

лирования стохастических систем // Известия СНЦ РАН, Самара, №1, 2003. – С. 119-126.

5. Тарасов В.Н., Бахарева Н.Ф. Компьютерное моделирование вычислительных систем. Теория, алгоритмы, программы. Оренбург: Изд. ОГУ, 2005. – 225 с.
6. Тарасов В.Н., Карташевский И.В. Программно реализованная имитационная модель массового обслуживания общего вида // ИКТ. Т.7, №2, 2009. – С. 63-68.

УДК 621.391

СРЕДНЯЯ ДОЛЯ НЕДООБСЛУЖИВАНИЯ ЗАЯВОК В СИСТЕМАХ МАССОВОГО ОБСЛУЖИВАНИЯ

Лихтциндер Б.Я., Макаров И.С.

Вводится характеристика СМО – средняя доля недообслуженных заявок. Показано, что указанная характеристика определяется разностью второго и первого начальных моментов числа заявок, поступающих в течение каждого интервала времени обслуживания.

Среднее значение времени дообслуживания заявок является одной из характеристик систем массового обслуживания (СМО). Она характеризует среднее время, необходимое для завершения обслуживания ранее выбранной заявки [1]. Это время учитывает, что на момент поступления очередной заявки в системе может не оказаться ни одной заявки, то есть предыдущая заявка была обслужена полностью и время ее дообслуживания равно нулю. Вместе с тем, имеется еще одна не менее важная характеристика, которой, на наш взгляд, в теории СМО уделяется недостаточное внимание. Это – средняя доля недообслуживания заявок. Она характеризуется средним числом недообслуженных заявок, за время обслуживания очередной заявки. При этом может оказаться, что в течение указанного времени в систему не поступает ни одной заявки и число недообслуженных заявок будет равно нулю.

В [2] показано, что основное влияние на процесс формирования очередей в СМО оказывают математическое ожидание $E(n)$ и второй начальный момент $E(n^2)$ числа заявок, приходящихся на одну обработанную заявку.

$$q = \frac{E(n^2) - E(n)}{2(1 - E(n))}, \quad (1)$$

где q – средняя длина очереди в одноприборной СМО. При анализе СМО поток заявок обычно характеризуется плотностями вероятностей распределения интервалов времени между соседними заявками $W(\mathcal{G})$ и распределения интервалов времени обработки указанных заявок $W(\tau)$. В данной работе мы попытаемся установить зависимости между указанными временными характеристиками и средней долей недообслуженных заявок в СМО.

Анализ проведем с использованием функций распределения вероятностей интервалов между соседними заявками для потоков различного порядка, получаемых путем разреживания исходного потока.

На рис. 1. показана временная диаграмма исходного стационарного рекуррентного потока заявок, для которого случайная величина \mathcal{G}_1 интервалов между соседними заявками распределена по закону, имеющему плотность вероятностей $W(\mathcal{G})$ и функцию распределения вероятностей

$$F_1(\tau) = P(\mathcal{G}_1 \leq \tau) = \int_0^{\tau} W(\mathcal{G}) d\mathcal{G}. \quad (2)$$

На том же рис. 1 показаны потоки заявок более высоких порядков, для которых

$$F_n(\tau) = P(\mathcal{G}_n \leq \tau), \quad (3)$$

где $\mathcal{G}_n = \sum_{k=1}^n \mathcal{G}_k$. Функции распределения вероятностей для потоков n -го порядка назовем образующими функциями.

Поток, соответствующий образующей функции n -го порядка, получается путем удаления из исходного потока подряд $(n-1)$ заявок, с сохранением каждой n -ой заявки. Следовательно, функция распределения вероятностей интервалов для исходного потока имеет первый порядок $F_1(\tau) = F(\tau)$, поскольку из него не удаляется ни одной заявки.

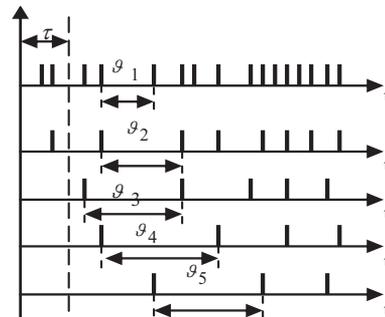


Рис. 1. Разреживание потока

Интервалы времени между моментами поступления двух соседних заявок для потока n -го порядка получаются в результате суммирования независимых случайных величин, распределенных по закону исходного потока. Следовательно, математическое ожидание $E(\mathcal{G}_n)$ интервалов между моментами времени поступления соседних заявок и соответствующая дисперсия $D(\mathcal{G}_n)$ определяются соотношениями:

$$\begin{aligned} E(\mathcal{G}_n) &= n E(\mathcal{G}_1); \\ D(\mathcal{G}_n) &= n D(\mathcal{G}_1). \end{aligned} \quad (4)$$

Коэффициент вариации интервалов времени для потока n -го порядка

$$v^2(\mathcal{G}_n) = \frac{D(\mathcal{G}_n)}{[E(\mathcal{G}_n)]^2} = \frac{v^2(\mathcal{G}_1)}{n} \quad (5)$$

уменьшается по сравнению с исходным потоком, с увеличением порядка n , что и следовало ожидать.

Рассмотрим распределение числа событий, попадающих на интервал τ для каждого из потоков, представленных на рис. 1. Обозначим через $P_i(n, \tau)$ вероятность поступления на интервале τ ровно i событий потока n -го порядка. Для реализации потока, показанной на рис. 1, следует, что в течение интервала времени τ поступили два события исходного потока, одно событие – потока второго порядка. Для потоков третьего, четвертого и пятого порядков не поступило ни одного события.

Вероятность отсутствия событий на интервале τ для потока n -го порядка равна вероятности того, что интервал между соседними событиями в этом потоке превышает τ :

$$P_0(n, \tau) = P(\mathcal{G}_n > \tau) = 1 - F_n(\tau). \quad (6)$$

Событие, состоящее в том, что на интервале τ возникнет ровно i заявок исходного потока, произойдет лишь при одновременном выполнении двух условий: $(\mathcal{G}_i \leq \tau)$ и $(\mathcal{G}_{i+1} > \tau)$. Следовательно, вероятности указанных событий

$$\begin{aligned} P_i(1, \tau) &= P(\mathcal{G}_i \leq \tau) \cdot P(\mathcal{G}_{i+1} > \tau) = \\ &= F_i(\tau) [1 - F_{i+1}(\tau)]. \end{aligned}$$

После соответствующих преобразований, получим

$$\begin{aligned} P_i(1, \tau) &= F_i(\tau) - F_{i+1}(\tau). \\ i &= 0; 1; 2 \dots \end{aligned} \quad (7)$$

Здесь необходимо учесть, что $F_0(\tau) = P(0 \leq \tau) = 1$.

Очевидно, что (7) соответствует (6) при $i = 0$, $n = 1$ и обобщает закон Пуассона на неэкспоненциальные потоки.

Учитывая, что представляет интерес лишь распределение вероятностей числа событий для исходного потока, $P_i(1, \tau)$ будем обозначать через $P_i(\tau)$, опуская 1.

Определим значение суммы вероятностей поступления от 0 до n заявок на интервале τ :

$$\begin{aligned} \sum_{i=0}^n P_i(\tau) &= \sum_{i=0}^n F_i(\tau) - \sum_{i=0}^n F_{i+1}(\tau) = \\ &= F_0(\tau) + \sum_{i=1}^n F_i(\tau) - \sum_{j=1}^{n+1} F_j(\tau) = \\ &= F_0(\tau) + \sum_{i=1}^n F_i(\tau) - \sum_{j=1}^n F_j(\tau) - F_{n+1}(\tau) = \\ &= 1 - F_{n+1}(\tau). \end{aligned}$$

Учитывая $\lim_{n \rightarrow \infty} F_{i+1}(\tau) = 0$, получим $\sum_{i=0}^{\infty} P_i(\tau) = 1$. Найдем математическое ожидание $E[n(\tau)]$

числа событий исходного потока, возникающих на постоянном интервале времени τ :

$$\begin{aligned} E[n(\tau)] &= \sum_{i=0}^{\infty} i \cdot P_i(\tau) = \sum_{i=1}^{\infty} i \cdot F_i(\tau) - \\ &- \sum_{i=1}^{\infty} i \cdot F_{i+1}(\tau) = F_1(\tau) + \sum_{i=2}^{\infty} i \cdot F_i(\tau) - \\ &- \sum_{i=1}^{\infty} (i+1) \cdot F_{i+1}(\tau) + \sum_{i=1}^{\infty} i \cdot F_{i+1}(\tau) = \\ &= F_1(\tau) + \sum_{i=2}^{\infty} i \cdot F_i(\tau) - \sum_{i=2}^{\infty} i \cdot F_i(\tau) + \\ &+ \sum_{i=2}^{\infty} F_i(\tau) = \sum_{i=1}^{\infty} F_i(\tau). \end{aligned}$$

Получили удивительный результат:

$$E[n(\tau)] = \sum_{i=1}^{\infty} F_i(\tau). \quad (8)$$

Определим второй начальный момент $E[n^2(\tau)]$ числа событий исходного потока, возникающих на постоянном интервале времени τ :

$$E[n^2(\tau)] = \sum_{i=0}^{\infty} i^2 \cdot P_i(\tau) = \sum_{i=0}^{\infty} i^2 \cdot F_i(\tau) - \sum_{i=0}^{\infty} i^2 \cdot F_{i+1}(\tau).$$

После преобразований получим

$$E[n^2(\tau)] = 2 \sum_{i=0}^{\infty} (i-1) \cdot F_i(\tau) + \sum_{i=1}^{\infty} F_i(\tau). \quad (9)$$

С учетом (8), получим окончательно

$$\frac{E[n^2(\tau)] - E[n(\tau)]}{2} = \sum_{i=1}^{\infty} (i-1) F_i(\tau). \quad (10)$$

Обозначив через $E[\Delta n(\tau)]$ выражение в левой части (10), получим

$$E[\Delta n(\tau)] = \sum_{i=1}^{\infty} (i-1) F_i(\tau). \quad (11)$$

В дальнейшем будет показано, что именно эта величина в СМО характеризует среднее число недообслуженных заявок за постоянное время τ .

Рассмотрим, представленный на рис. 2а стационарный поток заявок на обслуживание в СМО. Допустим, что все заявки обслуживаются с постоянным временем τ . Распределим указанные интервалы обслуживания на оси времени, как это показано на рис. 2., причем заявки, после которых начинается отсчет интервалов времени τ , считаются не принадлежащим указанным интервалам. Началом каждого очередного интервала τ считается момент появления заявки, которая непосредственно следует за окончанием предыдущего интервала времени обработки. Эту операцию следует повторить начиная отсчет времени обработки от каждой заявки.

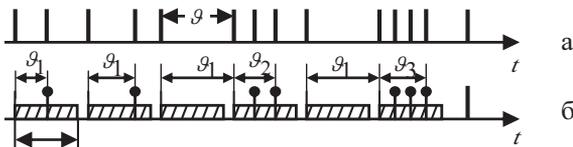


Рис. 2. Стационарный поток заявок

Из рис. 2б следует, что в течение постоянных интервалов обработки поступает различное число заявок i . Например, в течение первого и второго интервалов поступило по одной заявке, в течение последнего интервала – три заявки. В течение третьего и четвертого интервалов не поступило ни одной заявки. Из рис. 2б также следует, что

случайные величины \mathcal{G}_i получаются каждый раз в результате суммирования i интервалов \mathcal{G} между заявками исходного потока. При этом, всегда выполняется условие $\mathcal{G}_i \leq \tau$. Вероятность выполнения такого условия – $P(\mathcal{G}_i \leq \tau)$, как было показано в (3), равна $F_i(\tau)$. Каждому событию $\mathcal{G}_i \leq \tau$ соответствует появление на интервале обработки i заявок. Поскольку в течение интервала времени всегда полностью обрабатывается одна заявка, то недообслуженными будут оставаться каждый раз $(i-1)$ заявок. Из указанных рассуждений следует, что среднее число недообслуженных заявок, приходящихся на одну обслуженную заявку, определяется соотношением

$$\sum_{i=1}^{\infty} (i-1) F_i(\tau),$$

что полностью соответствует (10)-(11).

Таким образом, мы показали что выражения (10)-(11) определяют среднее число (долю) недообслуженных заявок, приходящееся на постоянное время τ обслуживания одной заявки. Именно указанной величине пропорциональна средняя длина очереди в СМО при любом законе распределения интервалов времени между соседними заявками.

Если интервалы времени обработки τ не постоянны и имеют плотность вероятностей $W(\tau)$, то

$$\begin{aligned} E(n) &= \int_0^{\infty} E[n(\tau)] \cdot W(\tau) dt \\ E(n^2) &= \int_0^{\infty} E[n^2(\tau)] \cdot W(\tau) dt = \\ &= \int_0^{\infty} E^2[n(\tau)] \cdot W(\tau) dt + \\ &+ \int_0^{\infty} D[n(\tau)] \cdot W(\tau) dt = [E(n)]^2 + D(n), \end{aligned} \quad (12)$$

где $D[n(\tau)]$ дисперсия числа заявок поступающих в течение постоянного интервала обслуживания τ .

Средняя доля недообслуженных заявок $E(\Delta n)$ определится соотношением

$$\begin{aligned} E(\Delta n) &= \int_0^{\infty} E[\Delta n(\tau)] W(\tau) dt = \\ &= \sum_{i=1}^{\infty} (i-1) \int_0^{\infty} F_i(\tau) W(\tau) d\tau. \end{aligned}$$

Если, например, каждое вновь возникающее событие обрабатывается с вероятностью P_k в течение времени τ_k ($k = 1; 2 \dots m$), то есть

$$W(\tau) = \sum_{k=1}^m P_k \cdot \delta(\tau_k - \tau), \text{ то}$$

$$E(\Delta n) = \sum_{k=1}^m E[\Delta n(\tau)] P_k \cdot \delta(\tau_k - \tau) =$$

$$= \sum_{k=1}^m E[\Delta n(\tau)] \cdot P_k = \sum_{k=1}^m P_k \times \quad (13)$$

$$\times \sum_{k=1}^m (i-1) \cdot F_i(\tau_k).$$

Соотношение (13) устанавливает зависимость характеристики СМО $-E[\Delta n(\tau)]$ от значений семейства образующих функции $F_i(\tau)$.

Получение аналитических выражений для функции распределение i -го порядка из функции $F_1(\tau)$ распределения вероятностей исходного потока в общем виде не всегда возможно. Однако, для потоков, интервалы между соседними заявками в которые подчиняются Γ -распределению, такие зависимости устанавливаются весьма просто. Допустим, что функция распределения интервалов между соседними заявками в исходном потоке подчиняется Γ -распределению

$$F(\tau) = \frac{\int_0^{\tau} x^{\eta-1} e^{-x} dx}{\Gamma(\eta)}, \quad (14)$$

где $x(\tau) = \frac{\eta \tau}{E(\mathcal{G}_1)}$; $\eta = \frac{1}{v^2(\mathcal{G}_1)}$ — величина, обратная квадрату коэффициента вариации интервалов между заявками в исходном потоке.

Поскольку случайные интервалы \mathcal{G}_i между соседними заявками в потоке i -го порядка получаются в результате последовательного суммирования i интервалов \mathcal{G}_1 исходного потока, функция распределения интервалов между соседними заявками, в потоке i -го порядка, также будет подчиняться Γ -распределению. В соответствии с (4)-(5) указанная функция имеет вид:

$$F_i(\tau) = \frac{\int_0^{\tau} x^i \eta^{-1} e^{-x} dx}{\Gamma(i \eta)}. \quad (15)$$

Подставляя значения из (15) в (7) для Γ -распределения интервалов, получим

$$P_i(1, \tau) = \frac{\int_0^{\tau} x^i \eta^{-1} e^{-x} dx}{\Gamma(i \eta)} - \frac{\int_0^{\tau} x^{(i+1) \eta^{-1}} e^{-x} dx}{\Gamma[(i+1) \eta]}. \quad (16)$$

Указанное соотношение является интегральным обобщением закона Пуассона на потоки с Γ -распределением интервалов между соседними заявками.

Поскольку экспоненциальное распределение интервалов есть частичный случай Γ -распределения при $\eta = 1$, следует ожидать, что в результате такой подстановки мы должны прийти к формуле закона Пуассона.

В указанном частном случае

$$P_i(1, \tau) = \frac{\int_0^{\tau} x^{i-1} e^{-x} dx}{\Gamma(i)} - \frac{\int_0^{\tau} x^i e^{-x} dx}{\Gamma(i+1)},$$

здесь $\eta = 1$.

Известно, что

$$\int_0^{\tau} x^m e^{a x} dx = \frac{1}{a} \cdot x^m e^{a x} - \int_0^{\tau} x^{m-1} \cdot e^{a x} dx,$$

с учетом этого соотношения

$$P_i(1, \tau) = \int_0^{\tau} \frac{x^{i-1} e^{-x} dx}{\Gamma(i)} + \frac{x(\tau)^i e^{-x(\tau)}}{\Gamma(i+1)} -$$

$$- \int_0^{\tau} \frac{x^i e^{-x} dx}{\Gamma(i+1)} = \frac{x(\tau)^i e^{-x(\tau)}}{i!} =$$

$$= \frac{\left[\frac{\tau}{E(\mathcal{G}_1)} \right]^i e^{-\frac{\tau}{E(\mathcal{G}_1)}}}{i!},$$

что и следовало ожидать. Получили закон Пуассона для исходного потока

$$P_i(1, \tau) = \frac{E[n(\tau)]^i e^{-E[n(\tau)]}}{i!}, \quad (17)$$

где

$$E[n(\tau)] = D[n(\tau)] = \frac{\tau}{E(\mathcal{G}_1)}. \quad (18)$$

Если интервалы времени обслуживания не постоянны, а имеют плотность вероятностей $W(\tau)$, то с учетом (12) и (18) для пуассоновского потока получим:

$$E(n) = \frac{E(\tau)}{E(\mathcal{G}_1)} = \rho;$$

$$E(n^2) = \frac{E(\tau^2)}{E^2(\mathcal{G}_1)} = \left[\frac{E(\tau)}{E(\mathcal{G}_1)} \right]^2 (1 + v_\tau^2) = \rho^2 (1 + v_\tau^2),$$

где $v_\tau^2 = \frac{D(\tau)}{[E(\tau)]^2}$ – коэффициент вариации интервалов времени обработки. Путем подстановки в (1), получим формулу Хинчина-Поллячека для экспоненциального распределения интервалов между заявками

$$q = \frac{\rho^2 (1 + v_\tau^2)}{2(1 - \rho)}, \quad (19)$$

где величина $E(\Delta n) = \frac{\rho^2 (1 + v_\tau^2)}{2}$ характеризует среднюю долю недообслуженных заявок, приходящуюся на одну обслуженную заявку в СМО.

Заключение

В соответствии с (1), средняя доля недообслуженных заявок определяется разностью второго и первого начальных моментов случайной величины числа заявок, поступающих в течение случайных интервалов времени обслуживания заявок. Именно указанные параметры целесообразно определять при экспериментальном исследовании СМО.

Литература

1. Основы теории вычислительных систем. Под ред. С.А. Майорова. М.: Высшая школа, 1978. – 408 с.
2. Клейнрок Л. Вычислительные системы с очередями. Пер. с англ. М.: Мир, 1979. – 600 с.

УДК 621.3:681.3

УНИФИЦИРОВАННОЕ ПРЕДСТАВЛЕНИЕ ДАННЫХ ДЛЯ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

Зияутдинов В.С., Слепцов Н.В.

Рассматривается применение генетических алгоритмов для оптимизации многомерных, плохо определенных функций. Рассмотрены два родственные невоичных преобразования – унифицированное преобразование разбиения, инкрементное преобразование объединения и их инверсии.

Введение

Эффективность работы элементов биологических систем вызывает пристальный интерес к использованию принципов биологической эволюции для задач оптимизации систем, практически важных для самых различных областей [1].

По сравнению с обычными оптимизационными методами методы эволюционно-генетических вычислений (ЭГВ) имеют следующие особенности: параллельный поиск, случайные мутации и рекомбинации уже найденных хороших решений. Они хорошо подходят как простой эвристический метод оптимизации многомерных, плохо определенных функций.

Наибольшее распространение из ЭГВ получили генетические алгоритмы (ГА) [2]. ГА – это компьютерная модель эволюции популяции искусственных «особей». Каждая особь характеризуется своей хромосомой S_k , хромосома ин-

терпретируется как «геном» особи и определяет приспособленность особи $f(S_k)$; $k = 1 \dots n$; n – численность популяции. Хромосома есть цепочка символов $S_k = (S_{k1}, S_{k2} \dots S_{kN})$, N – длина цепочки. Символы интерпретируются как «гены» особи, расположенные в хромосоме S_k . Задача алгоритма состоит в максимизации функции $f(S_k)$.

Эволюция состоит из последовательности поколений. Для каждого поколения отбираются особи с большими значениями приспособленности. Хромосомы отобранных особей рекомбинируются и подвергаются мутации. Формально, схема ГА может быть представлена следующим образом (популяция t -го поколения обозначается как $\{S_k(t)\}$):

A0. Создать случайную начальную популяцию $\{S_k(0)\}$.

A1. Вычислить приспособленность $f(S_k)$ каждой особи S_k популяции $\{S_k(t)\}$.

A2. Производя отбор особей S_k в соответствии с их приспособленностями $f(S_k)$ и применяя генетические операторы к отобранным особям, сформировать популяцию следующего поколения $\{S_k(t+1)\}$.

A3. Повторить шаги A1, A2 для $t = 0; 1; 2 \dots$ до тех пор, пока не выполнится условие окончания эволюционного поиска.