Analysis of Closed Queueing Networks with Batch Service

E. P. Stankevich, I. E. Tananko, V. I. Dolgov

Elena P. Stankevich, https://orcid.org/0000-0003-0630-4550, Saratov State University, 83 Astrakhanskaya St., Saratov 410012, Russia, StankevichElena@mail.ru

Igor E. Tananko, https://orcid.org/0000-0001-8960-9709, Saratov State University, 83 Astrakhanskaya St., Saratov 410012, Russia, TanankoIE.sgu@gmail.com

Vitalii I. Dolgov, https://orcid.org/0000-0003-2413-7880, Saratov State University, 83 Astrakhanskaya St., Saratov 410012, Russia, mail@vidolgov.ru

We consider a closed queuing network with batch service and movements of customers in continuous time. Each node in the queueing network is an infinite capacity single server queueing system under a RANDOM discipline. Customers move among the nodes following a routing matrix. Customers are served in batches of a fixed size. If a number of customers in a node is less than the size, the server of the system is idle until the required number of customers arrive at the node. An arriving at a node customer is placed in the queue if the server is busy. The batch service time is exponentially distributed. After a batch finishes its execution at a node, each customer of the batch, regardless of other customers of the batch, immediately moves to another node in accordance with the routing probability. This article presents an analysis of the queueing network using a Markov chain with continuous time. The qenerator matrix is constructed for the underlying Markov chain. We obtain expressions for the performance measures. Some numerical examples are provided. The results can be used for the performance analysis manufacturing systems, passenger and freight transport systems, as well as information and computing systems with parallel processing and transmission of information.

Keywords: queueinq networks, batch service, Markov chains.

Received: 15.06.2019 / Accepted: 23.07.2020 / Published: 30.11.2020

DOI: https://doi.org/10.18500/1816-9791-2020-20-4-527-533

INTRODUCTION

Queueing models are widely used for system performance evaluation and prediction for different classes of real systems. Some examples of these systems include telecommunication, computing, traffic engineering, health care, and much more. There is a large number of works focused on queueing networks where customers are served one at a time [1–3].

However, there are a lot of systems where customers are served in batches (transport vehicles, busses, ship locks). M. L. Chaudhry and J. G. C. Templeton [4] presents an overview of the main results for queueing systems with batch service. More results can be found in [5–7]. Furthermore, in telecommunications, packets are grouped in batches and these batches are transmitted instead of each packet individually. The most natural models for this systems are queueing networks with batch service and movements. Thus batch service and movements queueing networks are a generalization of queueing networks and allow to model more complicated systems.

Discrete-time open queueing networks with batch service were considered in [8–10]. The networks are realistic and practical for modeling wireless sensor networks, ATM, slotted ALOHA. It is worth mentioning, there is a product form stationary distribution for the queueing networks. For example, C. Malchin and H. Daduna [9] extends product form results for discrete-time open queueing networks to include availability of unreliable nodes and state dependent arrival rates.

Customer coalescence was considered in [10]. At a service completion the entire batch coalesces into a single unit, and it either leaves the system or goes to another node according to given routing probabilities. When the batch sizes are identical to one, the network reduces to a classical Jackson network.

A queueing network with triggered concurrent batch arrivals and batch services was considered in [11]. The batch movement can also be triggered by arrivals and/or departures. Specifically, an arriving or departing batch may induce another event to occur before they are routed. This triggered event may either be the addition of a batch of customers to the network, or the removal of a batch of customers from the network. For this network, its stationary distribution has a product form.

Non-product form queueing networks are considered in [12]. For this networks the decomposition method for ordinary single class open queueing networks is extended to queueing networks with batch processing. A comparison study with discrete event simulation as benchmark shows that the approach provides fairly good results for a wide range of applications. The model has been integrated into a software system for analysing large scale semiconductor manufacturing systems.

In this article, we consider a closed continuous-time queuing network with batch service and movements of customers. Each node in the queueing network is an infinite capacity single server queueing system under a RANDOM discipline. Customers move among the nodes following a routing matrix. Customers are served in batches of a fixed size. If a number of customers in a node is less than the size, the server of the system is idle until the required number of customers arrive at the system. An arriving at a node customer is placed in the queue if the server is busy. The batch service time is exponentially distributed. After a batch finishes its execution at a node, each customer of the batch, regardless of other customers of the batch, immediately moves to another node in accordance with the routing probability.

The remainder of the paper is organized as follows. Section 2 describes the queueing network under consideration. In section 3 we obtain the stationary distribution and performance measures. Section 4 provides a numerical example. Finally, a section of conclusions commenting the main research contributions of this paper is presented.

1. THE MODEL

Consider a continuous-time closed queueing network consisting of L nodes S_i, $i \in I = \{1, \dots, L\}$. There are H customers in the network.

Each node S_i, $i = 1, \dots, L$, operates like an infinite capacity single-server queueing system under a RANDOM discipline. An arriving at a node customer is placed in the queue if the server is busy. Customers are served in batches, let g_i be the customer batch size for node S_i. If a number of customers in the queue is less than the size, the server is idle until the required number of customers arrive at the node. The service of a batch is started immediately after there are at least g_i customers in the queue.

After a batch finishes its execution at a node, the customers leave the node and move independently between nodes according to routing matrix $\Theta = (\theta_{ij})$, $i, j = 1, \dots, L$. Upon service completion at node S_i, a customer may go to node S_j with probability θ_{ij}, $i, j = 1, \dots, L$. The service times of batches at node S_i have an exponential distribution with parameter μ_i, $i = 1, \dots, L$.

Let s_i be the number of customers at node S_i, $i = 1, \dots, L$. The network state is described by a vector $s = (s_1, \dots, s_L)$. Denote by $X = \left\{s:\ s_i \geqslant 0,\ \sum_{i=1}^{L} s_i = H\right\}$ the state space of the queueing network, $|X| = \binom{H+L-1}{L-1}$. By V_i denote the set $V_i = \{j \in I\colon \theta_{ij} > 0\}$, $i = 1, \dots, L$.

Consider a transition from state $s \in X$ to state $s' \in X$, $s \neq s'$.

1. Let a batch complete its service at node S_i, $i \in I$, and thus g_i customers leave node S_i. Denote by $d = (d_1, \dots, d_L)$ a vector representing departing customers, all components of the vector equal to 0, except the ith, which is g_i. Let D be the set of the departing vectors, $|D| = L$.
2. Each of d_i customers goes independently to nodes according to the routing matrix.
3. Let the size of a batch arriving at node S_j be a_j, $j = 1, \dots, L$, then $d_i = \sum_{j \in V_i} a_j$. Thus vector $a = (a_1, \dots, a_L)$ represents the entering customers. Denote by A the set of the entering vectors, $|A| = \sum_{i=1}^{H} \binom{i+L-1}{L-1}$.
4. Thus for the transition $s' = s - d + a$.

2. PERFORMANCE MEASURES

The process $\{s(t), t > 0\}$ is a continuous time Markov chain on the state space X. It is known, the transition rate $q(s, s')$ from state s to state s' has the following form [13]:

$$q(s, s') = \sum_{\substack{s' \in X, \\ s' = s - d + a}} u(s, d)\rho(d, a), \quad d \in D, \quad a \in A, \quad s \in X, \tag{1}$$

$u(s, d)$ is a function associated with the service rates of the network, $\rho(d, a)$ is a function associated with the routing probabilities of the network.

According to the service policy we have

$$u(s, d) = \sum_{i=1}^{L} \mu_i \mathbf{1}(s_i \geqslant d_i), \tag{2}$$

where $\mathbf{1}(s_i \geqslant d_i) = 1$, if $s_i \geqslant d_i$, and $\mathbf{1}(s_i \geqslant d_i) = 0$ otherwise.

Let a batch finish its service at node S_i, customers arrive at nodes according to vector a. Denote by ζ_j the random variable representing the size of the arriving batch at node S_j. As customers go between nodes independently on each other, the random variables are independent random variables with the multinomial distribution. Thus we can write

$$P_{d_i}(\zeta_1 = a_1, \dots, \zeta_L = a_L) = \binom{d_i}{a_1, \dots, a_L} \prod_{j=1}^{L} \theta_{ij}^{a_j}.$$

In this case, $\rho(d, a)$ represents the probability distribution for the sizes of arriving batches, where

$$\rho(d, a) = \sum_{i=1}^{L} \binom{d_i}{a_1, \dots, a_L} \prod_{j=1}^{L} \theta_{ij}^{a_j}, \quad d_i = \sum_{j \in V_i} a_j, \quad d \in D, \quad a \in A. \tag{3}$$

Substituting $u(s,d)$ and $\rho(d,a)$ in (1) into (2) and (3), we get

$$q(s,s') = \sum_{\substack{s'\in X,\\ s'=s-d+a}} \sum_{i=1}^{L} \mu_i \mathbf{1}(s_i \geqslant d_i) \binom{d_i}{a_1,\dots,a_L} \prod_{j=1}^{L} \theta_{ij}^{a_j}, \ s\in X, \ d_i = \sum_{j\in V_i} a_j, \ d\in D, \ a\in A.$$

The stationary distribution $\pi = (\pi(s))$, $s \in X$, for the queueing network can be obtained as a solution of the following equations

$$\pi Q = 0, \qquad \sum_{s\in X} \pi(s) = 1,$$

where Q is the generator matrix, $Q = (q(s,s'))$, $s, s' \in X$.

Once the stationary distribution is computed, a variety of other performance measures may be obtained.

The average number $\bar{s}_i$ of customers at the node S_i

$$\bar{s}_i = \sum_{k=1}^{H} k \sum_{\substack{s\in X,\\ s_i=k}} \pi(s), \quad i = 1,\dots,L,$$

the arrival rate λ_i to node S_i

$$\lambda_i = \mu_i d_i \Big(1 - \sum_{s\in X} \sum_{s_i=0}^{d_i-1} \pi(s)\Big), \quad i = 1,\dots,L,$$

the average response time $\bar{u}_i$ for node S_i

$$\bar{u}_i = \frac{\bar{s}_i}{\lambda_i}, \quad i = 1,\dots,L,$$

the average idle time $\bar{v}_i$ for node S_i

$$\bar{v}_i = \frac{\sum\limits_{k=0}^{d_i-1} (d_i - k) \sum\limits_{\substack{s\in X,\\ s_i=k}} \pi(s)}{\lambda_i \sum\limits_{k=0}^{d_i-1} \sum\limits_{\substack{s\in X,\\ s_i=k}} \pi(s)}, \quad i = 1,\dots,L,$$

the average waiting time $\bar{w}_i$ for node S_i

$$\bar{w}_i = \bar{u}_i - \frac{1}{\mu_i}, \quad i = 1,\dots,L,$$

the average number $\bar{b}_i$ of customers in the queue for node S_i

$$\bar{b}_i = \bar{w}_i \lambda_i, \quad i = 1,\dots,L.$$

3. NUMERICAL RESULTS

Consider a queueing networks which consists of $L = 5$ nodes with service rates $\mu = (0.5, 0.8, 0.4, 0.5, 0.6)$ and routing matrix Θ, where

$$\Theta = \begin{bmatrix} 0.0 & 0.3 & 0.5 & 0.0 & 0.2 \\ 0.4 & 0.0 & 0.1 & 0.5 & 0.0 \\ 0.2 & 0.2 & 0.0 & 0.3 & 0.3 \\ 0.0 & 0.6 & 0.2 & 0.0 & 0.2 \\ 0.7 & 0.0 & 0.1 & 0.2 & 0.0 \end{bmatrix}.$$

There are $H = 10$ customers in the network.

Table represents the performance measures for several service batch sizes.

Table

Performance measures for different values of batch sizes

g	$(2,2,2,2,2)$	$(1,3,2,3,1)$	$(1,1,1,1,1)$
$\bar{s}$	$(2.51, 1.50, 2.62, 2.08, 1.28)$	$(3.61, 1.66, 1.85, 1.97, 0.91)$	$(2.95, 0.96, 3.37, 2.01, 0.71)$
λ	$(0.61, 0.60, 0.51, 0.53, 0.38)$	$(0.46, 0.45, 0.38, 0.40, 0.29)$	$(0.41, 0.40, 0.34, 0.35, 0.25)$
$\bar{u}$	$(4.11, 2.49, 5.16, 3.93, 3.37)$	$(7.87, 3.66, 4.83, 4.93, 3.16)$	$(7.24, 2.39, 9.90, 5.67, 2.80)$
$\bar{v}$	$(2.24, 2.31, 2.68, 2.55, 3.68)$	$(2.18, 4.00, 3.57, 4.28, 3.49)$	$(2.46, 2.48, 2.94, 2.82, 3.93)$
$\bar{w}$	$(2.11, 1.24, 2.66, 1.93, 1.71)$	$(5.87, 2.41, 2.33, 2.93, 1.49)$	$(5.24, 1.14, 7.40, 3.70, 1.13)$
$\bar{b}$	$(1.29, 0.75, 1.36, 1.02, 0.65)$	$(2.70, 1.09, 0.90, 1.17, 0.43)$	$(2.13, 0.46, 2.52, 1.30, 0.29)$

Note that for $g = (1, 1, 1, 1, 1)$, it is a Gordon – Newell network. We see that the arrival rates for the network with $g = (2, 2, 2, 2, 2)$ are greater than for the network where $g = (2, 2, 2, 2, 2)$. However, the average response times for $g = (2, 2, 2, 2, 2)$ are less, expecting the average response time for S_5. There are large batch sizes for S_2 and S_4 in the network with $g = (1, 3, 2, 3, 1)$ whereas there are $H = 10$ customers in the network. We have the similar values for arrival rates, but other performance measures differ significantly.

CONCLUSION

This paper analyzed a closed queueing network with batch service. We obtain the stationary distribution of the network and its performance measures. At the end, some examples are presented. The results can be used for the performance analysis of transport vehicles, telecommunication systems, manufacturing systems and in the design of factories, shops, offices and hospitals.

Acknowledgements: This work was supported by the Ministry of Science and Higher Education of the Russian Federation in the framework of the state assignment (project No. FSRR-2020-0006).

References

1. Basharin G. P., Bocharov P. P., Kogan A. M. *Analiz ocheredei v vychislitel'nykh setyakh. Teoriya i sposoby rascheta* [Analysis of queues in computer networks. Theory and methods of calculation]. Moscow, Nauka, 1989. 336 p. (in Russian).
2. Vishnevskii V. M. *Teoreticheskie osnovy proektirovaniya komp'iuternykh setei* [Theoretical foundations for the design of computer networks]. Moscow, Tekhnosfera, 2003. 512 p. (in Russian).

3. Mitrophanov Yu. I. *Analiz setei massovogo obsluzhivaniya* [Analysis of queueing networks]. Saratov, Nauchnaya kniga, 2005. 175 p. (in Russian).
4. Chaudhry M. L., Templeton J. G. C. *A First Course in Bulk Queues*. New York, John Wiley & Sons Inc., 1983. 372 p.
5. Krishnamoorthy A., Ushakumari P. A Queueing System with Single Arrival Bulk Service and Single Departure. *Mathematical and Computer Modelling*, 2000, vol. 31, no. 2, pp. 99–108. DOI: https://doi.org/10.1016/S0895-7177(99)00226-5
6. Bountali O., Economou A. Equilibrium Joining Strategies in Batch Service Queueing Systems. *European Journal of Operational Research*, 2017, vol. 260, iss. 3, pp. 1142–1151. DOI: https://doi.org/10.1016/j.ejor.2017.01.024
7. Pershakov N. V., Samouylov K. E. M|G|1 Queues with batch service and its aplication to the stream control trasmission protocol performance analysis. Part I. *Vestnik RUDN. Ser. Matematika. Informatika. Fizika* [Bulletin of Peoples' Friendship. Ser. Maths. Informatics. Physics], 2009, no. 1, pp. 34–44 (in Russian).
8. Walrand J. A Discrete-Time Queueing Network. *Journal of Applied Probability*, 1983, vol. 20, no. 4, pp. 903–909. DOI: https://doi.org/10.2307/3213603
9. Malchin C., Daduna H. Discrete Time Queueing Networks with Product Form Steady State. Availability and Performance Analysis in an Integrated Model. *Queueing Systems*, 2010, vol. 65, no. 4, pp. 385–421. DOI: https://doi.org/10.1007/s11134-010-9181-2
10. Chao X., Pinedo M., Shaw D. Networks of Queues with Batch Services and Customer Coalescence. *Journal of Applied Probability*, 1996, vol. 33, no. 3, pp. 858–869. DOI: https://doi.org/10.2307/3215364.
11. Chao X., Zheng S. Triggered Concurrent Batch Arrivals and Batch Departures in Queueing Networks. *Discrete Event Dynamic Systems: Theory and Applications*, 2000, vol. 10, pp. 115–129. DOI: https://doi.org/10.1023/A:1008339216447
12. Hanschke Th., Zisgen H. Queueing Networks with Batch Service. *European Journal Industrial Engineering*, 2011, vol. 5, no. 3, pp. 313–326. DOI: https://doi.org/10.1504/EJIE.2011.041619
13. Boucherie R. J., Dijk N. M. Product Forms for Queueing Networks with State-Dependent Multiple Job Transitions. *Advances in Applied Probability*, 1991, vol. 23, no. 1, pp. 152–187. DOI: https://doi.org/10.2307/1427516

Cite this article as:
Stankevich E. P., Tananko I. E., Dolgov V. I. Analysis of Closed Queueing Networks with Batch Service. *Izv. Saratov Univ. (N. S.), Ser. Math. Mech. Inform.*, 2020, vol. 20, iss. 4, pp. 527–533. DOI: https://doi.org/10.18500/1816-9791-2020-20-4-527-533

УДК 519.872

Анализ замкнутых сетей массового обслуживания с групповым обслуживанием

Е. П. Станкевич, И. Е. Тананко, В. И. Долгов

Станкевич Елена Петровна, кандидат физико-математических наук, старший научный сотрудник научно-образовательного математического центра «Математика технологий будущего», Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, Россия, 410012, г. Саратов, ул. Астраханская, д. 83, StankevichElena@mail.ru

Тананко Игорь Евстафьевич, кандидат физико-математических наук, заведующий кафедрой системного анализа и автоматического управления, Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, Россия, 410012, г. Саратов, ул. Астраханская, д. 83, TanankoIE.sgu@gmail.com

Долгов Виталий Игоревич, кандидат физико-математических наук, доцент кафедры системного анализа и автоматического управления, Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, Россия, 410012, г. Саратов, ул. Астраханская, д. 83, mail@vidolgov.ru

Рассматривается замкнутая сеть массового обслуживания с групповым обслуживанием, групповыми переходами требований и непрерывным временем. Каждая система обслуживания сети состоит из одного прибора и очереди бесконечной длины. В соответствии с маршрутной матрицей сети между системами массового обслуживания осуществляются переходы требований одного класса. Длительности обслуживания требований приборами систем являются экспоненциально распределенными случайными величинами. Обслуживание требований в системах производится группами фиксированного размера. Если число требований, находящихся в системе обслуживания, меньше заданного размера группы, то обслуживающий прибор системы простаивает до момента прибытия в систему необходимого числа требований. Если же прибор занят обслуживанием группы требований, то вновь приходящие требования становятся в очередь системы. Выбор требований из очереди осуществляется согласно дисциплине RANDOM. После завершения обслуживания в системе каждое требование группы независимо от других требований в соответствии с маршрутной вероятностью мгновенно переходит в другую систему обслуживания. Предложен метод анализа сети обслуживания данного вида с использованием цепи Маркова с непрерывным временем. Для модельной цепи Маркова построена матрица интенсивностей переходов. Получены выражения для вычисления стационарных характеристик систем массового обслуживания рассматриваемой сети. Приведен пример численного анализа сети массового обслуживания. Полученные результаты могут быть использованы для решения задач распределения ресурсов, анализа производственных систем, систем пассажирских и грузовых перевозок, а также информационных и вычислительных систем с параллельной обработкой и передачей информации.

Ключевые слова: сети массового обслуживания, групповое обслуживание, цепи Маркова.

Поступила в редакцию: 15.06.2019 / Принята: 23.07.2020 / Опубликована: 30.11.2020

Благодарности. Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках выполнения государственного задания (проект № FSRR-2020-0006).

Образец для цитирования:
Stankevich E. P., Tananko I. E., Dolgov V. I. Analysis of Closed Queueing Networks with Batch Service [*Станкевич Е. П., Тананко И. Е., Долгов В. И.* Анализ замкнутых сетей массового обслуживания с групповым обслуживанием] // Изв. Сарат. ун-та. Нов. сер. Сер. Математика. Механика. Информатика. 2020. Т. 20, вып. 4. С. 527–533. DOI: https://doi.org/10.18500/1816-9791-2020-20-4-527-533