Bank Queuing Optimization Based on Markov Process

Yuqi He, Haoxuan Li*
College of Mathematics, Sichuan University, Chengdu 610065, China
lihaoxuan_2002@126.com

Abstract. Queuing optimization is a troubling problem in the field of operations research and it is difficult to solve completely due to the complexity and uncertainty of the queuing environment. Among them, bank queuing optimization is the most widespread and representative in the queuing theory. This article focuses on using customer arrival and service of cashiers to establish birth-death process and its transfer rate matrix. By introducing service intensity and customer satisfaction index, it is concluded that the optimal number of open cashiers in different time period can be calculated and predicted.

1. Introduction
Under the finance-oriented society, bank queuing phenomenon is particularly serious due to the complicated business and consulting process to control financial risk. Although banking business has been expended gradually, such as introducing credit card loans and online banking, queuing problem still bothers people who seek for the bank service.

The determination of the number of cashiers can be attributed to the problem of queuing optimization [1-4]. In 1996, Kutsc and Christopher introduced the crowding indicator $K$ to compare the relationship between the number of cashiers and customers, determining whether to increase the number of cashiers or not [5]. Based on the David analysis in 1998 of cost for cashiers and cost of customers’ loss, an optimization model was established to calculate the number of cashiers that benefited the most [6-7]. In 2007, Bielen, F. and Demoulin, N proposed that waiting time had a huge effect on the customers’ satisfaction, which is critical to service evaluation [8-10].

This article focuses on comparing service intensity and customer satisfaction, then analyzing the survey data to get the appropriate cashier number in each time period. In Chapter 2, the distribution of customer flow and service time is constructed, in which the time period of customer flow significantly affects the sample value. Maximum likelihood estimation is used for parameter estimation of the distribution, and the estimation is uniformly minimum variance unbiased estimate by the unbiasedness and variance of the estimation reaching the Cramer-Rao bound. In Chapter 3, $M/M/C$ model is introduced through establishing birth-death process and its transfer rate matrix, we use steady birth-death process to establish the bank queuing model and introduce service intensity and customer satisfaction to calculate and predict the optimal number of open cashiers in different time periods.

2. Distribution of Customer Flow and Service Time

2.1. Distribution of Customer Flow
This article selects the passenger flow and service time data on a bank in Chengdu from April and December. Counting in hours as a period of time, the customer flow of these selected days’ banks is shown as Table 1.
Table 1. Customer Flow in April and December.

| Date | 9:00-10:00 | 10:00-11:00 | 11:00-12:00 | 12:00-13:00 | 13:00-14:00 | 14:00-15:00 | 15:00-16:00 | 16:00-17:00 |
|------|------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| April 4th | 34 | 24 | 33 | 46 | 24 | 30 | 38 | 29 |
| April 5th | 26 | 19 | 26 | 42 | 22 | 25 | 30 | 21 |
| April 6th | 36 | 16 | 29 | 39 | 19 | 22 | 35 | 24 |
| April 7th | 23 | 19 | 41 | 51 | 26 | 29 | 41 | 27 |
| April 8th | 32 | 23 | 28 | 47 | 19 | 27 | 46 | 35 |
| April 9th | 27 | 31 | 39 | 45 | 19 | 33 | 39 | 31 |
| April 10th | 33 | 26 | 29 | 41 | 19 | 24 | 33 | 21 |
| April 11th | 36 | 28 | 20 | 35 | 17 | 27 | 36 | 26 |
| April 12th | 30 | 26 | 31 | 44 | 29 | 30 | 31 | 19 |
| April 13th | 21 | 22 | 34 | 45 | 26 | 29 | 41 | 33 |
| December 4th | 47 | 33 | 46 | 65 | 34 | 42 | 53 | 41 |
| December 5th | 36 | 26 | 37 | 59 | 31 | 35 | 42 | 29 |
| December 6th | 50 | 23 | 40 | 54 | 26 | 31 | 49 | 34 |
| December 7th | 32 | 27 | 58 | 72 | 37 | 40 | 57 | 38 |
| December 8th | 45 | 32 | 39 | 66 | 26 | 38 | 64 | 49 |
| December 9th | 38 | 43 | 54 | 63 | 27 | 46 | 55 | 43 |
| December 10th | 46 | 37 | 41 | 57 | 31 | 34 | 46 | 30 |
| December 11th | 51 | 39 | 28 | 49 | 24 | 38 | 50 | 36 |
| December 12th | 42 | 36 | 44 | 61 | 40 | 42 | 44 | 26 |
| December 13th | 29 | 31 | 47 | 63 | 37 | 41 | 57 | 46 |

Note that there are more customers at noon than in the morning, it is necessary to determine whether the time period has a significant impact on the number of customers per unit time, and we use one-way analysis of variance to verify. The eight different time intervals from the earliest to the latest are denoted as $A_1, A_2 \ldots, A_8$, respectively. The data from the 4th to the 13th of the month are regarded as independent repeated tests, and the corresponding data is recorded as $Y_{ij}$ ($i = 1, 2, \ldots, 8; j = 1, 2, \ldots, 10$).

Suppose the sample comes from the model $Y_{ij} = \mu_i + e_{ij}$ ($i = 1, 2, \ldots, 8; j = 1, 2, \ldots, 10$), where $\{e_{ij}\}$ is independent of each other, and $e_{ij} \sim N(0, \sigma^2)$. The null hypothesis is $H_0: \mu_1 = \mu_2 = \cdots = \mu_8$.

Let

$$\bar{Y}_i = \frac{1}{10} \sum_{j=1}^{10} Y_{ij}, \quad \bar{Y} = \frac{1}{80} \sum_{i=1}^{8} \sum_{j=1}^{10} Y_{ij} = \frac{1}{8} \sum_{i=1}^{8} \bar{Y}_i. \quad (1)$$

The residual decomposition formula is as follows:

$$S_T = \sum_{i=1}^{8} \sum_{j=1}^{10} (Y_{ij} - \bar{Y})^2 = \sum_{i=1}^{8} \sum_{j=1}^{10} [(Y_{ij} - \bar{Y}_i) + (\bar{Y}_i - \bar{Y})]^2$$
where \( S_E \) represents the influence of random error, while \( S_A \) represents the influence of factor \( A \) on the value of \( Y \).

Obviously, when \( H_0 \) holds, \( S_A \) should be relatively low to \( S_E \). At the same time statistics

\[
F = \frac{S_A / (8 - 1)}{S_E / 8(10 - 1)} \sim F(8 - 1, 8(10 - 1)) = F(7, 72).
\]

Hence the rejection region is

\[
W = \{ (Y_{ij})_{8 \times 10} ; F > F_{\alpha = 0.05}(7, 72) = 2.14 \}.
\]

The AVONA results are shown in Table 2. For April customer flow data, the F statistic value is \( 21.50 > 2.14 \), and for December customer flow data, the F statistic value is \( 21.45 > 2.14 \). Therefore, the time period has a significant impact on the number of customers in April and December.

| Table 2. AVONA Results in April and December. |
| April                   | December              |
|-------------------------|-----------------------|
| \( S_T \)               | 5208.2                |
| \( S_E \)               | 1685.4                |
| \( S_A \)               | 3522.8                |
| \( F\)-Value            | 21.499                |

Empirical evidence asserts that the **average arrival rate of customers in a unit of time should obey the Poisson distribution.** Through using the maximum likelihood estimation (MLE) to estimate the parameter \( \lambda \) of the Poisson distribution, it shows that the MLE is a Uniformly Minimum Variance Unbiased Estimate (UMVUE), then the validity of the estimation can be verified by using the chi-square test.

The customer's arrival is obey to Poisson distribution, the probability function is

\[
P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0, 1, 2, ...
\]

The likelihood function of the Poisson distribution is

\[
L(x_1, x_2, ..., x_n; \lambda) = \prod_{i=1}^{n} P(X = x_i) = \frac{\lambda^{\sum_{i=1}^{n} x_i}}{x_1! \cdot \ldots \cdot x_n!} e^{-\lambda n}.
\]

Take the logarithm and find the partial derivative about \( \lambda \) to get the likelihood equation

\[
\frac{\partial \ln L(\lambda)}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n = 0.
\]

Hence, the MLE of parameter \( \lambda \) of the Poisson distribution is

\[
\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i.
\]

This article takes the least customer flow in April (idle period) 13: 00-14: 00 and the most customer flow in December (busy period) 12: 00-13: 00 as examples for the remaining analysis. From the data in Table 1 and Equation (8), the average arrival rate of customers in the unit time of April 13: 00-14: 00 follows the Poisson distribution of parameter \( \lambda_1 = 0.372 \ customers/min \), while the unit time of
December 12: 00-13: 00 The average customer arrival rate follows the Poisson distribution of parameter $\lambda_2 = 1.015$ customers/min.

The above estimate is obviously an unbiased estimate, Cramer-Rao inequality gives the lower bound of the unbiased estimate

$$\text{Var}_2(\psi(X_1, X_2, ..., X_n)) \geq \frac{[g'(\lambda)]^2}{nI(\lambda)}, \quad (9)$$

where $\psi(X_1, X_2, ..., X_n)$ is an unbiased estimate of $g(\lambda)$ and $I(\lambda)$ is the Fisher information

$$I(\lambda) = E \left( \frac{d \ln f(X; \lambda)}{d\lambda} \right)^2 = E \left( \frac{X}{\lambda} - 1 \right)^2 = \frac{1}{\lambda}. \quad (10)$$

Hence, the Cramer-Rao lower bound of the unbiased estimate of the Poisson distribution is

$$\text{Var}_2(\psi(X_1, X_2, ..., X_n)) \geq \frac{\lambda}{n} = \text{Var}_\lambda \left( \frac{1}{n} \sum_{i=1}^{n} x_i \right). \quad (11)$$

Therefore, the MLE of the parameter $\lambda$ of the Poisson distribution is actually a UMVUE.

Then, the chi-square test method is used to verify the validity of the Poisson distribution. Taking the customer flow in April (idle period) 13:00-14:00 as an example, the null hypothesis is that the sample is derived from a Poisson distribution with a parameter of $\lambda_1 = 22.30$ customers/hour. As shown in Table 3, the value of the sample is divided into 5 groups, and calculate the statistic

$$V = \sum_{k=1}^{5} \frac{(v_k/n - p_k)^2}{n p_k^2}, \quad (12)$$

where $v_k$ is the frequency of the sample falling into the $k$-th interval, $p_k$ is the theoretical probability of the sample falling into the $k$-th interval, and $n = 10$ is the total number of samples.

It can be proved that when the null hypothesis holds, the statistic $V$ approximately obeys the chi-square distribution with degree of freedom ($df$) = 4. The relevant data is shown in Table 3.

| Interval       | $v_k/n$ | $p_k$ | $(v_k/n - p_k)^2/n p_k^2$ | $n(v_k/n - p_k)^2/p_k$ |
|----------------|---------|-------|---------------------------|------------------------|
| $k = 1$        | [0, 17] | 0.10  | 0.003                      | 0.189                  |
| $k = 2$        | [18, 21]| 0.30  | 0.000                      | 0.002                  |
| $k = 3$        | [22, 25]| 0.30  | 0.000                      | 0.004                  |
| $k = 4$        | [26, 28]| 0.20  | 0.003                      | 0.211                  |
| $k = 5$        | [29, +∞]| 0.10  | 0.000                      | 0.000                  |
| $\text{Sum}$   | [0, +∞]| 1.00  | 0.006                      | 0.406                  |

The rejection region is

$$W = \{ X : V(X) \geq \lambda_{df=4,a=0.05} = 9.49 \}. \quad (13)$$

Hence it is reasonable to suppose that the customer flow in April 13: 00-14: 00 obeys the Poisson distribution with parameter $\lambda_1 = 22.30$ customers/hour.

2.2. Distribution of Service Time

The service time data obtained by the survey is shown in Table 4.

| Time Consuming | # of Services | Average Service Rate | Time Consuming | # of Services | Average Service Rate |
|----------------|---------------|----------------------|----------------|---------------|----------------------|
| 1              | 4.5           | 1                    | 4.50           | 14            | 10.0                 |
| 2              | 2.5           | 1                    | 2.50           | 15            | 17.0                 |
| 3              | 3.0           | 2                    | 1.50           | 16            | 14.0                 |

4.00
Empirical evidence indicates that the average service rate should obey a negative exponential distribution. It is worth noting that the distribution of the average service rate is not related to the idle or busy period, that is, the April sample and the December sample have the same average service rate distribution.

The maximum likelihood estimate of the negative exponential distribution parameter $\lambda$ is $\hat{\lambda} = 1/X$. From Table 4, the average service rate follows a negative exponential distribution with a parameter of $\lambda = 0.252 \text{ person/min}$. Then use the Blackwall-Lehmann-Scheffe Theorem to confirm that the estimate is a UMVUE. Finally, the chi-square test is used to verify the validity of the distribution.

3. Optimization of Bank Queuing

3.1. Construction of Model and Solution

In a bank queuing system, the waiting area of the bank lobby is generally ample, and there is no need to limit the customer capacity of the system, that is, assuming that the bank lobby can accommodate an unlimited number of customers. Customer arrival is independent of each other and follows Poisson distribution, and service of cashiers is also independent and follows negative exponential distribution. Based on the M/M/C model, the birth-death process and transfer rate matrix are established as

$$
\begin{align*}
C &= \begin{pmatrix}
-\lambda & \lambda & 0 & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\
\mu & -(&\mu+\lambda) & \lambda & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\
0 & 2\mu & -(&2\mu+\lambda) & \lambda & 0 & \ldots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & c\mu & -(&c\mu+\lambda) & \lambda & 0 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & c\mu & -(&c\mu+\lambda) & \lambda & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\end{pmatrix}.
\end{align*}
$$

(14)

The limit distribution of the positive recurrence and irreducible equivalence class in Markov chain is equal to Stationary distribution, hence the balance equation is

$$
\begin{align*}
\{ & \mu P_1 = \lambda P_0, & n = 0 \\
& (n + 1) \mu P_{n+1} + \lambda P_{n-1} = (\lambda + n\mu)P_n, & n \leq c, \\
& c\mu P_{n+1} + \lambda P_{n-1} = (\lambda + c\mu)P_n, & n > c
\end{align*}
$$

(15)

Solution of the balance equation (15) is

$$
\begin{align*}
P_0 &= \left[ \sum_{i=0}^{c-1} \frac{\lambda^i}{i!} \mu + \frac{1}{c!} \left( 1 - \rho \frac{\lambda}{\mu} \right) \right]^{-1}, \\
P_n &= \begin{cases} \\
\frac{1}{n!} \rho^n P_0, & n \leq c \\
1 & \frac{1}{c!} \rho^n P_0, & n > c
\end{cases}
\end{align*}
$$

(16)

where $c$ is the number of cashiers, $n$ is the number of customers in the system, $P_0$ is the probability of waiting at any time in the system, $P_n$ is the probability that there are $n$ customers in the system and
\( \rho = \frac{\lambda}{cu} \) is the service intensity. According to \( P_0 \) and \( P_n \), the average number of customers in line \( L_q \) and the average number of customers in system \( L_s \) can be obtained as

\[
L_q = \frac{(np)^n \rho}{n! (1 - \rho)^2} P_0, \quad L_s = L_q + \frac{\lambda}{\mu}. \tag{17}
\]

The average waiting time \( W_q \) and average stay time \( W_s \) of customers in the system are

\[
W_q = \frac{L_q}{\lambda}, \quad W_s = W_q + \frac{1}{\mu}. \tag{18}
\]

3.2. Model Optimization and Data Analysis

In 1991, Wolff assumed that the customer satisfaction decreases as the waiting time goes on and follows negative exponential distribution [11]. Thus, the customer satisfaction index can be written as

\[ C_s = e^{-\beta W_q}, \tag{19} \]

where \( C_s \) is the customer satisfaction index and \( \beta (0 < \beta < 1) \) is the customer sensitivity to the waiting time, assuming that the customer sensitivity to waiting time is the constant 1.

| \( \rho \) | \( L_s \) | \( L_q \) | \( W_s \) | \( W_q \) | \( C_s \) |
|---|---|---|---|---|---|
| 2 | 0.833 | 5.455 | 3.788 | 14.676 | 10.192 | 0.844 |
| 3 | 0.556 | 2.041 | 0.375 | 5.492 | 1.008 | 0.983 |
| 4 | 0.417 | 1.740 | 0.073 | 4.681 | 0.197 | 0.997 |
| 5 | 0.333 | 1.682 | 0.015 | 4.525 | 0.041 | 0.999 |
| 6 | 0.278 | 1.670 | 0.003 | 4.492 | 0.008 | 1.000 |

| \( \rho \) | \( L_s \) | \( L_q \) | \( W_s \) | \( W_q \) | \( C_s \) |
|---|---|---|---|---|---|
| 2 | 2.276 | -1.089 | -5.641 | -1.073 | -5.557 | 1.097 |
| 3 | 1.517 | -1.606 | -6.157 | -1.582 | -6.066 | 1.106 |
| 4 | 1.138 | -6.362 | -10.913 | -6.268 | -10.752 | 1.196 |
| 5 | 0.910 | 12.526 | 7.975 | 12.341 | 7.857 | 0.877 |
| 6 | 0.759 | 5.926 | 1.375 | 5.839 | 1.354 | 0.978 |

Table 7. Optimal Indicators of Different Time Period in April

| Time Period | \( \rho \) | \( L_s \) | \( L_q \) | \( W_s \) | \( W_q \) | \( C_s \) |
|---|---|---|---|---|---|---|
| 9:00-10:00 | 3 | 0.742 | 3.830 | 1.602 | 7.711 | 3.226 | 0.948 |
| 10:00-11:00 | 2 | 0.874 | 7.431 | 5.682 | 19.053 | 14.569 | 0.784 |
| 11:00-12:00 | 3 | 0.772 | 4.361 | 2.044 | 8.441 | 3.956 | 0.936 |
| 12:00-13:00 | 4 | 0.813 | 5.941 | 2.690 | 8.195 | 3.710 | 0.940 |
| 13:00-14:00 | 2 | 0.833 | 5.455 | 3.788 | 14.676 | 10.192 | 0.844 |
| 14:00-15:00 | 3 | 0.688 | 3.107 | 1.044 | 6.754 | 2.269 | 0.963 |
| 15:00-16:00 | 4 | 0.691 | 3.695 | 0.930 | 5.993 | 1.508 | 0.975 |
| 16:00-17:00 | 3 | 0.663 | 2.850 | 0.862 | 6.429 | 1.945 | 0.968 |

Table 8. Optimal Indicators of Different Time Period in December

| Time Period | \( \rho \) | \( L_s \) | \( L_q \) | \( W_s \) | \( W_q \) | \( C_s \) |
|---|---|---|---|---|---|---|
| 9:00-10:00 | 4 | 0.777 | 5.050 | 1.941 | 7.284 | 2.799 | 0.954 |
| 10:00-11:00 | 3 | 0.815 | 5.394 | 2.950 | 9.897 | 5.413 | 0.914 |
| 11:00-12:00 | 4 | 0.811 | 5.886 | 2.643 | 8.138 | 3.653 | 0.941 |
| 12:00-13:00 | 6 | 0.759 | 5.926 | 1.375 | 5.839 | 1.354 | 0.978 |
| 13:00-14:00 | 3 | 0.780 | 4.515 | 2.176 | 8.656 | 4.171 | 0.933 |
In 2007, Xuan Li and Qiang Su proposed that the service can maintain the best work efficiency when service intensity $\rho$ is between 70% ~ 85% [5]. Two groups with the lowest arrival in April and the highest arrival in December are extracted to calculate the value of each indicator under different numbers of cashiers. From Table 5, in the case that customer satisfaction is within the acceptable range, service intensity is most appropriate when $n = 2$. Otherwise, service intensity is too low so that cashiers cannot be effectively used. Similarly, from Table 6, during the peak hours in December, if $\rho$ is greater than 1, the indicator will be a negative value, which means the cashiers will work overload. This circumstance won’t be advocated because of the low efficiency. The service intensity and customer satisfaction are both optimal when $n = 6$.

From Table 7 and Table 8, by comparing the two indicators of service intensity and customer satisfaction, optimal numbers of open cashiers can be calculated in different time of the bank's idle and busy periods.

4. Conclusion

Based on the results and discussions presented above, the conclusions are obtained as below:

(1) The average arrival rate of customers per unit time is subject to Poisson distribution, however, the time period factor can significantly affect the sample value. The average service rate follows a negative exponential distribution, and its parameters have nothing to do with idle or busy periods.

(2) The bank queuing model is established according to the distribution of Customer arrival and service of cashiers, the balance equation is derived from the steady birth-death process, then the service intensity of the system can be solved.

(3) The index of customer satisfaction is introduced to optimize the number of cashiers during idle and busy periods. Empirical evidence shows that the use of service intensity and customer satisfaction on birth-death process can effectively calculate and predict the optimal number of open cashiers in different periods. This method can be further optimized by assuming customer's sensitivity as stochastic variable which is much in line with reality.

References

[1] Cassandras, C G, and Lafortune, S. Introduction to Discrete Event Systems. *IEEE transactions on automatic control* 46.2(2002):353-354.

[2] Leonard Kleinrock. *Theory, Volume 1, Queueing Systems*. Wiley-Interscience, 1975.

[3] Donald Gross. *Fundamentals Of Queueing Theory*. *Fundamentals of queueing theory*. J. Wiley & Sons, 2008.

[4] Mi-Sha, Yang, and Y. I. Kun-Nan. Optimization of bank counters based on the queuing process. *Journal of Zhejiang University(ence Edition)* 36.5(2009):514-407.

[5] So, Kut C., and C. S. Tang. On managing operating capacity to reduce congestion in service systems. *European Journal of Operational Research* 92.1(1996):83-98.

[6] Pepyne, David L., and C. G. Cassandras. Modeling, Analysis, and Optimal Control of a Class of Hybrid Systems. *Discrete Event Dynamic Systems* 8.2(1998):175-201.

[7] Noémi Kalló, and Tamás Koltai. Increasing Customer Satisfaction in Queueing Systems with Rapid Modelling. (2011).

[8] Frédéric Bielen, and N. Demoulin. Waiting time influence on the satisfaction-loyalty relationship in services. *Journal of Service Theory & Practice* 17.2(2007):174-193.

[9] Davis, Mark M., and T. E. Vollmann. A Framework for Relating Waiting Time and Customer Satisfaction in a Service Operation. *Journal of Services Marketing* 4.1(1990):61-69.

[10] Satoshi Shimada, Kei Taira, Tatsunori Hara, and Tamio Arai. Customers' Satisfaction on Estimates of Queue Waiting Time in Service Delivery. *Functional Thinking for Value*
8

[11] Stephen G. Strickland. "Stochastic modeling and the theory of queues, by R. W. WOE, Rentice Hall, Engle- wood Cliffs, NJ, 1989, 556 pp." *Networks.* 21.3(1991):371-371.