УДК 519.872

О.В. Коба, канд. техн. наук

СТАЦІОНАРНІ ХАРАКТЕРИСТИКИ СИСТЕМИ МАСОВОГО ОБСЛУГОВУВАННЯ *GI/G/*1 ІЗ *T*-ПОВЕРНЕННЯМ ПРИ ОБСЛУГОВУВАННІ В ПОРЯДКУ ЧЕРГИ

Національний авіаційний університет, elvin@svitonline.com

Узагальнено модель системи масового обслуговування, що введена Л. Лакатошем. Розглянуто систему обслуговування GI/G/1, в якій цикли заявок на орбіті дорівнюють сталій T і прийнята дисципліна обслуговування "першим прийшов, першим обслужився". Для випадку, коли час обслуговування завжди менше або дорівнює T, знайдено стаціонарний розподіл відповідного ланцюга Маркова. Визначено показник ефективності функціонування системи – середнє число заявок на орбіті.

Постановка проблеми

Теорія масового обслуговування виникла на початку XX ст. Ця теорія із самого початку стимулювалася проблемами інженерних розрахунків у галузі телефонії, транспорту (авіаційного, автомобільного тощо) та інших прикладних дисциплін. Класичні задачі теорії масового обслуговування обмежувалися моделями систем із чеканням та систем із відмовами [1]. Протягом останніх двох–трьох десятиліть з'явилася низка нових відгалужень теорії масового обслуговування: теорія стохастичних мереж, теорія систем із рухомими об'єктами та каналами обслуговування й багато інших.

Серед таких дисциплін, що фактично являють собою розділи науки зі своїми специфічними методами, протягом двох останніх десятиліть значний розвиток одержала теорія систем масового обслуговування (СМО) з заявками, що повертаються. Головним мотивуванням розвитку цієї теорії послужили телекомунікаційні системи. Дійсно, класична схема Ерланга припускала, що виклик, який одержав відмову, зникає. Тим часом, у випадку зайнятості лінії він може повторюватися через сталий чи випадковий час. Іншою мотивацією є обчислювальні системи, де у випадку переповнення буфера заявки, що надходять у процесор, блокуються і повертаються через визначений час. Нарешті, істотною причиною розвитку теорії систем обслуговування з поверненням заявок є система обслуговування, що служить моделлю роботи аеропорту [2]. У кожній із цих реальних систем, крім первинного потоку, створюється ще один – вторинний, який складається з заявок, що вже отримали відмову в обслуговуванні і спробують обслужитися знову.

Аналітичні результати для СМО з поверненням заявок досить складні для отримання за рахунок складності систем, тому на сьогодні розроблено значну кількість числових методів, що моделюють роботу подібних систем обслуговування. Аналітичні дослідження з теорії СМО з поверненням в Україні проводять І.М. Коваленко [3], В.В. Анісімов [4], О.В. Коба [5].

Аналіз досліджень і публікацій

Крім класичних систем M/G/1 з поверненням і M/G/S з поверненням, останнім часом активно вивчаються моделі СМО з поверненням заявок і місцями чекання, оскільки відповідні їм реальні системи широко використовують на практиці. Сфера їх застосування – комп'ютерні мережі, системи з замовленням у телефонії, комп'ютерні системи. Місця чекання (або буфер) використовують для підвищення ефективності обслуговування і зменшення впливу на систему заявок, що повертаються.

Система масового обслуговування з поверненням і груповим надходженням заявок– широко розповсюджена модель, оскільки описує функціонування комп'ютерних комутаційних мереж (колективний доступ). Проте в науковій літературі цьому типу моделей приділялося відносно мало уваги.

Багатоканальні СМО з поверненням заявок привертають увагу багатьох дослідників у зв'язку з важливістю їх додатків, насамперед, у телефонних комутаційних системах.

Основні аналітичні результати з теорії СМО з поверненням заявок були отримані з використанням апарату вкладеного ланцюга Маркова.

Традиційно в теорії СМО з поверненням розглядається показниково розподілений інтервал часу, через який заявка знову повертається в систему [6; 7]. У класифікації за Кендаллом для позначення закону розподілу часу, через який відбувається повернення заявки, що одержала відмову (за прийнятою термінологією, заявки, яка знаходилася на орбіті), навіть не виділяється позиція [6], оскільки за умовчанням цей час вважається показниково розподіленим. Таке припущення створює найбільш прості умови для математичного дослідження, оскільки в будь-який момент t для прогнозування подальшої поведінки системи достатньо знати лише кількість заявок, що одержали відмову і отже знаходяться на орбіті. Тим часом у багатьох додатках дане припущення суперечить фізичному процесу. Так, повітряне судно, яке не прийняте на посадку, може повертатися тільки через сталий час [2]. При загальному розподілі часу перебування заявки на орбіті аналіз системи більш складний і умова ергодичності такої системи в загальному випадку невідома. Це відноситься і до системи зі сталим часом перебування на орбіті.

Для систем з поверненням заявок і загальною функцією розподілу часу, через який заявка знову повертається в систему (зокрема, через детермінований час), були проведені досить незначні дослідження. При цьому завдяки розвитку інформаційних і телекомунікаційних технологій існує багато реальних систем, в яких час між поверненням заявок є детермінованим.

Автором були досліджені деякі моделі СМО з поверненням заявок і довільним (у більшості випадків сталим) часом перебування заявки на орбіті. Деякі результати наведено в роботах [5; 8; 9].

Описання системи

Розглянемо одноканальну СМО з дисципліною обслуговування в порядку черги *FCFS* (first come, first served), загальним розподілом часу між моментами надходження заявок вхідного потоку, загальним розподілом часу обслуговування заявки і сталим часом одного циклу перебування заявки на орбіті.

Нехай $A(t)$ – функція розподілу часу між моментами надходження заявок, $B(t)$ – функція розподілу часу обслуговування Y_n n-ї заявки, t_n – момент надходження n-ї заявки, $t_n + V_n$ – момент закінчення її обслуговування, T – час перебування заявки на одному циклі орбіти.

Як правило, модель СМО з поверненням заявок передбачає таку дисципліну обслуговування: за наявності вільного каналу на обслуговування приймається та заявка, яка першою застала канал вільним, незалежно від того, звідки вона потрапляє до каналу – з орбіти чи з первинного каналу.

У розглянутій моделі аналогічно моделі Лакатоша [2], заявка може бути прийнята до обслуговування в момент її надходження чи повернення з орбіти тільки в тому випадку, якщо ні в каналі, ні на орбіті заявок, що надійшли в систему раніше даної, немає.

Отже, обслуговування ведеться в порядку черги з урахуванням моментів надходження заявок, що перебувають на орбіті.

Метою дослідження СМО є виведення деяких стаціонарних характеристик наведеної системи, а саме: стаціонарного розподілу відповідного ланцюга Маркова і середньої кількості заявок на орбіті.

Враховуючи порядок формування черги і дисципліну обслуговування в системі, маємо, якщо

$$t_n \geq t_{n-1} + V_{n-1},$$

то n-а заявка надходить до обслуговування в момент t_n і тоді

$$V_n = Y_n.$$

Якщо для деякого цілого $k \geq 1$ виконується нерівність

$$(k-1)T < V_{n-1} - (t_n - t_{n-1}) \leq kT,$$

то n-а заявка надходить до обслуговування в момент $t_n + kT$, тобто

$$V_n = kT + Y_n.$$

Отже, n-а заявка приймається до обслуговування в момент $t_n + k_n T$, де k_n – мінімальне ціле число, при якому в даний момент система вільна від усіх попередніх заявок. Час $k_n T$ є часом перебування заявки на орбіті.

Легко бачити, що (k_n) є однорідний ланцюг Маркова. У загальному випадку різниця $k_n - k_{n-1}$ може бути як завгодно великою. Однак у відомих нам задачах (наприклад, деякі системи аеродромного обслуговування) природно прийняти умову, що завжди час обслуговування $Y_n \leq T$ або

$$B(T+) = 1. \qquad (1)$$

Умова (1) забезпечує нерівність

$$k_{n+1} - k_n \leq 1.$$

Дійсно, якщо $k_{n-1} = k$, то n-а заявка буде прийнята на обслуговування не пізніше, ніж $t_{n-1} + (k+1)T$, тобто

$$k_n \leq k + 1.$$

Позначимо

$$f_k = \mathrm{P}\{(k-1)T < Y_{n-1} - (t_n - t_{n-1}) \leq kT\} =$$

$$= \int_0^T (A(y - (k-1)T) - A(y - kT))dB(y).$$

Тоді за умови (1)

$$\sum_{k=-\infty}^{1} f_k = 1 .$$

Нехай p_{ij} – імовірності переходу ланцюга Маркова (k_n), тоді

$$p_{ij} = f_{j-i} \quad \text{при} \quad 1 \le j \le i+1; \qquad (2)$$

$$p_{i0} = \sum_{k=-\infty}^{-i} f_k . \qquad (3)$$

Умову ергодичності ланцюга Маркова (k_n) виведено в роботі [5]. Завдяки умові (1) вона перепишеться так:

$$\sum_{k=-\infty}^{1} k f_k < 0. \qquad (4)$$

Стаціонарний розподіл ланцюга Маркова

Припустимо, що умова (4) виконана, і позначимо через $\upsilon = (\upsilon_j)$ стаціонарний розподіл ланцюга Маркова (k_n). Із рівнянь (2), (3) знаходимо систему рівнянь

$$\upsilon_j = \sum_{i=j-1}^{\infty} \upsilon_i f_{j-i} \quad \text{при} \quad j \ge 1; \qquad (5)$$

$$\upsilon_0 = \sum_{i=0}^{\infty} \upsilon_i \sum_{k=-\infty}^{-i} f_k . \qquad (6)$$

До рівнянь (5), (6) додається умова нормування

$$\sum_{j=0}^{\infty} \upsilon_j = 1 .$$

Безпосередньо підстановкою можна пересвідчитися, що за умови (4) система (5)–(6) має такий імовірнісний розв'язок:

$$\upsilon_j = (1-z) z^j , \quad j \ge 1, \qquad (7)$$

де z – корінь рівняння

$$\sum_{k=-\infty}^{1} f_k z^{-k} = 1, \qquad (8)$$

що лежить в інтервалі (0,1).

Ліва частина рівняння (8) являє собою опуклу в напівінтервалі (0, 1] функцію, що прямує до нескінченності при $z \to 0$, оскільки $f_1 > 0$, і дорівнює одиниці при $z = 1$. Крім того, ліва похідна цієї функції в точці $z = 1$ додатна через умови (4). Отже, в інтервалі $(0, 1)$ існує єдиний корінь рівняння (8).

Цей факт також випливає з теорії неперервних справа випадкових блукань [10].

Розглянемо важливий випадок, коли час обслуговування $Y_n = \tau$, $t_n - t_{n-1}$ розподілений експоненціально з параметром λ.

У цьому випадку

$$f_1 = 1 - e^{-\lambda\tau};$$

$$f_k = e^{-\lambda\tau + \lambda k T} (1 - e^{-\lambda T}) \quad \text{при} \quad k \le 0 .$$

Умова ергодичності (4) набуває вигляду $z < 1$, де $z = e^{\lambda T}(1 - e^{-\lambda\tau})$.

Нехай потік заявок, що надходять до системи є груповий пуассонівський, причому число ζ заявок в одній групі – геометрично розподілена випадкова величина:

$$\mathrm{P}\{\zeta = k\} = (1-\theta)\theta^{k-1}, \quad k \ge 1,$$

а інтервали між групами є експоненціально розподілені випадкові величин із параметром λ. Тоді маємо таку формулу:

$$f_k = \theta 1_{\{k=1\}} + (1-\theta) f_k^0 ,$$

де f_k^0 – значення f_k при ординарному (негруповому) потоці Пуассона з параметром λ.

Величина $1_{\{k=1\}}$ дорівнює одиниці при $k = 1$ та нулю – в іншому разі.

Для f_k^0 справджується така формула:

$$f_k^0 = e^{-\lambda(\tau - kT)} (1 - e^{-\lambda T}) \quad \text{при} \quad k \le 0;$$

$$f_1^0 = 1 - e^{-\lambda\tau}.$$

Для параметра z є рівняння

$$\frac{\theta}{z} + (1-\theta)\left(\frac{1-a}{z} + \frac{a(1-b)}{1-bz} \right) = 1$$

або

$$\theta + (1-\theta)\left(1 - a + az \frac{1-b}{1-bz} \right) = z ,$$

де для спрощення запису позначено

$$a = e^{-\lambda\tau}; \; b = e^{-\lambda T}.$$

Розв'язавши його, отримаємо таку формулу для z за умови, що $z < 1$:

$$z = \frac{1 - a + a\theta}{b} = \frac{1 - e^{-\lambda\tau} + e^{-\lambda\tau}\theta}{e^{-\lambda T}} .$$

Якщо $\theta \to 0$, то $z \to \dfrac{1-a}{b}$, що відповідає випадкові пуассонівського потоку заявок.

Середнє число заявок на орбіті

Нехай $N(t)$ – число заявок на орбіті в момент t. Тоді інтеграл $\int_0^T N(t)dt$ є сумарний час перебування на орбіті тих заявок, що надійшли в систему в інтервалі $(0, T)$, за винятком залишкового часу чекання тих із них, що не були прийняті до обслуговування до моменту T.

З ергодичних міркувань [11] при великих значеннях s

$$\int_0^s \mathrm{E}[N(t)]dt \sim \lambda s \mathrm{E}[KT],$$

$$\lambda = \frac{1}{\int_0^\infty x dA(x)},$$

де K – стаціонарна версія k_n.

Звідси ергодичне середнє число заявок на орбіті

$$\lim_{s\to\infty} \frac{1}{s}\int_0^s \mathrm{E}[N(t)] = \lambda T \sum_{k=0}^{\infty} k(1-z)z^k = \frac{\lambda Tz}{1-z}.$$

Стаціонарне середнє число $\overline{K}$ циклів заявки на орбіті визначають за формулою

$$\overline{K} = \frac{z}{1-z}.$$

Висновки

Для узагальненої моделі Л. Лакатоша *GI/G/*1 при певному співвідношенні часу обслуговування і часу перебування на орбіті знайдено стаціонарний розподіл відповідного ланцюга Маркова. Як показник ефективності функціонування системи розглянуто середню кількість заявок, що перебувають на орбіті.

Система Лакатоша, не зважаючи на своєрідну організацію черги, має своє самостійне значення, оскільки вона служить тестуванням більш реальної системи обслуговування, в якій до каналу приймається та заявка, що першою до нього підійшла незалежно з первинного чи вторинного потоку.

Список літератури

1. *Гнеденко Б.В., Коваленко И.Н.* Введение в теорию массового обслуживания. – 2-е изд. – М.: Наука. – 1987. –336 с.
2. *Lakatos L.* A probability model connected with landing of airplanes. – Safety and Reliability. Vol. 1. – A. Balkema / Rotterdam/ Brookfield / 1999. – P. 151–154.
3. *Коваленко И.Н.* Вероятность потери в системе обслуживания *M/G/*1 с *T*-повторением вызовов в режиме малой нагрузки // Доп. НАНУ. – 2002. – №5. – С. 77–80.
4. *Anisimov V.V.* Switching stochastic models and applications in retrial queues // Sociedad de Estadística e Investigatión operativa. – Top. (1999. – Vol.7. – №2. – P. 169–186.
5. *Коба Е.В.* О системе обслуживания *GI/G/*1 с повторением заявок при обслуживании в порядке очереди // Доп. НАНУ. – 2000. – №6. – С.101–103.
6. *Falin G.I.* A survey of retrial queues // Queueing Systems.– 1990. – 7. – P. 127–168.
7. *Falin G.I., Templeton J.G.C.* Retrial Queues. – Chapman & Hall. London–Weinheim–New York–Tokyo–Melbourne–Madras. – 328 p.
8. *Koba O.V.* On a *GI/G/*1 retrial queueing system with *FIFO* queueing discipline // Theory of Random Processes. – №8(24). – K. – 2002. – Vol. 1–2.
9. *Коба Е.В.* Вероятность потери заявки в системе *M/D/*1 с постоянным временем возвращения // Доп. НАНУ. – 2002. – №4. – С. 61–66.
10. *Королюк В.С., Боровских Ю.В.* Аналитические ассимптотики вероятностных распределений. – К.: Наук. думка. – 1981. –348 с.
11. *Бочаров П.П., Печинкин А.А.* Теория массового обслуживания. – М.: Изд-во Рос. ун-та дружбы народов. – 1995. – 528 с.

Стаття надійшла до редакції 05.04.03.

Е.В.Коба

Стационарные характеристики системы массового обслуживания *GI/G/*1 с *T*-возвращением при обслуживании в порядке очереди

Обобщая модель системы массового обслуживания, введенную Л.Лакатошем, автор рассматривает систему обслуживания *GI/G/*1, в которой циклы заявок на орбите равны постоянной T и принята дисциплина обслуживания «первым пришел, первым обслужился». Для случая, когда время обслуживания всегда меньше или равно T, найдено стационарное распределение соответствующей цепи Маркова. Определен показатель эффективности функционирования системы – среднее число заявок на орбите.

O.V.Koba

Stationary parameters of a retrial queueing system with a *T*-returns of customers under a first come, first served discipline

Generalizing a model introduced by L.Lakatos, the author considers a *GI/G/*1 retrial queueing system in which orbit cycles of customers equal a constant T and a *FIFO* (first come, first out) queueing discipline is fulfilled. If a service time does not exceed T then a steady-state distribution of an imbedded Markov chain is obtained. Mean number of customers in the orbit is derived.