

Analysis of the age Process in discrete-time

$SM[K]/PH[K]/2/FCFS$ Queueing System

Gao Zhuo

(Guangzhou Civil Aviation College, China)

Gao Zhuo(1981-),Male, Associate Professor, Main Research Direction: Stochastic Operations Research.Supported by the University-level Research Cadre Training Program Special Project(Project No.24X4421).

Abstract: Queueing models with multiple servers are widely used in communication networks. This paper focuses on the common discrete-time $SM[K]/PH[K]/2/FCFS$ queueing system with two servers. The age process of the system is then analyzed. To reduce computational complexity, the age of a certain batch of customers in the system is selected as the system's age, and some auxiliary variables are introduced to construct a Markov chain associated with the age process.

Keywords: Multiple servers; Queueing system; Age process; Markov chain

Date of Submission: 01-08-2025

Date of acceptance: 11-08-2025

I. Introduction

The problem of multiple servers has attracted significant attention due to its wide range of applications, especially in the design of communication networks, where multiple-server queueing models are indispensable. However, current communication network designs predominantly utilize single-arrival multiple-server models. With the increasing complexity of communication network applications, single-arrival models are gradually becoming insufficient. Therefore, batch-arrival models have been introduced to make the models more realistic. The generalized age process we construct does not have a boundary at level 0. [1] provides a detailed analysis of the $SM[K]/PH[K]/1/FCFS$ type of queueing system and has obtained some valuable results. We believe that the methods used for single-server queueing systems can be extended to multiple-server systems, although the complexity increases significantly.

II. $SM[K]/PH[K]/2/FCFS$ Queueing System

Due to the practical value and universality of queueing systems, we only introduce $SM[K]/PH[K]/2/FCFS$ here. The customer arrival process is the same as in the single-server case,

which is detailed in [1], and will not be repeated here. The difference from $SM[K]/PH[K]/1/FCFS$ lies in the queuing process of customers. After a batch of customers arrives, all customers join the queue according to their order within the batch. All batches are served by two servers in a "first-come, first-served" manner. It is assumed that each batch of customers is served by the same server, and for each batch of customers, customers are served in the order they appear in the batch. Let $q(t)$ be the sequence of customer types in the queue at time t , which can be obtained from the customers who may complete service and the arriving customers at time $t-1$. If $q(t) = j_1 j_2 \cdots j_n$, then there are n customers in the system at time t , with j_1 type customers being served at one server, and the situation for the other server is as follows:

(1) If $q(t)$ is exactly one batch of customers, then the other server is idle. j_2 type customers are the first in the queue, \cdots , and j_n type customers are the last in the queue. These n customers will be served in the order of $j_1 j_2 \cdots j_n$.

(2) If the number of batches in $q(t)$ is greater than 1, i.e., $j_1 j_2 \cdots j_i$ is the first batch, $j_{i+1} j_{i+2} \cdots j_k$ is the second batch, \cdots , then j_{i+1} type customers are being served at the other server. The first server serves in the order of $j_1 j_2 \cdots j_i$, and the second server serves in the order of $j_{i+1} j_{i+2} \cdots j_k$.

If a J type customer batch arrives next, the queue will become $q(t) + J$. If a customer completes service next, assuming it is a j_1 type customer, then the queue will become $j_2 j_3 \cdots j_n$, and the j_2 type customer will begin to be served. At this time, j_2 must belong to the same batch as j_1 . Otherwise, this server will serve the first customer in the queue who does not belong to the same batch as j_2 . Of course, if all customers after j_2 belong to the same batch as j_2 , then this server will be idle at this time.

III. Analysis of the Generalized Age Process

For two servers, the age analysis is obviously more complex than for one server. Each server may have a batch of customers, and each batch has an age. Of course, analyzing the age of each batch of customers being served can, like in the single-server case, analyze other system metrics, but this would increase the computational complexity. The Markov process of the age process has a high dimensionality, so we must analyze the system clearly and find an age process that is conducive to analyzing other system metrics and reduces computational complexity. Referring to [1], we can consider the age of the younger batch of customers being served, denoted as the system age $a_g(t)$ at time t . When $a_g(t) \geq 0$, the system is in a busy period (i.e., both servers have batches of customers being served), and there are at least two batches of customers in the

system. When $a_g(t) < 0$, there is at most one batch of customers in the system, and $a_g(t)$ is the age of the first batch of customers arriving at the system after time t . This batch of customers will arrive at the system after $-a_g(t)$ time units. We denote $a_g(t)$ as the age of the $N(t) = n$ -th batch of customers at time t , and τ_n is the time between the arrival of the $n-1$ -th and n -th batches.

Let's analyze this age process $\{a_g(t), t \geq 0\}$:

(1) When $a_g(t) \geq 0$, at time $t+1$:

a) If neither of the two batches of customers has completed service, then $N(t+1) = N(t)$, and thus

$$a_g(t+1) = a_g(t) + 1;$$

b) If one batch of customers completes service while the other continues to be served:

i. If the waiting batch of customers begins to be served, then $N(t+1) = N(t) + 1$, and thus

$$a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1};$$

ii. If there are no waiting batches of customers in the system, then $N(t+1) = N(t) + 1$, and thus

$$a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1};$$

c) If both batches of customers complete service and leave the system:

i. If two waiting batches of customers begin to be served, then $N(t+1) = N(t) + 2$, and thus

$$a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1} - \tau_{N(t)+2};$$

ii. If one waiting batch of customers is served while the other batch has not yet arrived at the system, then

$$N(t+1) = N(t) + 2, \text{ and thus } a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1} - \tau_{N(t)+2};$$

iii. If there are no waiting batches of customers in the system and the batches have not yet arrived, then

$$N(t+1) = N(t) + 1, \text{ and thus } a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1}.$$

(2) When $a_g(t) < 0$ and there is only one batch of customers in the system, at time $t+1$:

a) If the batch of customers being served at time t has not completed service, then $N(t+1) = N(t)$,

$$\text{and thus } a_g(t+1) = a_g(t) + 1;$$

b) If the batch of customers being served at time t completes service and leaves the system:

i. If the $N(t)$ -th batch of customers has not yet arrived at the system, then $N(t+1) = N(t)$, and thus

$$a_g(t+1) = a_g(t) + 1;$$

ii. If the $N(t)$ -th batch of customers arrives at the system exactly at time $t+1$, then

$$N(t+1) = N(t) + 1, \text{ and thus } a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1}.$$

(3) When $a_g(t) < 0$ and there are no batches of customers in the system, at time $t+1$:

a) If the $N(t)$ -th batch of customers arrives at the system exactly at time $t+1$, then

$$N(t+1) = N(t) + 1, \text{ and thus } a_g(t+1) = a_g(t) + 1 - \tau_{N(t)+1};$$

b) If the $N(t)$ -th batch of customers has not yet arrived at the system, then

$$N(t+1) = N(t), \text{ and thus } a_g(t+1) = a_g(t) + 1.$$

Now that we have analyzed the age process, we need to construct a Markov chain associated with the age process. We refer to $a_g(t)$ as the level variable, which can take any integer value. We also introduce some auxiliary variables. Since there are three cases in the system (see the analysis below), we introduce variables for each case.

(1) When $a_g(t) \geq 0$, the system has at least two batches of customers, with two batches being served.

$a_g(t)$ is the age of the younger batch of customers being served. We introduce the auxiliary variable

$\{(I_a(t), J_1(t), I_{s1}(t), J_2(t), I_{s2}(t)), t \geq 0\}$, where $I_a(t)$ represents the state of the semi-Markov chain

when the $N(t)$ -th batch of customers arrives, and $J_i(t), I_{si}(t)$ represents the type and service phase of the

batch of customers being served at server i at time t , ($i = 1, 2$). The values of 1 and 2 have no order; they

are just used to distinguish the vectors. The range of the auxiliary variable is the set:

$\{(i, J_1, j_1, J_2, j_2) : 1 \leq i \leq m_a, J_k \in \mathbb{N}, 1 \leq j_k \leq m_j, k = 1, 2\}$. In this set, the states are sorted

lexicographically.

(2) When $a_g(t) < 0$, there is only one batch of customers being served in the system, and the next batch of

customers (i.e., the $N(t)$ -th batch) will arrive at the system after $-a_g(t)$ time units. We still introduce the

auxiliary variable $\{(I_a(t), J_1(t), I_{s1}(t), J_2(t), I_{s2}(t)), t \geq 0\}$, where $I_a(t)$ represents the state of the

semi-Markov chain when the $N(t)$ -th batch of customers arrives, and $J_i(t), I_{si}(t)$ represents the type and service phase of the batch of customers being served at time t , as well as the type and initial phase of the $N(t)$ -th batch of customers. The range of the auxiliary variable is the same as above.

(3) When $a_g(t) < 0$, there are no batches of customers in the system, and the next batch of customers (i.e., the $N(t)$ -th batch) will arrive at the system after $-a_g(t)$ time units. We introduce the auxiliary variable $\{(I_a(t), J(t), I_s(t)), t \geq 0\}$, where $I_a(t)$ represents the state of the semi-Markov chain when the $N(t)$ -th batch of customers arrives, and $J(t), I_s(t)$ represents the type and initial phase of the $N(t)$ -th batch of customers. The range of the auxiliary variable is the set: $\{(i, J, j) : 1 \leq i \leq m_a, J \in \aleph, 1 \leq j \leq m_j\}$. In this set, the states are also sorted lexicographically. Like the single-server case, we can analyze that the constructed process is a Markov process associated with the age, which provides a theoretical basis for studying and calculating the $SM[K]/PH[K]/2/FCFS$.

References:

- [1]. Gao Zhuo. On the computation of transfer matrices for age processes in discrete-time $SM[K]/PH[K]/1/FCFS$ queue, American Journal of Engineering Research[J].2025,14(2)19-24.
- [2]. Zhao N. Huang Xiaofeng, Liu Wenqi. Approximate analysis of queuing system $GI/G/1$, Journal of Nanjing University of Science and Technology (Natural Science Edition) [J].2022,46(2)211-218.
- [3]. Xiaoli Cai, Yao Chen, Jun Li. Pricing and service capacity strategies for queuing systems oriented to loss averse customers, Operations and Management[J].2024,33(8) 135-140.
- [4]. Su Yan, Li Junping, Yu Yi. Research on optimal dynamic strategies for multi-type customer feedback queuing system, Systems Engineering Theory and Practice[J]. 2024,44(7) 2352-2361.
- [5]. Xinyi Zhao, Liwei Liu. Research on equilibrium strategies for repairable retry queuing systems with N-strategies and negative customers, Applied Mathematics[J].2024,37(3):589-600.
- [6]. Wu, X. P., Lan, S. J., Tang, Y. F. Interruptible Vacation Queuing System $M/G/1$ with Tolerant Attendant and N-Policy. Applied Mathematics[J].2024,37(2):563-578.