

Александр Григорьевич ИВАШКО —
зав. кафедрой информационных систем,
профессор, доктор технических наук
ivashco@mail.ru

Иван Юрьевич КАРЯКИН —
ассистент кафедры информационных систем
Ivan_NYS@mail.ru

Антон Александрович ЗЮРКАЛОВ —
студент 4 курса
Kadio90@mail.ru

*Институт математики и компьютерных наук
Тюменского государственного университета*

УДК 519.688

**АНАЛИЗ МЕТОДОВ ХУКА–ДЖИВСА И НЕЛДЕРА–МИДА ПРИМЕНИ-
ТЕЛЬНО К ЗАДАЧЕ НАХОЖДЕНИЯ ПАРАМЕТРОВ
МОДЕЛИ ФАЗОВЫХ ПРЕВРАЩЕНИЙ***

**ANALYSIS OF HUKA-DZHIVS AND NELDER-MID METHODS
WITH REFERENCE TO THE TASK OF FINDING THE PARAMETRES
OF PHASE TRANSFORMATIONS MODEL**

АННОТАЦИЯ. В работе рассматривается применение метода Нелдера–Мида для решения задачи нахождения кинетических параметров модели фазовых превращений аустенита при термической обработке стали. Приводится сравнительный анализ методов Хука–Дживса и Нелдера–Мида в задаче нахождения параметров модели.

SUMMARY. The given paper considers the application of Nelder-Mid method in finding kinetic parameters of phase models of austenite transformations at steel thermal processing. The comparative analysis of Huka-Dzhivs and Nelder-Mid methods in finding model parameters is presented.

КЛЮЧЕВЫЕ СЛОВА. Система массового обслуживания (СМО), аустенит, метод Хука–Дживса, метод Нелдера–Мида, технология параллельных вычислений.

KEY WORDS. Queuing system, austenite, method of Huka-Dzhivs, method of Nelder-Mid, technology of parallel computings.

Говоря об алгоритмах решения задач оптимизации нельзя не коснуться вопроса о качестве этих алгоритмов. К числу важных характеристик алгоритма могут быть отнесены время выполнения и простота реализации при решении практических задач [1].

В работе [2] разработана модель фазовых превращений аустенита при термической обработке стали на основе теории систем массового обслуживания.

Рассматривается задача нахождения параметров модели:

* Работа выполнена в рамках АВЦП «Развитие научного потенциала высшей школы (2000-2010 гг.)», НИР «Прогнозирование кинетики распада аустенита в порошковых сталях при непрерывном охлаждении». № 2.1.2/6498.

$$f(x_1, x_2, \dots, x_k) \rightarrow \min, \quad (1)$$

$$f(x_1, x_2, \dots, x_N) = \sum_{i,j=1}^{n,m} (Vm(x_1, x_2, \dots, x_N, \tau_{i,j}) - V(T_i, \tau_{i,j}))^2, \quad (2)$$

где n — число экспериментальных точек; m — количество температурных режимов; T_i — температура ($^{\circ}C$); $\tau_{i,j}$ — время после начала выдержки стальной детали (в секундах); V — доля распавшегося аустенита после времени $\tau_{i,j}$, полученная в результате натурального эксперимента (%); x_1, x_2, \dots, x_N — параметры модели, характеризующие $\lambda_{create}(T)$ — интенсивность возникновения зародышей аустенита, $\lambda_{raspada}(T)$ — интенсивность появления событий проверки на свертывание зародышей аустенита, sk_{rosta} — линейная скорость роста зародышей, R_0 — начальный радиус зародыша, R_{krit} — критический радиус. $Vm(x_1, x_2, \dots, x_N, \tau_{i,j})$ — доля распавшегося аустенита после времени $\tau_{i,j}$, полученная в результате компьютерного эксперимента.

Исходными данными задачи (1) являются значения V , полученные в результате натурального эксперимента. Для решения поставленной задачи возможно применить методы Хука–Дживса и Нелдера–Мида.

Использование метода Хука–Дживса для нахождения параметров модели рассмотрено в работе [3]. При применении этого метода выбирается начальная базисная точка $x^0 = (x_1^0, x_2^0, \dots, x_k^0)$ и шаги длиной h_1, h_2, \dots, h_k соответственно, исследуется локальное поведение функции $f(x_1, x_2, \dots, x_k)$ в точке x^0 и определяется направление, в котором наблюдается уменьшение функции. В результате определяется новая базисная точка или производится уменьшение шага. Процесс завершается, когда длины шагов будут уменьшены до заданной величины ε .

Метод Нелдера–Мида, также известный как метод деформируемого многогранника, является методом безусловной оптимизации функции нескольких переменных, не использующий производной функции, а поэтому легко применим к негладким функциям [4].

Суть метода заключается в последовательном перемещении и деформировании симплекса вокруг точки экстремума. Под симплексом понимается выпуклый многогранник с числом вершин $N+1$, где N — количество параметров модели.

Для изложения алгоритма метода Нелдера–Мида введем следующие обозначения: коэффициент отражения $\alpha > 0$; коэффициент сжатия $\beta > 0$; коэффициент растяжения $\gamma > 0$. Решение задачи начинается с выбора начальной точки $x^0 = (x_1^0, x_2^0, \dots, x_k^0)$. Изменяя каждую из координат точки x^0 на шаг h , находятся точки $x^i = (x_1^i, x_2^i, \dots, x_k^i)$, $i = 1, \dots, (N+1)$, образующие симплекс

N -мерного пространства. В этих точках вычисляются значения функции: $f_1 = f(x^1)$, $f_2 = f(x^2)$, ..., $f_{N+1} = f(x^{N+1})$.

Шаг 1. Из вершин симплекса выбираются три точки: x^h — с наибольшим значением функции f_h , x^g — со следующим по величине значением f_g и x^l — с наименьшим значением функции f_l . При сравнении значений функции необходимо учитывать стохастическую природу процесса фазовых превращений, поэтому для поиска точек x^h , x^g и x^l применяется статистический критерий.

Для сравнения значений функции в точках $x^1 = (x_1^1, x_2^1, \dots, x_N^1)$, $x^2 = (x_1^2, x_2^2, \dots, x_N^2)$ проводятся несколько серий экспериментов и вычисляются средние значения $\bar{f}(x_1^1, x_2^1, \dots, x_N^1)$ и $\bar{f}(x_1^2, x_2^2, \dots, x_N^2)$. Если $\bar{f}(x_1^1, x_2^1, \dots, x_N^1) < \bar{f}(x_1^2, x_2^2, \dots, x_N^2)$, то выдвигается гипотеза $H_0: M(\bar{f}(x_1^1, x_2^1, \dots, x_N^1)) < M(\bar{f}(x_1^2, x_2^2, \dots, x_N^2))$, для проверки которой используется t -критерий [5] при уровне значимости $\alpha = 0,05$. Учитывая, что средние квадратические отклонения $\sigma(\bar{f}(x_1^1, x_2^1, \dots, x_N^1))$ и $\sigma(\bar{f}(x_1^2, x_2^2, \dots, x_N^2))$ неизвестны (имеются только их оценки, найденные по выборке), а также то, что в общем случае нет оснований предполагать $\sigma(\bar{f}(x_1^1, x_2^1, \dots, x_N^1)) = \sigma(\bar{f}(x_1^2, x_2^2, \dots, x_N^2))$, для проверки гипотезы используется статистика:

$$t' = \frac{(\bar{f}(x_1^2, x_2^2, \dots, x_N^2) - \bar{f}(x_1^1, x_2^1, \dots, x_N^1))\sqrt{N}}{\sqrt{s^2(f(x_1^2, x_2^2, \dots, x_N^2)) + s^2(f(x_1^1, x_2^1, \dots, x_N^1))}}. \quad (3)$$

Наблюдаемое значение t' (вычисленное по данным выборки) сравнивается со значением $t_{1-\alpha}$ с ν степенями свободы, где

$$\nu = \frac{(s^2(f(x_1^2, x_2^2, \dots, x_N^2)) + s^2(f(x_1^1, x_2^1, \dots, x_N^1)))^2 (L+1)}{(s^2(f(x_1^2, x_2^2, \dots, x_N^2)))^2 + (s^2(f(x_1^1, x_2^1, \dots, x_N^1)))^2} - 2, \quad (4)$$

L — число значений в каждой серии.

При $t'_{\text{набл}} > t_{1-\alpha}$ гипотеза принимается, в противном случае — отвергается.

В случае принятия гипотезы H_0 имеем $f_1 < f_2$. В дальнейшем при каждом сравнении значений функции в вершинах симплекса будут проверяться аналогичные гипотезы.

Целью дальнейших манипуляций является уменьшение f_h .

Шаг 2. Находится центр тяжести всех точек, за исключением x^h : $x^c = \frac{1}{N} \sum_{i \neq h} x^i$.

Шаг 3. Точка x^h отражается относительно x^c с коэффициентом α (при $\alpha = 1$ применяется центральная симметрия), в результате имеем точку x^r , вычисляется значение функции $f_r = f(x^r)$. Координаты новой точки определяются по формуле $x^r = (1 + \alpha)x^c - \alpha x^h$.

Шаг 4. Значение f_r сравнивается со значениями f_h, f_g, f_l :

1. Если $f_r < f_l$, то производится растяжение. Новая точка $x^e = (1 + \gamma)x^c + \gamma x^r$ и значение функции $f_e = f(x^e)$.

Если $f_e < f_l$, то точка x^h заменяется на x^e и заканчивается итерация (переход на шаг 8).

Если $f_e > f_l$, то точка x^h заменяется на x^r и заканчивается итерация (переход на шаг 8).

2. Если $f_l < f_r < f_g$, то точка x^h заменяется на x^r и осуществляется переход на шаг 8.

3. Если $f_h > f_r > f_g$, то меняются обозначения x^r, x^h (и соответствующие значения функции) местами и осуществляется переход на шаг 5.

4. Если $f_r > f_h$, то осуществляется переход на шаг 5.

Шаг 5. Строится точка $x^s = \beta x_h + (1 - \beta)x_c$ и в ней вычисляется значение f_s .

Шаг 6. Если $f_s < f_h$, то точка x^h заменяется на x^s и осуществляется переход на шаг 8.

Шаг 7. Если $f_s > f_h$, то производится сжатие симплекса — гомотетия к точке с наименьшим значением $x^0 : x^i \rightarrow x^0 + (x^i - x^0) / 2$ для всех требуемых точек x^i .

Шаг 8. Проверка сходимости. Суть проверки заключается в том, чтобы проверить взаимную близость полученных вершин симплекса, что предполагает их близость к искомому минимуму. Если требуемая точность не достигнута, действия продолжаются с шага 1.

По изложенным алгоритмам созданы программные модули, реализующие поиск минимума функции методом Хука–Дживса и Нелдера–Мида, которые позволяют находить параметры модели по существующим кривым и получать изотермические кривые при различных параметрах имитационной модели.

Проведенные компьютерные эксперименты показали, что поиск необходимых параметров вышеуказанными методами требует значительных временных затрат.

Для сокращения этих затрат оказалось целесообразным применение технологии параллельных вычислений. Были выделены области распараллеливания в каждом из методов.

В методе Хука–Дживса основная вычислительная нагрузка приходится на блок, в котором вычисляются значения целевой функции (2) в базисной точке и точках, принадлежащих ее окрестности, количество точек равно удвоенному числу параметров модели. В схеме параллельных вычислений каждое значение функции определяет отдельный второстепенный процесс. Главный процесс в качестве исходных данных использует график изотермической

кривой, полученной в результате натурального эксперимента, и обобщает результаты второстепенных процессов. Для вычисления данных компьютерного эксперимента используется DLL-файл с функцией моделирования.

В методе Нелдера–Мида технология параллельных вычислений применяется при расчете значений функции в каждой из вершин симплекса, а также в точках отражения, растяжения и сжатия.

Большое количество программного времени приходится на имитацию процесса фазовых превращений при определении значений Vm функции (2). Для нахождения значений $Vm(x_1, x_2, \dots, x_N, \tau)$ проводится ряд экспериментов, имитирующих рассматриваемый процесс, результатом которых являются значения $Vm_1(x_1, x_2, \dots, x_N, \tau)$, $Vm_2(x_1, x_2, \dots, x_N, \tau)$, ..., $Vm_N(x_1, x_2, \dots, x_N, \tau)$, конечные значения Vm получаются путем их усреднения. В силу вышесказанного для снижения времени на вычисление значений компьютерного эксперимента также целесообразно применение технологии параллельных вычислений. В этом случае второстепенные процессы проводят компьютерную имитацию процесса фазовых превращений, а главный процесс усредняет их результаты.

С целью организации параллельных вычислений создана GRID-система, представляющая собой сеть из компьютеров, один из которых выполняет роль главного, остальные — роль второстепенных. Для корректной работы параллельных процессов разработанные алгоритмы были модифицированы с использованием MPI библиотек, которые реализуют обмен данными между компьютерами и осуществляют вычисления, необходимые для распараллеливания [6].

Для сравнения эффективности применения методов Хука–Дживса и Нелдера–Мида были проведены вычислительные эксперименты. В табл. 1 приведены показатели, характеризующие методы без применения и с применением параллельных вычислений.

Таблица 1

Показатели работы алгоритмов

Метод	Время работы алгоритма (чч:мм:сс)		Показатель уменьшения времени (%)
	Без применения параллельных вычислений	С применением параллельных вычислений	
Хука–Дживса	00:58:43	00:37:19	36,45%
Нелдера–Мида	00:41:17	00:28:59	29,79%

Из приведенных в таблице данных очевидно, что применение процесса параллельных вычислений более эффективно для метода Хука–Дживса. Однако фактическое время работы алгоритма Нелдера–Мида меньше, чем время работы алгоритма Хука–Дживса. В случае применения параллельных вычислений — на 22,33%, а без применения — на 29,69%. Поэтому для решения задачи (1) с целью уменьшения временных затрат рекомендуется использовать метод Нелдера–Мида с применением параллельных вычислений.

В табл. 2 приводятся сравнительные данные методов Хука–Дживса и Нелдера–Мида при различных исходных данных.

Таблица 2

Сопоставление результатов применения методов Хука–Дживса и Нелдера–Мида

	Метод Хука–Дживса	Метод Нелдера–Мида	Метод Хука–Дживса	Метод Нелдера–Мида
Исходные параметры	$\lambda_{create}=10; \lambda_{raspada}=10;$ $sk_{rosta}=10; R_{krit}=10; R_0=10$		$\lambda_{create}=30; \lambda_{raspada}=30;$ $sk_{rosta}=30; R_{krit}=30; R_0=30$	
Результат	$\lambda_{create}=4,9231$ $\lambda_{raspada}=3,218$ $sk_{rosta}=2,01$ $R_{krit}=6,031$ $R_0=4,989$	$\lambda_{create}=4,78$ $\lambda_{raspada}=3,07$ 1 $sk_{rosta}=2,299$ $R_{krit}=7,172$ $R_0=4,6$	$\lambda_{create}=5,1003$ $\lambda_{raspada}=3,31$ 7 $sk_{rosta}=3,8$ $R_{krit}=8,979$ $R_0=5,29$	$\lambda_{create}=3,139$ $\lambda_{raspada}=4,1$ $sk_{rosta}=10$ $R_{krit}=11,2$ $R_0=7,076$
Время вычислений	00:37:19	00:28:59	1:22:31	00:53:01
Количество итераций	527	189	1372	574
Количество операций при имитации	197631	162039	153017	147201
Отклонение результатов компьютерного и натурального экспериментов	23,51%	21,30%	22,47%	24,74%

Результаты проведенных экспериментов показали лучшую сходимость метода Нелдера–Мида по сравнению с методом Хука–Дживса.

СПИСОК ЛИТЕРАТУРЫ

1. Хохлюк В.И. Параллельные алгоритмы целочисленной оптимизации. М.: Радио и связь, 1987. 224 с.
2. Ивашко А.Г., Цыганова М.С., Карякин И.Ю. Имитационное моделирование фазовых превращений // Математическое и информационное моделирование. Сб. науч. тр. Вып. 10. Тюмень: Вектор Бук, 2008. С. 106-113.
3. Ивашко А.Г., Цыганова М.С., Карякин И.Ю. Модифицированный метод Хука–Дживса для нахождения параметров модели фазовых превращений // Вестник Тюменского государственного университета. 2009. № 6. С. 197-202.
4. Аттетков А.В., Галкин С.В., Зарубин В.С.. Методы оптимизации: Учеб. для вузов. 2-е изд. М.: Изд-во МГТУ им. Н. Э. Баумана, 2003. 440 с.
5. Химмельблау Д. Анализ процессов статистическими методами. М.: Мир, 1973.
6. Шпаковский Г. И., Стецюренко В. И., Верхотуров А. Е., Серикова Н. В. Применение технологии MPI в Грид. Минск.: БГУ, 2008. 137 с.