http://dx.doi.org/10.15407/dopovidi2016.05.030
УДК 519.872

І. М. Кузнєцов

Інститут кібернетики ім. В. М. Глушкова НАН України, Київ
E-mail: sea_hawk@icloud.com

Використання методу Монте-Карло для статистичної перевірки асимптотичної нормальності стаціонарного розподілу кількості вимог у системі $GI/G/\infty$ у випадку великого завантаження

(Представлено академіком НАН України І. М. Коваленком)

Розглядається система обслуговування $GI/G/\infty$ в умовах великого завантаження. Для перевірки гіпотези про асимптотично нормальний розподіл кількості вимог у системі пропонується застосовувати статистичний підхід, який грунтується на методі Монте-Карло. Використання χ^2-критерію дозволяє встановити мінімальне завантаження системи, починаючи з якого статистичні дані не суперечать цій гіпотезі.

Ключові слова: система обслуговування, велике завантаження, метод Монте-Карло, гіпотеза, нормальній розподіл, χ^2-критерій.

Останніми роками підвищена увага приділяється розвитку, дослідженню та впровадженню телекомунікаційних мереж. Це є однією з найбільш перспективних областей використання новітніх інформаційних технологій. Відомо, що процес, який описує кількість вимог у мережі, моделюється тією чи іншою системою масового обслуговування з одним чи декількома вхідними потоками. Як правило, ці потоки вважаються пуасонівськими. Це істотно спрощує розрахунки, оскільки тоді стаціонарні ймовірності станів системи обслуговування вдається обчислити в явному вигляді за допомогою формул Ерланга. Але далеко не завжди припущення про пуасонівський характер вхідних потоків підтверджується статистичними даними. У цьому випадку застосовують наближені методи, зокрема асимптотичні методи та методи прискореного моделювання. Значна частина досліджень зосереджена на системах обслуговування в умовах малого завантаження [1–4]. Але для сучасних телекомунікаційних мереж більш характерним є випадок високого завантаження, коли в мережі одночасно присутні декілька сотень викликів [5–8]. Останнім часом все більш поширеними стають спеціальні методи Монте-Карло (методи прискореного моделювання), які дозволяють на декілька порядків зменшити дисперсію оцінки, а тим самим досягти бажаної точності при відносно невеликих затратах часу на обчислення [9–11].

У даному повідомленні розглядається система обслуговування $GI/G/\infty$ (рекурентний вхідний потік, який визначається функцією розподілу $F(x)$, нескінчена кількість обслуговуючих пристроїв, час обслуговування має функцію розподілу $G(x)$). Метою дослідження є вивчення асимптотичної поведінки стаціонарного розподілу кількості вимог у системі за умови, що відношення середнього часу обслуговування до середнього часу між моментами надходження вимог необмежено зростає. При цьому зростає і середня кількість вимог у системі. Академік НАН України І.М. Коваленко висунув гіпотезу, що за відповідного нормування розподіл кількості вимог у системі буде наближатися до нормального розподілу. На чисельних прикладах покажемо, як метод Монте-Карло може бути застосованим не тільки до перевірки цієї гіпотези, але й до з'ясування мінімального завантаження системи, починаючи з якого статистичні дані не суперечать гіпотезі. При цьому використовується χ^2-критерій.

Збір статистичних даних методом Монте-Карло. Загальну схему збору статистичних даних та їх перевірки на відповідність нормальному розподілу сформулюємо наступним чином. Позначимо через $\{\xi_i, i = 1, 2, \ldots\}$ та $\{\eta_i, i = 1, 2, \ldots\}$ незалежні послідовності незалежних у кожній послідовності випадкових величин, які мають розподіли $F(x)$ та $G(x)$ відповідно. Припустимо, що $F(+0) < 1$, $\tau_F = \int\limits_0^\infty [1-F(u)]\,du < \infty$ та $\tau_G = \int\limits_0^\infty [1-G(u)]\,du < \infty$.

Нехай в момент часу $t = 0$ система обслуговування $GI/G/\infty$ знаходиться у стаціонарному режимі. Якщо момент $t = 0$ розглядати як момент надходження вимоги, то кількість вимог Z в цей момент визначається рівністю

$$Z = \sum_{n=1}^{\infty} I(\eta_n > \xi_1 + \ldots + \xi_n),$$

де $I(\cdot)$ — індикатор відповідної події. Якщо ж досліджується кількість вимог $Z^{(0)}$ в довільний стаціонарний момент часу, то відповідний рекурентний процес надходження вимог (який направлений у минуле) є стаціонарним, тобто у цьому випадку

$$Z^{(0)} = \sum_{n=1}^{\infty} I(\eta_n > \xi_1^{(0)} + \xi_2 + \ldots + \xi_n), \tag{1}$$

де $\mathbf{P}\{\xi_1^{(0)} < x\} = 1/\tau_F \int\limits_0^x [1 - F(u)]\,du$, $x > 0$. Саме на дослідженні випадкової величини $Z^{(0)}$ зосередимо основну увагу.

Очевидно, що зі збільшенням n кожний доданок ряду (1) все менш впливає на загальне його значення. Нехай $\varepsilon > 0$ — це деякий пороговий рівень, який визначає значення ймовірностей, якими будемо нехтувати. Позначимо

$$\beta = \inf\{t\colon 1 - G(t) < \varepsilon\}, \qquad N = \min\{n\colon \xi_1^{(0)} + \xi_2 + \ldots + \xi_n > \beta\}.$$

Замість випадкової величини $Z^{(0)}$, яка визначається рядом (1), будемо досліджувати випадкову величину

$$Y^{(0)} = \sum_{n=1}^{N} I(\eta_n > \xi_1^{(0)} + \xi_2 + \ldots + \xi_n), \tag{2}$$

яка визначається сумою із випадковою кількістю доданків (кожна подія $\{\eta_n > \xi_1^{(0)} + \xi_2 + \ldots + \xi_n\}$, $n \geqslant N$, має ймовірність, меншу за ε).

Моделюючи методом Монте-Карло послідовності випадкових величин $\{\eta_n\}$, $\xi_1^{(0)}$, $\{\xi_i, i \geqslant 2\}$, за формулою (2) знаходимо реалізації $y_1^{(0)}, \ldots, y_m^{(0)}$ випадкової величини $Y^{(0)}$, а також вибіркові значення математичного сподівання та дисперсії:

$$\overline{y}(m) = \frac{1}{m} \sum_{i=1}^{m} y_i^{(0)}, \qquad \sigma^2(m) = \frac{1}{m-1} \left\{ \sum_{i=1}^{m} [y_i^{(0)}]^2 - m\overline{y}^2(m) \right\}.$$

Утворимо нову послідовність $x_i^{(0)} = (y_i^{(0)} - \overline{y}(m))/\sigma(m)$, $i = 1, 2, \ldots, m$. Метою дослідження є перевірка гіпотези H_0: спостереження $\{x_i^{(0)}\}$ — це реалізації нормально розподіленої випадкової величини з математичним сподіванням 0 та дисперсією 1.

Для перевірки цієї гіпотези скористаємось χ^2-критерієм Пірсона. Нехай K — це деяке натуральне число. Область значень випадкової величини $Y^{(0)}$ розіб'ємо на області таким чином, щоб у кожну область попало щонайменше K спостережень. Якщо в останню область попало менш, ніж K спостережень, то цю область об'єднуємо з попередньою. Припустимо, що таких областей виявилось r:

$$A_1 = \{0, 1, \ldots, l_1\}, \qquad A_i = \{l_{i-1}+1, \ldots, l_i\}, \qquad i = 2, 3, \ldots, r-1,$$
$$A_r = \{l_{r-1}+1, l_{r-1}+2, \ldots\},$$

де $0 \leqslant l_1 < l_2 < \ldots < l_{r-1} < \infty$. Позначимо через ν_i кількість спостережень, які містить область A_i. Обчислимо теоретичну ймовірність потрапляння випадкової величини в область A_i за умови, що виконується гіпотеза H_0 :

$$p_1 = \Phi(v_1), \qquad p_i = \Phi(v_i) - \Phi(v_{i-1}), \qquad i = 2, \ldots, r-1, \qquad p_r = 1 - \Phi(v_{r-1}), \tag{3}$$

де $v_i = (l_i + 0{,}5 - \overline{y}(m))/\sigma(m)$, $i = 1, \ldots, r-1$, $\Phi(z) = 1/\sqrt{2\pi} \int\limits_{-\infty}^{z} e^{-u^2/2} du$. Далі утворюємо міру Пірсона

$$\Delta_m^{(r)} = \sum_{i=1}^{r} \frac{(\nu_i - mp_i)^2}{mp_i}. \tag{4}$$

Згідно з теоремою Пірсона асимптотичний (при $m \to \infty$) розподіл випадкової величини $\Delta_m^{(r)}$ є χ^2-розподілом з $r-1$ ступенем свободи. Позначимо цей розподіл $T_{r-1}(z)$. Якщо γ — це рівень значущості критерія, то за таблицями знаходимо $z_\gamma^{(r-1)}$ з рівняння: $T_{r-1}(z) = 1 - \gamma$. Якщо за результатами обчислення $\Delta_m^{(r)}$ згідно з (4) виявилось, що

$$\Delta_m^{(r)} \leqslant z_\gamma^{(r-1)}, \tag{5}$$

то статистична гіпотеза H_0 не суперечить отриманим статистичним даним. Якщо ж нерівність (5) не виконується, то гіпотезу H_0 слід відкинути.

Результати числового експерименту. Продемонструємо запропонований підхід на прикладі розподілів Вейбулла. Нехай

$$F(x) = 1 - \exp\{-(\rho x)^\alpha\}, \qquad G(x) = 1 - \exp\{-x^\alpha\},$$

де $\rho > 0$, $\alpha > 0$, $\tau_F = 1/\rho\Gamma(1 + 1/\alpha)$, $\tau_G = \Gamma(1 + 1/\alpha)$, $\Gamma(\mu) = \int\limits_0^\infty x^{\mu-1} e^{-x} dx$. Дослідимо асимптотичну поведінку стаціонарного розподілу кількості вимог у системі, коли завантаження $\tau_G/\tau_F = \rho$ зростає. Розглянемо три випадки: a) $\alpha = 1$ (марковська модель; стаціонарний розподіл знаходиться за явними аналітичними формулами); b) $\alpha = 2$; c) $\alpha = 0{,}5$. З'ясуємо, при яких значеннях ρ отримані статистичні дані не суперечать гіпотезі H_0. Виберемо $\varepsilon = 10^{-6}$, $m = 100000$, $K = 2000$ та $\gamma = 0{,}05$. Значення m та K підібрані таким чином, щоб ймовірність потрапляння у кожну з областей A_i була не нижчою за 0,02 за умови, що виконується гіпотеза H_0. Результати обчислення наведено у табл. 1–3. У випадку $\alpha = 1$, коли стаціонарний розподіл знаходиться у явному вигляді (розподіл Пуасона), обчислюються також точні значення ймовірностей $\{p_i\}$. Якщо їх підставити у формулу (4), то отримаємо відповідне значення $\Delta_m^{*(r)}$, яке має задовольняти нерівність (5) для будь-якого $\rho > 0$.

Якщо відомий стаціонарний розподіл $\{p_i\}$, то χ^2-критерій цілком підтверджує відповідність статистичних спостережень цьому розподілу (див. табл. 1), оскільки $\Delta_m^{*(r)} \leqslant z_\gamma^{(r-1)}$ для всіх розглянутих значень ρ. Якщо ж використовувати наближення нормальним розподілом згідно з формулами (3), то можна стверджувати, що швидкість збіжності суттєво залежить від виду функцій розподілу $F(x)$ і $G(x)$. Найбільшою швидкість збіжності є у випадку $\alpha = 2$ (так звані "старіючі" розподіли). Вже при $\rho \geqslant 300$ статистичні дані не суперечать гіпотезі H_0 (див. табл. 2). У випадку $\alpha = 0{,}5$ (так звані "молодіючі" розподіли) швидкість збіжності

Таблиця 1. Випадок $\alpha = 1$

ρ	$\overline{y}(m)$	$\sigma^2(m)$	$r-1$	$\Delta_m^{*(r)}$	$\Delta_m^{(r)}$	$z_\gamma^{(r-1)}$
100	100,0	100,2	30	29,46	108,6	43,77
200	200,0	200,9	36	29,03	77,72	50,00
300	300,0	299,8	38	36,09	88,49	53,38
400	400,1	396,7	33	29,44	51,03	47,40
500	500,1	499,2	34	24,03	51,28	48,60
600	600,0	598,0	37	29,18	38,31	52,19
700	700,0	698,5	38	25,85	36,66	53,38
800	800,0	800,3	39	29,76	35,89	54,57
900	899,7	901,2	40	39,41	37,28	55,76
1000	1000,0	1000,6	42	38,38	54,31	58,12

Таблиця 2. Випадок $\alpha = 2$

ρ	$\overline{y}(m)$	$\sigma^2(m)$	$r-1$	$\Delta_m^{*(r)}$	$\Delta_m^{(r)}$	$z_\gamma^{(r-1)}$
100	100,0	49,09	24	73,49	36,42	
200	200,7	97,14	31	63,49	44,99	
300	300,6	145,7	35	43,78	49,80	
400	400,6	193,7	37	39,69	52,19	
500	500,6	241,5	38	39,95	53,38	
600	600,7	293,1	39	46,43	54,57	
700	700,6	339,7	37	40,91	52,19	
800	800,7	390,4	33	34,97	47,40	
900	900,6	437,5	33	37,71	47,40	
1000	1000	486,5	34	28,42	48,60	

Таблиця 3. Випадок $\alpha = 0{,}5$

ρ	$\overline{y}(m)$	$\sigma^2(m)$	$r-1$	$\Delta_m^{*(r)}$	$\Delta_m^{(r)}$	$z_\gamma^{(r-1)}$
100	102,7	194,8	37	184,2	52,19	
200	202,7	393,5	33	80,80	47,40	
300	302,8	591,8	36	87,50	50,00	
400	402,7	793,7	39	87,79	54,57	
500	502,8	988,7	41	65,88	56,94	
600	602,8	1190	42	81,82	58,12	
700	702,8	1380	42	56,68	58,12	
800	803,0	1578	40	55,03	55,76	
900	902,9	1796	40	47,83	55,76	
1000	1003	1998	41	48,16	56,94	

є найменшою: пороговим значенням є $\rho = 700$. У випадку марковської моделі ($\alpha = 1$) гіпотезу H_0 можна приймати при $\rho \geqslant 600$ (див. табл. 1).

Цитована література

1. *Kovalenko I. N.* Approximation of queues via small-parameter method // Advances in Queueing. – Boca Raton: CRC Press, 1995. – P. 481–506.
2. *Blaszczyszyn B., Frey A., Schmidt V.* Light-traffic approximations for Markov-modulated multi-server queues // Commun. Statist. Stoch. Models. – 1995. – **11**, No 3. – P. 423–445.
3. *Wang C. L.* Light-traffic approximations for regenerative queueing processes // Adv. Appl. Probab. – 1997. – **29**, No 3. – P. 532–541.
4. *Kovalenko I. N., Atkinson J. B., Mikhalevich K. V.* Three cases of light-traffic insensitivity of the loss probability in a $GI/G/M/0$ loss system to the shape of the service time distribution // Queueing Systems. – 2003. – **45**, No 3. – P. 245–271.
5. *Ross K. W.* Multiservice loss models for broadband telecommunication networks. – London: Springer, 1995. – 288 p.
6. *Simonian A., Roberts J. W., Theberge F., Mazumdar R.* Asymptotic estimates for blocking probabilities in a large multi-rate loss network // Adv. Appl. Probab. – 1997. – **29**, No 4. – P. 806–829.
7. *Ливинская А. В., Лебедев Е. А.* Предельная теорема для перегруженных многоканальных сетей // Кибернетика и систем. анализ. – 2012. – № 6. – С. 106–113.
8. *Lebedev E., Livinska G.* Gaussian approximation of multi-channel networks in heavy traffic // Communications in Computer and Information Science. – 2013. – No 356. – P. 122–130.
9. *Mandjes M.* Fast simulation of blocking probabilities in loss networks // Europ. J. Oper. Res. – 1997. – **101**, No 2. – P. 393–405.
10. *Lassila P. E., Virtamo J. T.* Efficient importance sampling for Monte Carlo simulation of loss systems // Proc. of the ITC – **16**, Teletraffic Engineering in a Competitive World. – Edinburgh: Elsevier, 1999. – P. 787–796.
11. *Шумская А. А.* Оценка стационарной вероятности потери в системе массового обслуживания с рекуррентными потоками требований // Кибернетика и систем. анализ. – 2004. – No 2. – С. 133–145.

References

1. *Kovalenko I. N.* Advances in Queueing, Boca Raton, CRC Press, 1995: 481–506.
2. *Blaszczyszyn B., Frey A., Schmidt V.* Commun. Statist. Stoch. Models, 1995, **11**, No 3: 423–445.
3. *Wang C. L.* Adv. Appl. Probab., 1997, **29**, No 3: 532–541.
4. *Kovalenko I. N., Atkinson J. B., Mikhalevich K. V.* Queueing Systems, 2003, **45**, No 3: 245–271.
5. *Ross K. W.* Multiservice loss models for broadband telecommunication networks, London, Springer, 1995.
6. *Simonian A., Roberts J. W., Theberge F., Mazumdar R.* Adv. Appl. Probab., 1997, **29**, No 4: 806–829.
7. *Livinskaya A. V., Lebedev E. A.* Cybernetics and System Analysis, 2012, No 6: 106–113.
8. *Lebedev E., Livinska G.* Communications in Computer and Information Science, 2013, No 356: 122–130.
9. *Mandjes M.* Europ. J. Oper. Res., 1997, **101**, No 2: 393–405.
10. *Lassila P. E., Virtamo J. T.* Proc. of the ITC-16, Teletraffic Engineering in a Competitive World, Edinburgh, Elsevier, 1999: 787–796.
11. *Shumskaya A. A.* Cybernetics and System Analysis, 2004, No 2: 133–145.

Надійшло до редакції 17.11.2015

И. Н. Кузнецов

Институт кибернетики им. В. М. Глушкова НАН Украины, Киев

E-mail: sea_hawk@icloud.com

Использование метода Монте-Карло для статистической проверки асимптотической нормальности стационарного распределения количества требований в системе $GI/G/\infty$ в случае большой загрузки

Рассматривается система обслуживания $GI/G/\infty$ в условиях большой загрузки. Для проверки гипотезы об асимптотически нормальном распределении количества требований в системе предлагается использовать статистический подход, основанный на методе Монте-Карло. Использование χ^2-критерия позволяет установить минимальную загрузку системы, начиная с которой статистические данные не противоречат этой гипотезе.

Ключевые слова: система обслуживания, большая загрузка, метод Монте-Карло, гипотеза, нормальное распределение, χ^2-критерий.

I. N. Kuznetsov

V. M. Glushkov Institute of Cybernetics of the NAS of Ukraine, Kiev

E-mail: sea_hawk@icloud.com

Application of the Monte-Carlo simulation for the statistical testing of the hypothesis that a steady-state distribution of the number of customers in the queueing system $GI/G/\infty$ tends to the normal distribution in the case of heavy traffic

The queueing system $GI/G/\infty$ in a heavy traffic is considered. A statistical testing of the hypothesis that a steady-state distribution of the number of customers tends to the normal distribution is proposed. The χ^2-criterion enables us to determine a minimal load, from which the statistical data don't contradict this hypothesis.

Keywords: queueing system, heavy traffic, Monte-Carlo method, hypothesis, normal distribution, χ^2-criterion.