Tạp Chí Tin Học và Điều Khiển Học
Tập IX(1993), Số 1, 19-27

Mô Phỏng Quá Trình Phục Vụ Người Đọc Ở Một Thư Viện

Đoàn Phan Tân
Khoa thư viện
Trường đại học Văn hóa Hà nội

I. Mô hình toán học của quá trình phục vụ người đọc ở một thư viện

Quá trình đến của một yêu cầu của người đọc tại một thưviện là quá trình ngẫu nhiên. Giả sử ký hiệu T_n là thời điểm đến của yêu cầu thứ n sau thời điểm ban đầu $T_0 = 0$.

Khoảng thời gian

$$\tau_n = T_n - T_{n-1}, \; n = 1, 2, ... \qquad (1.1)$$

gọi là khoảng thời gian giữa hai yêu cầu liên tiếp. Ta nhận thấy rằng các khoảng thời gian này là những biến ngẫu nhiên độc lập, không âm, tuân theo cùng một quy luật xác suất nào đó. Ta gọi đó là quá trình đến của dòng yêu cầu.

Giả sử phòng đọc của thư viện có một kênh phục vụ. Nếu tại thời điểm người đọc đưa yêu cầu kênh phục vụ rỗi, thì yêu cầu được đáp ứng ngay, còn nếu kênh phục vụ bận thì người đọc phải xếp hàng đợi. Ta luôn giả thiết rằng người đọc được phục vụ theo thứ tự đến, không có ưu tiên.

Thời gian cần thiết để phục vụ một yêu cầu gọi là thời gian phục vụ. Ta ký hiệu t_n là thời gian phục vụ người đọc thứ n, $n = 1, 2, \dots$. Đó cũng là những biến ngẫu nhiên độc lập, không âm tuân theo một quy luật xắc suất nào đó. Khoảng thời gian giữa thời điểm đến của một yêu cầu và thời điểm yêu cầu bắt đầu được phục vụ gọi là thời gian đợi. Thời gian đợi của người đọc thứ n ký hiệu là w_n. Nếu không còn người nào trước thời điểm đến T_n chưa được phục vụ thì $w_n = 0$.

Nếu gọi R_n là thời điểm người đọc thứ n nhận được tài liệu và rời hệ thì

$$R_n = T_n + w_n + t_n, \; n = 1, 2, ... \qquad (1.2)$$

Thời gian đợi của người đọc thứ n rõ ràng bằng 0 nếu anh ta đưa yêu cầu sau lúc người đọc thứ $n-1$ rời hệ tại thời điểm R_{n-1}. Vì thế ta có

$$w_n = \begin{cases} o, & \text{nếu } R_{n-1} \le T_n \\ R_{n-1} - T_n & \text{nếu } R_{n-1} > T_n. \end{cases} \tag{1.3}$$

Tổng thời gian đợi và thời gian phục vụ một yêu cầu gọi là thời gian lưu của người đọc trên hệ. Ký hiệu d_n là thời gian lưu của người đọc thứ n ta có

$$d_n = w_n + t_n. \tag{1.4}$$

Ta thấy rằng thời gian đợi của người đọc thứ n cũng bằng 0 nếu khoảng thời gian giữa anh ta và người đến trước là τ_n không nhỏ hơn khoảng thời gian d_{n-1} mà người đến trước lưu trên hệ. Vì vậy ta có

$$w_n = \begin{cases} 0 & \text{nếu } d_{n-1} \le \tau_n, \\ d_{n-1} - \tau_n & \text{nếu } d_{n-1} > \tau_n. \end{cases} \tag{1.5}$$

Ta cũng nhận thấy nếu người đọc thứ $n-1$ rời hệ mà người đọc thứ n chưa tới thì hệ rỗi. Ký hiệu thời gian rỗi này là f_n thì ta có

$$f_n = \begin{cases} T_n - R_n & \text{nếu } T_n > R_{n-1} \\ 0 & \text{nếu } T_n \le R_{n-1} \end{cases}$$

$$= \begin{cases} \tau_n - d_{n-1} & \text{nếu } \tau_n > R_{n-1} \\ 0 & \text{nếu } \tau_n \le d_{n-1}. \end{cases} \tag{1.6}$$

Khoảng thời gian mà hệ liên tục bận gọi là thời kỳ bận. Trong thời kỳ bận có thể có một số người đọc có mặt ở hệ đợi hoặc đang được phục vụ. Ta ký hiệu số người đọc có mặt trên hệ ở thời điểm t là x_t. Khi đó $x_t = 0$ có nghĩa là hệ rỗi tại thời điểm t, còn $x_t = n$ có nghĩa là tại thời điểm t hệ đang bận và có $n-1$ người đọc đang đợi được phục vụ.

Ta ký hiệu v_t là tổng thời gian phục vụ x_t người đọc tại thời điểm t thì ta có

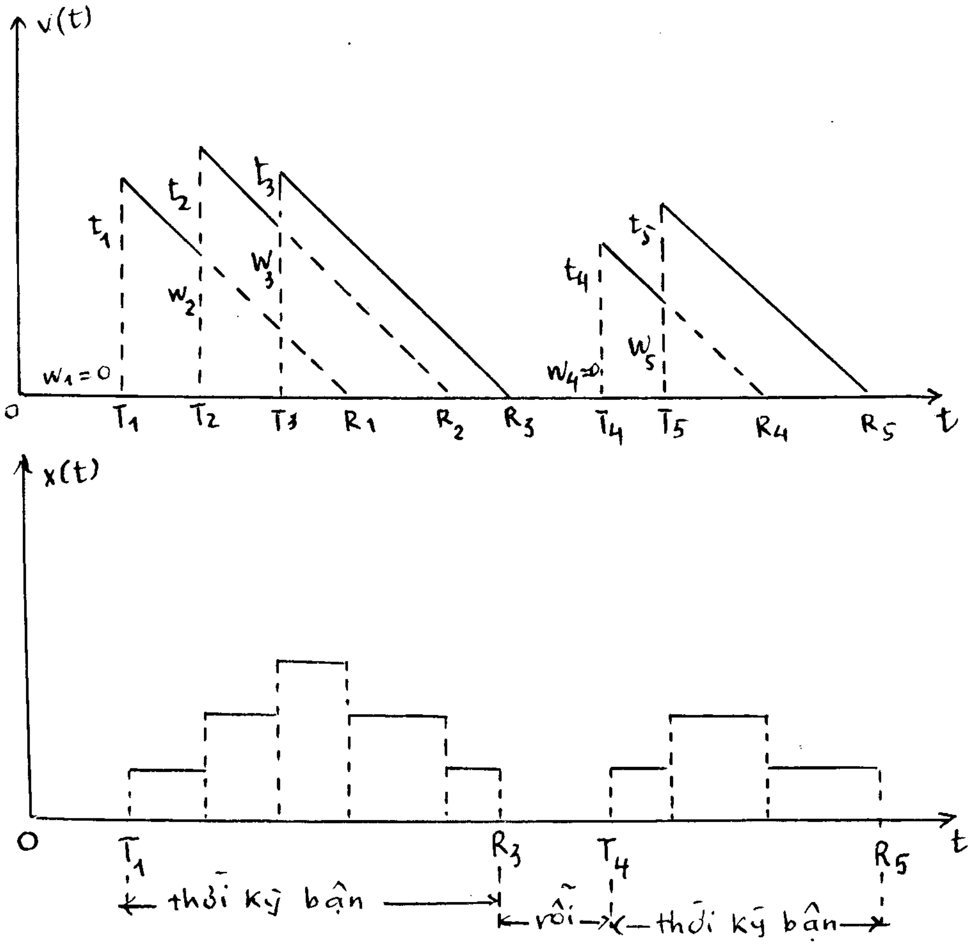
$$v_t = \begin{cases} 0 & \text{nếu } x_t = 0 \\ t_1' + t_2 + \dots + t_n & \text{nếu } x_t = n \ge 1. \end{cases} \tag{1.7}$$

trong đó t_1' là thời gian phục vụ còn lại của người đọc đang được phục vụ tại thời điểm t, còn $t_2, \dots, t_n$ là thời gian phục vụ của $n-1$ người đọc tại thời điểm t.

Biểu diễn đồ thị của v_t và x_t được trình bày ở hình vẽ dưới đây cho ta hình ảnh minh họa các khái niệm trên.

Các yếu tố chính tác động lên quá trình phục vụ người đọc ở một thư viện là:

- Quy luật đến của dòng yêu cầu của người đọc.
- Quy luật về thời gian ph ục vụ một yêu cầu.
- Quy luật về thời gian đợi của người đọc.



Hình 1

Nói chung người ta cần quan tâm đến tác động của hai yếu tố dẫn đến yếu tố thứ ba vì đó là thước đo tính hiệu quả của một hệ thống phục vụ.

Người ta chứng minh được rằng số yêu cầu của người đọc xuất hiện trong một khoảng thời gian t là một biến ngẫu nhiên N tuân theo quy luật Poisson và xác suất để có $N = n$ yêu cầu trong khoảng thời gian t là

$$P_n(t) = ((\lambda t)^n \, e^{-\lambda t})/n!, \tag{1.8}$$

trong đó

$$\lambda = E(N)/t$$

chính là số yêu cầu trung bình của người đọc trong một đơn vy thời gian.

Nếu dòng yêu cầu là dòng Poisson với trung bình λ thì khoảng thø2wh i gian giữa hai yêu cầu liên tiếp là biến ngẫu nhiên τ tuân theo quy luật phân phối mũ với tham số λ, tức là hàm phân phối xáx suất của nó có dạng

$$F(t) = \begin{cases} 0 & \text{nếu } t \leq 0 \\ 1 - e^{-\lambda t} & \text{nếu } t > 0 \end{cases} \tag{1.9}$$

và

$$E(\tau) = 1/\lambda$$

là trung bình của khoảng thời gian giữa hai yêu cầu.

Còn với thời gian phục vụ một yêu cầu, người ta cũng chứng minh được rằng đó là một biến ngẫu nhiên T tuân theo quy luật phân phối mũ với hàm phân phối là

$$F(t) = \begin{cases} 0 & \text{nếu } t \leq 0 \\ 1 - e^{-\mu t} & \text{nếu } t > 0. \end{cases} \tag{1.10}$$

Kỳ vọng toán học của T là

$$E(T) = 1/\mu.$$

Do đó

$$\mu = 1/E(T)$$

chính là số yêu cầu trung bình được phục vụ trong một đơn vị thời gian.

Đối với bài toán đặt ra, phương pháp mô phỏng đem lại nhiều lợi ích. Nhờ đó ta có thể phân tích đánh giá được hiện trạng hoạt động của hệ thống phục vụ, và xác định được phương án phục vụ tốt hơn.

2. Mô phỏng quá trình phục vụ người đọc.

Trong [8] chúng tôi đã trình bày các kết quả điều tra chọn mẫu và nghiên cứu hoạt động phục vụ ở phòng đọc của thư viện Khoa học kỹ thuật Trung ương, chúng tôi đã thu được các kết quả sau đây:

- Dòng yêu cầu của người đọc là dòng Poisson với trung bình là $\lambda = 0.08$ yêu cầu / phút. Khi đó khoảng thời gian giữa hai yêu cầu là biến ngẫu nhiên τ tuân theo quy luật phân phối mũ với trung bình là $E(\tau) = 1/\lambda = 1/0.08 = 1.25$ phút.
- Thời gian phục vụ T tuân theo quy luật phân phối mũ với trung bình là $E(T) = 1/\mu = 6.72$ phút và $\mu = 1/6.72 = 0.15$ là số yêu cầu trung bình được phục vụ trong một phút.

Ta biết rằng nếu x là một biến ngẫu nhiên liên tục, có hàm phân phối $F(x)$ và u là một thể hiện của biến ngẫu nhiên U tuân theo quy luật phân phối đều trên khoảng $[0,1]$ thì $x = F^{-1}(u)$ là một thể hiện mô phỏng của biến ngẫu nhiên X.

Ở đây hàm phân phối của thời gian phục vụ T có dạng

$$F(t) = \begin{cases} 0 & \text{với } t \leq 0 \\ 1 - e^{-\mu t} & \text{với } t > 0. \end{cases} \tag{2.1}$$

Hàm ngược của F là

$$F^{-1}(u) = -1/\mu.ln(1-u). \tag{2.2}$$

Do đó với số ngẫu nhiên u_1 thì giá trị mô phỏng của thời gian phục vụ T là

$$T_1 = -1/\mu.lnu_1 = 6.72(-lnu_1).$$

Tương tự với số ngẫu nhiên v_1, giá trị mô phỏng của khoảng thời gian giữa hai yêu cầu liên tiếp τ là

$$\tau_1 = -1/\lambda lnv_1 = 12.5(-lnv_1).$$

Từ mô hình toán học của quá trình phục vụ người đọc trình bày ở trên ta có thể xây dựng được thuật toán mô phỏng quá trình phục vụ người đọc ở thư viện Khoa học Kỹ thuật Trung ương.

Giả sử $T_1 = 0$ là thời điểm đến của yêu cầu thứ nhất. Thời điểm đến của yêu cầu tiếp theo được tính theo công thức

$$T_i = T_{i-1} + \tau_i,\ i = 1, 2, \ldots$$

trong đó τ_i là khoảng thời gian giữa hai yêu cầu liên tiếp thứ $i-1$ và i.

Từ bảng số ngẫu nhiên, ta lấy ra một số ngẫu nhiên, chẳng hạn lấy từ dưới lên ta được số $v_2 = 0.69$. Khi đó khoảng thời gian đến giữa yêu cầu thứ nhất và thứ hai có giá trị mô phỏng là $\tau_2 = 12.5(-ln0.69) = 12.50.37 = 4.64$ phút.

Lấy tiếp số ngẫu nhiên theo thứ tự từ trên xuống, ta được số $u_1 = 0.10$. Khi đó giá trị mô phỏng thời gian phục vụ yêu cầu thứ nhất là $t_1 = 6.72(-ln10) = 6.722.3 = 15,47$ phút.

Thời gian đợi của yêu cầu thứ nhất là $w_1 = 0$ nên thời gian lưu và thời điểm yêu cầu thứ nhất rời hệ là

$$d_1 = w_1 + t_1 = 0 + 15,47 = 15,47$$

$$R_1 = T_1 + w_1 + t_1 = 0 + 0 + 15,47 = 15,17.$$

Vì yêu cầu thứ hai đến vào thời điểm

$$T_2 = T_1 + \tau_2 = 0 + 4,64 = 4,64$$

trước thời điểm R_1 nên yêu cầu thứ hai phải xếp hàng và thời gian đợi là

$$w_2 = R_1 - T_2 = 15,47 - 4,64 = 10,83$$

Lấy số ngẫu nhiên tiếp theo $u_2 \doteq 0.09$. Thời gian phục vụ yêu cầu thứ hai có giá trị mô phỏng là

$$t_2 = 6.72(-ln0.09) = 6.722.41 = 16,18$$

yêu cầu này rời hệ tại thời điểm

$$R_2 = T_2 + w_2 + t_2 = 4.64 + 10.83 + 16.18 = 31.65$$

Thời gian lưu của yêu cầu thứ hai là

$$d_2 = w_2 + t_2 = 10.83 + 16.18 = 27.01$$

Lấy số ngẫu nhiên tiếp theo $v_3 = 0.07$, khoảng thời gian giữa yêu cầu thứ hai và thứ ba được mô phỏng bởi

$$\tau_3 = 12.5(-ln0.07) = 12.52.66 = 33.24$$

Thời điểm đến của yêu cầu thứ ba là

$$T_3 = T_2 + \tau_3 = 4.64 + 33.24 = 37.88$$

Ta nhận thấy yêu cầu thứ ba đến sau lúc yêu cầu thứ hai rời hệ ($T_3 > R_2$) nên trong khoảng thời gian này hệ rỗi và thời gian rỗi trước lúc y—6eu cầu thứ ba đến hệ là

$$f_3 = T_3 - R_2 = 37.88 - 31.65 = 6.23$$

Trước lúc đến hệ rỗi, nên yêu cầu thứ ba không phải đợi ($w_3 = 0$). Thời gian phục vụ yêu cầu thứ ba, ứng với số ngẫu nhiên $u_3 = 0.73$ được mô phỏng bởi

$$t_3 = 12.5(-ln0.73) = 2.11$$

Do đó yêu cầu thứ ba rời hệ vào thời điểm

$$R_3 = T_3 + w_3 + t_3 = 37.88 + 0 + 2.11 = 39.99$$

Thời gian lưu của yêu cầu thứ ba là

$$d_3 = w_3 + t_3 = 0 + 2.11 = 2.11$$

Tiếp tục quá trình trên ta được bảng tính sau đây mô phỏng quá trình phục vụ cho $n = 10$ người dọc liên tiếp

yêu cầu i	Số ng. nhiên v_i	K.t.g.giữa 2 yêu cầu $\tau_i = 12.5$ (-lnv_i)	Thời đ. đến T_i $= T_{i-1}$ $+\tau_i$	Số ng. nhiên u_i	T.g. ph. vụ $t_i =$ 6.72ln($-u_i$)	T.g đ. $w_i =$ R_{i-1} $-T_i$	T.đr. hệ R_i $= T_i$ $+w_i + t_i$	T.g l. $d_i =$ w_i $+t_i$	T.g. rỗi $f_i =$ T_i $-R_{i-1}$
1			0	0.10	15.47	0	15.47	15.47	0
2	0.69	4.64	4.61	0.09	16.18	10.83	31.65	27.01	0
3	0.07	33.24	37.88	0.73	2.11	0	39.99	2.11	6.23
4	0.49	8.92	46.80	0.25	9.32	0	56.12	9.32	6.81
5	0.41	11.14	57.94	0.33	7.45	0	65.39	7.45	1.82
6	0.38	12.09	70.13	0.76	1.84	0	71.97	1.84	4.74
7	0.87	1.74	71.87	0.52	4.39	0.10	76.36	4.49	0
8	0.63	5.78	77.65	0.01	30.95	0	108.60	30.95	1.29
9	0.79	2.95	80.60	0.35	7.05	28.00	115.65	35.05	0
10	0.19	20.76	101.36	0.86	1.01	14.29	116.66	15.30	0
$\sum$					S=95.77 (f)	C=53.22 (f)	=Q	P=149.99 (g)	L=20.89 (f)

Từ bảng trên ta có thể rúr ra kết luận sau đây đối với quá trình phục vụ người yêu cầu:

- Tổng thời gian phục vụ là
$$S = \sum t_i = 95.77 \text{ phút.}$$
- Tổng thời gian rỗi của hệ là
$$L = \sum f_i = 21.20 \text{ phút.}$$
- Thời gian hoạt động của hệ là
$$Q = L + S = 95.77 + 20.89 = 116.66 \text{ phút.}$$

hay là
$$Q = T_{10} + d_{10} = 101.36 + 15.30 = 116.66 \text{ phút.}$$
- Tổng thời gian đợi của người đọc là
$$C = \sum w_i = 53.22 \text{ phút.}$$
- Tổng thời gian lưu của người đọc trên là
$$P = S + C = 95.77 + 53.22 = 148.99 \text{ phút.}$$
- Thời gian đợi trung bình là
$$C/N = 53.22/10 = 5.32 \text{phút.}$$
- Thời gian lưu trung bình là
$$P/N = 148.99/10 = 14.89 \text{phút.}$$
- Tỷ lệ giữa thời gian rỗi và thời gian hoạt động của hệ là
$$L/Q = 21.20/148.99 = 0.18.$$

Sau đây ta xây dựng một sơ đồ khối thể hiện thuật toán mô phỏng quá trình phục người đọc đã được trình bày ở trên.

Ta sử dụng các ký hiệu sau đây

τ - khoảng thời gian giữa hai yêu cầu liên tiếp.

t - Thời gian phục vụ một yêu cầu.

w - Thời gian đợi của một yêu cầu.

f - Thời gian rỗi của hệ trước lúc yêu cầu đến.

T - Thời điểm đến của một yêu cầu.

R - Thời điểm yêu cầu rời hệ.

S - Tổng thời gian phục vụ N yêu cầu.

C - Tổng thời gian đợi của N yêu cầu.

P - Tổng thời gian lưu của N yêu cầu.

Q - Tổng thời gian hoạt động của hệ.

L - Tổng thời gian rỗi của hệ.

Ta có các quan hệ sau
$$P = S + C, \ Q = S + L.$$

Dựa vào các điều đã trình bày trên đây chúng tôi đã viết chương trình mô phỏng quá trình phục vụ người đọc ở một thư viện Khoa học và Kỹ thuật trung ương. Chương trình được viết bằng ngôn ngữ BASIC. Hoạt động của chương trình được tổ chức theo chế độ menu. Các tùy chọn là như sau:

1 - Mô phỏng số yêu cầu của người đọc.
2 - Mô phỏng thời gian phục vụ.
3 - Mô phỏng quá trình phục vụ.
4 - Ra khỏi chương trình.

Chương trình con 1 và 2 hoạt động như những chương trình độc lập, cung cấp cho ta những thông tin về dòng yêu cầu và thời gian phục vụ như: số yêu cầu trung bình trong khoảng thời gian định trước, khoảng thời gian trung bình giữa thời điểm đến của hai yêu cầu liên tiếp, thời gian trung bình phục vụ một yêu cầu. Nhưng khi chương trình con 3 hoạt động thì chương trình con 1 và 2 cung cấp các tham số λ và μ cho chương trình con 3. Khi chạy chương trình này, ta chỉ cần đưa vào số yêu cầu N cần mô phỏng.

Kết quả mô phỏng sẽ cung cấp cho ta những thông tin sau đây về quá trình phục vụ N yêu cầu của người đọc ở một thư viện:

- Tổng thời gian phục vụ.
- Tổng thời gian rỗi.
- Tổng thời gian sắp hàng.
- Tổng thời gian hoạt động của hệ.
- Tỷ lệ giữa thời gian rỗi và thời gian hoạt động của hệ.
- Trung bình của thời gian đợi.
- Trung bình của trhời gian lưu.

Đó là những tham số đặc trưng cho một quá trình phục vụ.

Nhưng dữ liệu ban đầu cho chương trình này có hai loại:

- Phân phối thực nghiệm của số yêu cầu.
- Phân phối thực nghiệm của thời gian phục vụ.

Khi tiến hành cho chương trình chạy với những dữ liệu khác nhau, phân tích các kết quả mô phỏng nhận được, ta có thể xác định được một hình thức tổ chức phục vụ thích hợp theo yêu cầu đề ra.

Tài liệu tham khảo

1. Lee A.M., Les files d'attent, Theorie et application, Paris, Dunod, 1970.
2. Kaufmann A., Cours modern de calcul des probabilite's, Paris Albin Michel, 1965.
3. Bernad Grais., Méthodes statistiques, Paris, Dunod, 1981.
4. Leroudier J., La simulation à événements discrets, Surenes, Hommes et Technique, 1980.
5. Lebart Ludevic & Fénelon Jean Pierre, Stastiques et informatique appliques, Paris, Dunod, 1975.
6. Cohen J.M., The single server queue, Amsterdam, North Holland Publ. Co. 1969.

7. Boisgontier Jaques, Le basic et ses fichiers, Laguysur - marne cedex, P.S.I. 1984.
8. Đoàn Phan Tân, *Mô hình hóa quá trình phục vụ người đọc ở một thư viện KHKT,* Tập san Công tác Thư viện Khoa học Kỹ thuật, (1988), n° 2, 14-24. an

Abstract

Simulation of the process of service in a library

This article presents an application of queuing theory in order to establish a mathematical model for the process of service in a library or in information center.

On the basic of the model mentioned and applying the simulation method of Monte - Carlo with properties of random numbers, the author proposes an algorithm and a flowchart for writing a simulation program of this process. The input data of this program is composed:

- Distribution of the number of andience's requests.

- Distribution of the service time.

Afftcr running, this program will give the following information on service system: Total of service time, total of writing time, total of free time, average of service time, average of writing time...

Thus are sharacteristics features of the service system.