutionary Computing: Proceedings of EvoWorkshops

- 2002 / Cagnoni S., Gottlieb J., Hart E., Middendorf M., Raidl G. Lecture Notes in Computer Science, Vol. 2279. Berlin : Springer-Verlag, 2002. P. 31–40.
11. Eremeev A.V., Reeves C.R. On confidence intervals for the number of local optima // Applications of Evolutionary Computing, EvoWorkshop 2003: EvoBIO, EvoCOP, EvoIASP, EvoMUSART, EvoROB, and EvoSTIM / Raidl G.R., Meyer J.-A., Middendorf M., Cagnoni S., Cardalda J.J.R., Corne D., Gottlieb J., Guillot A., Hart E., Johnson C.G., Marchiori E. Lecture Notes in Computer Science, Vol. 2611. Heidelberg : Springer, 2003. P. 224–235.
  12. Garnier J., Kallel L. How to detect all maxima of a function? // Proceedings of the Second EVONET Summer School on Theoretical Aspects of Evolutionary Computing (Anvers, 1999). Berlin : Springer, 2001. P. 343–370.
  13. Garthwait P.H., Buckland S.T. Analysis of multiple recapture census by computing conditional probabilities // Biometrics. 1990. V. 46, N. 1. P. 231–238.
  14. Glover F. and Laguna M. Tabu Search. Norwell, MA: Kluwer, 1997.
  15. Golay M.J.E. Series for low-autocorrelation binary sequences. IEEE Trans. Inform. Theory. 1977. V. 23. P. 43–51.
  16. Hansen P., Mladenovic N. Variable neighborhood search: Principles and applications // European Journal of Operational Research. 2001. V. 130. P. 449–467.
  17. Hernando L., Mendiburu A., Lozano J.A. An evaluation of methods for estimating the number of local optima in combinatorial optimization problems // Evol. Comput. 2013. V. 21, N. 4. P. 625–658.
  18. Johnson D.S., Trick M.A. Introduction to the second DIMACS challenge: cliques, coloring, and satisfiability // DIMACS Series in Discrete Math. and Theoretical Comput. Sci. / Johnson D., Trick. M. Providence, RI : American Math. Soc., 1996. Vol. 26. P. 1–10.
  19. Kauffman S.A. Adaptation on rugged fitness landscapes // Lectures in the Sciences of Complexity, Vol. I of SFI Studies. Addison-Wesley, 1989. P. 619–712.
  20. Kochetov Yu., Mikhailova A., Plyasunov A. A Genetic Local Search Algorithm for the Graph Partitioning Problem with Cardinality Constraints // Preprints of the 13th IFAC Symposium on Information Control Problems in Manufacturing (INCOM09), Moscow, Russia, June 3–5, 2009. P. 1991–1996.
  21. Mao C.X. On the nonidentifiability of population sizes // Biometrics. 2008. V. 64. P. 977–981.
  22. Mood A.M., Graybill F.A., Boes D.C. Introduction to the theory of statistics. Third Ed. New York : McGraw-Hill, 1973.
  23. Pickands J., Raghavachari M. Exact and asymptotic inference for the size of a population // Biometrika. 1987. V. 74, N. 2. P. 355–363.
  24. Pledger S. The performance of mixture models in heterogeneous closed population capture-recapture // Biometrics. 2005. V. 61, N. 3. P. 868–876.
  25. Reeves C. Genetic algorithm for the operations researcher // INFORMS Journal on Computing. 1997. V. 9, N. 3. P. 231–250.
  26. Reeves C.R. The "Crossover Landscape" and the Hamming Landscape for Binary Search Spaces // Foundations of Genetic Algorithms 7 / De Jong K., Poli R., Rowe J. San Francisco, CA: Morgan Kaufmann. 2003. P. 81–98.
  27. Reeves C.R. Estimating the Number of Optima in a Landscape, Part I: Statistical Principles. Coventry University Technical Report SOR#01-03. 2001.

28. Reeves C.R. Estimating the Number of Optima in a Landscape, Part II: Experimental Investigations. Coventry University Technical Report SOR#01-04. 2001.
29. Reeves C.R., Eremeev A.V. Statistical analysis of local search landscapes // J. Oper. Res. Soc. 2004. V. 55, N. 7. P. 687–693.
30. Resende M.G.C. and Ribeiro C.C. Greedy randomized adaptive search procedures // Handbook of metaheuristics / Glover F. and Kochenberger G. N.Y.: Kluwer Academic Publishers, 2003. P. 219-249.
31. Sahoo S., Albrecht A.A. Approximating the set of local minima in partial RNA folding landscapes // Bioinformatics. 2012. V. 28, Issue 4. P. 523–530.
32. Schnabel Z.E. The estimation of the total fish population of a lake // American Math. Monthly. 1938. V. 45. P. 348–352.
33. Seber G.A.F. The estimation of animal abundance. London : Charles Griffin, 1973.
34. Yagiura M., Kishida M., Ibaraki. T. A 3-flip neighborhood local search for the set covering problem // Eur. J. Oper. Res. 2006. V. 172. P. 472–499.

## ON CONFIDENCE INTERVALS FOR THE NUMBER OF LOCAL OPTIMA

**A.V. Eremeev**<sup>1</sup>

Associate Professor, Dr. Sci. (Phys.-Math.), e-mail: eremeev@ofim.oscsbras.ru

**C.R. Reeves**<sup>2</sup>

Professor Emeritus, e-mail: c.reeves@coventry.ac.uk

<sup>1</sup>Omsk Branch of Sobolev Institute of Mathematics SB RAS

<sup>2</sup>Coventry University

**Abstract.** The number of local optima is an important indicator of optimization problem difficulty for local search algorithms. Here we will discuss some methods of finding the confidence intervals for this parameter in problems where the large cardinality of the search space does not allow exhaustive investigation of solutions. Computational results are reported that were obtained by using these methods for NK landscapes model of S.A. Kauffman, for the low autocorrelation binary sequence, for buffer allocation problems in production line, and vertex cover problems.

**Keywords:** local search, combinatorial optimization problem, Schnabel census, conservative confidence interval.

*Дата поступления в редакцию: 02.02.2017*