ФИЗИКО-МАТЕМАТИЧЕСКИЕ НАУКИ

УДК 519.2:004.421.5:004.7

В. Н. ЗАДОРОЖНЫЙ

Омский государственный технический университет

КАСКАДНЫЙ МЕТОД РЕАЛИЗАЦИИ РАСПРЕДЕЛЕНИЙ С ТЯЖЕЛЫМИ ХВОСТАМИ

Исследуется и решается проблема искажения распределений с тяжелыми хвостами, реализуемых в статистическом моделировании. Предлагаются методы решения этой проблемы. Разрабатывается эффективный метод генерации случайных величин с тяжелыми хвостами распределений. Оценивается точность разработанного метода.

Ключевые слова: фрактальный трафик, теория массового обслуживания, имитационное моделирование, генераторы случайных чисел.

1. Введение. Обнаружение у трафика информационных сетей фрактальных свойств (масштабной инвариантности, долговременной зависимости и «взрывных» пульсаций) привело к существенной корректировке теории телетрафика и математических моделей, используемых в области проектирования компьютерных сетей [1–3]. Для анализа и моделирования фрактального трафика теперь широко используются распределения с тяжелыми хвостами (РТХ). В частности, узлы сетей передачи данных (СПД) представляются системами массового обслуживания (СМО), в которых интервалы поступления сообщений и/или время их обработки задаются РТХ. И поскольку исследование таких СМО классическими методами теории массового обслуживания затруднено, широко используется метод статистического моделирования (СМ).

В СМ для генерации случайных величин (с.в.) с РТХ используется метод обращения. При генерации любой с.в. x с заданной функцией распределения (ф.р.) $F(t)$ метод обращения сводится к реализации значений x по формуле $x = F^{-1}(u)$, где u — базовая с.в. (БСВ), равномерно распределенная в промежутке от 0 до 1, а F^{-1} — это обратная к F функция. Метод обращения является точным методом. Поэтому предполагается, что использование в методе обращения качественных программных генераторов БСВ гарантирует качественную генерацию требуемой с.в. x.

Однако, как показано в [4, 5], при генерации с.в. с РТХ это далеко не так. Выясняется, что реализуемые с.в. с РТХ имеют смещенные моменты. Ниже будет показано, что, соответственно, искажается и форма распределения. Причиной такого искажения является ограниченная разрядность генериру-

Выполненные расчеты и анализ полученного общего соотношения (6) приводят к выводу, что смещения моментов РП при использовании любого фиксированного ε (т.е. при любой разрядности генератора БСВ) могут быть сколь угодно велики. Очевидно, аналогичные особенности имеют место и при реализации других РТХ.

4. Проявление смещений моментов в СМ ФС. В разных задачах СМ ФС отличие дискретного РП от РП проявляется по-разному.

4.1. Свойства генерируемых степенных выборок. Если при $\alpha\in(1,\ 2]$ по независимым реализациям $\tilde{x}_1, ..., \tilde{x}_n$ д.с.в. $\tilde{x}\in \mathrm{DPa}(K;\ \alpha;\ \varepsilon)$ вычисляется оценка $\tilde{M}_x = (\tilde{x}_1 + ... + \tilde{x}_n)/n$, то $\tilde{M}_x \to \mathrm{E}(\tilde{x})$ при $n\to\infty$. Сходимость оценки $\tilde{M}_x$ к $\mathrm{M}(\tilde{x})$ очень медленная из-за большого КВ $C(\tilde{x})$. Так, если $\alpha=1{,}1$, то даже при $\varepsilon=10^{-6}$ для получения приемлемого приближения 8,028 к точному значению $\mathrm{M}(\tilde{x})=8{,}0297$ (см. (3)) потребовалась выборка длиной $n=1$ млрд.

При этом разные способы оценки параметров распределения по выборке $\tilde{x}_1, ..., \tilde{x}_n$ приводят к разным результатам. Поэтому нужно тщательно обосновывать применяемые оценки и осторожно их интерпретировать. Например, если по выборке $\tilde{x}_1, ..., \tilde{x}_n$ параметры K и α оценивать методом максимального правдоподобия, который приводит к формулам $\tilde{\alpha} = n\Big/\left(\sum_{i=1}^{n}\ln\tilde{x}_i - n\ln\tilde{K}\right)$, $\tilde{K} = \min\{\tilde{x}_i\}$, то такие оценки сходятся к точным значениям K и α достаточно быстро, и даже при $\varepsilon=10^{-6}$ оказываются практически несмещенными. По найденным оценкам $\tilde{K}, \tilde{\alpha}$ можно получить косвенную оценку м.о. в виде $\tilde{M}_x = \tilde{K}\tilde{\alpha}/(\tilde{\alpha}-1)$. Эта оценка, соответственно, оказывается достаточно точной. Но сходится она не к м.о. $\mathrm{M}(\tilde{x})$ элементов выборки, а к м.о. $\mathrm{M}(x)=\alpha K/(\alpha-1)$ с.в. $x\in\mathrm{Pa}(K;\ \alpha)$, подлежащей реализации.

В рассмотренном примере путем обработки выборки непосредственно обнаруживается смещение м.о. $\tilde{M}_x \approx \mathrm{M}(\tilde{x}) \neq \mathrm{M}(x)$. Искажение формы распределения, определяемой параметрами K, α, непосредственно не обнаруживается: статистические оценки этих параметров близки к заданным в генераторе значениям.

4.2. Особенности СМ ФС Pa|M|1|¥. В этой системе [7] среднее время ожидания может быть найдено точно:

$$W = \frac{\sigma}{\mu(1-\sigma)}, \qquad (7)$$

где $\mu=1/b$ — интенсивность обслуживания,
σ — единственный в области $0\le\sigma<1$ корень уравнения

$$\sigma = A^*(\mu-\mu\sigma), \qquad (8)$$

$A^*(s)=\int_0^\infty e^{-st}dA(t)$ — преобразование Лапласа от ф.р. $A(t)$.

В данном случае $A(t)$ есть ф.р. Парето (1).

Влияние дискретности датчиков БСВ на результаты СМ системы Pa|M|1|∞ оценим на следующем примере. Пусть при $\alpha=1{,}1$, $K=1$ (тогда среднее время между приходами заявок $a=11$) требуется найти $W=W(\rho)$ при коэффициенте загрузки $\rho=b/a=0{,}4$. Такая загрузка имеет место при $b=0{,}4a=4{,}4$, т.е. при $\mu=0{,}22727$. Решая уравнение (8) численным методом, находим $\sigma=0{,}964199$, $W=\mathbf{118{,}5}$.

При СМ этой СМО с шагом генератора БСВ $\varepsilon=10^{-6}$ среднее время между приходами заявок равно 8,0297 (см. (3)). Коэффициент загрузки $\rho=0{,}4$ достигается при $b=0{,}4\times8{,}0297=3{,}21188$, т.е. при $\mu=0{,}311344$. Но СМ данной СМО при таком μ (при прогоне около 125 млн заявок), действительно дающее для коэффициента загрузки ρ оценку $\rho_{СМ}=0{,}401\approx0{,}4$, для W дает оценку $W_{СМ}=\mathbf{19{,}657}$. Таким образом, решение методом СМ поставленной задачи о величине $W(\rho)$ при $\rho=0{,}4$ отличается от точного решения более чем впятеро.

Здесь уже проявляется не смещение моментов РТХ (моменты интервалов поступления заявок, согласно (7), на время W не влияют), а искажение формы распределения Парето, которою определяется решение уравнения (8). Действительно, с учетом дискретности реализуемого распределения $\mathrm{DPa}(1;\ 1{,}1;\ 10^{-6})$ уравнение (8) принимает вид

$$\sigma = 10^{-6}\sum_{i=1}^{10^6} 1/\exp[(0{,}3113441(1-\sigma))(10^{-6}i)^{-1/1{,}1}], \qquad (9)$$

его численное решение дает $\sigma=0{,}85953$ (что легко проверить) и, в соответствии с (8) (при $\mu=0{,}311344$), $W=\mathbf{19{,}653}$. Этому точному решению вполне соответствует полученная путем СМ оценка $W_{СМ}=\mathbf{19{,}657}$.

Этот пример показывает, как может влиять на результаты СМ искажение формы реализуемого РТХ, обусловленное дискретностью генератора БСВ.

4.3. Особенности СМ ФС M|Pa|1|¥. В данной системе время обслуживания $x\in\mathrm{Pa}(K;\ \alpha)$, и если $\alpha\in(1,\ 2]$, то стационарная средняя длина $L(\rho)$ очереди заявок бесконечна при любом $\rho\in(0,\ 1)$. Действительно, согласно формуле Полячека – Хинчина [7], здесь имеем

$$L(\rho) = \frac{\rho^2[1+C^2(x)]}{2(1-\rho)} = \infty, \quad \rho\in(0,\ 1), \qquad (10)$$

так как коэффициент вариации $C(x)$ при $\alpha\in(1,\ 2]$ бесконечен. Но если эту ФС мы будем исследовать методом СМ, то вместо с.в. $x\in\mathrm{Pa}(K;\ \alpha)$ реализуется д.с.в. $\tilde{x}\in\mathrm{DPa}(K;\ \alpha;\ \varepsilon)$ и, например, при $K=1$, $\alpha=2$, $\varepsilon=10^{-15}$ мы получим:

$$L(\rho) = \frac{\rho^2[1+C^2(\tilde{x})]}{2(1-\rho)} \approx \frac{\rho^2(1+2{,}79^2)}{2(1-\rho)} \approx \frac{4{,}39\rho^2}{1-\rho} \qquad (11)$$

(табл. 1). Особенностью рассмотренного случая является принципиальное отличие статистического решения (11) от точного (10) из-за смещения моментов, обусловленного дискретностью генератора БСВ.

5. Критическая зона значений БСВ. При заданной ф.р. $F(t)$ с.в. x генерируется по формуле $x=F^{-1}(u)$, либо по формуле $x=\bar{F}^{-1}(u)$, где $\bar{F}^{-1}$ — функция, обратная к дополнительной ф.р. (ДФР) $\bar{F}$, $\bar{F}(t)=1-F(t)$. Частным случаем формулы обращения ДФР является (2). Выполненные расчеты и анализ показывают, что с.в. с РТХ, генерируемые методом обращения ДФР, искажаются тем больше, чем ближе к нулю оказываются реализованные значения БСВ u. При использовании обращения не ДФР, а ф.р. $F(t)$, этому условию вполне симметричным образом соответствует условие близости значений БСВ к единице, поэтому достаточно рассмотреть случай обращения ДФР.

Если значение u приближается к нулю, то (так устроены генераторы БСВ) старшие разряды, начиная с первого после десятичной точки, становятся нулями, и общее число значащих цифр уменьшается. Это и является источником погрешностей. В достаточно вероятной *критической зоне* значений u, близких к нулю, эти значения превращаются при реализации РТХ в очень большие значения x, вносящие существенный вклад в формирование моментов, тогда как при реализации обычных распределений соответствующие x относительно невелики, и ими можно пренебречь.

6. Каскадный метод реализации РП. Пусть имеется генератор Rand1( ), реализующий БСВ с точностью $r\geq 12$ десятичных разрядов и Rand2( ), дающий хотя бы 6 разрядов. Разобьем область $t\geq K$ значений с.в. $x\in \mathrm{Pa}(K;\ \alpha)$ на интервалы точками $K_1=K$, $K_2=K_1\cdot 10^{6/\alpha}$, $K_3=K_2\cdot 10^{6/\alpha}$, ... Тогда с учетом (1) вероятности интервалов составят $p_i=\mathrm{P}(K_i\leq x<K_{i+1})=$ $=0{,}999999\times 10^{-6(i-1)}$, $i=1,\ 2,\ \ldots$. Представим ДФР $\overline{F}(t)=(K/t)^{\alpha}$ линейной комбинацией соответствующих условных ДФР:

$$\overline{F}(t)=\sum_{i=1}^{\infty}p_i\overline{F}_i(t)\,, \qquad (12)$$

где

$$\overline{F}_i(t)=\begin{cases}1, & if\ t<K_i\,,\\ \dfrac{\overline{F}(t)-\overline{F}(K_{i+1})}{p_i}, & if\ K_i\leq t<K_{i+1}\,,\\ 0, & if\ K_{i+1}\leq t.\end{cases} \qquad (13)$$

Случайную величину x с ДФР $\overline{F}(t)=(K/t)^{\alpha}$ будем генерировать как смесь с.в., имеющих ДФР (13) и выбираемых с соответствующими вероятностями p_i.

В соответствии со сказанным процедуру генерации с.в. $x\in \mathrm{Pa}(K;\ \alpha)$ можно представить так, как показано на рис. 1. Цикл на шаге Step2 (не считая проверки условия) выполняется однократно приблизительно один раз за 10^6 обращений к процедуре (двукратно — один раз за 10^{12} обращений, и т.д.), что обеспечивает достаточно быструю генерацию с.в. x. Шаг Step3 реализует метод обращения условной ДФР с учетом (13).

Описанный метод нетрудно модифицировать и для реализации других РТХ.

Процедуру на рис. 1 нетрудно модифицировать и для генераторов Rand( ) с другой разрядностью, с другими разбиениями оси t на интервалы условных распределений, с другой, не десятичной системой счисления и для других РТХ.

7. Оценка точности каскадного метода. Для оценки точности предложенного каскадного метода реализации РТХ рассчитаем м.о. $\mathrm{M}(\tilde{x})$ с.в. $\tilde{x}$, которая фактически реализуется процедурой на рис. 1 при $K=1$, $\alpha=1{,}1$, и сравним это м.о. с точным значением 11 м.о. с.в. $x\in \mathrm{Pa}(1;\ 1{,}1)$, подлежащей реализации. Поскольку используемый генератор Rand1( ) имеет $r\geq 12$ десятичных разрядов, рассмотрим наихудший случай $r=12$. Применим использованный в (3) метод вычисления $\mathrm{M}(\tilde{x})$ как м.о. дискретной с.в. Для условного распределения на интервале $(K_i,\ K_{i+1})$ м.о. фактически реализованной с.в. составляет:

$$\mathrm{M}_i(\tilde{x})=10^{-12}\sum_{n=1}^{10^{12}}(0{,}999999\cdot 10^{-6i}\,n+10^{-6i})^{-1/1{,}1}\,. \qquad (14)$$

Этот интервал выбирается с вероятностью $p_i=$ $=0{,}999999\cdot 10^{-6(i-1)}$. Результаты вычисления условных м.о. (14), соответствующих вероятностей интервалов и безусловного м.о. представлены в табл. 2. Эти расчеты показывают, что безусловное м.о. $\mathrm{M}(\tilde{x})$ фактически реализуемой д.с.в. $\tilde{x}$ с точностью до 6 значащих цифр совпадает с точным м.о. $\mathrm{M}(x)$. Аналогично численными методами устанавливается, что форма РП, реализуемого каскадной процедурой, определяющая решение уравнения (8), практически не искажается (с точностью до $(r-6)$ значащих цифр).

```
Input: K, α
output: x
Step1:  p ← (1 – 10^–6)
        i ← 1
Step2: If Rand2( ) < 10^–6
        then p ← p·10^–6
             i ← i + 1
             goto Step2
Step3: u ← Rand1( )
        x ← K(up + 10^–6i)^–1/α
return: x
```

Рис. 1. Каскадная процедура генерации с.в. с РП

Таблица 2

Расчет безусловного м.о. реализуемой с.в.

i	$M_i=\mathrm{M}_i(\tilde{x})$	p_i	p_iM_i	Частичная сумма p_iM_i
1	7,86717E+00	9,99999E-01	7,867162	7,867162
2	2,24060E+06	9,99999E-07	2,240598	10,10776
3	6,38130E+11	9,99999E-13	0,638129	10,74589
4	1,81742E+17	9,99999E-19	0,181742	10,92763
5	5,17607E+22	9,99999E-25	0,051761	10,97939
6	1,47416E+28	9,99999E-31	0,014742	10,99413
7	4,19847E+33	9,99999E-37	0,004198	10,99833
8	1,19574E+39	9,99999E-43	0,001196	10,99953
9	3,40551E+44	9,99999E-49	0,000341	10,99987
10	9,69901E+49	9,99999E-55	9,7E-05	10,99997
11	2,76231E+55	9,99999E-61	2,76E-05	10,99999

Выполненный расчет убедительно показывает, что каскадный генератор полностью решает выявленную проблему смещения моментов реализуемых РТХ.

8. Заключение. Исследование свойств реализуемых при СМ РТХ, показывает, что они реализуются в общем случае со значительными искажениями. Это может приводить к существенным, иногда принципиальным ошибкам в результатах СМ и их интерпретации. В большинстве практических случаев эти ошибки скрыты от исследователя-экспериментатора.

Причиной искажения РТХ является дискретность используемых генераторов стандартных случайных чисел. Но «устранение» дискретности за счет перехода к «длинной арифметике» неэффективно из-за возникающих аппаратных затрат или потерь производительности, а также из-за того, что заранее невозможно определить, какая длина разрядной сетки будет достаточна.

В статье исследованы проблемы влияния искажений РТХ на результаты СМ и предложен эффективный каскадный метод реализации РТХ. Он не требует использования «длинной арифметики» и приводит к минимальным потерям производительности.

В инженерной практике СМ редко требуются выборки такого объема, при котором отличие предложенного генератора от обычных генераторов было бы заметно. Однако в научных исследованиях он может быть весьма полезен. По крайней мере, он снимает с исследователя-экспериментатора заботу об оценке погрешностей СМ, связанных с дискретностью используемых генераторов.

Библиографический список

1. Crovella, M. E., Taqqu M., Bestavros A. Heavy Tailed-Probability distributions in the World Wide Web, 5 (6), December 1997, pp. 835—846.
2. Resaul, K. M., Grout, V. A Comparison of Methods for Estimating the Tail Index of Heavy-tailed Internet Traffic, in Innovative Algorithms and Techniques in Automation, Industrial Electronics and Telecommunications, Springer, Dordrecht, 2007, pp. 219—222.
3. Задорожный, В. Н. Предпосылки создания фрактальной теории массового обслуживания / В. Н. Задорожный // Омский научный вестник. Сер. Приборы, машины и технологии. — 2010. — № 2 (90). — С. 182—187.
4. Задорожный, В. Н. Проблемы генерации случайных величин с фрактальными распределениями / В. Н. Задорожный, О. И. Кутузов // Омский научный вестник. Сер. Приборы, машины и технологии. — 2012. — № 3 (113). — С. 20—24.
5. Zadorozhnyi, V. N. Simulation modeling of fractal queues, in Dynamics of Systems, Mechanisms and Machines (Dynamics), 2014, December, 2014, pp. 1—4. — DOI: 10.1109/Dynamics. 2014.7005703.
6. GPSS World reference manual. Minuteman Software, Fifth Edition, Holly Springs, NC, U.S.A., 2009 [Электронный ресурс]. — Режим доступа : http://www.minutemansoftware.com/reference/rpreface.htm (дата обращения: 30.03.2015).
7. Клейнрок, Л. Вычислительные системы с очередями / Л. Клейнрок ; пер. с англ. Б. С. Цыбакова. — М. : Мир. — 1979. — 600 с.
8. Вишневский, В. М. Теоретические основы проектирования компьютерных сетей / В. М. Вишневский. — М. : Техносфера. — 2003. — 512 с.
9. Мандельброт, Б. Фрактальная геометрия природы / Б. Мандельброт ; пер. с англ. — М. : Ин-т компьютерных исследований, 2002. — 656 с.

ЗАДОРОЖНЫЙ Владимир Николаевич, доктор технических наук, доцент (Россия), профессор кафедры автоматизированных систем обработки информации и управления.
Адрес для переписки: zwn@yandex.ru

Статья поступила в редакцию 31.03.2015 г.

УДК 519.2+621.391

А. А. РОМАНОВА

Омская юридическая академия

ИССЛЕДОВАНИЕ ЗАДАЧИ ВЫПОЛНЕНИЯ ЗАКАЗОВ ПАРТИЯМИ С КРИТЕРИЕМ МИНИМИЗАЦИИ ВЗВЕШЕННОЙ СУММЫ МОМЕНТОВ ЗАВЕРШЕНИЯ

В работе рассматривается задача составления расписания выполнения заказов клиента одним производителем партиями ограниченного размера с критерием минимизации взвешенной суммы моментов завершения. Доказана NP-трудность задачи, выделены полиномиально разрешимые случаи, предложен алгоритм нахождения приближенного решения, проведен вычислительный эксперимент.

Ключевые слова: расписание, выполнение требований партиями, сложность задачи, приближенное решение.

Задачи теории расписаний состоят в определении оптимальной очередности обработки изделий на различных станках или других рабочих местах, составлении программы-диспетчера для управления работой ЭВМ в мультипрограммном режиме и т. п. Подобные задачи возникают во многих сферах человеческой деятельности: образовании, транспорте, управлении, информатике, производстве, сельском хозяйстве.

Многие современные производственные системы обслуживают схожие либо идентичные требования партиями. Требования одной и той же партии обслуживаются прибором непосредственно друг за другом либо одновременно. Такая технология приводит к снижению транспортных издержек и затрат на переналадки оборудования, что особенно актуально в настоящее время. Монография [1] посвящена обзору подобного рода задач в различных условиях.

В данной работе исследуется задача составления расписания выполнения заказов некоторого клиента одним производителем. Заказы выполняются последовательно и доставляются клиенту партиями для уменьшения транспортных издержек. В силу того, что заказы доставляются с помощью транспортных средств, вместимость которых чаще всего ограничена, количество заказов в одной партии не должно превышать заданной величины. Число транспортных средств будем считать заданным, поэтому в случае одного клиента транспортные издержки на перевозку продукции можно считать постоянными. Кроме этого, суммарное время доставки заказов клиенту также можно считать постоянным. Поэтому под моментом завершения выполнения заказа будем понимать время завершения обслуживания партии, в которую данный заказ размещен. В качестве критерия оптимизации рассматривается классический критерий минимизации взвешенной суммы моментов завершения выполнения заказов.

Задача в аналогичной постановке с критерием минимизации суммы взвешенного числа не выполненных в срок заказов исследовалась в [2]; показана полиномиальная разрешимость этой задачи для одного потребителя и NP-трудность задачи в случае большего числа потребителей. В [3] показана полиномиальная разрешимость задачи с критерием минимизации максимального запаздывания.

Наиболее близкими к рассматриваемой задаче среди задач из монографии [1] по режиму выполнения заказов (последовательное выполнение, одновременное завершение обслуживания) являются задачи с наличием переналадок между партиями и неограниченным размером партий.

В первом параграфе дана постановка задачи, во втором — приведен алгоритм динамического программирования решения задачи распределения заказов по партиям при известной последовательности их выполнения. Третий параграф посвящен исследованию сложности задачи: показана полиномиальная разрешимость задачи для критерия минимизации суммы моментов завершения заказов; во взвешенном случае показана NP-трудность задачи. В четвертом параграфе предложен полиномиальный алгоритм нахождения приближенного решения, приведены результаты вычислительного эксперимента.

1. Постановка задачи. Опишем формальную постановку задачи. Производителю необходимо выполнить N заказов некоторого клиента. Для каждого заказа j известны длительность p_j его выполнения и весовой коэффициент (показатель относительной важности) w_j. Выполнение заказов производителем происходит последовательно. Последовательность заранее неизвестна, и ее требуется определить. Для доставки клиенту готовые заказы распределяются по партиям. Моментом завершения выполнения заказа считается время завершения выполнения всех заказов партии, в которую данный заказ размещен. Количество партий ограничено числом K, в каждой партии должно быть не более m заказов (предполагается, что mK^3N). Эти условия связаны с ограниченным количеством транспортных средств у производителя для доставки заказов клиенту, а также с ограниченной вместимостью транспортных средств.

Расписание определяется перестановкой $\pi=(\pi_1, \pi_2, ..., \pi_N)$, задающей последовательность выполнения заказов, и вектором $n=(n_1, n_2, ..., n_K)$, где n_k — количество заказов в k-ой партии, $k=1, ..., K$. При этом набор векторов (π, n) задает допустимое расписание, если $n_k \le m$, $k=1, ..., K$, и $\sum_{k=1}^{K} n_k = N$. Обозначим через c_{π_j} момент завершения выполнения заказа π_j в решении (π, n).

В качестве критерия рассматривается минимизация взвешенной суммы $C_{\Sigma w}(\pi, n)=\sum_{j=1}^{N} w_{\pi_j} c_{\pi_j}$ моментов завершения выполнения заказов. При известной последовательности выполнения заказов задача может быть решена за полиномиальное время алгоритмом динамического программирования [4], который описан в следующем параграфе.

2. Случай известной последовательности выполнения заказов. Приведем алгоритм динамического программирования для нахождения оптимального вектора распределения заказов по партиям при условии, что последовательность π выполнения заказов известна [4]. Без ограничения общности, будем считать, что $\pi=(1, 2, ..., N)$.

Идеей алгоритма является последовательное вычисление оптимального значения критерия $C_{\Sigma w}$ для первых i заказов, отправленных в k партиях. Состояниями динамической системы являются векторы (k, i), где i — количество заказов, которые необходимо распределить по k партиям. Управлением n_k, переводящим систему в состояние (k, i), является количество заказов в партии k. То есть под воздействием управления n_k система переходит из состояния $(k-1, i-n_k)$ в состояние (k, i). Пусть $C_{\Sigma w}(k, i)$ — оптимальное значение целевой функции для задачи с заказами 1, 2,..., i при использовании k партий. Выпишем рекуррентные соотношения:

$$C_{\sum w}(k, i)=\min_{1\le n_k \le \min\{m,i\}}\left(\sum_{q=i-n_k+1}^{i} w_q \sum_{l=1}^{i} p_l + C_{\sum w}(k-1, i-n_k)\right),$$

$$i=1, ..., N;\ k=1, ..., K;$$

$$C_{\sum w}(0, 0)=0\ ;\ C_{\sum w}(0, i)=+\infty,\ i=1, ..., N.$$

Алгоритм реализует стандартную схему динамического программирования. Начиная свою работу с состояния (0,0), алгоритм перебирает все состояния (k, i) в порядке лексикографического возрастания и вычисляет значения $C_{\Sigma w}(k, i)$ до тех пор, пока не дойдет до состояния (K, N). Полученное значение $C_{\Sigma w}(K, N)$ является оптимальным значением целевой функции задачи. Обратный ход производится стандартным образом.

Трудоёмкость алгоритма составляет $O(NKm)$ операций. Длина входа задачи зависит полиномиально от величин $\log_2 m$ и $\log_2 K$, но так как по смыслу задачи величины K и m можно оценить сверху числом N заказов, алгоритм является полиномиальным.

Далее будем обозначать через n^π оптимальный вектор распределения заказов по партиям, соответствующий последовательности π.

3. Сложность задачи. В данном параграфе исследуется вычислительная сложность рассматриваемой задачи. Установлено, что при $w_j=1$, то есть в невзвешенном случае, задача полиномиально разрешима, а во взвешенном случае — NP-трудна.

Теорема 1. *Последовательность* $\alpha=(\alpha_1, \alpha_2, ..., \alpha_N)$, *для которой выполняется* $p_{\alpha_1} \le p_{\alpha_2} \le ... \le p_{\alpha_N}$, *является оптимальной для задачи составления расписания выполнения заказов партиями ограниченного размера с критерием минимизации суммы моментов завершения.*

Доказательство. Пусть указанная в теореме последовательность α не является оптимальной. Рассмотрим некоторую оптимальную последовательность $\beta=(\beta_1, \beta_2, ..., \beta_N)$ и соответствующий оптимальный вектор n^β распределения заказов по партиям. Для последовательности β не выполняется условие теоремы, поэтому существует такой номер j, для которого $p_{\beta_j} > p_{\beta_{j+1}}$. Определим последовательность $\overline{\beta}=(\beta_1, \beta_2, ..., \beta_{j-1}, \beta_{j+1}, \beta_j, \beta_{j+2}, ..., \beta_N)$.

Обозначим через $C_\Sigma(\pi, n)$ значение целевой функции некоторого допустимого решения (π, n). Здесь вектор n — не обязательно оптимальный для последовательности π. Заметим, что $C_\Sigma(\pi, n)=\sum_{k=1}^{K} n_k C_k(\pi, n)$, где $C_k(\pi, n)$ — момент завершения партии k в решении (π, n).

Оценим разность $C_\Sigma(\beta,\ n^\beta)-C_\Sigma(\overline{\beta},\ n^\beta)$. Если заказы β_j и β_{j+1} согласно n^β попали в одну партию, то эта разность равна нулю. Пусть данные заказы попадают в разные партии k и $k+1$. Очевидно, перестановка заказов повлияет лишь на момент завершения партии k, а именно $C_k(\overline{\beta},\ n^\beta)=C_k(\beta,\ n^\beta)+p_{\beta_{j+1}}-p_{\beta_j}$. Имеем: $C_k(\beta,\ n^\beta)-C_k(\overline{\beta},\ n^\beta)=n_k^\beta(p_{\beta_j}-p_{\beta_{j+1}})>0$, таким образом, в любом случае выполняется $C_\Sigma(\beta,\ n^\beta)\geq C_\Sigma(\overline{\beta},\ n^\beta)\geq$ $\geq C_\Sigma(\overline{\beta},\ n^{\overline{\beta}})$. Следовательно, последовательность $\overline{\beta}$ также определяет оптимальное решение, но заказы β_j и β_{j+1} стоят в порядке возрастания длительностей. Повторяя рассуждения для последовательности $\overline{\beta}$, в итоге получим последовательность α, указанную в теореме, которая также должна быть оптимальной. Противоречие. Теорема 1 доказана.

Таким образом, задача с критерием минимизации суммы моментов завершения является полиномиально разрешимой. Для получения оптимального решения достаточно упорядочить заказы по неубыванию их длительностей, затем по данной последовательности вычислить оптимальный вектор распределения заказов по партиям с помощью алгоритма динамического программирования.

В отличие от невзвешенного критерия задача минимизации взвешенной суммы моментов завершения заказов NP-трудна.

Теорема 2. *Задача составления расписания выполнения заказов партиями ограниченного размера с критерием минимизации взвешенной суммы моментов завершения является NP- трудной.*

Доказательство. Докажем, что к рассматриваемой задаче полиномиально сводится NP-полная задача «Разбиение», которая состоит в следующем. Имеется $2n$ чисел $a_i\in\mathbf{Z}^+$, сумма которых равна $2A$. Вопрос задачи заключается в существовании подмножества S' множества номеров $S=\{1,\ ...,\ 2n\}$ такого, что $\sum_{i\in S'}a_i=\sum_{i\in S\setminus S'}a_i$. Данная задача остается NP-полной, даже если потребовать $|S'|=n$ [5].

Построим класс примеров рассматриваемой задачи составления расписания. Пусть количество заказов $N=2n$, количество партий $K=2$, размер партии $m=n$. Для каждого заказа $j=1,\ ...,\ N$ зададим длительность выполнения $p_j=a_j$ и весовой коэффициент $w_j=a_j$.

Докажем, что в задаче «Разбиение» будет ответ «да» тогда и только тогда, когда в соответствующей задаче построения расписания существует допустимое решение со значением целевой функции, не превышающим $3A^2$.

Пусть в задаче «Разбиение» существует подмножество S' такое, что $\sum_{i\in S'}a_i=\sum_{i\in S\setminus S'}a_i$. Рассмотрим допустимое решение в задаче построения расписания, в котором в первой партии выполняются заказы $j\in S'$, а во второй — все остальные. Значение целевой функции данного решения $C_{\Sigma w}=\sum_{j=1}^{N}w_jc_j=\sum_{j\in S'}a_jA+$ $+\sum_{j\in S\setminus S'}2a_jA=3A^2$, что и требовалось установить.

Докажем в обратную сторону. Пусть удалось построить допустимое расписание, для которого $C_{\Sigma w}\leq 3A^2$. Пусть Y — множество номеров заказов первой партии, а T — момент завершения ее выполнения. Тогда

$$C_{\Sigma w}=\sum_{j\in Y}a_jT+\sum_{j\in S\setminus Y}2a_jA=T^2+2A\cdot(2A-T)=T^2-2AT+4A^2.$$

Легко проверить, что из неравенства $C_{\Sigma w}\leq 3A^2$ следует $T=A$. В силу допустимости решения $|Y|=$ $=|S\setminus Y|=n$. Положив в задаче «Разбиение» $S'=Y$, получим требуемое разбиение множества чисел a_i, $i=1,\ ...,\ N$. Полиномиальность сводимости очевидна. Теорема 2 доказана.

В силу труднорешаемости рассматриваемой задачи актуальными направлениями ее исследования являются разработка алгоритмов нахождения приближенного решения и выделение малотрудоемких частных случаев. В следующем параграфе предлагается алгоритм нахождения приближенного решения задачи.

4. Приближенное решение задачи. Идея алгоритма состоит в локальном улучшении текущей последовательности заказов. Начиная с последовательности, в которой заказы упорядочены по неубыванию отношений длительностей к весам, пытаемся с помощью перестановки двух заказов соседних партий либо перемещения некоторого заказа в соседнюю партию улучшить значение целевой функции.

Пусть S_k — множество номеров заказов партии k, $C_{\Sigma w}^{rec}$ — текущий рекорд целевой функции (вначале полагаем равным $+\infty$) и $(\pi^{rec},\ n^{rec})$ — соответствующее решение.

Алгоритм.

Шаг 1. Упорядочиваем заказы по неубыванию отношений длительностей к весам. Получаем последовательность π, в которой $\frac{p_{\pi_1}}{w_{\pi_1}}\leq\frac{p_{\pi_2}}{w_{\pi_2}}\leq...\leq\frac{p_{\pi_N}}{w_{\pi_N}}$. Переходим на шаг 2.

Шаг 2. С помощью алгоритма динамического программирования (п. 2) вычисляем для перестановки π соответствующий оптимальный вектор n^π распределения заказов по партиям и значение целевой функции $C_{\Sigma w}(\pi,\ n^\pi)$. Если $C_{\Sigma w}^{rec}>C_{\Sigma w}(\pi,n^\pi)$, то обновляем $C_{\Sigma w}^{rec}$, π^{rec} и n^{rec}. Переходим на шаг 3.

Шаг 3. Для каждой партии $k=1,...,\ K-1$ и для всех q, l, где $q\in S_k$, $l\in S_{k+1}$, проверяем условие: $\left(\sum_{j\in S_k\setminus\{q\}}w_j+w_l\right)(p_q-p_l)<\left(\sum_{i\in S_{k+1}}p_i\right)(w_q-w_l)$. Если для некоторых q и l условие не выполняется, меняем их местами и переходим на шаг 2.

Шаг 4. Для каждой партии $k=2,...,\ K$, где $|S_k|<m$, для всех заказов $q\in S_{k-1}$, кроме последнего в партии, проверяем условие $p_q\left(\sum_{j\in S_{k-1}\setminus\{q\}}w_j\right)<w_q\left(\sum_{i\in S_k}p_i\right)$. Если для некоторого q условие не выполняется, то перемещаем заказ q перед первым заказом партии k и переходим на шаг 2.

Шаг 5. Для каждой партии $k=1,...,\ K-1$, где $|S_k|<m$, для всех заказов $l\in S_{k+1}$, кроме первого в партии, проверяем условие $p_l\left(\sum_{j\in S_k}w_j\right)>w_l\left(\sum_{i\in S_{k+1}\setminus\{l\}}p_i\right)$.

Если для некоторого l условие не выполняется, то перемещаем заказ l после последнего заказа партии k и переходим на шаг 2.

Конец алгоритма.

Описанный алгоритм находит приближенное решение для рассматриваемой задачи за $O(NK^3m^3)$ операций и, таким образом, является полиномиальным.

Данный алгоритм был реализован. Для проверки его работоспособности и анализа погрешности получаемых решений проведен вычислительный эксперимент. Для получения оптимального решения был реализован алгоритм полного перебора последовательностей заказов с последующим вычислением оптимального вектора распределения заказов по партиям. Вследствие этого сравнение удалось провести лишь для небольших размерностей (8 – 10 заказов).

Были сгенерированы случайным образом 32 задачи. Эксперимент показал, что в 26 задачах алгоритм нашел оптимальное решение. В остальных задачах относительная погрешность найденного решения составила от 1,53 % до 4,15 %. В ходе эксперимента установлено, что алгоритм не находит оптимального решения, если имеются заказы, длительности и весовые коэффициенты которых принимают близкие друг к другу значения.

Заключение. В работе исследована сложность задачи выполнения заказов партиями ограниченного размера с критерием минимизации взвешенной суммы моментов завершения. Установлено, что в невзвешенном случае задача полиномиально разрешима. Для получения оптимальной последовательности выполнения заказов достаточно их упорядочить по неубыванию длительностей. Данная последовательность является оптимальной и для известной задачи директора $1\|C_\Sigma$, которая близка к рассматриваемой по критерию оптимизации. Однако для взвешенного случая доказана NP-трудность задачи, что отличает ее от задачи директора $1\|C_{\Sigma w}$, которая и во взвешенном случае полиномиально разрешима. В работе также выявлена структура трудных примеров исследуемой задачи.

Для взвешенного случая предложен алгоритм построения приближенного решения, показавший хорошие результаты на случайных данных.

Библиографический список

1. Танаев, В. С. Теория расписаний. Групповые технологии / В. С. Танаев, М. Я. Ковалев, Я. М. Шафранский. – Минск : Ин-т техн. кибернетики НАН Беларуси, 1998. – 290 с.
2. Wang, Y. Models and Algorithms for Some Combinatorial Optimization Problems: University Course Timetabling, Facility Layout and Integrated Production // Grado Department of Industrial and Systems Engineering. – Virginia Tech, Blackburg, USA, 2007. – 140 p.
3. Здорова, И. В. Исследование одной задачи планирования производства и доставки продукции партиями / И. В. Здорова // Молодёжь третьего тысячелетия : материалы XXXV Регион. науч.-практ. студ. конф. – Омск : ОмГУ, 2011. – С. 353.
4. Романова, А. А. Исследование сложности одной задачи выполнения заказов партиями с различными критериями / А. А. Романова // Дискретная оптимизация и исследование операций : материалы Междунар. конф. – Новосибирск : Изд-во Ин-та математики, 2013. – С. 98.
5. Гэри, М. Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон. – М. : Мир, 1982. – 416 с.

РОМАНОВА Анна Анатольевна, кандидат физико-математических наук, заведующая кафедрой математики и информационных технологий Омской юридической академии; доцент кафедры прикладной и вычислительной математики Омского государ-ственного университета им. Ф. М. Достоевского.

Адрес для переписки: anna.a.r@bk.ru

Статья поступила в редакцию 19.01.2015 г.

Книжная полка

Кудинов, И. В. Аналитические решения параболических и гиперболических уравнений тепломассопереноса : учеб. пособие для вузов по направлению подготовки бакалавров и магистров в обл. техн. наук и по направлению подготовки дипломир. специалистов в обл. техники и технологии / И. В. Кудинов, В. А. Кудинов ; под ред. Э. М. Карташова. – М. : ИНФРА-М, 2015. – 389. – ISBN 978–5–16–006724–7.

Излагаются инженерные методы построения решений задач нестационарной теплопроводности, позволяющие получать эффективные точные и приближенные аналитические решения. При определении собственных чисел вводятся дополнительные граничные условия, получаемые из основного дифференциального уравнения путем его дифференцирования в граничных точках. С помощью интегрального метода теплового баланса на основе введения фронта температурного возмущения и при использовании дополнительных граничных условий получены аналитические решения задач теплопроводности с переменными начальными условиями, с переменными во времени граничными условиями и внутренними источниками теплоты, нелинейных задач теплопроводности, а также задач теплообмена в жидкостях, включая динамический и тепловой пограничные слои. Представлены результаты получения и анализа точных аналитических решений гиперболических уравнений, описывающих распространение тепловой и гидравлической волны с конечной скоростью.

Юрьева, А. А. Математическое программирование : учеб. пособие для вузов по направлению и специальности «Прикладная математика» / А. А. Юрьева. – 2-е изд., испр. и доп. – СПб. [и др.] : Лань, 2014. – 431 с. – ISBN 978–5–8114–1585–4.

Учебное пособие состоит из семи разделов. Три раздела посвящены математическому программированию, два — теории игр, два — теории графов и сетей. Основное внимание уделено прикладному аспекту. Все методы решения иллюстрируются типовыми примерами, а в конце каждой главы приведены упражнения (25 – 30 вариантов) для самостоятельной работы студентов. Задачи данных упражнений в основном оригинальны и лишь некоторые взяты из источников, указанных в списке литературы. По объему информации учебное пособие соответствует курсу математического программирования, читаемому во всех технических и экономических вузах страны. Рекомендовано для студентов, обучающихся по направлениям подготовки «Математика и компьютерные науки», «Прикладная математика и информатика», «Прикладная информатика», «Информационные системы и технологии», «Информатика и вычислительная техника».