ЛИТЕРАТУРА

1. Абдулаев Ш.-С. О., Беседин Б. А. О синтезе оптимальных фильтрующих и сглаживающих информационно-измерительных систем.— «Автометрия», 1974, № 2, с. 10—18.
2. Борисов В. И. Проблемы векторной оптимизации.— В кн.: Исследование операций. Методологические аспекты. М., «Наука», 1972, с. 72—91.
3. Kalman R. E., Bucy R. S. New results in linear filtering and prediction theory.— "Trans. ASME, J. Basic Engrg.", 1961, vol. D83, N 1, p. 95—107.
4. Kailath T. An innovation approach to least — squares estimations. Part I: Linear-filtering in additive white noise.— "Trans. IEEE", vol. AC-13, N 6, p. 646—655.
5. Аоки М. Оптимизация стохастических систем. М., «Наука», 1971.
6. Поддубный В. В., Якупов Р. Т. Об оптимизации измерительного комплекса для динамических систем.— «Труды СФТИ», 1974, вып. 60, ч. 2, с. 180—185.
7. Беседин Б. А. Об оптимальных структурах информационно-измерительных систем.— Тезисы докладов III Всесоюзной конференции по проблемам теоретической кибернетики. Новосибирск, Изд. ИМ СО АН СССР, 1974, с. 55—56.
8. Демьянов В. Ф. К решению целочисленных задач выпуклого программирования.— В кн.: Оптимальные системы автоматического управления. М., «Наука», 1967, с. 14—26.
9. Червак Ю. Ю. Метод лексикографического поиска решения для дискретных задач выпуклого программирования.— «Укр. матем. журн.», 1974, т. 26, № 2, с. 269—272.
10. Bank B., Kleinmann P. Note on the reduction of nonlinear integer programming to a sequence of linear integer optimization problems.— "Ekonomicko-matematcky Obzor", 1975, vol. 11, N 2, p. 162—164.
11. Коробкова З. В. Программа для решения целочисленной задачи выбора вариантов.— В кн.: Вопросы численного решения оптимизационных экономико-математических задач. Новосибирск, Изд. ИЭ и ОПП СО АН СССР, 1973, с. 17—40.
12. Болдырев В. И., Хохлюк В. И. Три процедуры симплекс-метода.— В кн.: Вычислительные системы. Вып. 59. Новосибирск, «Наука», 1974, с. 119—138.

Поступила в редакцию 27 апреля 1976 г.;
окончательный вариант — 10 сентября 1976 г.

УДК 62-506 : 519.8

Н. Д. ВАСИЛЬЕВ, В. Г. ИВКИН, И. В. МОЗИН, В. А. ШЕЛЕХОВ

(Ленинград)

СРАВНИТЕЛЬНАЯ ОЦЕНКА ЭФФЕКТИВНОСТИ НЕКОТОРЫХ МЕТОДОВ МНОГОПАРАМЕТРИЧЕСКОЙ ОПТИМИЗАЦИИ

Задача сравнения методов поиска экстремума возникла в связи с построением системы автоматической оптимизации режима работы циклического ускорителя заряженных частиц — протонного синхротрона ИТЭФ. Методы поиска, сравнительная оценка которых дается в настоящей работе, — последовательное симплексное планирование (ПСП) [1] и случайный поиск [2] в модификациях с наказанием случайностью (СПНС), с оценкой градиента (СПОГ) и с самообучением (СПС) — были признаны ранее [3] наиболее пригодными для автоматической оптимизации ускорителя.

С точки зрения теории управления поиск максимальной интенсивности пучка ускоренных протонов классифицируется как многопараметрическая оптимизация в условиях высокого уровня шумов и неизвестного вида поверхности критерия оптимальности. Поэтому важно знать характеристики эффективности априорно выбранных методов оптимизации, полученные в условиях, близких к существующим на объекте.

Характерной особенностью настоящей работы отличающей ее от

$$I(X)=E_{\xi}\{Q(X,\ \xi)\},$$

где $X=(X_1,\ X_2,\ \ldots,\ X_n)$ — вектор управляемых параметров, подлежащих определению. Измеряемым в процессе оптимизации считается функционал

$$Q(X,\ \xi)=I(X)+\xi$$

(ξ — случайная величина, распределенная нормально, с нулевым математическим ожиданием и дисперсией σ^2). Экстремум функции $I\ (X)$ определяется в заданной области

$$X_i^{\min}\leqslant X_i\leqslant X_i^{\max},\quad i=1,2,\ldots,n.$$

Пусть измерения функции качества и измерения вектора X производятся лишь в дискретные моменты времени $t^j=j\Delta t$ ($j=0,\ 1,\ 2,\ \ldots$); значения вектора управляемых параметров X на каждом j-м шаге поиска обозначаются через X^j; $I^j=I(X^j)$; $Q^j=I(X^j)+\xi^j$. Способ определения величины шага $\Delta X^{j+1}=X^{j+1}-X^j$ задается алгоритмом поиска.

Кратко опишем сравниваемые методы.

Случайный поиск с наказанием случайностью. Очередной шаг в пространстве управляемых параметров определяется по правилу

$$\Delta X^{j+1}=\begin{cases}\Delta X^j, & \text{если}\quad Q^j>Q^{j-1};\\ a\Xi^{j+1}, & \text{если}\quad (Q^j\leqslant Q^{j-1})\vee(X_i^{\max}\leqslant X_i^{j-1})\vee(X_i^{\min}\geqslant X_i^{j+1}),\end{cases}$$

где $\Xi=(\xi_1,\ \xi_2,\ \ldots,\ \xi_n)$ — случайный вектор, равномерно распределенный на сфере единичного радиуса; a — масштаб шага.

Масштаб шага адаптируется в процессе поиска по зависимости

$$a=A_0\exp\{-10^{-L}[N^2(1+\Delta n^2/T^2)+$$
$$+(\sup N)^2(1-\Delta n^2/T^2)]\},$$

где A_0 — начальный масштаб шага; N — число неудачных шагов, совершенных из последней точки; $\sup N$ — наибольшее число неудачных шагов, совершенных из какой-либо точки за весь процесс поиска; Δn — разность между номером текущей точки поиска и номером точки, где было совершено $\sup N$ неудачных шагов; T — максимально допустимое удаление текущей точки поиска от точки, из которой было совершено $\sup N$ неудачных шагов; L — параметр, определяющий динамику адаптации шага.

Случайный поиск с оценкой градиента. Алгоритм поиска имеет вид

$$\Delta X^{j+1}=\begin{cases}\Delta X^j, & \text{если}\quad Q^j>Q^{j-1};\\ a^{j+1}\dfrac{S^j}{|S^j|}, & \text{если}\quad Q^j\leqslant Q^{j-1},\end{cases}$$

где a — масштаб рабочего шага; S^j — стохастическая оценка градиента, определяемая по следующему алгоритму:

$$S^j=\frac{1}{m2g^j}\sum_{\psi=1}^{m}[Q(X^j+g^j\Xi^{\psi})-Q(X^j-g^j\Xi^{\psi})]\,\Xi^{\psi},$$

где $m\leqslant n$ — число пар проб для оценки градиента; g^j — величина рабочего шага; Ξ^{ψ} — вектор, равномерно распределенный на сфере единичного радиуса.

Масштабы рабочих и пробных шагов адаптируются в процессе поиска по следующим зависимостям:

$$a^{j+1} = \begin{cases} a^j, & \text{если} \quad Q^j > Q^{j-1}; \\ a^j/N + 1, & \text{если} \quad Q^j \leqslant Q^{j-1}; \end{cases}$$

$$g^j = \begin{cases} g^{j-1}, & \text{если} \quad Q^j > Q^{j-1}; \\ g^j/\sqrt[4]{N+1}, & \text{если} \quad Q^j \leqslant Q^{j-1} \end{cases}$$

(N — число неудачных шагов за весь предыдущий процесс поиска).

Кроме того, по сравнению с [1] в алгоритм включены следующие добавки, повышающие его гибкость и эффективность:

а) Если в процессе оценки градиента выявляются положительные приращения функции качества, то совершается переход в пробную точку, т. е.

$$X^j := \begin{cases} X^j + g^j\Xi^{\psi}, & \text{если} \quad \Delta Q^{j,\psi} > 0; \\ X^j, & \text{если} \quad \Delta Q^{j,\psi} \leqslant 0. \end{cases}$$

б) После совершения шага вдоль оценки градиента итоговое состояние определяется максимальным приростом функции качества при пробных и рабочем шагах, т. е.

$$X^j := X^j + \Delta X^{j+1} = \begin{cases} X^j + a^{j+1}\dfrac{S^j}{|S^j|}, & \text{если } \Delta Q^{j+1} = \max\left(\Delta Q^{j+1}, \Delta Q^{j,\psi}\right); \\ X^j + g^j\Xi^{\psi}, & \text{если } \Delta Q^{j,\psi} = \max\left(\Delta Q^{j+1}, \Delta Q^{j,\psi}\right). \end{cases}$$

Случайный поиск с самообучением. Правило расчета шага в этом алгоритме следующее:

$$\Delta X^{j+1} = (W^{j+1} + R\Xi^{j+1})/|W^{j+1} + R\Xi^{j+1}|,$$

где $R = \text{const} > 0$ — радиус направляющей сферы; Ξ — вектор, равномерно распределенный на сфере единичного радиуса $W^j = (W_1^j, W_2^j, \ldots, W_n^j)$ — вектор памяти, $|W^j| \leqslant C$, $C = \text{const} > 0$; $W^0 = C \operatorname{grad} I(X^0)$; $W^{j+1} = lW^j - \delta\Delta Q^j \Delta X^j$, где $0 \leqslant l \leqslant 1$ — параметр забывания; $\delta \geqslant 0$ — параметр самообучения; $\Delta Q^j = Q^j - Q^{j-1}$ — приращение функционала на j-м шаге.

Последовательное симплексное планирование. Сущность оптимизации этим методом состоит в следующем.

В пространстве управляемых переменных строится регулярный симплекс с центром в точке старта и оценивается функция качества во всех его вершинах: $Q(X^j)$, $j = 0, 1, \ldots, n$. Затем совершается пробный шаг — зеркальное отражение худшей вершины, в которой функция качества минимальна, через центр противоположной грани

$$X^{\text{отр}} = 2X^{\text{ц}} - X^{\text{х}},$$

где $X^{\text{отр}}$ — вектор координат отраженной вершины; $X^{\text{ц}}$ — вектор координат центра грани; $X^{\text{х}}$ — вектор координат худшей вершины.

За пробным следует рабочий шаг в точку, определяемую по правилу:

$$X^{\text{р}} = \begin{cases} (1+\gamma)\,X^{\text{отр}} - \gamma X^{\text{ц}}, & \text{если } Q\left(X^{\text{отр}}\right) > Q\left(X^{\text{л}}\right) — \text{растяжение}; \\ \beta X^{\text{х}} + (1-\beta)\,X^{\text{ц}}, & \text{если } Q\left(X^{\text{отр}}\right) < Q\left(X^{\text{х}}\right) — \text{сжатие}; \\ X^{\text{отр}}, & \text{если } Q\left(X^j\right) < Q\left(X_{\text{отр}}\right) < Q\left(X^{\text{л}}\right),\ j = x. \end{cases}$$

(Если $Q(X^{\text{х}}) < Q(X^{\text{отр}}) < Q(X^j)$, $j \neq x$, то $X^{\text{х}}$ заменяется на $X^{\text{отр}}$ и производится сжатие.) Здесь приняты следующие обозначения: $X^{\text{р}}$ — вектор координат конца рабочего шага; $X^л$ — вектор координат лучшей вершины; $\gamma>0$ — коэффициент растяжения; $0<\beta\leqslant 1$ — коэффициент сжатия.

В зависимости от результата рабочего шага производится замена худшей вершины по правилу:

$$X^x:=\begin{cases} X^p, & \text{если } Q(X^p)>Q(X^{отр}); \\ X^{отр}, & \text{если } Q(X^{отр})>Q(X^p); \end{cases} \text{ при растяжении;}$$
$$\quad X^p, \quad \text{если } Q(X^p)>Q(X^x) \text{ при сжатии.}$$

Если после сжатия $Q(X^p)\leqslant Q(X^x)$, то исходный симплекс стягивается к лучшей вершине:

$$X^j:=0{,}5(X^j+X^л),\quad j=0,\ 1,\ \ldots,\ n.$$

Далее итерационный процесс повторяется с момента определения худшей вершины.

Исследование описанных методов проводилось на тестовой нелинейной функции в двух модификациях:

1) сепарабельная квадратичная форма, осложненная шумом:

$$Q_1(X,\ \xi)=X^тBX+B_0+\xi,$$

где B — диагональная матрица коэффициентов размерностью 11×11, причем $B_{ii}=-0{,}9$, $i=1,\ 2,\ \ldots,\ 11$; $B_{ij}=0$, $i\neq j$; $B_0=1{,}6$ — свободный член; ξ — шумовая добавка, распределенная псевдонормально, с дисперсией $\sigma^2\left(\frac{\sigma}{Q_{\max}}100=3\%\right)$;

2) несепарабельная функция с рельефом типа «двумерный хребет»:

$$Q_2(X,\xi)=\sum_{i=1}^{9}B_{ii}X_i^2-100\left(X_{11}-X_{10}^2\right)^2-(1-X_{10})^2+B_0+\xi.$$

При сопоставлении различных методов поиска, как правило, придерживаются общепринятой точки зрения на необходимые условия такого сравнения [2]. Корректность постановки вопроса о равенстве условий сравнения методов при заданном виде модели заключается в том, чтобы были тождественны начальные условия для всех выбранных методов и оптимальны параметры каждого метода с точки зрения заданной характеристики-быстродействия, определяемого числом замеров функции качества до достижения заданной зоны экстремума.

Все выбранные методы предварительно оптимизированы по параметрам. Начальными условиями (точкой старта) для всех восхождений принят вектор

$$X^0=\{X_i^0\},\quad i=1,2,\ldots,11;\quad X_i^0=-0{,}9.$$

Аналитический анализ характеристик методов для принятых функций качества крайне сложен, поэтому исследование проводилось эмпирически путем моделирования процесса оптимизации на ЭВМ БЭСМ-4. Причем результаты осреднялись по десяти восхождениям для детерминированного метода (симплексного) и по пятидесяти восхождениям для разновидностей случайного поиска.

Характеристики сравнения методов. Все характеристики делятся на два класса: локальные и интегральные. Первый класс связан с одним элементарным этапом поиска — рабочим шагом, а второй — со всем процессом оптимизации от момента старта до срабатывания правила остановки. В настоящей работе получены лишь основные характеристики методов, наилучшим образом отражающие свойства применительно к автоматической оптимизации ускорителя.

Потери на поиск. Потери определяют среднюю локальную скорость оптимизации и вычисляются по следующей формуле:

$$\Pi^j = E(K^j)/E(\delta Q^j),$$

где K^j — число проб (замеров функции качества) на j-м шаге; $\delta Q^j = \Delta Q^j / Q^{j-1}$ — относительное изменение показателя качества.

Вероятность ошибки. Такая вероятность также относится к классу локальных характеристик и определяет вероятность появления ошибочного рабочего шага, который считается ошибочным, если значение функции качества не увеличивается после совершения этого шага. Таким образом, вероятность ошибки определяется выражением

$$P = P\{Q(X^j + \Delta X^j) \leqslant Q(X^j)\}.$$

Предполагается, что ошибочное решение принимается либо за счет неточностей в оценке направления движения, либо за счет действия случайной помехи, не позволяющей точно определить значение функции качества.

Неточность. Характеристика неточности метода является интегральной. Она определяет невязку ε полученного и искомого решений

$$E(\varepsilon) = E|X^* - X^{ext}|,$$

где X^* — решение, полученное после окончания работы метода; X^{ext} — решение, соответствующее экстремуму функции качества.

Число замеров функции качества для достижения заданного уровня. Эта характеристика зависит от начальных условий X^0. Поэтому для удобства сопоставления методов следует выбрать определенные начальные условия или осреднить результат по всем начальным условиям. Первый вариант более прост в реализации и приемлем для проведенных исследований, хотя и носит менее общий характер.

Надежность. Под надежностью метода $P(\varepsilon)$ принято понимать вероятность попадания в заданную χ-окрестность экстремума за фиксированное число замеров функции качества. Численно надежность оценивается следующим образом:

$$P(\varepsilon) = 1 - P_\chi = \int_0^\chi p(\varepsilon^j)\, d\varepsilon^j,$$

где $\varepsilon^j = |X^j - X^{ext}|$ — невязка на j-м шаге, распределенная с плотностью $p(\varepsilon^j)$; $P_\chi = \int_\chi^\infty p(\varepsilon^j)\, d\varepsilon^j$ — вероятность недостижения требуемой точности решения.

Помехоустойчивость. Один из важнейших критериев применимости методов поиска к решению задачи оптимизации ускорителя — способность ориентироваться в зашумленной ситуации. Нестабильность уровня помех на ускорителе выдвигает требование устойчивой сходимости процедуры поиска в некотором диапазоне шумов. Поэтому важно исследовать зависимость быстродействия методов от величины шума и оценить верхнюю границу уровня шума, при которой еще возможна правильная ориентация.

Для набора статистики по помехоустойчивости методов уровень шума изменяется от 0 до 25%, и для каждого конкретного уровня фиксировалось среднее число шагов, необходимое для попадания в зону, ограниченную поверхностью

$$\{X : I(X) = C\}, \quad C = 0{,}95 I_{max}.$$

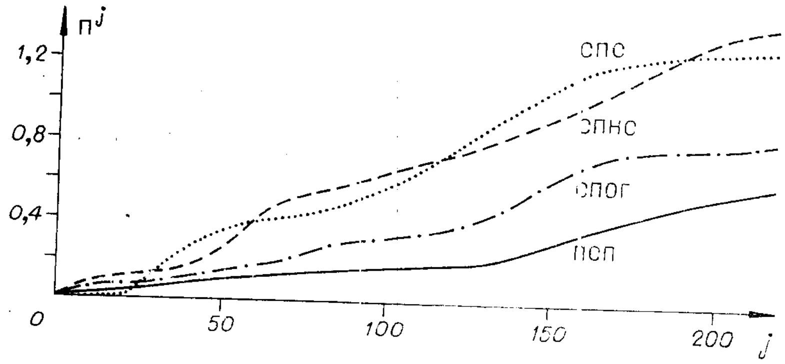

Рис. 1.

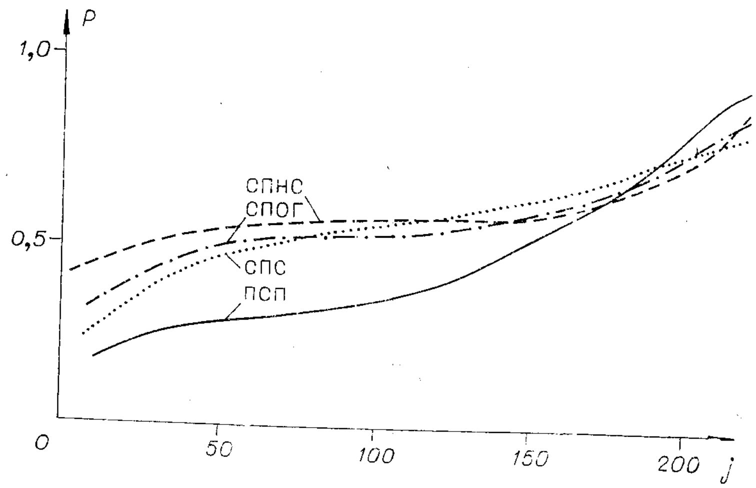

Рис. 2.

Результаты исследований. Основные выводы получены при поиске на сепарабельной модели (кроме тех случаев, которые отмечены особо). Графики зависимости средних потерь от стадии поиска представлены на рис. 1.

В начальной стадии поиска наименьшими потерями характеризуется случайный поиск с самообучением, что объясняется его полной «обученностью» на первых шагах. Потери симплексного метода обусловлены пробами на построение исходного симплекса. На всех последующих стадиях симплексный метод характеризуется минимумом потерь. К нему близок случайный поиск с оценкой градиента. Методы случайного поиска с самообучением и с наказанием случайностью оказались наихудшими по этому критерию.

Графики зависимости вероятности ошибки от номера шага приведены на рис. 2, из которого следует, что симплексному методу соответствует наименьшая вероятность ошибки. Значительное увеличение вероятности ошибки на завершающем этапе ($j>100$) объясняется попа-

Таблица 1

Q_i	Метод			
	СПНС	ПСП	СПОГ	СПС
Q_1	0,29	0,17	0,22	0,27
Q_2	0,64	0,42	0,62	0,57

Таблица 2

Q_i	Метод			
	ПСП	СПОГ	СПНС	СПС
Q_1	55	90	120	110
Q_2	110	140	140	175

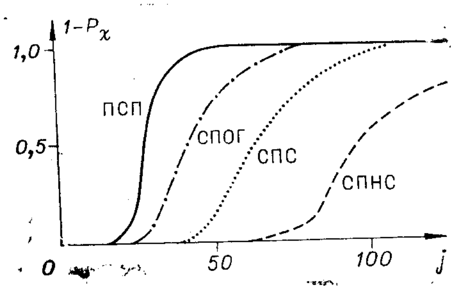

Рис. 3.

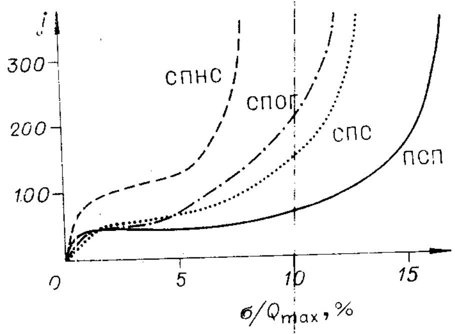

Рис. 4.

данием симплекса в достаточно малую окрестность максимума, где возможный прирост функции качества сравним с уровнем шума. Из разновидностей метода случайного поиска наибольшая вероятность ошибки у поиска с наказанием случайностью, который использует минимальное накопление.

Результаты сравнения методов по неточности сведены в табл. 1.

Хотя симплексный метод показал несколько большую точность отыскания максимума, однако это преимущество относительное, так как точка финиша зависит и от параметра правила остановки. Иначе говоря, при оптимизации алгоритмов по параметрам оказалось возможным подобрать для методов, при всем их качественном различии, такие константы остановки, которые обеспечивают почти одинаковую точность поиска. Во всех методах остановка производится по заданному числу следующих друг за другом неудачных шагов.

При сравнении методов по быстродействию (табл. 2) симплексный метод потребовал наименьшего числа замеров функции качества для достижения 95%-ного уровня максимума независимо от рельефа тестовой функции. Случайный поиск с самообучением оказался наиболее инерционным в окрестности экстремума.

Иными словами, самообучение на траекториях со значительным изменением градиентного направления неэффективно.

Графики зависимости надежности методов от стадии поиска приведены на рис. 3. Наиболее надежными оказались методы симплексного планирования и случайного поиска с оценкой градиента, что объясняется, по-видимому, детерминированностью выбора направления движения (полной в симплексном методе и частичной в случайном поиске с оценкой градиента). Метод случайного поиска с наказанием случайностью, не содержащий в стратегии элементов детерминизма, оказался наименее надежным.

Помехоустойчивость методов отражена на рис. 4. Симплексный метод оказался наименее чувствительным к шумам и сохранил способность к верной ориентации вплоть до 15%-ного уровня шумов. Несколько худшие свойства показал случайный поиск с самообучением, устойчивость к помехе у которого падает, начиная с шума в 5%. Случайный поиск с наказанием случайностью наиболее чувствителен к помехам и перестает сходиться при шумах более 7%.

ВЫВОДЫ

Резюмируя изложенное, можно сказать, что в результате проведенного исследования поисковых методов оптимизации выявлены преимущества метода последовательного симплексного планирования перед разновидностями случайного поиска. Наиболее близок к нему по локальным и интегральным характеристикам метод случайного поиска с оценкой градиента.

ЛИТЕРАТУРА

1. Nedler J. A., Mead R. A simplex method for function minimization.— "Computer J", 1965, vol. 7, № 4, p. 308.
2. Растригин Л. А. Случайный поиск в задачах оптимизации многопараметрических систем. Рига, «Зинатне», 1965.
3. Васильев Н. Д., Гусев О. А., Мозин И. В., Тункин А. А., Шелехов В. А. О выборе и исследовании методов оптимизации управления протонным синхротроном на 7 ГЭВ.— В кн.: Электрофизическая аппаратура. Вып. 13. М., Атомиздат, 1976.
4. Захаров В. В., Филипченко Н. М. Сравнение случайного поиска со схемой скорейшего спуска на тестовых функциях. (Редколлегия журн. «Автоматика и вычислительная техника» АН Латв. ССР), Рига, 1971, № 3527—71 Деп.
5. Грудачев В. Г., Костромин Г. Я., Половинкин А. И. Экспериментальное исследование модификаций алгоритма случайного поиска оптимальных форм.— «Автоматика и вычислительная техника», Рига, «Зинатне», 1974, № 3, с. 73.

Поступила в редакцию 29 июня 1976 г.

Редактор *Г. А. Кузнецова*
Художественный редактор *Э. С. Филонычева*
Технический редактор *Ф. Ф. Орлова*
Корректоры *Л. А. Гуринович, Н. В. Клопотная*

Сдано в набор 15 марта 1977 г. Подписано в печать 8 июня 1977 г. МН 02053. Формат 70×108 1/16. Бумага машиномелованная. 9,5 печ. л., 13,3 усл. печ. л., 13,3 уч.-изд. л. Тираж 3300 экз. Заказ № 470.
Цена 1 руб.

Издательство «Наука». Сибирское отделение. 630099, Новосибирск, 99, Советская, 18.
4-я типография издательства «Наука». 630077, Новосибирск, 77, Станиславского, 25.