Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 28 (2025), No. 5, pp. 913–919
DOI : 10.47974/JSMS-1399

On the queueing theory : Problems and applications

Stefan M. Stefanov
Department of Mathematics
South-West University "Neofit Rilski"
2700 Blagoevgrad
Bulgaria

Abstract

In this paper, the queueing theory and its applications are considered. Some basic concepts and results are presented, and some important queue disciplines and mass service systems are considered.

Subject Classification (2020): *90B22, 68M20, 60K25, 60K30.*

Keywords: *Queueing theory, Queue, Service requests, Mass service systems, Queue disciplines, Service models.*

1. Introduction

In the public life and in the everyday life we often encounter a "queue". Typical examples are queues of customers to buy a transport ticket or another type of ticket, queues at a grocery or at other store, queues in banks, queues of TV/Internet subscribers, etc. Queues occur because the flow of service requests is random and generally uncontrollable, and the service points number is limited. If the service points number is large enough, then queues will rarely occur and the average waiting time in a queue will not be long, however, in this case, the long downtime of the employees and of the equipment are possible. If the service points number is small, downtime is almost impossible, however, long queues and losses due to waiting occur.

In order to analyze the functioning of complex systems (for example, automatic telephone exchanges, computer systems, automated information systems, dispatch services, etc.), at which it is necessary to perform mass service of objects, taking into consideration some random factors, it is convenient to study special models, the so called *mass service systems* (MSS). Due to their nature, these are systems for serving queues of requests. The concept of "request"

E-mail: **stefm@swu.bg**

here is understood in the widest possible sense. The theory that studies such models is called the *mass service theory* or the *queueing theory*.

One of the possible formulations of the problem of the queueing theory is as follows: determine the number n of service points, such that the sum $F(n) = F_1(n) + F_2(n)$ of the expected losses $F_1(n)$ due to delayed service and of the equipment downtime $F_2(n)$ be minimal.

The description of a mass service system includes information about:

- incoming flow of requests (queries, calls, customers, etc.);
- devices (lines, channels) and duration of service of the requests;
- order (algorithm, mechanism) of such a service.

In order to increase the efficiency of the real systems, especially important are these characteristics of the mass service system, which take into account the presence of queues, the wait due to absence of requests, and the loss of requests, that is, leaving the system (logout) without service. These characteristics of mass service systems are described by random processes and random variables.

The conditions for existence and the concrete expressions for stationary distributions of the characteristics of mass service systems are of main interest.

The random variables, which describe the parameters and characteristics of the mass service system, are characterized by their nonnegativity.

An important characteristic of the mass service system is the *system occupancy period* (*service occupancy period*). This is the time interval from the moment when the request is received by a vacant (unoccupied) system to the first subsequent moment when the system is free of requests.

The considered topics and problems and related to them are discussed in references [1]-[31], etc.

Queueing theory is also connected with the scheduling theory and its applications (see, for example, [Stefanov 27]), Markov processes and Markov chains, etc.

In Section Two, some probability results are presented. Some queue disciplines and mass service systems are considered in Section Three.

2. Basic concepts of probability theory

Recall some basic concepts and results of probability, which are used in the next section.

Distribution function (DF) $F_X(x)$ of the random variable X is defined as

$$F_X(x) = \begin{cases} P(X \leq x), & x \geq 0, \\ 0, & x < 0, \end{cases}$$

where $P(X \leq x)$ is the probability with which random variable X satisfies inequality $X \leq x$.

The random variable X has a *uniform distribution* (U) in $[0,2a]$ if the distribution function $U(x)$ of X is defined as follows

$$U(x)=\begin{cases}0, & x\le 0,\\ \frac{x}{2a}, & x\in[0,2a],\\ 1, & x\ge 2a.\end{cases}$$

The random variable X has a *deterministic distribution* (D) if

$$P(X=x)=1.$$

The random variable X has an *exponential distribution* (E) with parameter $a>0$ if

$$P(X\le x)=\begin{cases}1-e^{-ax}, & x\ge 0,\\ 0, & x<0.\end{cases}$$

If the random variable X has an exponential distribution then for all values of $x\ge 0$ and $t\ge 0$ we have

$$P(X>x+t|X>x)=P(X>t).$$

The random variable X has *Erlang distribution* (Er) of order r if

$$X=\sum_{k=1}^{r+1} X_k,$$

where X_k, $k=1,\dots,r+1$, is a sequence of independent random variables, distributed uniformly, that is, exponentially distributed random variables with parameter $a(r+1)$.

The random variable X has *Poisson distribution* (P) with a parameter $a>0$ if

$$P(X=k)=\frac{a^k}{k!}e^{-a}, k=0,1,\dots$$

Let X be a discrete random variable, taking nonnegative integer values, that is, $X\in\{0,1,2,\dots\}$. The *probability generating function* G of X is defined as follows

$$G(y)=E(y^X)=\sum_{x=0}^{\infty} p(x)y^x,$$

where E denotes the mathematical expectation, and $p(x)$ denotes the mass probability function of X.

3. Some important mass service systems

The most common case of a mass service system is that with *recurrent input flow*, where elements of the sequence $\{z_k, k\ge 1\}$ of intervals between arrival of requests in the service system are independent in the totality of uniformly distributed random variables with DF $F(x)$. If $F(x)=1-e^{-ax}$ with $a>0$, the flow is called *simple* with intensity a or *Poisson* with parameter a. If requests arrive in the mass service system at the moments that form a recurrent flow (with distributed function $F(x)$ of the length of the intervals between these moments) and at any such moment, a random number of requests n can arrive in the mass service system (with a generating function $\Phi(z)=Ez^n$), then the flow, incoming in a mass service system, is called the *quasi-recurrent flow* of requests and it is denoted by (F,Φ). If the recurrent moments form a simple flow with parameter a, we have a *quasi-Poisson flow* (a,Φ).

The following notation is introduced for the basic types of mass service systems: $A|B|n|m, C$ or $(A|B|n): (C|m|q)$, where

- A is the distribution of the intervals between consecutive moments of request arrivals (the arrival time distribution);
- B is the distribution of the duration of service of the requests (the service time distribution);
- n is the number of servers (devices, channels) in the service system;
- m is the number of the wait places for requests (either finite or infinite);
- C is the order for serving the requests (the queue discipline);
- q is the size of the calling source (either finite or infinite).

The most commonly used types of queue disciplines C are

- FIFO (First In, First Out) or FCFS (First Come, First Served);
- LIFO (Last In, First Out) or LCFS (Last Come, First Served);
- GD (General Discipline, service in any order);
- SIRO (Service In Random Order); etc.

The following mass service systems can be considered:

1. $\mathrm{E}|\mathrm{E}|3|2$, FIFO : a three-channel (three-server) mass service system with simple flow, exponential service (E), with two wait places, requests are served in the arrival order (FIFO).
2. $\mathrm{E}r_2|\mathrm{D}|1|\infty$, LIFO : a single channel (single server) mass service system with recurrent Erlang flow of order $r = 2$ $(\mathrm{E}r_2)$, deterministic service (D), without losses (infinite number of servers, ∞), with reverse order of service (LIFO).
3. $\mathrm{M}_2|\mathrm{U}_2|1|\infty$, PR, FIFO : a single channel (single server) mass service system with two simple flows of requests of different priorities (M_2), uniformly distributed (U_2) durations of service, without losses (∞), requests are served in the arrival order (FIFO).
4. $\mathrm{G}I|\mathrm{E}|\infty$: a mass service system with recurrent income flow (GI, General (generic) distribution of inter-arrival time), infinite number of channels (servers) (∞) and exponential time of request service (E).
5. $(\mathrm{E}|\mathrm{E}|1): (\mathrm{F}IFO|\infty|\infty)$: a single channel (single server) mass service system with simple flow, exponential service of requests (E) in the order of arrival (FIFO) and infinite wait places for requests (∞).
6. $(\mathrm{E}|\mathrm{E}|\infty): (\mathrm{G}D|\infty|\infty)$: a self-service model, the servers number is infinite (∞) since the customers serve requests in general discipline (GD) as well.

4. Conclusions

This paper is devoted to the queueing theory and its applications. Some basic concepts and results are presented, and some important queue disciplines and mass service systems are considered.

References

[1] F. Baskett, K. M. Chandy, R. R. Muntz, and F. G. Palacios, "Open, closed, and mixed networks of queues with different classes of customers", *Journal of the ACM*, vol. 22, no. 2, pp. 248-260 (1975).

[2] A. Bobbio, M. Gribaudo, and M. Telek, "Analysis of large scale interacting systems by mean field method", 2008 Fifth International Conference on Quantitative Evaluation of Systems, pp. 215-224 (2008).

[3] S. K. Bose, An Introduction to Queueing Systems, Kluwer Academic Publishers/Plenum Publishers, Dordrecht-Boston-London (2001).

[4] M. Bramson, "A stable queueing network with unstable fluid model", *The Annals of Applied Probability,* vol. 9, no. 3, pp. 818-853 (1999).

[5] J. P. Buzen, "Computational algorithms for closed queueing networks with exponential servers", *Communications of the ACM*, vol. 16, no. 9, pp. 527-531 (1973).

[6] H. Chen, and W. Whitt, "Diffusion approximatioms for open queueing networks with service interruptions", *Queueing Systems,* vol. 13, no. 4, pp. 335-359 (1993).

[7] E. Gelenbe, "G-networks with triggered customer movement", *Journal of Applied Probability*, vol. 30, no. 3, pp. 742-748 (1993).

[8] W. J. Gordon, and G. F. Newell, "Closed queueing systems with exponential servers", *Operations Research,* vol. 15, no. 2, pp. 254-265 (1967).

[9] J. F. Shortle, J. M. Thompson, D. Gross, and C. M. Harris, Fundamentals of Queueing Theory, 5th ed., Wiley Series in Probability and Statistics, John Wiley & Sons, New York (2018).

[10] F. P. Kelly, "Networks of queues with customers of different types", *Journal of Applied Probability*, vol. 12, no. 3, pp. 542-554 (1975).

[11] L. R. Lipsky, Queueing Theory: A Linear Algebraic Approach, Macmillan, New York (1992).

[12] S. M. Stefanov, "On the implementation of stochastic quasigradient methods to some facility location problems", *Yugoslav Journal of Operations Research,* vol. 10, no. 2, pp. 235-256 (2000).

[13] S. M. Stefanov, "Convex separable minimization subject to bounded variables", *Computational Optimization and Applications. An International Journal*, vol. 18, no. 1, pp. 27-48 (2001).

[14] S. M. Stefanov, Separable Programming: Theory and Methods, Applied Optimization, vol. 53, Kluwer Academic Publishers, Dordrecht-Boston-London (2001).

[15] S. M. Stefanov, "Convex quadratic minimization subject to a linear constraint and box constraints", A*pplied Mathematics Research Express AMRX,* vol. 2004, no. 1, pp. 17-42 (2004).

[16] S. M. Stefanov, "Minimization of a strictly convex separable function subject to convex separable inequality constraint and box constraints", *Journal of Interdisciplinary Mathematics,* vol. 12, no. 5, pp. 647-673 (2009).

[17] S. M. Stefanov, "Solution of some convex separable resource allocation and production planning problems with bounds on the variables", *Journal of Interdisciplinary Mathematics*, vol. 13, no. 5, pp. 541-569 (2010).

[18] S. M. Stefanov, "On the solution of multidimensional convex separable continuous knapsack problem with bounded variables", *European Journal of Operational Research,* vol. 247, no. 2, pp. 366-369 (2015).

[19] S. M. Stefanov, "Strictly convex separable optimization with linear equality constraints and bounded variables", *Journal of Statistics and Management Systems,* vol. 21, no. 2, pp. 261-272 (2018).

[20] S. M. Stefanov, "Characterization of the optimal solution of the convex separable continuous knapsack problem and related problems", *Journal of Information and Optimization Sciences*, vol. 42, no. 1, pp. 1-16 (2021).

[21] S. M. Stefanov, "On the numerical solution of separable stochastic inventory control problems", *Journal of Information and Optimization Sciences,* vol. 42, no. 3, pp. 533-561 (2021).

[22] S. M. Stefanov, Separable Optimization: Theory and Methods, Springer Optimization and Its Applications, vol. 177, Springer, Cham (2021).

[23] S. M. Stefanov, "On the solution of quadratic programming problems", *Journal of Information and Optimization Sciences,* vol. 44, no. 2, pp. 243-253 (2023).

[24] S. M. Stefanov, "Continuous linear knapsack problems revisited", *Journal of Information and Optimization Sciences*, vol. 44, no. 5, pp. 909-922 (2023).

[25] S. M. Stefanov, "Numerical solution of box constrained separable convex quadratic programming problems", *Journal of Information and Optimization Sciences,* vol. 45, no. 1, pp. 57-71 (2024).

[26] S. M. Stefanov, "Linear programming problems and matrix games", *Journal of Interdisciplinary Mathematics*, vol. 27, no. 1, pp. 57-65 (2024).

[27] S. M. Stefanov, "Scheduling theory and applications revisited", *Journal of Interdisciplinary Mathematics,* to appear.

[28] M. Tanner, Practical Queueing Analysis, McGraw-Hill, New York (1995).

[29] W. Whitt, "The queueing network analyzer", *The Bell System Techical Journal,* vol. 62, no. 9, pp. 2779-2815 (1983).

[30] W. Whitt, Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues, Springer Series in Operations Research, Springer-Verlag, New York-Berlin-Heidelberg (2002).

[31] K. Yamada, "Diffusion approximation for open state-dependent queueing networks in the heavy traffic situation", *The Annals of Applied Probability,* vol. 5, no. 4, pp. 958-982 (1995).

Received April, 2024