ЭКОНОМИКА И МАТЕМАТИЧЕСКИЕ МЕТОДЫ 1968 том IV, вып. 6

НЕКОТОРЫЕ УСЛОВИЯ ПАРАЛЛЕЛЬНОЙ РАБОТЫ УСТРОЙСТВ В СИСТЕМЕ ОБРАБОТКИ ИНФОРМАЦИИ

Ю. В. ГЕРОНИМУС, Н. Б. ДЕМИДОВИЧ

(Москва)

Рассматривается система обработки информации, состоящая из центрального устройства (ЦУ) и нескольких внешних устройств (ВУ), среди которых могут быть и однотипные.

Обозначим через ВУ$_{pq}$ p-е устройство q-го типа, $q = 1, 2, \ldots, m$; $p = 1, 2, \ldots, l_q$. Каждому ВУ$_{pq}$ приписан приоритет П$_{pq}$, определяемый формулой *

$$\Pi_{pq} = \sum_{i=1}^{q-1} l_i + p.$$

Каждое ВУ имеет местную систему управления, позволяющую ему работать независимо от других ВУ и от ЦУ. Однако любое работающее внешнее устройство ВУ$_{pq}$ через каждые t_q единиц времени заканчивает свой «квант» автономной работы и прерывает ЦУ. В результате этого прерывания на ЦУ должна быть выполнена некоторая программа обслуживания, использующая полученную порцию информации (или подготавливающая очередную порцию информации). Эта программа обозначается в дальнейшем через ОП$_{pq}$ **. На выполнение ОП$_{pq}$ затрачивается τ_q единиц времени. После выхода из ОП$_{pq}$ ЦУ возобновляет свою работу по прерванной программе.

Если включено одно лишь ВУ$_{pq}$, то условием правильной работы системы, состоящей из ЦУ и ВУ$_{pq}$, является выполнение неравенства

$$\tau_q \leqslant T_q. \tag{1}$$

Значение T_q зависит от технических характеристик ВУ q-го типа. Если, например, ВУ q-го типа есть устройство ввода, то T_q — интервал времени между окончанием поступления очередной порции информации с внешнего носителя в буфер ВУ$_{pq}$ и началом поступления следующей порции информации.

При $\tau_q > T_q$, в зависимости от конструкции ВУ q-го типа, имеет место либо потеря информации, либо простаивание ВУ.

Предположим теперь, что одновременно включена некоторая группа Г устройств. В ходе выполнения ОП$_{rj}$ (ВУ$_{rj} \in$ Г) может появиться сигнал прерывания от другого устройства. Внешние устройства упорядочены в системе в том смысле, что сигнал прерывания от ВУ$_{pq}$, ВУ$_{pq} \in$ Г, появившийся во время выполнения ОП$_{rj}$, не приводит при $\Pi_{rj} > \Pi_{pq}$ к немедленному прерыванию ОП$_{rj}$; сигнал прерывания от ВУ$_{pq}$ запоминается,

---

* Принципы упорядочивания типов ВУ в статье не рассматриваются.

** Все ОП$_{pq}$ при фиксированном q одинаковы.

и $ОП_{pq}$ начинает выполняться не раньше, чем закончится выполнение $ОП_{rj}$. Если же $П_{pq} > П_{rj}$, то выполнение $ОП_{rj}$ прерывается немедленно и управление передается $ОП_{pq}$. Выполнение $ОП_{rj}$ будет продолжено лишь после того, как выполняется все $ОП_{pq}$, $ВУ_{pq} \in Г$, $П_{rj} < П_{pq}$, вызванные к работе своими сигналами прерывания. Такого рода вмешательства в работу $ОП_{rj}$ могут происходить в ходе ее выполнения несколько раз. Таким образом, время между началом работы $ОП_{rj}$ и моментом выхода из нее оказывается, вообще говоря, в силу возможных вмешательств большим, чем τ_q.

Обозначим через θ_{rj} интервал времени между моментом возникновения некоторого сигнала прерывания от $ВУ_{rj}$ и моментом выхода из $ОП_{rj}$, вызванной к действию этим сигналом. Необходимое и достаточное условие правильной работы системы, состоящей из ЦУ и группы ВУ, заключается в том, что для всех ВУ из Г и для всех сигналов прерывания, возникающих в период параллельной работы,

$$\theta_{rj} \leqslant T_j, \quad ВУ_{rj} \in Г. \tag{2}$$

Выразим достаточные условия правильной работы системы через величины t_q, T_q и τ_q.

Для любого сигнала прерывания от $ВУ_{rj} \in Г$ значение θ_{rj} удовлетворяет уравнению

$$\theta_{rj} = \tau_j + \sum_{\substack{П_{rj}<П_{pq}\\ ВУ_{pq}\in Г}} \alpha_{rj}^{pq}(\theta_{rj}). \tag{3}$$

Через $\alpha_{rj}{}^{pq}(x)$ обозначено время, затраченное на выполнение $ОП_{pq}$ за время x, отсчитанное от момента возникновения рассматриваемого сигнала прерывания от $ВУ_{rj}$. Значение $\alpha_{rj}{}^{pq}(x)$ зависит не только от x, но и от

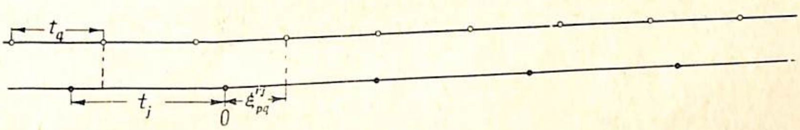

Рис. 1

сдвига $\xi_{rj}{}^{pq}$ последовательности сигналов прерывания от $ВУ_{pq}$ относительно момента возникновения рассматриваемого сигнала прерывания от $ВУ_{rj}$. Сдвиг $\xi_{rj}{}^{pq}$ различен, вообще говоря, для различных сигналов прерывания от $ВУ_{rj}$ и может считаться в нашей задаче случайной величиной, распределенной на интервале $(0, t_q)$ (см. рис. 1; светлые кружки означают сигналы прерывания от $ВУ_{pq}$, темные — от $ВУ_{rj}$).

Условие (2) будет заведомо выполнено, если для всех сигналов прерывания от ВУ будет иметь место неравенство

$$T_j \geqslant \tau_j + \sum_{\substack{П_{rj}<П_{pq}\\ ВУ_{qp}\in Г}} \alpha_{rj}{}^{pq}(T_j); \quad ВУ_{rj} \in Г. \tag{4}$$

Найдем достаточные условия выполнения (4) для любого ВУ из Г при любом его сигнале прерывания.

Определим сначала числа k_{jq}, $j = 1, 2, \ldots, m$; $q \geqslant j$, удовлетворяющие неравенствам

$$\alpha_{rj}{}^{pq}(T_j) \leqslant k_{jq}; \quad П_{pq} > П_{rj} \tag{5}$$

при любых сдвигах последовательности сигналов прерывания от ВУ$_{pq}$ относительно момента возникновения сигнала прерывания от ВУ$_{rj}$.

На рис. 2 через 0 обозначен момент возникновения некоторого произвольно выбранного сигнала прерывания внешнего устройства ВУ$_{rj}$. На прямой, параллельной оси $0t$ и проходящей через точку A_ξ, изображена последовательность сигналов прерывания от ВУ$_{pq}$, сдвинутая на ξ относи-

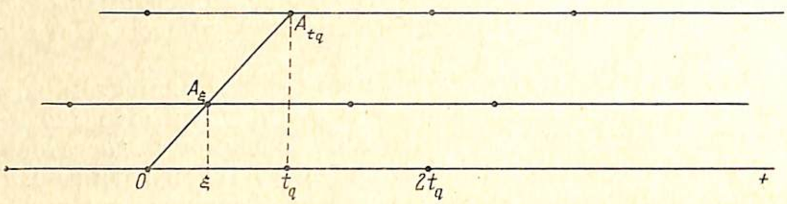

Рис. 2

тельно момента 0. Таким образом, на прямых, параллельных $0t$ и пересекающих отрезок $[0, A_{t_q}]$, изображены все возможные расположения последовательности сигналов прерывания от ВУ$_{pq}$ относительно 0. Ордината точки A_{t_q} выбрана равной t_q. Прямая, проходящая через A_{t_q}, несет на себе последовательность, соответствующую $\xi = t_q$, которую можно отождествить с последовательностью для $\xi = 0$.

Введем обозначения *

$$e = E(T_j / t_q); \qquad \beta = T_j - t_q \cdot e; \tag{6}$$

$$\lambda = t_q - T_r; \qquad h = \beta - \lambda.$$

Пусть

$$\beta < T_q. \tag{7}$$

Этот случай изображен на рис. 3. Лежащие на прямых, параллельных оси $0t$, интервалы времени продолжительностью T_q, начинающиеся в мо-

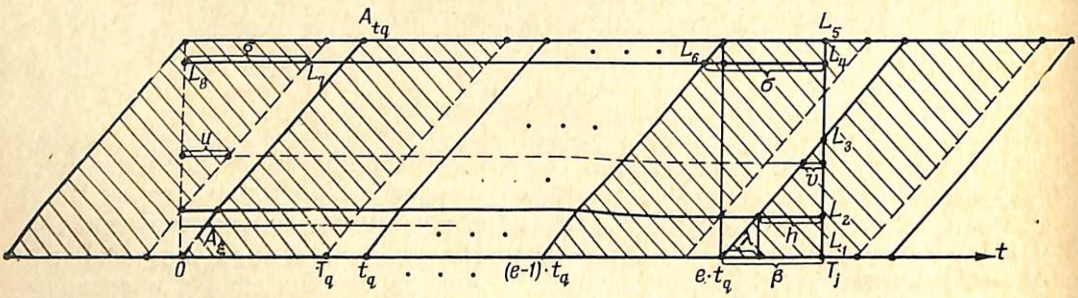

Рис. 3

менты возникновения сигналов прерывания от ВУ$_{pq}$, образуют заштрихованные области. Нас будут интересовать заштрихованные области, расположенные над отрезком $[0, T_j]$.

Дальнейшие оценки основаны на том соображении, что на всяком интервале времени из заштрихованной области продолжительностью ζ ОП$_{pq}$ может выполняться не более чем ζ единиц времени, если $\zeta < \tau_q$ и не более чем τ_q единиц времени, если $\zeta \geqslant \tau_q$. Неравенство (2) для ВУ$_{pq}$ предполагается при этом выполненным. Из условия (7) вытекает, что существует горизонтальная прямая (она отмечена на рис. 3 буквой L_4), для которой

* Индексы у переменных, определяемых формулами (6) опущены для сокращения письма.

отрезки L_6L_4 и L_8L_7 имеют одинаковую длину σ, меньшую T_q и определяемую, как легко увидеть из чертежа, формулой

$$\sigma = {}^1/_2(\beta + T_q). \tag{8}$$

(Этой формулой удобно пользоваться и при $\beta > T_q$, хотя при этом первоначальный смысл σ теряется.)

Рассмотрим два случая. Первый из них определяется неравенством

$$\sigma < \tau_q. \tag{9}$$

Из (8) следует, что в этом случае и

$$\beta < \tau_q. \tag{10}$$

Рассмотрим последовательно все варианты сдвигов серии сигналов прерывания от ВУ$_{pq}$ относительно момента 0 (характеризуемые различными горизонтальными прямыми, отмеченными точками отрезка L_1L_5). Рассмотрение проводится для изображенного на рис. 3 подслучая, когда $h > 0$. При $h < 0$ рассуждения и результаты сохраняются.

Для сдвигов, соответствующих прямым, расположенным между L_1 и L_2, в силу (10) имеет место неравенство

$$\alpha_{rj}{}^{pq}(T_j) \leqslant e\tau_q + \beta.$$

Для сдвигов, соответствующих прямым, расположенным между L_2 и L_3, в силу того, что $h < \beta < \tau_q$, имеет место оценка

$$\alpha_{rj}{}^{pq}(T_j) \leqslant e\tau_q + h.$$

Для прямых, расположенных между L_3 и L_4 и между L_4 и L_5, выполняется неравенство

$$\alpha_{rj}{}^{pq}(T_j) \leqslant e\tau_q + \sigma.$$

Ввиду того, что $\sigma > \beta > h$, в случае (9) следует положить

$$k_{jq} = e\tau_q + \sigma. \tag{11}$$

Пусть теперь

$$\sigma > \tau_q. \tag{12}$$

Если при этом

$$h < \tau_q, \tag{13}$$

то максимум $\alpha_{rj}{}^{pq}(T_j)$ достигается на прямой L_4. Таким образом, k_{jq} можно определить по формуле

$$k_{jq} = (e + 1)\tau_q; \tag{14}$$

этот же результат сохраняется и при $h < 0$.

Рассмотрим случай

$$\tau_q \leqslant h < 2\tau_q. \tag{15}$$

Для прямых, лежащих между L_1 и L_2, имеет место оценка

$$\alpha_{rj}{}^{pq}(T_j) \leqslant (e + 1)\tau_q, \tag{16}$$

так как β теперь заведомо больше τ_q.

Относительно прямых, лежащих между L_2 и L_3, заметим, что если $u \geqslant \tau_q$, то $v < \tau_q$, и наоборот, так как

$$u + v = h < 2\tau_q.$$

Поэтому

$$\alpha_{rj}{}^{pq}(T_j) \leqslant e\tau_q + h.$$

На участке L_3L_5 имеет место оценка (16). Таким образом, при условии (15)

$$k_{jq} = e\tau_q + h. \tag{17}$$

Если

$$h \geqslant 2\tau_q,$$

то, как нетрудно показать, надо положить

$$k_{jq} = (e + 2)\tau_q. \tag{18}$$

При $\beta > T_q$ выполняется условие (12). Все оценки, выведенные только что для этого случая, сохраняются.

Воспользовавшись языком АЛГОЛ-60, приведенные выше результаты можно выразить следующей последовательностью формул

$e: = E(T[j]\,/\,t[q])$; $\beta: = T[j] - t[q] \times e$; $\lambda: = t[q] - T[q]$; $h: = \beta - \lambda$; $\sigma: = (\beta + T[q])\,/\,2$; $k[j, q]: =$ **если** $\sigma > \tau[q]$, **то** $e \times \tau[q] + \sigma$ **иначе если** $h < \tau[q]$ **то** $(e + 1) \times \tau[q]$ **иначе если** $h < 2 \times \tau[q]$ **то** $e \times \tau[q] + h$ **иначе** $(e + 2) \times \tau[q]$

Группе Г параллельно работающих устройств соответствует вектор y с целочисленными компонентами

$$y = \{y_1, y_2, \ldots, y_m\},$$

где y_j равен числу устройств j-го типа, включенных в Г. Компоненты y_j удовлетворяют неравенствам

$$0 \leqslant y_j \leqslant l_j, \quad j = 1, 2, \ldots, m.$$

В силу (4) и (5) условия (4) будут выполнены для любого устройства j-го типа из Г, если выполняются неравенства

$$T_j \geqslant \tau_j + k_{jj}(y_j - 1) + \sum_{i=j+1} y_i k_{ji}. \tag{19}$$

Если положить

$$p_j = T_j - \tau_j + k_{jj},$$

то неравенство (19) приобретает вид

$$p_j \geqslant \sum_{i=j}^{m} k_{ji} y_i. \tag{20}$$

При $y_j = 0$ требовать выполнения неравенства (20), разумеется, не следует. Отметим, что задание вектора y эквивалентно заданию Г (с точностью до несущественного для нас выбора одного из $C_{l_j}^{y_j}$ сочетаний устройств j-го типа; $j = 1, 2, \ldots, m$; $y_j \neq 0$).

Рассмотрим два частных случая:

а) пусть Г определена вектором

$$y = \{y_1, \ldots, y_m\}, \quad l_j = 1; \quad j = 1, 2, \ldots, m.$$

Пусть для любой пары y_i, y_j, такой, что $y_i = y_j = 1$, $i > j$, отношение $T_j\,/\,t_i$ — большое число. Тогда можно полагать приближенно

$$k_{ji} = (T_j\,/\,t_i)\tau_i,$$

и условие (20) приобретает вид

$$\frac{\tau_j}{T_j} \leqslant 1 - \sum_{i=j+1}^{m} y_i \frac{\tau_i}{t_i};$$

б) пусть все компоненты вектора y, кроме y_j, равны 0; $y_j \geqslant 2$, тогда $E(T_j / t_j) = 0$ и $\beta = T_j$. Поэтому $\sigma > T_j$. При $h = 2T_j - t_j < \tau_j$ имеем $k_{jj} = \tau_j$, и условие (20) принимает вид

$$T_j \geqslant y_j \tau_j.$$

При $\tau_j \leqslant 2T_j - t_j < 2\tau_j$ получаем $k_{jj} = 2T_j - t_j$ и условие (20) имеет вид

$$T_j \geqslant \tau_j + (y_j - 1)(2T_j - t_j).$$

При $h \geqslant 2\tau_j$ $k_{jj} = 2\tau_j$, и условие (20) имеет вид

$$T_j \geqslant (2y_j - 1)\tau_j.$$

На группу Г, определяемую вектором y, кроме неравенств (20) следует наложить еще ограничения, связанные с работой ЦУ. Эти ограничения зависят от назначения системы. Обозначим через $T_{\text{ЦУ}}(t)$ часть времени из интервала $(0, t)$, остающуюся у ЦУ для выполнения работ, отличных от обслуживания ВУ из Г. Если t достаточно велико, то можно положить

$$T_{\text{ЦУ}}(t) = t - \sum_{j=1}^{m} \frac{t}{t_j} \tau_j y_j.$$

Простейшее ограничение, связанное с ЦУ, состоит в требовании, чтобы в среднем за единицу времени доля, приходящаяся на выполнение ОП, не превосходила $\varepsilon < 1$. Это требование выполняется, если

$$\sum_{j=1}^{m} \frac{\tau_j}{t_j} y_j \leqslant \varepsilon. \tag{21}$$

Можно представить себе и более жесткие ограничения на $T_{\text{ЦУ}}(t)$. Пусть, например, известно, что за время t работы Г на обработку каждой единицы информации, введенной через ВУ j-го типа, требуется в среднем λ_j единиц времени работы ЦУ. Для того чтобы в памяти системы не накапливалась введенная, но еще не переработанная информация, требуется, чтобы выполнялось неравенство

$$\sum_{j=1}^{m} \left(\lambda_j I_j + \frac{\tau_j}{t_j} \right) y_j \leqslant 1, \tag{21$'$}$$

где I_j — количество единиц информации, поступающих в систему через устройство j-го типа за единицу времени. (Для выходных устройств полагается $I_j = 0$.)

Аналогично, пусть известно, что за каждую единицу времени работы ЦУ к выводу через устройство j-го типа подготавливается в среднем μ_j единиц информации. Для того чтобы память системы не переполнялась невыведенными результатами, следует потребовать выполнения условия

$$O_j y_j \geqslant \mu_j \left(1 - \sum_{k=1} \frac{\tau_k}{t_k} y_k \right); \quad j = 1, 2, \ldots, m, \tag{21$''$}$$

где O_j — число единиц информации, выводимых за единицу времени устройством j-го типа. (Для входных устройств $\mu_j = O_j = 0$.) Можно представить себе несколько ситуаций, в которых играют роль условия (20) и условия типа (21). Остановимся на двух таких ситуациях.

1°. Пусть система предназначена для выполнения случайного потока работ. В процессе выполнения ее состав, вообще говоря, не постоянен во времени. Набор параллельно работающих устройств зависит от текущего потока заявок от выполняемых программ и поведения потребителей, подающих на вход системы свои исходные данные. Поток заявок на устройства регулируется специальной управляющей программой (супервайзором). Супервайзор должен содержать блок, проверяющий перед каждым предполагаемым включением одного или нескольких новых устройств, будут ли при намеченном составе Г выполняться условия (20) и условия типа (21). Если какие-либо из этих условий не выполняются, то супервайзор откладывает выполнение заявок, вызывающих расширение состава Г, до момента, когда одно или несколько устройств из Г выключаются (вследствие, например, окончания работы над очередным заказом). В этот момент следует снова при помощи условий (20) и условий типа (21) проверить, какими устройствами можно «догрузить» Г.

2°. Пусть система предназначена для обработки данных, характеризуемой устойчивыми значениями λ_j и μ_j. Такой характер обработки данных дает возможность включить в постоянную параллельную работу некоторую группу внешних устройств Г и ЦУ. Возникает вопрос о наилучшем комплектовании Г. Можно, например, выбирать Г так, чтобы доставить максимум количеству единиц информации, перерабатываемой системой за единицу времени. Другими словами, надо найти целые неотрицательные y_j, $j = 1, 2, \ldots, m$, дающие линейной форме

$$I = \sum_{j=1}^{m} I_j y_j \tag{22}$$

максимальное значение при ограничениях (20) и (21) *. Для того чтобы свести поставленную задачу к задаче линейного целочисленного программирования, надо еще несколько видоизменить условия (20). В первоначальной формулировке они нелинейны, так как должны выполняться только для индексов j, удовлетворяющих неравенству $y_j \geqslant 1$. Для «линеаризации» условий (20) введем m групп вспомогательных целочисленных переменных z_j, σ_{j1}, σ_{j2}, $\ldots$, $\sigma_{j(l_j-1)}$; $j = 1, 2, \ldots, m$, удовлетворяющих ограничениям:

$$\begin{gathered} 0 \leqslant z_j \leqslant 1; \quad 0 \leqslant \sigma_{jk} \leqslant 1; \, z_j \leqslant 1 - \sigma_{jk}; \\ y_j = 1 - z_j + \sigma_{j1} + \ldots + \sigma_{j(l_j-1)}; \\ j = 1, 2, \ldots, m; \quad k = 1, 2, \ldots, l_j - 1. \end{gathered} \tag{23}$$

Нетрудно видеть, что если $y_j = 0$, то $z_j = 1$, а если $y_j > 0$, то $z_j = 0$. Определим M_j формулами

$$M_j = \sum_{s=j}^{m} k_{js} l_s, \quad j = 1, 2, \ldots, m.$$

Ограничения (20) эквивалентны линейным неравенствам

$$Mz + p \geqslant Ky, \tag{24}$$

где M — диагональная матрица с элементами M_j на главной диагонали, K — треугольная матрица с элементами k_{jq}, $j = 1, 2, \ldots, m$; $j \leqslant q \leqslant m$, а p и z — векторы с компонентами p_j и z_j. Решение (22) дает оптимальное распределение исходной информации по типам носителей.

Поступила в редакцию
12 V 1967

---

* Можно включить в рассмотрение еще и другие ограничения, например, ограничения на стоимость Г. Ничего нового в постановку задачи эти ограничения не вносят.

ЭКОНОМИКА И МАТЕМАТИЧЕСКИЕ МЕТОДЫ — 1968, том IV, вып. 6

НОВЫЕ ПОДХОДЫ К СТОХАСТИЧЕСКОМУ ПРОГРАММИРОВАНИЮ

Д. Б. ЮДИН

(Москва)

1. О ПОСТАНОВКАХ ЗАДАЧ СТОХАСТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Исходная информация для планирования, проектирования и управления в экономике, технике и военном деле, как правило, недостаточно достоверна. Планирование производства обычно производится в условиях неполной информации об обстановке, в которой будет выполняться план и реализовываться произведенная продукция. Работа автоматических устройств сопровождается непредвиденными случайными помехами, статистические закономерности которых не всегда могут быть определены и учтены при вычислении управляющих воздействий. Еще сложней обстоит дело в военных задачах, где помимо естественного недостатка информации возникает и возможность дезинформации. Как правило, эффективная стратегия дезинформации является смешанной стратегией, организуемой при помощи случайных механизмов, статистические характеристики которых удовлетворяют определенным закономерностям.

Таким образом, в моделях математического программирования, к исследованию которых сводятся задачи планирования, проектирования и управления, отдельные или все параметры (или характеристики) показателя качества и ограничений могут оказаться неопределенными или случайными. В одних случаях опыт, статистика и исследование процессов, определяющих изменение исходных данных и формирующих условия, в которых реализуется план, проект или система управления, позволяют устанавливать те или иные вероятностные характеристики параметров задач. В других случаях нет оснований для каких бы то ни было суждений о статистических особенностях явлений, способных изменить предполагаемые значения параметров условий задачи. И те и другие ситуации являются предметом исследования так называемого стохастического программирования.

Стохастическим программированием называют раздел математического программирования, изучающий методы решения условных экстремальных задач при неполной информации.

В стохастическом программировании больше, чем в других разделах математического программирования, значительные трудности подстерегают исследователя не только (а может быть и не столько) при разработке методов решения задач, но и при постановке задач, в которых необходимо отразить подчас довольно тонкие ситуации планирования и управления в условиях риска и неопределенности.

Естественный, на первый взгляд, путь анализа задач стохастического программирования — замена случайных параметров их средними значениями и вычисление оптимальных планов полученных таким образом детерминированных задач — далеко не всегда приводит к приемлемому решению. Отсюда необходимость в разработке моделей планирования, проекти-