ATLANTIS PRESS
Advances in Economics, Business and Management Research, volume 100
International Conference of Organizational Innovation (ICOI 2019)

USING STATISTICS ANALYSIS TO STUDY SERVICE SYSTEM OPTIMIZATION OF CAPACITY- LIMITED READING PAVILION

Kaijun Yu (Library, Shanghai University of Medicine& Health Sciences, Shanghai, China)
Ruiyi Gong (Library, Shanghai University of Medicine& Health Sciences, Shanghai, China)
Longjie Sun (Library, Shanghai University of Medicine& Health Sciences, Shanghai, China)
Rui Wang (Micro Mechatronics Laboratory, Yamaguchi University, Ube, Japan)

Email：sunlj@sumhs.edu.cn

Abstract—In order to allow more teachers and students to make good use of the library, it is necessary to optimize the number of existing reading pavilion, reasonably plan the number of future reading pavilion, improve the efficiency of use, reduce the idle rate and waiting time in line, and avoid the loss of readers and waste of resources. Based on statistics, queuing theory, time series prediction and other related theories. Select the Nanyuan Library of Shanghai University of Medicine & Health Sciences as evaluation object to collect data and analyse the key parameters and operational indicators. According to actual operating index parameters and their corresponding ranges, the current optimal configuration description is given. Develop a reasonable budget plan based on analysis and forecast data of time series related statistics.

Keywords—Stratified Sampling, Reading Pavilion, Time Series Analysis, Library

I. INTRODUCTION

The reading pavilion of domestic university libraries has risen around 2000. In just over 10 years, many domestic libraries have regarded the construction of reading pavilion as a mandatory item. Reading Pavilion can provide readers with relatively independent space, which is convenient for searching literature, deep learning, academic discussion, etc.. Nowadays, the management of the Reading Pavilion in the university library has gradually spread from the manual management stage to the network reservation. The user selects the room through the network to make an appointment. The system automatically reviews the user, prompts whether the appointment is successful, and the reservation result is timely. Students only need to take the campus card and open the door at the appointed time to the agreed room [1-5]. Many university libraries have encountered some problems in their actual operation. For example, the long waiting time in the peak period and the low efficiency in some time periods have affected the service quality of the library. Therefore, it is imperative to study and develop a scientific and reasonable and feasible optimization plan. Based on the above problems combined with the queuing optimization theory, this paper conducted an in-depth study on the queuing problem of reading pavilion, and found that the reading pavilion system should belong to the M/M/C/N model, and solved the parameters to find the allocation scheme of reading pavilion that met the actual operation requirements[6-11].

II. SERVICE SYSTEM MODEL OF READING PAVILION

A. *M/M/C/N model*

The standard form of the queuing model is X/Y/Z/A/B/C, where X represents the distribution of successive arrival intervals, Y represents the distribution of service time, Z represents the number of Reading Pavilions, A represents the capacity limit of the system represents the number of readers, and C represents the service rules[12-13].

The arrival of the reader corresponds to the Poisson flow. Both the reader arrival interval X and the reading pavilion service time Y follow the negative exponential distribution, so X=M, Y=M. Each Reading Pavilion room can accommodate up to N people, so the queuing model is M/M/C/N.

When the status of the system is n, the service rate between each service is μ. When $0 < n < c$, the total service rate of the system is nμ; When $n \geq c$, the total service rate of the system is cμ. Let $\rho = \left(\frac{\lambda}{c\mu}\right)$ be the average service strength of the system, where λ is the average number of readers arriving at the reading pavilion per hour, and μ is the number of readers per reading pavilion service per hour. It can be seen that the values of λ and μ are decisive for the queuing model.

Therefore, the calculation formula of each state probability of the system can be obtained.

$$P_0 = \begin{cases} \left[\sum_{k=0}^{c-1} \frac{1}{k!} \left(\frac{\lambda}{\mu}\right)^k + \frac{1}{C!} \frac{1-\rho^{N-C+1}}{1-\rho} \left(\frac{\lambda}{\mu}\right)^c \right]^{-1} & \rho \neq 1 \\ \left[\sum_{k=0}^{c-1} \frac{c^k}{k\,!} + \frac{c^C}{C!} (N - C + 1) \right]^{-1} & \rho = 1 \end{cases} . \quad (1)$$

$$P_n = \begin{cases} \frac{1}{n!} \left(c\rho\right)^n P_0 \text{，当} 0 \leq n \leq c \text{，} \\ \frac{c^c}{c!} \rho^n P_0 \text{，当} c < n \leq N. \end{cases} \quad (2)$$

P_0 is the reading probability of the reading pavilion, and$P_n$is the probability of each state. When the reading pavilion is at full load, there are C receiving services, N-C are queued, and readers will be rejected. The system loss rate is PL. It can be seen that the arrival rate λ obtained from the initial data stratified sampling statistics needs to be further analyzed and solved as λe (λe < λ)due to the loss state[14]. According to the analysis, λe and μ determine the size of another important operating index parameter L of the system. L is the average occupancy of the reading pavilion, which determines the average usage rate P_U of the reading pavilion.

When C =1, ie M/M/1/N, the system is a single reading pavilion model.

$$P_0 = \begin{cases} \frac{1-\rho}{1-\rho^{N+1}} & \text{，} \rho \neq 1 \\ \frac{1}{1+N} & \text{，} \rho = 1 \end{cases} . \quad (3)$$

When N=C, that is, M/M/C/C, the system is a real-time service, and no readers are allowed to queue in the system.

$$P_0 = \left[\sum_{k=0}^{c-1} \frac{(c\rho)^k}{k!} \right]^{-1} \quad (4)$$

$$P_L = \frac{(c\rho)^n}{c!} P_0 \ . \quad (5)$$

$$L = c\rho\ (1 - P_L\ )\ . \quad (6)$$

B. Overall design process

The optimal design of Reading Pavilion service in is not limited to the study of the main parameters such as loss rate, idle rate and use efficiency. Moreover, this research is further extended to the prediction model based on time series, which can calculate the operation index data for a period of time in the future and provide a more scientific and reasonable theoretical basis for the library's future planning and budget evaluation[15].

III. DATA MINING

A. Sample data mining

The initial data collected in this paper is from Nanyuan Library of Shanghai University of Medicine & Health Sciences. The preprocessing of data mining is needed, which mainly includes lifting and sorting, unique inspection and percentage sampling. As the cyclical and seasonal factors of college holidays had a great impact on the use of reading pavilion, this paper chose the general semester for analysis, for example, sample data for March-May, October-December. Other selected months are similar. At the same time, in the normal working hours, it is also necessary to consider the influence of teaching and other factors, we use time-series stratified sampling statistics.

C. λ estimate

After the sample data is mined, the overall sample data (90 days) is randomly selected for 10 days according to the peak characteristics, and the number of arrivals is counted. See Table 1 and Table 2, respectively.

TABLE 1 READING PAVILION ARRIVALS SAMPLING

Layer（*h*）	**Randomly selected for 10 days**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
h1(09：00-11：00)	23	16	13	12	8	16	17	9	23	19
h2(11：00-13：00)	32	25	14	34	46	37	33	21	23	45
h3(13：00-15：00)	56	45	30	28	60	35	15	36	37	59

TABLE 2 READING PAVILION ARRIVAL STATISTICS

H	**N_h**	**n_h**	**f_h**	**w_h**	**Min value**	**Max value**	**Mean value（$\bar{y}_h$）**	**$N_h\bar{y}_h$**	**Variance（S_h^2）**	**$\bar{\lambda}_1$**
h1	90	10	0.11	0.33	8	23	15.6	1404	42.865	16
h2	90	10	0.11	0.33	14	46	0	2700	86.568	$\bar{\lambda}_1$ range
h3	90	10	0.11	0.33	15	60	40.1	3609	117.453	14~18
										$\bar{\lambda}_2$
Total	270	30		1.00	——	——	——	7713	——	21~23

Note: $\bar{\lambda}_1$ is the normal number of arrivals per hour. $\bar{\lambda}_2$ is the number of arrivals per hour during peak hours(13:00-15:00). The range of $\bar{\lambda}_1$and $\bar{\lambda}_2$ takes a 95% confidence interval.

D. μ estimate

15 days were randomly selected from the overall sample data, and 2 different services were counted. The average usage time of deep learning and group discussion is shown in table 3 and table 4 respectively.

Note: μ is the average number of people served by the reading pavilion per hour, with a 95% confidence interval.

TABLE 3 SINGLE READING PAVILION USAGE TIME SAMPLING

Service layer type	Randomly selected for 10 days														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Deep learning	35	27	18	28	45	36	31	26	17	17	26	15	27	18	25
Group discussion	15	40	32	21	18	45	29	57	39	48	50	52	20	17	43

TABLE 4 SINGLE READING PAVILION USAGE TIME STATISTICS

H	N_h	n_h	f_h	w_h	Min value	Maxv alue	Mean value ($\bar{y}_h$)	$N_h \bar{y}_h$	Variance (S_h^2)	$\bar{\lambda}_1$
h1	90	15	0.167	0.50	15	45	26.1	2349	8.341	11
h2	90	15	0.167	0.50	15	57	35.1	3159	23.652	μ range
Total	180	30	——	1.00	——	——	——	5508	——	10~13

E. Operational indicator

From the analysis results in the previous section, it isappropriate to take the intermediate amount μof 11. Therefore, it is necessary to observe the numerical changes of P_0, P_L and P_U corresponding to different λ values, as shown in Table 5.

TABLE 5 SYSTEM OPERATION REFERENCE TABLE WHEN M=11

System parameters / Arrival parameterλ	C=1	C=2	C=3	C=4
	N=6	N=12	N=18	N=24
12	P_0=13.5% P_L=15.6% P_U = 82.3%	P_0=16.7% P_L=5.2% P_U = 65.9%	P_0=18.6% P_L=2.1% P_U = 55.2%	P_0=21.3% P_L=0.84% P_U = 41.7%
16	P_0=6.67% P_L=34.8% P_U = 82.3%	P_0=8.2% P_L=16.1% P_U = 71.3%	P_0= 9.4% P_L=9.7% P_U = 67.8%	P_0=11.6% P_L=3.2% P_U = 52.3%
24	P_0=0.08% P_L=68.9% P_U = 82.3%	P_0=0.07% P_L=33.7% P_U = 85.6%	P_0= 0.26% P_L=28.6% P_U = 80.3%	P_0=0.89% P_L=19.2% P_U = 67.8%

This shows that there are 3 principles for the configuration of the Reading Pavilion.1) The probability of loss is controlled within a reasonable range;2) The average utilization rate of the reading pavilion during the general period is not lower than the predetermined standard, and the peak period is not higher than the predetermined standard. Reading pavilion idle rate is not higher than the predetermined minimum value. 3) Reading pavilion idle rate is not higher than the predetermined minimum value. If the loss rate is below 30%, the average utilization rate is above 60%, and the general idle rate is less than 10% as the reference standard. Obviously, only c=2, 3, the parameters are basically in compliance with the standard. After a comprehensive analysis of the calculation parameters in table 2 and table 5, the optimal reading pavilion configuration can be obtained when c=2.

F. Time series prediction model

The above operation results give the library the current optimal configuration plan. However, as time goes on and the impact expands, the library must plan for the future number of people in order to increase or decrease the number of reading pavilions according to the above configuration. According to the statistics of the access control statistics, the time series-based exponential smoothing model is used to predict the monthly forecast value and the 95% confidence floating interval value for the next year[16]. The statistical data is shown in table 6.

TABLE 6 NUMBER OF READERS FROM 2017 TO 2018

Month	2017								2018					
	5	6	7	8	9	10	11	12	1	2	3	4	5	6
No.	2479	2891	962	738	1123	2062	1908	1923	1945	975	1854	2922	2532	3448

For the data in table 6, further observation is needed to reasonably predict the future model data. First, the distribution of residual autocorrelation (ACF) and partial autocorrelation (PACF) is observed. The predictions are all within the allowable tolerance scope, and none of them exceeds the set range. Set the upper limit of confidence interval (UCL) and lower limit (LCL) to get the prediction graph, as shown in figure 1.

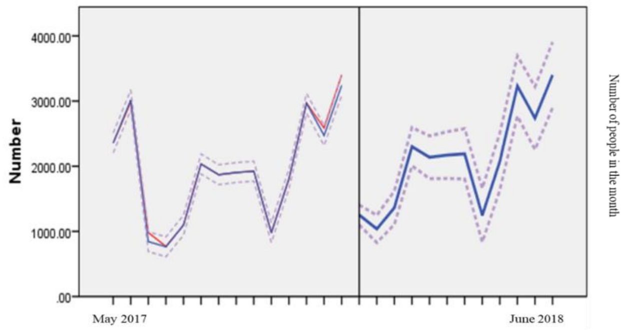

Fig. 1 Prediction fit map

As can be seen from figure 1, the predicted variation range is basically consistent with the observed value, showing the characteristics of seasonal periodic variation as a whole. The above can be further accurately explained by model fitting statistics such as R square value and sig value.

The R can be calculated as 0.993, and the fitting degree of this model and observation statistics is very high. The value of sig is 0, indicating that the model is remarkably reliable. Therefore, the upper limit value and the lower limit value of the confidence interval can be taken as the predicted value, as shown in table 7.

TABLE 7 FORECAST FIGURES FOR EACH MONTH IN THE NEXT YEAR

2018		July	August	September	October	October	Dec
Number of readers	Forecast	1145.72	1003.52	1298.51	2278.95	2073.47	2126.39
	UCL	1315.33	1187.33	1589.43	2520.16	2401.87	2499.73
	LCL	998.65	854.61	1088.27	1985.35	1821.36	1799.47

IV. CONCLUSION

This paper gives the theoretical basis and index parameters of the optimized configuration of reading pavilion. Through the statistical estimation of the number of visitors in different time periods and the usage time of different purposes during the library opening period, the parameters required in the optimized configuration scheme are given. By comparing the conditional probability under different parameters, the optimal configuration scheme was found, which provided a powerful basis for reservation management of Reading Pavilion and a scientific guarantee for the improvement of library service quality.

REFERENCES

Feng Xiangchun, Yang Xiaofeng. Investigation and analysis of the current situation of the construction of university library training room. Modern Intelligence, 2017, 37 (10): 110-113.

Wang Zhengjun, Dong Xiaomei, Yu Xiaoyi, Jin Yuling. Design and implementation of the i.ntelligent training room system of the library. Modern Intelligence, 2016, 36(8): 79-83

Jia Baoxia. Discussion on the scientific management of university library training--Taking the library of Northeast Agricultural University as an example. University Library and Information Science, 2014, 32(2): 9-11.

Lu Wei, Liu Jingqin. Investigation and Analysis of the Service and Management of the Library of Renmin University of China. Library Tribune,2013,33(3):165-169.

Liu Yingying. Research on the Status and Management of Reading Pavilion and Discussion Room in University Libraries. Library Tribune, 2010, 30(1):148-150.

Kirpichnikov A, Titovtsev A. Physical and mathematical queues in the applied queuing theory. 2016, 108(2):409-419.

Priyangika J S K C, Cooray T M J A. Analysis of the Sales Checkout Operation in Supermarket Using Queuing Theory. 2016, 4(7):393-396.

Sembakutti D, Sasmito A P, Kumral M. Analysing equipment allocation through queuing theory and Monte-Carlo simulations in surface mining operations. International Journal of Mining & Mineral Engineering, 2017, 8(1):56.

Koyuncu E, Başpınar B. Demand and Capacity Balancing Through Probabilistic Queuing Theory and Ground Holding Program for European Air Transportation Network. 2017, 18(2):1-1.

Cai Keping, Guo Juncang, Su Ning. Analysis and Solution of Modeling Conditions for Application of Queuing Theory in Circulation Links of University Libraries, Science and Technology Information Development and Economy, 2014(23): 6-8.

Jiang Chen, Yuan Yulin, Liu Li. Research on the configuration model of circulation service between university libraries. Library and Information Work, 2017, 61(20): 97-104.

Nandhini S, Sanjay M S. A novel optimal routing protocol with implementation of M/M/1 queuing model. International Journal of Civil Engineering & Technology, 2017, 8(11):306-311.

orregrosa J R, Jodar L, Cortés J C, et al. Mathematical Modeling and Computational Methods. Journal of Computational & Applied Mathematics, 2017, 309:320-324.

Jin Yongjin, Du Zifang, Jiang Wei. Sampling Technology. Beijing: China Renmin University Press, 2015: 65-72.

Ronald E, Walpole, Jia Junping, et al. Probability and Statistics. Beijing: China Renmin University Press, 2016: 146-173.

Messina A R, Vittal V. A structural time series approach to modeling dynamic trends in power system data. IEEE Conferences: 2012 IEEE Power and Energy Society General Meeting, 2012:1-8.