УДК 519.872

АНАЛИЗ И ОПТИМИЗАЦИЯ ХАРАКТЕРИСТИК НЕОДНОРОДНЫХ СИСТЕМ ОБСЛУЖИВАНИЯ В ТОРГОВЛЕ

ГИБКИНА Н.В.

Рассматривается неоднородная система массового обслуживания в торговле. Исследуются методы определения оптимальных характеристик работы этой системы. Решается задача нахождения оптимального плана обслуживания.

Введение

Математические модели, построенные на основе систем массового обслуживания (СМО), используются при описании широкого класса технических, биологических и других систем, таких как сети связи, сети ЭВМ, предприятия торговли, станции обслуживания и др. [1, 2]. Одной из основных характеристик СМО являются входящие потоки событий, которые поступают на обслуживающие приборы в случайные моменты времени. Оценка интенсивности входного потока заявок, а также скорости обслуживания этих заявок является актуальной технической проблемой. Сложности в решении этой задачи вызваны тем, что в реальных системах интенсивность входного потока заявок существенно меняется с течением времени. В том случае, если сам поток событий, поступающий в систему массового обслуживания, доступен наблюдению, задача оценки его характеристик сильно упрощается [3].

Целью данного исследования является разработка общих методов определения таких характеристик системы, которые обеспечивают оптимальный режим ее работы. Оценивание проводится на основе статистического материала, полученного в ходе непосредственного наблюдения за процессом. Решается задача определения оптимального объема и оптимальной скорости обслуживания в системе при условиях минимизации затрат, связанных с работой СМО.

Постановка задачи

Рассмотрим n-канальную неоднородную СМО вида $M/M/n/m$. Процесс обслуживания описывается неоднородным марковским процессом с непрерывным временем и дискретным множеством состояний S_j ($j = 0, 1, \ldots, m+n$), где j — число требований, находящихся в системе в рассматриваемый момент времени, m — число мест в очереди. Переходы системы из одного состояния в другое происходят в случайные моменты времени.

Оптимизация функционирования СМО предполагает выбор критерия качества работы, относительно которого определяются оптимальные значения показателей работы системы, причем в разные периоды времени важность различных показателей работы может существенно меняться.

Построим функцию, которая определяет общие издержки системы в момент времени t:

$$C_{сист}(t) = C_{экспл}\bar{z} + C_{пр}(n - \bar{z}) + C_{отк}P_{отк}\lambda + C_{оч}\bar{r}, \quad (1)$$

где $C_{экспл}$ — издержки, связанные с эксплуатацией одного канала системы за время t; $C_{пр}$ — издержки, связанные с простоем одного канала за время t; $C_{отк}$ — издержки, связанные с одним отказом за время t; $C_{оч}$ — издержки, связанные с пребыванием одного требования в очереди за время t; $C_{сист}$ — общесистемные издержки за время t; $\bar{z}$ — среднее число занятых каналов; $\bar{r}$ — среднее число требований в очереди в момент времени t.

Поскольку определение точных значений издержек часто представляет собой достаточно сложную задачу, то вместо них могут быть использованы весовые коэффициенты, которые отражают значимость соответствующих издержек для каждого конкретного случая.

Разобьем временной интервал $[t_0, t_T]$, на котором исследуется поведение СМО, на T частей $[t_k, t_{k+1}]$, $k = 0, \ldots, T-1$ (число T выбирается индивидуально в каждой задаче). Будем считать, что на каждом из $[t_k, t_{k+1}]$ параметры системы, характеризующие значения издержек, постоянны. Таким образом, на каждом из частичных интервалов $C_{сист}$ является функцией переменных λ, μ, n.

Тогда задача нахождения оптимального режима функционирования СМО сводится к минимизации функции общих издержек:

$$C_{сист}(\lambda, \mu, n) \to \min . \quad (2)$$

Решение задачи (2) может быть найдено с использованием методов теории игр с природой, поскольку в работе СМО присутствует неопределенность, вызванная отсутствием информации об условиях, в которых функционирует система; один из игроков (например, покупательский спрос и т.д.) действует случайным образом. Значения, полученные для каждого из отрезков $[t_k, t_{k+1}]$, $k = 0, \ldots, T-1$, в совокупности определяют оптимальный режим функционирования системы на всем исследуемом временном интервале $[t_0, t_T]$.

Выбор оптимальных характеристик СМО

Рассмотрим систему массового обслуживания, в которой для обслуживания входящего потока требований возможно использовать одну из $n_1, n_2, \ldots, n_r$-канальных обслуживающих систем с интенсивностью обслуживания $\mu(t)$. Используя критерий качества (2), определяем оптимальное число каналов n_k, $k = 1, 2, \ldots, r$, необходимых для бесперебойной работы СМО и обеспечивающих минимум общих издержек и потерь, которые связаны с функционированием СМО.

Пусть на временном интервале $[t_k, t_{k+1}]$, $k = 0, \ldots, T-1$ в систему поступает случайный поток требований, интенсивность которого $\lambda(t)$ может находиться в одном из возможных состояний $\lambda(t) = \lambda_i$, $i = 1, 2, \ldots, s$, с заданной вероятностью p_i. Интенсивность μ обслуживания требований постоянна.

Матрица выигрышей $A = \| a_{ij} \|$ для данной задачи имеет вид:

$$a_{ij} = -C_{сист}(\lambda_j, \mu, n_i), \quad i = 1, \ldots, q, \quad j = 1, \ldots, s,$$

т.е. a_{ij} противоположна величине издержек от функционирования СМО с интенсивностью входящего потока требований λ_j и n_i-канальной обслуживающей системой.

Для поиска оптимального решения воспользуемся критерием Байеса [4, 5], согласно которому из множества вариантов $1, \ldots, q$ выбирается вариант i^***,** который обеспечивает максимум математического ожидания выигрыша:

$$i^* \in \{1, \ldots, q\} : B_{i^*} = \max_{i=1,\ldots,q} \sum_{j=1}^{s} a_{ij} p_j . \tag{3}$$

Необходимо учитывать, что, вообще говоря, полученное оптимальное решение может быть не единственным.

Исследуем, как изменится решение для адаптивной СМО, т.е. в том случае, когда интенсивность обслуживания требований каналами зависит от интенсивности входящего потока требований, $\mu = \mu(\lambda)$.

Пусть (M, Λ) – дискретная двумерная случайная величина, которая принимает значения из наборов $\mu_1, \mu_2, \ldots, \mu_r$; $\lambda_1, \lambda_2, \ldots, \lambda_s$. Если условные вероятности $p(\mu_i \mid \lambda_j)$, $i = 1, \ldots, r$, $j = 1, \ldots, s$, вычисленные в предположении, что событие $\Lambda = \lambda_j$ уже наступило, известны, то условные математические ожидания случайной величины M при $\Lambda = \lambda_j$, $j = 1, \ldots, s$ равны

$$M_j = M\{M \mid \Lambda = \lambda_j\} = \sum_{i=1}^{q} \mu_i p(\mu_i \mid \lambda_j), \quad j = 1, \ldots, s.$$

Элементы матрицы выигрышей в данном случае имеют вид:

$$a_{ij} = -C_{сист}(\lambda_j, M_j, n_i), \quad i = 1, \ldots, q, \quad j = 1, \ldots, s.$$

Для решения задачи оптимизации, как и ранее, может быть использован критерий (3).

Выбор оптимальной интенсивности обслуживания для n-канальной СМО

Пусть, как и ранее, в период времени $[t_k, t_{k+1}]$, $k = 0, \ldots, T-1$ на вход обслуживающей системы поступает случайный поток требований, интенсивность которого $\lambda(t)$ может находиться в одном из возможных состояний $\lambda(t) = \lambda_i$, $i = 1, 2, \ldots, s$, с заданной вероятностью p_i.

Предположим, что для обслуживания поступающего потока требований возможно использовать r различных n-канальных систем с интенсивностями обслуживания соответственно $\mu_1, \mu_2, \ldots, \mu_r$.

Элементы матрицы выигрышей A определяются как $a_{ij} = -C_{сист}(\lambda_i, \mu_j, n)$, $i = 1, \ldots, r$, $j = 1, \ldots, s$. Оптимальное решение, как и ранее, может быть найдено с помощью критерия Байеса (3).

Определение оптимального объема обслуживания неоднородной СМО

Пусть на временном интервале $[t_0, t_T]$ в систему поступает случайный поток требований с интенсивностью $\lambda(t)$.

Рассмотрим эволюцию СМО на системе частичных временных отрезков $[t_k, t_k + \Delta t_k]$, $k = 1, \ldots, T$, полагая на каждом из них интенсивность входящего потока требований случайной величиной, которая может находиться в одном из состояний $\lambda_j^{(k)}$, $j = 1, 2, \ldots, s$, с заданной вероятностью $p_j^{(k)}$, $\sum_{j=1}^{s} p_j^{(k)} = 1$, $k = 1, \ldots, T$.

Найдем оптимальный объем обслуживания $v^{(k)}$ на k-м частичном отрезке, $k = \overline{0, T-1}$. Элементы матрицы выигрышей $A^{(k)}$ на k-м интервале для данной задачи будут иметь вид:

$$a_{ij}^{(k)} = -(C_{сист})_{ij}^{(k)}, \quad i, j = 1, \ldots, s,$$

где
$$(C_{сист})_{ij}^{(k)} = \begin{cases} C_{нед}^{(k)}(\lambda_j^{(k)} - v_i^{(k)}), & v_i^{(k)} < \lambda_j^{(k)}, \\ C_{изб}^{(k)}(v_i^{(k)} - \lambda_j^{(k)}), & v_i^{(k)} > \lambda_j^{(k)}, \\ 0, & v_i^{(k)} = \lambda_j^{(k)}. \end{cases}$$

Здесь $C_{нед}^{(k)}$ и $C_{изб}^{(k)}$ – издержки, связанные с недостаточным (очередь) и избыточным (простои каналов) объемом обслуживания соответственно.

Таким образом, $a_{ij}^{(k)}$ соответствует издержкам системы на k-м частичном временном отрезке в случае, когда на вход поступает входящий поток требований с интенсивностью λ_j, а объем обслуживания равен $v_i = \lambda_i$, $i, j = 1, \ldots, s$. Поиск оптимального объема обслуживания на k-м интервале осуществляем с использованием критерия (3).

Аналогично находим оптимальные объемы обслуживания $v^{(k)}$ для всех $k = \overline{0, T-1}$.

Для дальнейшего уточнения полученных результатов воспользуемся методом наименьших квадратов.

Целевая функция данной задачи имеет вид:

$$Z(\vec{w}) = \alpha\sum_{i=1}^{T}(v^{(i)} - w^{(i)})^2 + \sum_{i=2}^{T}(w^{(i)} - w^{(i-1)})^2 ,$$

где $\vec{w} = \{w^{(1)}, w^{(2)}, \ldots, w^{(T)}\}$ — уточненный вектор объемов обслуживания; $| v^{(i)} - w^{(i)} |$ — погрешность вычисления оптимального объема обслуживания на i-м временном интервале, $i = 1, \ldots, T$; $| w^{(i)} - w^{(i-1)} |$ — величина, характеризующая колебания объема обслуживания на последовательных временных интервалах, $i = 2, \ldots, T$; α — весовой коэффициент значимости ошибки выбранного объема обслуживания.

Оптимальный план обслуживания соответствует минимуму целевой функции $Z(\vec{w})$:

$$Z(\vec{w}) \to \min . \qquad (4)$$

Решение задачи (4) будет удовлетворять условиям:

$$\frac{\partial Z(\vec{w})}{\partial w_j} = 0 , \quad j = 1, \ldots, T .$$

После преобразований получим систему из T линейных уравнений с T неизвестными:

$$\begin{cases} (\alpha + 1)w^{(1)} - w^{(2)} = \alpha v^{(1)}, \\ - w^{(j-1)} + (\alpha + 2)w^{(j)} - w^{(j+1)} = \alpha v^{(j)}, \quad j = 2, \ldots, T - 1; \\ - w^{(T-1)} + (\alpha + 1)w^{(T)} = \alpha v^{(T)}. \end{cases}$$

Решая эту систему любым из известных аналитических или численных методов, найдем элементы вектора $\vec{w}$, которые представляют собой оптимальные значения объемов обслуживания на каждом из частичных временных отрезков. Таким образом, полученное решение $\vec{w} = \{w^{(1)}, w^{(2)}, \ldots, w^{(T)}\}$ определяет оптимальный режим функционирования обслуживающей системы на всем исследуемом временном интервале.

Интенсивность потока заявок

$[t_i, t_{i+1}]$		$\lambda_j, \ j = 1, \ldots, 6$					
11^{00}-12^{00}	$\vec{\lambda}$	42,17	43,83	45,50	47,16	48,83	50,41
	$\vec{p}$	0,066	0,133	0,233	0,233	0,2	0,133
12^{00}-13^{00}	$\vec{\lambda}$	38,83	41,64	44,45	47,26	50,08	52,75
	$\vec{p}$	0,1	0,1	0,2	0,3	0,233	0,066
13^{00}-14^{00}	$\vec{\lambda}$	51,97	56,66	61,34	66,03	70,72	75,18
	$\vec{p}$	0,1	0,233	0,1	0,3	0,166	0,1
14^{00}-15^{00}	$\vec{\lambda}$	26,55	28,56	30,58	32,59	34,60	36,52
	$\vec{p}$	0,066	0,033	0,166	0,233	0,366	0,133
15^{00}-16^{00}	$\vec{\lambda}$	47,83	50,19	52,55	54,90	57,26	59,51
	$\vec{p}$	0,033	0,133	0,2	0,433	0,166	0,033
16^{00}-17^{00}	$\vec{\lambda}$	53,69	55,49	57,29	59,09	60,88	62,60
	$\vec{p}$	0,1	0,1	0,1	0,2	0,333	0,166
17^{00}-18^{00}	$\vec{\lambda}$	45,44	47,32	49,19	51,06	52,94	54,72
	$\vec{p}$	0,066	0,166	0,4	0,2	0,1	0,066

Пример. В качестве примера рассмотрим работу аптеки как системы массового обслуживания. Число каналов обслуживания $n = 3$; ограничения на длину очереди отсутствуют. В результате обработки статистического материала о проходимости торговой точки в разное время суток были получены следующие значения интенсивности входного потока заявок (таблица).

Для обслуживания требований может быть использована одна из трех n-канальных систем, $n = 3$ с интенсивностями обслуживания: $\mu = \{0.24, 0.30, 0.49\}$. Исходя из особенностей функционирования данного предприятия торговли, коэффициенты функции издержек были выбраны следующими: $C_{экспл} = 0{,}4$; $C_{пр} = 0{,}1$; $C_{отк} = 0{,}1$; $C_{оч} = 0{,}4$.

Для сокращения записи в расчетах будем задавать коэффициенты издержек системы в виде

$(C_{экспл}, C_{пр}, C_{отк}, C_{оч})$.

С учетом предложенного метода, была определена оптимальная интенсивность обслуживания покупателей в рассматриваемой СМО для заданного уровня значимости издержек $(C_{экспл}, C_{пр}, C_{отк}, C_{оч})$ и проведен анализ изменения μ в зависимости от их величины.

Для вектора издержек $(0.4, 0.1, 0.1, 0.4)$ оптимальной является обслуживающая система с интенсивностью обслуживания $\mu = 0.49$, что обеспечивает бесперебойную работу системы при минимальном износе оборудования.

Замечено, что при увеличении $C_{экспл}$ и одновременном уменьшении либо неизменности остальных издержек оптимальная интенсивность обслуживания уменьшается до минимального из возможных значений и становится равной $\mu = 0.24$ (для $(0.5, 0.0, 0.1, 0.3)$).

Увеличение издержек, связанных с наличием в системе очереди, влияет на рост μ, если несущественны эксплуатационные издержки. Так, были получены следующие значения оптимальной интенсивности обслуживания μ в зависимости от $(C_{экспл}, C_{пр}, C_{отк}, C_{оч})$:

$$\mu = 0.24 \ \text{для} \ (0.5, 0.3, 0, 0.2);$$

$$\mu = 0.49 \ \text{для} \ (0.05, 0.75, 0, 0.2).$$

Даже незначительный рост издержек $C_{пр}$, связанных с простоем оборудования обслуживающей системы, вызывает резкое увеличение оптимального значения μ при неизменных значениях остальных издержек:

$$\mu = 0.24 \ \text{для} \ (0.2, 0, 0.1, 0.7);$$

$$\mu = 0.49 \ \text{для} \ (0.2, 0, 0.2, 0.6).$$

Далее найдем оптимальный объем обслуживания в рассматриваемой торговой точке в течение дня,

$t_0 = 11$, $t_7 = 18$, $\Delta t = 1$. Нормированные значения издержек $C_{нед}$, вызванных наличием очереди в системе, и издержек $C_{изб}$, связанных с простоем оборудования, примем равными соответственно:

$$C_{нед} = \{0.21, 0.42, 0.5, 0.32, 0.34, 0.44, 0.65\};$$

$$C_{изб} = \{0.79, 0.58, 0.5, 0.68, 0.66, 0.56, 0.35\}.$$

Значения вектора $\vec{v}$, который является начальным планом обслуживания для задачи (4), были найдены с использованием метода (3) на каждом временном промежутке $[t_i, t_{i+1}]$, $i = 0, \ldots, 7$:

$$\vec{v} = \{45.5, 47.26, 66.03, 32.59, 52.55, 59.09, 51.06\}.$$

Уточненный план обслуживания, полученный из решения задачи нелинейного программирования (4), равен

$$\vec{w} = \{47.502, 49.504, 53.751, 45.72, 50.814, 54.174, 52.617\}.$$

Как видно, для рассматриваемого предприятия дополнительное уточнение оптимального плана обслуживания позволяет добиться более плавного изменение его объемов с течением времени (рис.1).

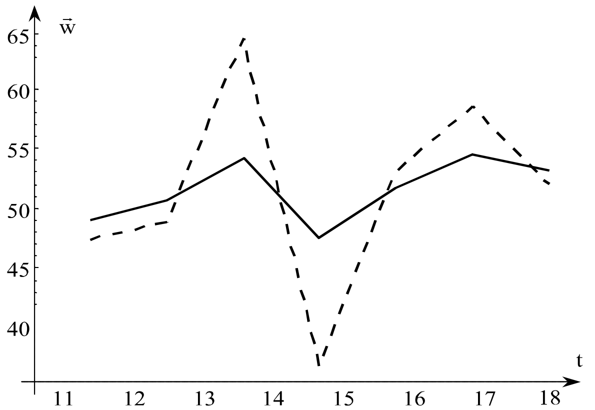

Рис. 1. Уточненный оптимальный план обслуживания при заданных издержках:
$C_{нед} = \{0.21, 0.42, 0.5, 0.32, 0.34, 0.44, 0.65\}$;
$C_{изб} = \{0.79, 0.58, 0.5, 0.68, 0.66, 0.56, 0.35\}$;
(пунктиром указан начальный оптимальный план)

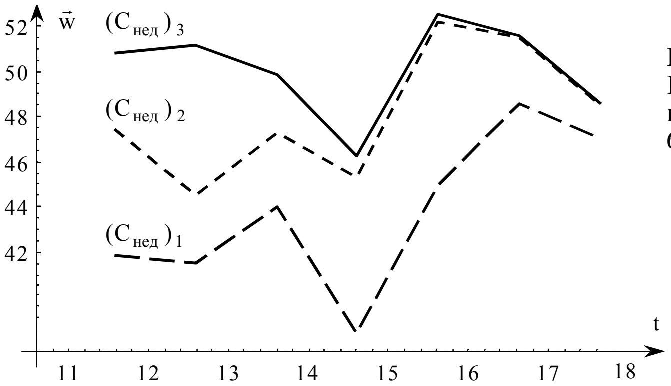

Рис. 2. Изменение оптимального плана обслуживания для разных уровней значимости издержек:
$(C_{нед})_1 = \{0.21, 0.32, 0.35, 0.32, 0.34, 0.44, 0.45\}$,
$(C_{нед})_2 = \{0.31, 0.42, 0.40, 0.38, 0.38, 0.49, 0.54\}$,
$(C_{нед})_3 = \{0.39, 0.49, 0.50, 0.45, 0.50, 0.61, 0.69\}$

На рис. 2 приведен график изменения оптимального плана обслуживания аптеки для различных значений издержек. Замечено, что с ростом величины $C_{нед}$ оптимальный объем обслуживания также возрастает.

Выводы. Традиционно подобные задачи решаются для однородного процесса, который описывает систему обслуживания [1, 2]. Научная новизна заключается в том, что в статье проведено обобщение на неоднородную ситуацию, при этом построенная модель является более адекватной для случая систем обслуживания в торговле.

Практическая ценность статьи состоит в том, что предложенные методы повышения качества обслуживания СМО позволяют определить оптимальные характеристики обслуживающей неоднородной системы и нагрузку на каналы обслуживания, а также минимизировать издержки, возникающие в ходе функционирования системы. Результаты, полученные в данном исследовании, были использованы при анализе работы сети аптек для оптимального распределения нагрузки на обслуживающий персонал и средства обслуживания. Сравнение начального плана обслуживания и его оптимизированного значения для примера, связанного с исследованием работы аптек, позволяет сделать вывод, что дополнительное решение задачи оптимизации улучшает конечные результаты не менее чем вдвое.

Все вычисления были проведены с помощью пакета программ, реализованного автором на языке программирования Delphi 6.0.

Литература: 1. *Саати Т.Л.* Математические методы исследования операций. М.: Воениздат, 1963. 520с. **2.** *Ивченко Г.И., Каштанов В.А., Коваленко И.Н.* Теория массового обслуживания. М.: Высш. шк., 1982. 256 с. **3.** *Гнеденко Б.В., Коваленко И.Н.* Введение в теорию массового обслуживания. М.: Наука, 1987. 336 с. **4.** *Вентцель Е.С.* Исследование операций. М.: Сов. радио, 1972. 552 с. **5.** *Дубров А.М., Лагоша Б.А., Хрусталев Е.Ю.* Моделирование рисковых ситуаций в экономике и бизнесе. М.: Финансы и статистика, 1999. 246 с.

Поступила в редколлегию 20.12.2003

Рецензент: д-р физ. мат. наук, проф. Дикарев В.А.

Гибкина Надежда Валентиновна, аспирантка кафедры ВМ ХНУРЭ. Научные интересы: теория вероятностей и ее приложения, программирование. Адрес: Украина, 61166, Харьков, пр. Ленина, 14.