

Т. И. АЛИЕВ

ЗАДАЧИ СИНТЕЗА СИСТЕМ С ПОТЕРЯМИ

Формулируются задачи синтеза систем с потерями при наличии различных ограничений на характеристики их функционирования. В процессе синтеза определяются емкость накопителя и производительность устройства, обеспечивающие минимальную стоимость системы.

***Ключевые слова:** система с потерями, накопитель ограниченной емкости, производительность, вероятность потери, время пребывания, стоимость системы.*

Введение. Проектирование дискретных систем со стохастическим характером функционирования, примерами которых могут служить информационно-управляющие системы, серверы локальных вычислительных сетей, сетевое оборудование (маршрутизаторы, коммутаторы) и т.п., осуществляется с использованием различных моделей и методов в зависимости от класса проектируемой системы, ее особенностей и требований, предъявляемых к качеству ее функционирования. При проектировании технических систем в качестве моделей целесообразно использовать системы и сети массового обслуживания с потерями, которые имеют накопители ограниченной емкости [1]. При некоторых предположениях могут быть получены аналитические зависимости для расчета характеристик дискретных систем со стохастическим характером функционирования, что позволяет эффективно решать задачи оптимального синтеза.

К основным характеристикам, определяющим эффективность функционирования систем с потерями, относятся: вероятность потери заявок π , среднее время пребывания заявок в системе u и стоимость системы S .

Требуемое качество функционирования системы, задаваемое в виде ограничений, налагаемых на характеристики, обеспечивается в процессе синтеза за счет выбора структурно-функциональной организации системы.

Описание системы. Положим, что исследуемая система содержит одно устройство, обрабатывающее заявки (запросы, команды, транзакции, детали и т.п.), поступающие в систему в случайные моменты времени из неограниченного источника заявок. Поступающие заявки образуют однородный поток и создают в системе однородную нагрузку. Средний интервал между заявками равен a . Средняя ресурсоемкость, измеряемая количеством работы, которое затрачивается на обработку одной заявки, равна θ . Скорость обработки одной заявки устройством (производительность, или быстродействие, устройства), измеряемая количеством работы, выполняемой устройством за единицу времени, равна V . В каждый момент времени устройство может обрабатывать только одну заявку. Заявки, поступившие в систему во время обработки устройством ранее поступившей заявки, располагаются в накопителе емкостью E , при этом обрабатываемая заявка также занимает одно место в накопителе. Заявка, поступившая в систему, теряется, если накопитель заполнен.

Расчет характеристик системы с потерями. Положим, что в систему поступает простейший поток заявок с интенсивностью $\lambda = 1/a$, длительность обработки которых распределена по экспоненциальному закону со средним значением b , причем нагрузка может принимать любые положительные значения:

$$y = \frac{\lambda\theta}{V} > 0. \quad (1)$$

В этом случае вероятность того, что в системе с накопителем емкостью E находится k заявок $k = 0, 1, 2, \dots, E$, рассчитывается в зависимости от значения нагрузки:

$$p_k = \begin{cases} \frac{y^k (1-y)}{1-y^{E+1}}, & y \neq 1, \\ \frac{1}{E+1}, & y = 1. \end{cases} \quad (2)$$

Тогда вероятность потери заявки из-за ограниченной емкости накопителя:

$$\pi = \begin{cases} \frac{y^E (1-y)}{1-y^{E+1}}, & y \neq 1, \\ \frac{1}{E+1}, & y = 1. \end{cases} \quad (3)$$

Соответственно загрузка системы составит

$$\rho = \lambda_0 b = \begin{cases} \left(1 - \frac{y^E (1-y)}{1-y^{E+1}}\right) y, & y \neq 1, \\ \frac{E}{E+1}, & y = 1, \end{cases} \quad (4)$$

где $\lambda_0 = (1-\pi)\lambda$ — интенсивность потока обслуженных заявок.

Среднее число заявок, находящихся в системе, и следовательно в накопителе, определяется через вероятности состояний (2):

$$m = \sum_{k=1}^E k p_k = \frac{1-y}{1-y^{E+1}} \sum_{k=1}^E k y^k,$$

из этого соотношения после некоторых преобразований получим:

$$m = \begin{cases} \frac{y}{1-y^{E+1}} \left[\frac{1-y^E}{1-y} - E y^E \right], & y \neq 1, \\ \frac{E}{2}, & y = 1. \end{cases} \quad (5)$$

Тогда среднее время пребывания заявок в системе:

$$u = \frac{m}{\lambda_0} = \begin{cases} \frac{b}{(1-\pi)(1-y^{E+1})} \left[\frac{1-y^E}{1-y} - E y^E \right], & y \neq 1, \\ \frac{E}{2(1-\pi)\lambda}, & y = 1. \end{cases} \quad (6)$$

Анализ свойств системы. Одной из основных характеристик, определяющих эффективность функционирования систем с потерями, является вероятность потери заявок π вследствие переполнения накопителя. Очевидно, что значение этого параметра зависит от емкости накопителя E и производительности устройства V : $\pi = f(E, V)$, чем больше E и V , тем меньше π .

На рис. 1 показан характер зависимостей вероятности потери заявок от емкости накопителя (рис. 1, а) и производительности устройства (рис. 1, б) для системы с низкой ($y_1 < 1$) и высокой нагрузкой ($y_2 > 1$).

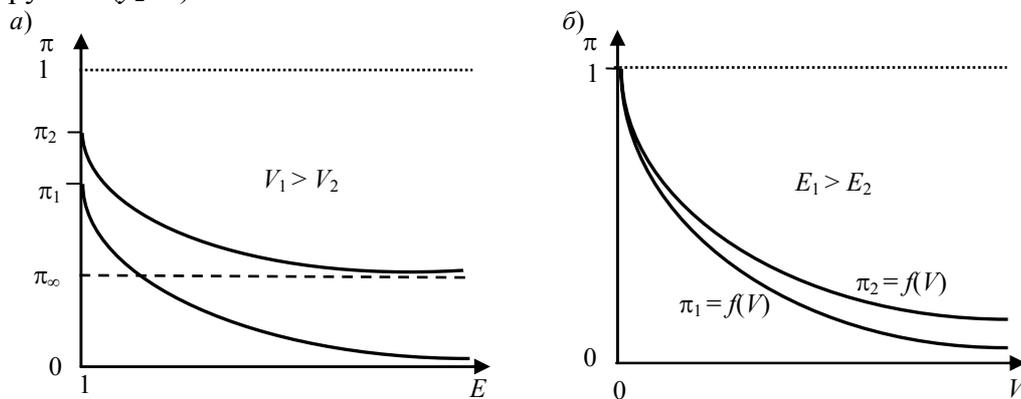


Рис. 1

Вероятность потери заявок в случае единичной емкости накопителя ($E=1$) определяется как $\pi_i = \frac{y_i}{1+y_i}$ ($i=1, 2$). Заметим, что для высоконагруженных систем ($y_2 > 1$) вероятность потери заявок с увеличением емкости накопителя ($E \rightarrow \infty$) стремится к некоторому пределу π_∞ . Этот предел обусловлен тем, что в случае $\lambda_2 > \mu_2$ даже при неограниченной емкости накопителя доля обработанных заявок за единицу времени, равная отношению интенсивности обработки $\mu_2 = 1/b_2$ к интенсивности поступления λ_2 , составит $\pi_0 = \mu_2/\lambda_2 = 1/y_2$. Следовательно, доля необработанных заявок $\pi_\infty = 1 - y_2^{-1}$ может рассматриваться как доля потерянных заявок.

Уменьшение вероятности потери заявок в системе за счет увеличения производительности V устройства и емкости E накопителя приводит к увеличению стоимости системы, которая может быть рассчитана как

$$S = S_y + S_n = \alpha V^\beta + s_0 E, \quad (7)$$

где $S_y = \alpha V^\beta$ — стоимость устройства (α и β — стоимостные коэффициенты пропорциональности и нелинейности соответственно); $S_n = s_0 E$ — стоимость накопителя (s_0 — стоимость единицы накопителя, предназначенной для одной заявки).

Постановка задачи синтеза. При отсутствии ограничений на характеристики функционирования системы в качестве обобщенного критерия эффективности может использоваться функция следующего вида:

$$F = \gamma_1 \pi + \gamma_2 u + \gamma_3 S, \quad (8)$$

где $\gamma_1, \gamma_2, \gamma_3$ — весовые коэффициенты, определяющие степень ценности (важности) соответствующей характеристики для синтезируемой системы и удовлетворяющие условию $\gamma_1 + \gamma_2 + \gamma_3 = 1$, а π, u и S — определяются соответственно выражениями (3), (6) и (7).

В этом случае задача синтеза формулируется следующим образом: определить емкость накопителя и производительность устройства, при которых критерий эффективности (8) принимает минимальное значение.

На рис. 2 иллюстрируется задача определения оптимальных значений емкости накопителя и производительности устройства. Здесь представлены графики, показывающие зависимость вероятности потерь $\pi(E, V)$, среднего времени пребывания $u(E)$ и $u(V)$, стоимости системы $S = \alpha V$ (предполагается, что $\beta = 1$) и критерия эффективности (8) от вектора оптимизируемых параметров (E, V) . Значения $(E, V)_{\text{opt}}$, соответствующие минимуму критерия эффективности F , являются оптимальными.

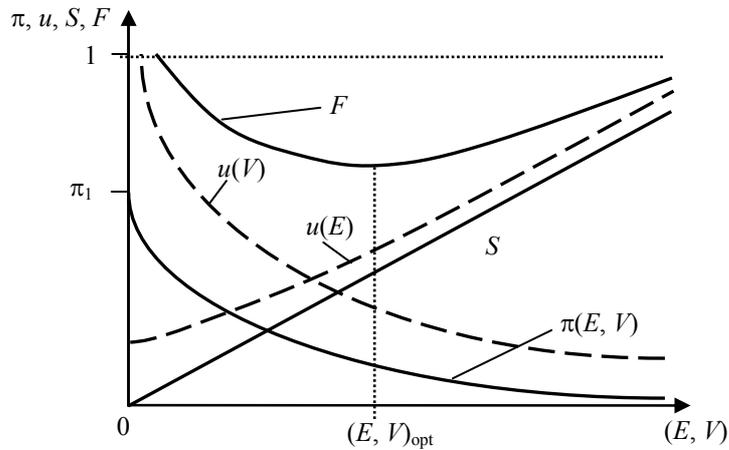


Рис. 2

Отметим, что зависимости среднего времени пребывания заявок в системе $u(E)$ и $u(V)$, представленные на рис. 2 пунктирными кривыми, ведут себя по-разному при увеличении соответственно емкости накопителя E и производительности устройства V : среднее время пребывания заявок в системе растет с увеличением емкости и снижается с увеличением производительности. При этом зависимость $u(E)$ может иметь разный характер — возрастающий, убывающий или любой другой.

При наличии ограничений на характеристики задача синтеза системы с одним устройством и накопителем ограниченной емкости может быть сформулирована в четырех постановках в зависимости от ограничений, а именно — определить производительность V устройства и емкость E накопителя, обеспечивающие выполнение ограничений (обозначено звездочкой):

- 1) на вероятность потери заявок $\pi \leq \pi^*$ при минимальной стоимости системы S ;

2) на вероятность потери заявок $\pi \leq \pi^*$ и среднее время пребывания заявок в системе $u \leq u^*$ при минимальной стоимости системы S ;

3) на вероятность потери $\pi \leq \pi^*$ и стоимость системы $S < S^*$ при минимальном значении среднего времени пребывания заявок в системе u ;

4) на среднее время пребывания заявок в системе $u \leq u^*$ и на стоимость системы $S \leq S^*$ при минимальном значении вероятности потери заявок π .

Алгоритм синтеза. Рассмотрим задачу синтеза применительно к первой и второй постановкам.

Выражение (2) позволяет определить емкость накопителя для заданного ограничения π^* при известной нагрузке y , т.е. при известной производительности V устройства:

$$E \geq \begin{cases} \log_y \frac{\pi^*}{1 - y(1 - \pi^*)}, & y \neq 1, \\ \frac{1 - \pi^*}{\pi^*}, & y = 1. \end{cases}$$

В общей постановке задачи синтеза производительность устройства неизвестна и подлежит определению. Поскольку получить решение задачи оптимального синтеза в явном аналитическом виде с использованием зависимостей (1)—(6) не представляется возможным, один из подходов к решению задачи ввиду небольшого числа оптимизируемых параметров состоит в последовательном переборе значений емкости накопителя E . При этом для каждого значения E с использованием выражения (3) может быть определена минимальная производительность устройства V , при которой выполняется заданное ограничение на вероятность потери заявок:

$$\frac{y^E(1-y)}{1-y^{E+1}} \leq \pi^*.$$

С учетом (1) преобразуем последнее неравенство к виду:

$$V^{E+1} - \frac{(\lambda\theta)^E}{\pi^*} V + (\lambda\theta)^{E+1} \left(\frac{1}{\pi^*} - 1 \right) \geq 0.$$

Решив полученное неравенство относительно V при заданном значении емкости E , например, с использованием одного из численных методов [2], можно определить минимальное значение производительности устройства V , при котором обеспечивается ограничение π^* .

В частности, при $E = 1$ последнее выражение примет вид:

$$V^2 - \frac{\lambda\theta}{\pi^*} V + (\lambda\theta)^2 \left(\frac{1}{\pi^*} - 1 \right) \geq 0,$$

откуда получим:

$$V \geq \frac{1 - \pi^*}{\pi^*} \lambda\theta.$$

Рассмотрим пример со следующими исходными данными: $\lambda = 1 \text{ с}^{-1}$; $\theta = 400$ единиц (ед.); $\alpha = 100$; $\beta = 0,5$; $s_0 = 50$; $\pi^* = 0,05$.

В таблице представлены результаты расчетов для разных значений емкости накопителя, из которой следует, что минимальная стоимость системы $S = 2534$ у.ед. достигается при $E = 7$ и $V = 477$ ед./с.

E	4	5	6	7	8	9	10	11	12
m	1,5	1,9	2,8	2,9	3,4	3,9	4,5	5,1	5,6
u, c	1,5	2,0	2,5	3,0	3,6	4,1	4,7	5,3	5,9
V, c^{-1}	594	537	501	477	459	446	436	428	422
$S, \text{у.ед.}$	2637	2567	2538	2534	2542	2562	2588	2619	2654

Во второй постановке задачи синтеза дополнительное ограничение налагается на среднее время пребывания заявок: $u \leq u^*$. Если в результате синтеза на предыдущем этапе это ограничение не выполняется, необходимо увеличить производительность устройства, что приведет к повышению стоимости системы. При этом вероятность потери заявок уменьшится и, следовательно, можно попытаться снизить стоимость системы за счет уменьшения емкости накопителя.

Положим, что в нашем примере дополнительно задано ограничение $u^* = 2,5$ с. Для того чтобы при $E = 7$ выполнялось это условие, необходимо увеличить производительность устройства до $V = 518$ ед./с, что приведет к увеличению стоимости системы до $S = 2626$ у.ед. В результате предпочтительным оказывается вариант построения системы с накопителем емкостью $E = 6$ (см. таблицу) и стоимостью $S = 2538$ у.ед.

Синтез систем с двумя устройствами. Предложенный подход может быть распространен на системы с двумя последовательно расположенными устройствами, в которых заявки после обработки в первом устройстве направляются ко второму устройству. Производительность 1-го и 2-го устройства и емкость накопителей обозначим как V_1, V_2 и E_1, E_2 . Тогда стоимость системы и вероятность потери заявок в системе:

$$S = \alpha_1 V_1^{\beta_1} + \alpha_2 V_2^{\beta_2} + s_0(E_1 + E_2) \text{ и } \pi = \pi_1 + \pi_2(1 - \pi_1).$$

С учетом ограничения на вероятность потери в системе π^* и последнего выражения получим: $\pi_2 \leq (\pi^* - \pi_1)/(1 - \pi_1)$, где правая часть неравенства представляет собой ограничение на вероятность потери во втором устройстве при известном значении вероятности потери в первом устройстве. Таким образом, если известна вероятность потери заявок в первом устройстве, то задача синтеза для второго устройства при некоторых условиях может быть сведена к задаче синтеза системы с одним устройством.

Положим, что в первое устройство поступает простейший поток заявок с интенсивностью λ , а длительности обработки в устройствах распределены по экспоненциальному закону со средним значением $b_i = \theta_i/V_i$ ($i = 1; 2$). При этих предположениях задача синтеза для первого устройства решается с использованием выражений (1)—(6) для $\pi_1^* = (0,25—0,75)\pi^*$. В качестве результирующего принимается значение π_1^* , при котором обеспечивается минимальная стоимость первого устройства. Тогда ограничение на вероятность потери для второго устройства: $\pi_2^* = (\pi^* - \pi_1^*)/(1 - \pi_1^*)$. Для того чтобы можно было воспользоваться выражениями (1)—(6) для второго устройства и решения задачи синтеза, необходимо выполнить анализ характера потока заявок, покидающих первое устройство и образующих входной поток заявок во второе устройство.

Многочисленные имитационные эксперименты показали, что коэффициент вариации v выходящего потока из первого устройства зависит от нагрузки u_1 и емкости E_1 накопителя и

принимает значения в интервале $v = 0,7—1$ при $E_1 = 1$ и $v = 0,95—1$ при $E_1 = 4$ (рис. 3), причем наименьшие значения коэффициента вариации достигаются при нагрузке $y = 1$. Другими словами, чем больше E_1 , тем ближе к единице коэффициент вариации выходящего потока, что при $E_1 \geq 4$ позволяет использовать выражения (1)—(6) с высокой степенью точности для расчета характеристик функционирования второго устройства.

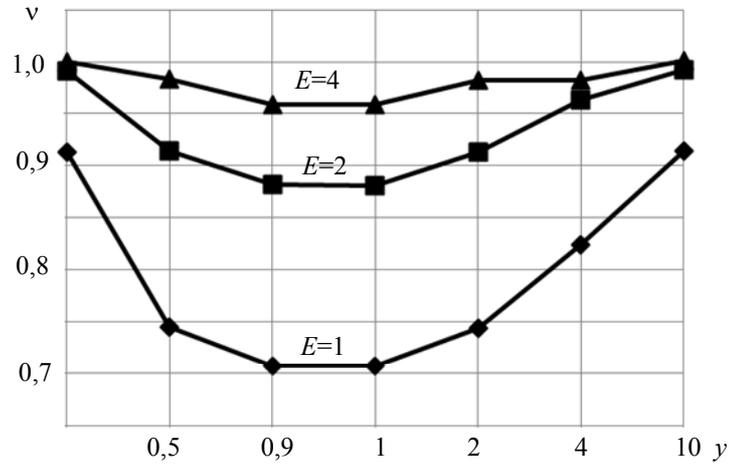


Рис. 3

Кроме того, предположение о простейшем потоке заявок ко второму устройству позволяет получить верхнюю оценку вероятности потери во втором устройстве.

Заключение. Рассмотренный подход к решению задач синтеза систем с потерями и полученные результаты позволяют определить не только емкость накопителей, но и оценить требования к производительности устройств, входящих в состав системы, с учетом совокупных затрат на построение системы для широкого диапазона параметров нагрузки.

СПИСОК ЛИТЕРАТУРЫ

1. Алиев Т. И. Основы моделирования дискретных систем. СПб: СПбГУ ИТМО, 2009. 363 с.
2. Формалев В. Ф., Ревизников Д. Л. Численные методы. М.: Физматлит, 2006. 398 с.

Сведения об авторе

Тауфик Измайлович Алиев — д-р техн. наук, профессор; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра вычислительной техники; заведующий кафедрой;
E-mail: aliev@d1.ifmo.ru

Рекомендована кафедрой
вычислительной техники

Поступила в редакцию
08.02.12 г.