Система обслуживания ветвящихся потоков с разделением времени

А.В. Зорин
andrei.zorine@itmm.unn.ru ✉

Кафедра теории вероятностей и анализа данных

Национальный исследовательский
Нижегородский государственный университет им. Н.И. Лобачевского

2-я Конференция матцентров
Москва • 11 ноября 2022

Научная тематика кафедры

- Конфликтные системы обслуживания являются идеализацией многих реальных процессов обслуживания в условиях изменяющихся стохастических факторов.
- Теория конфликтных систем обслуживания относится к одному из важных вопросов в общей теории управляемых случайных процессов — проблеме изучения предельных свойств распределений вероятностей многомерных последовательностей и процессов специального вида.

Идеальная модель СМО

Соглашение

Математическая модель системы массового обслуживания есть заданный на некотором вероятностном пространстве $(\Omega, \mathfrak{F}, \mathbf{P})$ случайный процесс $\{\Xi(t, \omega); t \in T\}$ с измеримым фазовым пространством $(S, \mathfrak{S})$, счётным или несчётным упорядоченным множеством T «моментов времени».

В настоящее время применяются следующие подходы:

- Описательное задание системы обслуживания и вывод некоторых свойств распределений изучаемого процесса $\{\Xi(t, \omega); t \in T\}$ из «физических соображений» (А.Я. Хинчин, Б.В. Гнеденко, Д. Кендалл, Ф. Поллячек, Т.Л. Саати и др.)
- Представление изучаемого процесса $\{\Xi(t, \omega); t \in T\}$ с помощью некоторого функционального уравнения или стохастической рекурсивной последовательности с априорно заданным «входным управляющим процессом» (С. Асмуссен, А.А. Боровков, Н. Прабху, С.Г. Фосс и др.)

Функционально-статистическое задание конфликтной СМО

Постулаты моделирования управляющей системы обслуживания

i) Дискретность актов функционирования управляющей системы обслуживания.

ii) Нелокальность (неполнота) описания поблочного строения управляющей системы обслуживания.

iii) Совместное рассмотрение поблочного строения управляющей системы обслуживания и её функционирования во времени.

О п р е д е л е н и е [М.А. Федоткин (1978, 1981)]. Пусть $\{\tau_i^{\text{н}}; i = 0, 1, \ldots\}$ — последовательность моментов наблюдения на оси Ot, $\eta_i^{\text{н}}$ и ν_i суть число требований и метка требований потока Π, поступивших на промежутке $(\tau_i^{\text{н}}, \tau_{i+1}^{\text{н}}]$. *Неполное описание* потока Π неоднородных требований есть векторная случайная последовательность

$$\{(\tau_i^{\text{н}}, \eta_i^{\text{н}}, \nu_i); i = 0, 1, \ldots\}. \tag{1}$$

Функционально-статистическое задание конфликтной СМО

Дискретная шкала времени и схема системы:

$$\tau_0 = 0,\ \tau_{i+1} = \tau_i + \mathfrak{v}_{i+1},\ i \in \mathfrak{I} = \{0, 1, \ldots\}\ (*)$$

$$\varkappa_{j,i+1} = \min\Big\{N_j, \varkappa_{j,i} + \sum_{\alpha=1}^{\widehat{m}} \partial_{\alpha,j}\eta_{\alpha,i} - \bar{\xi}_{j,i}\Big\}\ (**)$$

$$b \in \{0, 1, \ldots\}^{\widehat{m}}, \quad \bar{b}, w \in \{0, 1, \ldots\}^{m}$$

внешняя среда: $\{\chi_i : i \in \mathfrak{I}\}$ — марковская цепь, $\chi_i \in \mathfrak{E} = \{e^{(1)}, \ldots, e^{(d)}\}$

входные блоки: входящие потоки $\{\eta_i : i \in \mathfrak{I}\}$, потоки насыщения $\{\xi_i : i \in \mathfrak{I}\}$; $\varphi(b, \bar{b}, y; \Gamma^{(s)}, e^{(k)}, w)$ — совм. усл. распр. η_i, ξ_i, $\mathfrak{v}_{i+1}$; $\nu_i = (\Gamma_i, \chi_i, \varkappa_i)$

очереди: $\{\varkappa_i : i \in \mathfrak{I}\}$, соотношение $(**)$

блоки стратегии механизма обслуживания; $g_j(w_j, b, \bar{b}_j)$

обслуживающее устройство; $\Gamma_i \in \Gamma = \{\Gamma^{(1)}, \ldots, \Gamma^{(n)}\}$

блок реализации алгоритма управления потоками: $\Gamma_{i+1} = u(\Gamma_i, \varkappa_i, \eta_i)$

выходные блоки: выходящие потоки; $\bar{\xi}_{j,i} = g_j(\varkappa_{j,i}, \eta_i, \xi_{j,i})$

Функционально-статистическое задание конфликтной СМО

Определение

Поток насыщения — это виртуальный выходящий поток при неограниченном запасе требований в очереди и максимальном использовании возможностей обслуживающего устройства.

Определение

Потоки насыщения $\Pi_{j_1}^{\text{нас}}$, $\Pi_{j_2}^{\text{нас}}$, …, $\Pi_{j_\mu}^{\text{нас}}$ $(1 \leqslant j_1 < j_2 < \ldots < j_\mu \leqslant m, \mu \geqslant 2)$ называются конфликтными, если для каждого $i = 0, 1, \ldots$ вероятность одновременного выполнения неравенств $\xi_{j_1,i} > 0$, $\xi_{j_2,i} > 0$, …, $\xi_{j_\mu,i} > 0$ равна нулю.

Определение

Будем называть систему массового обслуживания *конфликтной*, если её потоки насыщения конфликтные.

Описание системы обслуживания I

Рассматривается следующая система обслуживания и управления конфликтными потоками.

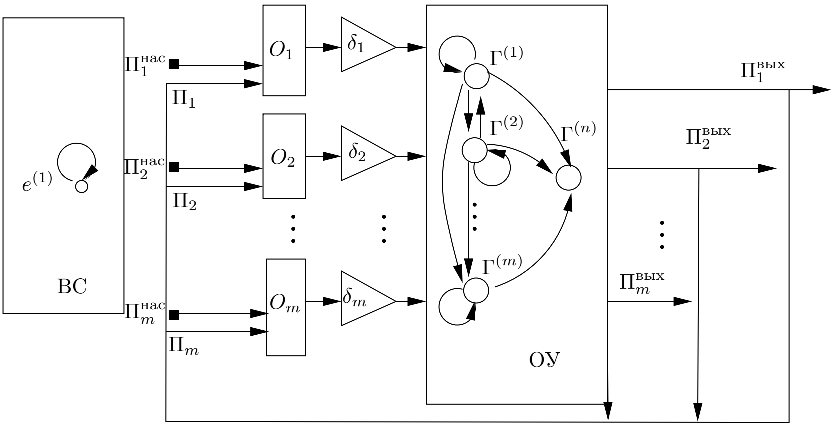
- В системе имеется $m < \infty$ узлов обслуживания $\Gamma^{(1)}, \Gamma^{(2)}, \ldots, \Gamma^{(m)}$ и узел консервации $\Gamma^{(m+1)}$.
- Одновременно обслуживание задания может происходить только на одном из узлов обслуживания. У каждого такого узла разрешена неограниченная очередь.
- Длительность обслуживания задания r-го типа узлом r, $r = 1, 2, \ldots, m$ — это случайная величина, имеющая функцию распределения $B_r(t)$, а длительность работы узла консервации имеет функцию распределения $B_{m+1}(t)$, при этом $\beta_{r,1} = \int\limits_0^\infty t\,dB_r(t) < \infty$, $r = 1, 2, \ldots, m+1$.
- Задание, обслуженное узлом r, $r \leqslant m$, порождает случайное число заданий каждого типа с заданным совместным законом распределения $p_r(x)$, $x = (x_1, x_2, \ldots, x_m) \in \mathfrak{X} = \{0, 1, \ldots\}^m$.

Описание системы обслуживания II

- Число заданий по m различным очередям узлов обслуживания в начальный момент описывается некоторым дискретным распределением.
- Если по окончании очередного акта обслуживания, акта консервации или в начальный момент работы системы количество заданий в очередях определяется ненулевым вектором $x = (x_1, x_2, \ldots, x_m)$, то немедленно начинается обслуживание задания в узле $r = h(x)$ и включается узел консервации при пустых очередях. Здесь черед $h(\cdot)$ обозначено некоторое заданное отображение множества $\mathfrak{X}$ на множество $\{1, 2, \ldots, m+1\}$, такое что $h(x) = r$ влечет $x_r > 0$ при $r = 1, 2, \ldots, m$ и прообразом точки $(m+1)$ является нулевой вектор $\mathbf{0}$.
- Данная постановка обобщает работу Высоцкий и Федоткин (1996), где общее число потомков одного задания могло быть ноль либо один.

Рассматриваемая система обслуживания как абстрактная стохастическая управляющая система I

Схема управляющей системы массового обслуживания:



Рассматриваемая система обслуживания как абстрактная стохастическая управляющая система II

На схеме присутствуют блоки:

1. внешняя среда с одним состоянием ($d = 1$);
2. входящие потоки Π_1, Π_2, …, Π_m вторичных требований и потоки насыщения $\Pi_1^{\text{нас}}$, $\Pi_2^{\text{нас}}$, …, $\Pi_m^{\text{нас}}$;
3. блоки очередей O_1, O_2, …, O_m;
4. устройства δ_1, δ_2, …, δ_m по организации дисциплины очереди в накопителях — блоки переработки информации внешней памяти;
5. обслуживающее устройство, которое имеет $n = m + 1$ состояние $\Gamma^{(1)}$, $\Gamma^{(2)}$, …, $\Gamma^{(n)}$ — внутренняя память;
6. блок реализации алгоритма управления потоками (граф смены состояния обслуживающего устройства);
7. выходные потоки $\Pi_1^{\text{вых}}$, $\Pi_2^{\text{вых}}$, …, $\Pi_m^{\text{вых}}$ в действительности обслуженных и покинувших систему требований.

Отметим, что в данной управляющей системе число $\hat{m}$ входящих потоков совпадает с числом очередей m.

Рассматриваемая система обслуживания как абстрактная стохастическая управляющая система III

Исходя из физической постановки задачи видим, что условная вероятность $\varphi(w, \bar{w}, y; \gamma, x)$ выполнения соотношений $\eta_i = w$, $\xi_i = \bar{w}$, $\mathfrak{v}_i < y$ при фиксированном значении метки $\nu_i = (\gamma, x)$ следует определить следующим образом:

$$\varphi(w, \bar{w}, y; \Gamma^{(s)}, x) = \begin{cases} p_r(w) B_r(y) & \text{при } x \neq y^{(n)},\ r = h(x),\ \bar{w} = y^{(r)}; \\ B_n(y) & \text{при } x = y^{(n)},\ w = y^{(n)},\ \bar{w} = y^{(n)}; \\ 0 & \text{в остальных случаях.} \end{cases} \qquad (2)$$

Для задания блока переработки информации внутренней памяти определим отображение

$$u(\Gamma^{(s)}, x, w) \colon \Gamma \times X^{(m)} \times X^{(m)} \to \Gamma$$

посредством равенства

$$u(\Gamma^{(s)}, x, w) = \Gamma^{(r)} \qquad \text{при } r = h(x). \qquad (3)$$

Рассматриваемая система обслуживания как абстрактная стохастическая управляющая система IV

Поскольку значение функции $u(\cdot,\cdot)$ не зависит от первого аргумента, $\Gamma^{(s)}$, можно ввести сокращенное обозначение: $u(x)$ для $u(\Gamma^{(s)},x)$. Далее, положим

$$g_j(x_j,w,\bar{w}) = \min\{x_j+w_j,\bar{w}_j\}. \tag{4}$$

Соотношения (2), (3), (4) позволяют построить вероятностное пространство $(\Omega,\mathfrak{F},\mathbf{P})$, на котором задаются все рассматриваемые случайные величины и случайные элементы, в том числе:

- τ_i — момент окончания i-го акта обслуживания, $\tau_0=0$,
- $\Gamma_i\in\{\Gamma^{(1)},\Gamma^{(2)},\ldots,\Gamma^{(m+1)}\}$ — состояние обслуживающего устройства (внутренней памяти) в момент τ_i,
- $\kappa_{j,i}$ — число заданий в очереди O_j в момент τ_i, векторы $\kappa_i=(\kappa_{1,i},\kappa_{2,i},\ldots,\kappa_{m,i})$ при $i=0,1,\ldots$.

При этом имеет место векторное рекуррентное по $i=0,1,\ldots$ равенство

$$(\Gamma_{i+1},\kappa_{i+1})=(u(\kappa_i),\kappa_i+\eta_i-\xi_i). \tag{5}$$

Рассматриваемая система обслуживания как абстрактная стохастическая управляющая система V

Выбирая теперь распределение вектора (Γ_0, κ_0), получаем многомерную стохастическую последовательность

$$\{(\Gamma_i, \kappa_i); i = 0, 1, \ldots\}. \tag{6}$$

Она является математической моделью процесса обслуживания потоков вторичных требований в системе с разделением времени.

Стохастическая многомерная последовательность $\{(\Gamma_i, \kappa_i); i = 0, 1, \ldots\}$ при заданном распределении случайного вектора (Γ_0, κ_0) будет однородной цепью Маркова.

Рассматриваемая система обслуживания как абстрактная стохастическая управляющая система VI

Пусть $v^x = v_1^{x_1} v_2^{x_2} \times \ldots \times v_m^{x_m}$ для $v = (v_1, v_2, \ldots, v_m)$ и $|v_j| \leqslant 1$, $j = 1, 2, \ldots, m$,

$$\Psi^{(i)}(\Gamma^{(s)}, v) = \sum_{x \in \mathfrak{X}} \mathbf{P}(\{\Gamma_i = \Gamma^{(s)}, \kappa_i = x\}) v^x, \tag{7}$$

$$\Phi^{(i)}(\Gamma^{(s)}, v) = \sum_{\substack{x \in \mathfrak{X} \\ h(x) = s}} \mathbf{P}(\{\Gamma_i = \Gamma^{(s)}, \kappa_i = x\}) v^x \text{для } s = 1, 2, \ldots, m+1, \tag{8}$$

$$R_j(v) = v_j^{-1} \sum_{x \in \mathfrak{X}} v^x p_j(v), \qquad j = 1, 2, \ldots, m; \qquad R_{m+1}(v) = 1. \tag{9}$$

Основные результаты I

Теорема 1.

Имеют место рекуррентные по $i = 0, 1, \ldots$ *соотношения*

$$\Psi^{(i+1)}(\Gamma^{(s)}, v) = R_s(v)\Phi^{(i)}(s, v). \tag{10}$$

Теорема 2.

Пусть для всех $j = 1, 2, \ldots, m$ *имеет место неравенство* $R_j(v) < 1$ *в одной точке* $v = v^* = (v_1^*, v_2^*, \ldots, v_m^*)$, *такой что* $v_1^* > 1$, $v_2^* > 1$, …, $v_m^* > 1$. *Тогда состояние* $(\Gamma^{(m+1)}, \mathbf{0})$ *достижимо отовсюду.*

Состояние $(\Gamma^{(m+1)}, \mathbf{0})$ — поглощающее, поэтому стационарное распределение всегда существует (возможно, не единственное).

Основные результаты II

Обозначим через

$$p_{s,r} = \sum_{x=(x_1,\ldots,x_m)\in\mathfrak{X}} x_r p_s(x) \tag{11}$$

среднее число потомков типа r у одного задания типа s, введем матрицу $Q = (p_{s,r})_{s,r=\overline{1,m}}$ и пусть E — единичная матрица размера $m \times m$.

Теорема 3.

Если выполнены условия теоремы 2,

$$\sum_{s=1}^{m+1} \Psi^{(0)}(\Gamma^{(s)}, v^*) < \infty, \qquad \det(E - Q) \neq 0 \tag{12}$$

и наибольшее по абсолютной величине собственное число матрицы Q *меньше единицы, то цепь Маркова* $\{(\Gamma_i, \kappa_i); i = 0, 1, \ldots\}$ *имеет предельное распределение.*

Основные результаты III

Поскольку переход к пределу в рекуррентных соотношениях для производящих функций определен не всегда, можно перейти к *суммированию по Чезаро*. Пусть марковская цепь $\{(\Gamma_i, \kappa_i); i = 0, 1, \ldots\}$ с начальным распределением $\mathcal{P}_0 = \{\mathbf{P}(\Gamma_0 = \gamma, \kappa_0 = x) \colon (\gamma, x) \in \Gamma \times X^{(m)}\}$ получена посредством функции $h(\cdot)$. Обозначим

$$d_s^x(\mathcal{P}_0, h(\cdot)) = \lim_{a\to\infty} a^{-1} \sum_{i=1}^{a} \mathbf{P}(\{\Gamma_{i-1} = \Gamma^{(s)}, \kappa_{i-1} = x\}), \tag{13}$$

$$d^x(\mathcal{P}_0, h(\cdot)) = \sum_{s=1}^{n} d_s^x(\mathcal{P}_0, h(\cdot)), \qquad x \in X^{(m)}, \tag{14}$$

$$d_s(\mathcal{P}_0, h(\cdot)) = \sum_{s\in X^{(m)}} d_s^x(\mathcal{P}_0, h(\cdot)), \qquad s = 1, 2, \ldots, n. \tag{15}$$

Основные результаты IV

Теорема

Набор $\{d_r^x(\mathcal{P}_0, h(\cdot))\colon (\Gamma^{(r)}, x) \in \Gamma \times X^{(m)}\}$ является стационарным распределением марковской цепи $\{(\Gamma_i, \kappa_i); i = 0, 1, \ldots\}$.

Обозначим

$$\Psi(\Gamma^{(s)}, v) = \sum_{x \in X^{(m)}} d_s^x(\mathcal{P}_0, h(\cdot)) v^x,$$

$$\Phi(\Gamma^{(s)}, v) = \sum_{x \in X_s} d^x(\mathcal{P}_0, h(\cdot)) v^x, \qquad s = 1, 2, \ldots, n.$$

Следствие

Выполняется равенство:

$$\Psi(\Gamma^{(s)}, v) = R_s(v) \Phi(\Gamma^{(s)}, v). \tag{16}$$

Постановка задачи оптимизации I

Обозначим через

$$\mathcal{I}_i(\mathcal{P}_0, h(\cdot))$$

стоимость средних потерь от простоя заявок на интервале $(\tau_{i-1}, \tau_i]$, $i = 1, 2, \ldots$. Тогда

$$\mathcal{I}_i(\mathcal{P}_0, h(\cdot)) = \sum_{r=1}^{m} c_r \mathbf{M}\zeta_{r,i},$$

где $\zeta_{r,i}$ — суммарное время пребывания заявок из очереди узла r в системе на интервале $(\tau_{i-1}, \tau_i]$, а c_r — стоимость пребывания в системе одной заявки из очереди узла r в единицу времени.

Постановка задачи оптимизации II

Предел по Чезаро

$$\lim_{a\to\infty} a^{-1}\sum_{i=1}^{a}\mathcal{I}_i(\mathcal{P}_0, h(\cdot)),$$

который будем обозначать через $\mathcal{I}(\mathcal{P}_0, h(\cdot))$, существует и равен:

$$\sum_{r=1}^{m} c_r \sum_{s=1}^{m}\beta_s \sum_{x\in X_s} x_r d^x(\mathcal{P}_0, h(\cdot)).$$

Определение (Высоцкий, Федоткин (1996))

Функцию $h^*(\cdot)\in H$ назовем оптимальной, если

$$\mathcal{I}(\mathcal{P}_0, h^*(\cdot)) = \inf\{\mathcal{I}(\mathcal{P}_0, h(\cdot))\colon\ h\in H\}$$

при любом начальном распределении.

Литература

1. *Ляпунов А.А., Яблонский С.В.* Теоретические проблемы кибернетики // Проблемы кибернетики. — М.: Физматгиз, 1963. Вып. 9. — 5-22.
2. *Федоткин М.А.* Процессы обслуживания и управляющие системы // Математические вопросы кибернетики. Вып. 6: Сборник статей / Под ред. С.В. Яблонского. — М.: Наука. Физматлит, 1996. — 51-70.
3. *Высоцкий А.А., Федоткин М.А.* Предельные свойства и оптимизация процессов с разделением времени // Доклады РАН, **350**:3, 295-297.