УДК 519.872
DOI: 10.17223/19988605/47/3

M.Tz. Dimitrov

HEAVY TRAFFIC ANALYSIS OF A QUEUE WITH BATCH MMAP

A single-server queue with work conserving FIFO discipline is considered. The input process is a multiple marked. Markovian arrival process governed by a continuous-time finite state Markovian chain. The service time distributions of customers may be different for different arrival streams. The virtual waiting time is considered under a heavy traffic. The probability distributions of virtual waiting time and state of the random environment are asymptotically independent. The virtual waiting time is asymptotically exponential with the mean depending on the characteristics of the modulated process.
Keywords: single-server queue; FIFO; multiple batch Markovian arrival process; virtual waiting time; heavy traffic analysis.

In this paper we consider a single-server work conserving queue with FIFO discipline and the following characteristics. The arrival process is a multiple batch arrival flows of customers governed by a continuous time finite state Markov chain. The service time distributions of customers may be different for arrival flows. This model was considered for the first time in [1]. The most popular Markovian arrival flows of customers is introduced in [2]. There are two extensions of arrival process. One is batch introduced in [3] that allows batch arrivals and the other is marked introduced in [4] that explicitly represents possibly correlated process considered in this paper and in paper [1] is a further extension of MAP introduced in [2–4]. We assume that the service time distributions of customers from respective arrival streams are different from one another.

In this paper we consider asymptotic behavior of a stationary distribution of the virtual waiting time and the state of a random environment under a heavy traffic. The asymptotic distribution of a stationary distribution of a virtual waiting time is an exponential with the mean depending on the parameters of random environment. In the present study we prove a heavy traffic limit theorem for steady state virtual waiting time for a model considered in [1]. We used analytical approach introduced in [6] which allows for a rigorous mathematical analysis of the stationary characteristics under a heavy traffic. The remainder of the paper is organized as follows. In section [2] we describe the model introduced in [1]. In section 3 we consider some preliminary results. Section 4 is devoted to main results. In section 4 we find the mean value of the virtual waiting time in a steady state, and there we show that the distribution of the environment becomes independent of the virtual waiting time in a steady state under heavy traffic conditions. The limit distribution of the virtual waiting time under heavy traffic conditions is an exponential distribution.

1. Model description

In this paper we consider a single sever queue with the following characteristics. Arrivals to the system are from *K* arrival streams. We call customers from the *k*-th ($k = 1, 2, \cdots, K$) arrival stream class *k* customers. Customer arrivals are governed by a continuous-time Markov chain $Z(t)$. We assume that Markov chain $Z(t)$ has finite state space $S = \{0, 1, 2, \cdots, M\}$ and is irreducible. The underlying Markov chain stais in state $i \in S$ for an exponential interval of time with a mean value μ_i^{-1} . When the sojourn time in state *i* has elapsed with probability $\sigma_{i,j}(0)$ the Markov chain $Z(t)$ changes its state *i* to state *j* without arrivals. The chain $Z(t)$ changes its state *i* to state *j* and *n* customers of class *k* arrive simultaneously with probability $\sigma_{k,i,j}(n)$, $(k \in \{1,2,\cdots,K\},\ n=1,2,\cdots)$ for convenience, let $\sigma_{i,i}=0$ for all $i \in S$. Then for all $i \in S$ $\sum_{j=0}^{j=M}(\sigma_{i,j}(0)+\sum_{k=1}^{k=K}\sum_{n=1}^{\infty}\sigma_{k,i,j}(n))=1$.We assume that service time of class *k* customers are i.i.d. according to a distribution $H_k(x)$ with a mean value h_k, and a second moment v_k. We need some notations to describe the arrival process. Let C denote an $(M+1)\times(M+1)$ matrix whose (i,j)-th element $(i,j \in \{0,1,2,\cdots,M\})$ is given by $c_{i,j}=-\mu_i$, if $i=j$ and $c_{i,j}=\sigma_{i,j}(0)\mu_i$, otherwise. Further, for $k \in \{1,2,\cdots,K\}$ we define $D_k(n),(n=1,2,\cdots)$ as an $(M+1)\times(M+1)$ matrix whose $(i,j)-$th element $(i,j \in \{0,1,2,\cdots M\})$ of $D_{k,i,j}(n)$ is given by $D_{k,i,j}(n)=\sigma_{k,i,j}(n)\mu_i$. Thus the counting process of arrivals is characterized by the set of matrices $\{C, D_1(n_1), D_2(n_2),\cdots,D_K(n_K)\}$. Customers arrive in the following way. When a state transition driven by $D_k(n)$ occurs, *n* customers of class *k* arrive simultaneously. On the other hand, when a state transition driven by C occurs, no customers arrive. We define $D_k(k \in \{1,2,\cdots,K\})$ and D as $D_k=\sum_{n=1}^{\infty}D_k(n)$, $D=\sum_{k=1}^{k=K}D_k$, respectively. Note that the infinitesimal generator of the underlying Markov chain is given by $C+D$ and $(C+D)e=0$, where *e* denotes a column vector whose elements are all equal to one. We denote by π, the stationary probability vector of the underlying Markov chain and therefore π satisfies the equation $\pi(C+D)=0$ and $\pi e=1$. Because of the finite state space and the irreducibility of the underlying Markov chain, π is uniquely determined. Note that the arrival rate λ_k of class *k* is $\lambda_k=\sum_{n=1}^{\infty}n\pi D_k(n)e.$ We assume that at least one element of D_k is positive, so that $\lambda_k>0$ for all $k \in \{1,2,\cdots,K\}$. Let ρ_k denote the utilization factor of class *k* customers, $\rho_k=\lambda_k h_k$. Furthermore, we denote the overall utilization factor by $\rho=\sum_{k=1}^{K}\rho_k$. If the utilization factor $\rho<1$ then all customers arriving to the system are eventually served. The goal of this paper is to derive the asymptotic distribution of virtual waiting time under a heavy traffic, i.e., when $\rho<1$ and $\rho\to 1$, or $\varepsilon=1-\rho\to 0$. The virtual waiting time is equivalent to the amount of work in a system. Let $V(t)$ denote a random variable representing the stationary amount of work in a system (the total amount of unfinished services of all customers in the system). We define $F(x)$ as a $(1,M+1)$ vector whose *j*-th element represents $P\{V(t)\le x, Z(t)=j\}$. The Laplace-Stiljes transforms (LST) of $H_k(x)$ and $F(x)$ are denoted by $h_k(s)$ and $\varphi(s)$ respectively. Applying the results in [5], we obtain the *LST* $\varphi(s)$ of a distribution function $F(x)$

$$\varphi(s)=sF(0)(sI+C+\bar{D}(s))^{-1}, \tag{1}$$

where $\bar{D}(s)=\int_0^{\infty}e^{-sx}dD(x)=\sum_{k=1}^{K}\sum_{n=1}^{\infty}D_k(n)h_k^n(s)$. We assume that $\mu_i=\lambda\mu_i^0$ and we will introduce the matrices $C^0, D_k(n), D_k^0, D^0, Q, Q^0$ by the equations $C=\lambda C^0$, $D_k(n)=\lambda D_k^0(n), D_k=\lambda D_k^0, D=\lambda D^0$, $Q=C+D, Q=\lambda Q^0$. Further we will also assume that all parameters except for λ are fixed, and λ increases in such a way that $\varepsilon\to 0$, or $\lambda\to\lambda^0=(\rho^0)^{-1}$, where $\rho^0=\pi\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k^0(n)h_k e.$

2.Some preliminary results

Consider the continuous Markov chain $Z(t)$ introduced in Section 1. Markov chain is with a finite space $S=\{0,1,2,\cdots,M\}$, the infinitesimal matrix $Q=C+D$ and a stationary distribution $\pi=(\pi_0,\pi_1,\cdots,\pi_M)$. The set of equations with respect to unknown $a_0,a_1,\cdots,a_M$

$$Qa=(\pi\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k(n)h_k e)e-\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k(n)h_k e \tag{2}$$

always has a solution, since the stationary distribution π is orthogonal to the right side of this equation. We will define matrices *A* and *R*. The matrix *A* is defined by the rows and columns with numbers $1,2,...,M$ of the matrix *Q*. The first raw and the first column of matrix *R* are equal to zero vectors, and next *K* rows and *K* columns are from matrix A^{-1}. Then a matrix *QR* looks as follows: the first raw is equal to $(0,-\frac{\pi_1}{\pi_0},...,-\frac{\pi_M}{\pi_0})$ and the first column is equal to zero vector, and next elements form the identity matrix. For any vector $x=(x_0,x_1,\cdots,x_M)$ we have

$$xQR=x-x_0\frac{\pi}{\pi_0}. \tag{3}$$

The next equation

$$(a_0,a_1,\cdots,a_M)=R[(\pi\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k(n)h_k e)e-\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k(n)h_k e] \tag{4}$$

define the solution to equation (2) when $a_0=0.$ Dividing both side of (2) by λ we get

$$Q^0a=(\pi\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k^0(n)h_k e)e-\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k^0(n)h_k e. \tag{5}$$

And the solution to equation (2) is also a solution to equation (5). From (4) follows

$$(a_0,a_1,\cdots,a_M)=R^0[(\pi\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k^0(n)h_k e)e-\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k^0(n)h_k e]\,. \tag{6}$$

3. The main results

The next theorem gives the mean value of a virtual waiting time in a steady state.

Theorem 1. The mean virtual waiting time *EV* is given by the formula

$$(1-\rho)EV=\frac{\pi}{2}\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k(n)v_{k,n}e+\pi(\sum_{k=1}^{K}\sum_{n=1}^{\infty}nD_k(n)h_k-I)a+F(0)a\,, \tag{7}$$

where *a* is a solution to equation (4) and

$$F(0)e=1-\rho. \tag{8}$$

Proof. Consider a LST $\varphi(s)$ of a vector $F(x)=(F_0(x),F_1(x),...,F_M(x))$ defined by the equation

$$\varphi(s)(sI+C+\bar{D}(s))=sF(0). \tag{9}$$

where $\operatorname{Re}s>0$. We denote by $\bar{D}(s)$ the next sum $\sum_{k=1}^{K}\sum_{n=1}^{\infty}D_k(n)h_k^n(s)$ and rewrite equation (9) to be

$$\varphi(s)Q=\varphi(s)(\sum_{k=1}^{K}\sum_{n=1}^{\infty}D_k(n)(s)(1-h_k^n(s))-sI)+sF(0). \tag{10}$$

Post-multiplying both sides of equation (10) by vector *e* we have

$$F(0)e=\varphi(s)(I-\sum_{k=1}^{K}\sum_{n=1}^{\infty}D_k(n)\frac{1-h_k^n(s)}{s})e\,. \tag{11}$$

Taking the limit $s \to +0$ of both sides of (11) and taking into account that $\varphi(0) = \pi$ we get equation (8). Differentiating both sides of (11) with respect to s at $s = 0$ yields

$$(EV_0, EV_1, ..., EV_M)(I - \sum_{k=1}^{K}\sum_{n=1}^{\infty} nD_k(n)h_k)e = \frac{\pi}{2}\sum_{k=1}^{K}\sum_{n=1}^{\infty} nD_k(n)v_{k,n}e \tag{12}$$

for the mean value of a virtual waiting time $EV_n = E(V(t); Z(t) = n)$. Post-multiplying both sides of the equation (10) by the vector a we obtain

$$\varphi(s)Qa = s\varphi(s)(\sum_{k=1}^{K}\sum_{n=1}^{\infty} nD_k(n)\frac{1-h_k^n(s)}{s} - I)a + sF(0)a.$$

After differentiating this equation with respect to s at the point $s = 0$ we have

$$-(EV_0, EV_1, ..., EV_M)Qa = \pi(\sum_{k=1}^{K}\sum_{n=1}^{\infty} nD_K(n)h_k - I)a + F(0)a. \tag{13}$$

Now, summing (13) with (12) we get equation (7). Thus the theorem is proved.

Theorem 2. Under heavy taffic assumption

a) the random variables $\varepsilon V(t)$ and $Z(t)$ are asymptotically independent;

b) the random variable $V(t)$ is asymptotically exponential with the mean

$$\frac{\pi}{2\rho^0}\sum_{k=1}^{K}\sum_{n=1}^{\infty} nD_k^0(n)v_{k,n}e + \frac{\pi}{\rho^0}(\sum_{k=1}^{K}\sum_{n=1}^{\infty} nD_k^0(n)h_k - \rho^0 I)a\,. \tag{14}$$

where a is a solution to equation (2).

Proof: We consider equation (10)

$$\varphi(s)Q = \varphi(s)(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI) + sF(0)\,.$$

Post-multiplying both sides of this equation by the vector $e = (1,1,...,1)^{\mathrm{T}}$, taking into account that $Qe = 0$ and from (8) we get

$$\varphi(s)(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)e + \varepsilon s = 0\,. \tag{15}$$

Now, post-multiplying equation (10) by the matrix R, which was introduced in section 3, from (3) follows the next equation

$$\varphi(s) = \frac{\varphi_0(s)}{\pi_0}\pi + \varphi(s)(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)R + sF(0)R\,. \tag{16}$$

Replacing the vector $\varphi(s)$ in the right hand side of equation (16) with the help of (16) we have the next equation

$$\varphi(s) = \frac{\varphi_0(s)}{\pi_0}\pi[I + (\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)R] + sY(s)\,, \tag{17}$$

where $Y(s) = s\varphi(s)[(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)\frac{1-h_k^n(s)}{s} - I)R]^2 + F(0)R\,[\mathrm{I}+(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)((1-h_k^n(s)) - sI)R]$. Substituting (17) into (15) we can express $\varphi_0(s)$ as

$$\varphi_0(s) = \pi_0\frac{A(s)}{B_1(s) + B_2(s)}, \tag{18}$$

Where $A(s) = s\varepsilon + sY(s)(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)e,$ $B_1(s) = -\pi(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)e$, $B_2(s) = -\pi(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)R(\sum_{k=1}^{K}\sum_{n=1}^{\infty} D_k(n)(1-h_k^n(s)) - sI)e$. Now, replace s by εs in

$A(s), B_1(s), B_2(s)$ and λ increases in such a way that $\varepsilon \to 0$, $\lambda \to \lambda^0 = (\pi \sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k^0(n) h_k e)^{-1} = (\rho^0)^{-1}$. Further more, we use the next limits $\lim_{\varepsilon \to 0} \frac{1 - h_k^n(\varepsilon s)}{\varepsilon s} = n h_k$, $\lim_{\varepsilon \to 0} \frac{h_k^n(\varepsilon s) - 1 + n h_k \varepsilon s}{(\varepsilon s)^2} = \frac{v_{k,n}}{2}$. Then

$$A(\varepsilon s) = \varepsilon^2 s + \varepsilon^2 s^2 Y(\varepsilon s)(\sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k(n) \frac{1 - h_k^n(\varepsilon s)}{\varepsilon s} - I)e,$$

$$B_1(\varepsilon s) = \varepsilon s - \pi \sum_{k=1}^{K} \sum_{n=1}^{\infty} n D_k(n) h_k e \varepsilon s + \pi \sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k(n)(h_k^n(\varepsilon s) + n h_k \varepsilon s - 1)e = \varepsilon^2 s + \pi \sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k(n)(h_k^n(\varepsilon s) + n h_k \varepsilon s - 1),$$

$$B_2(\varepsilon s) = -\pi \varepsilon^2 s^2 (\sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k(n) \frac{1 - h_k^n(\varepsilon s)}{\varepsilon s} - I) R (\sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k(n) \frac{1 - h_k^n(\varepsilon s)}{\varepsilon s} - I)e .$$

From (18) follows

$$\lim_{\varepsilon \to 0} \varphi_0(\varepsilon s) = \pi_0 \lim_{\varepsilon \to 0} \frac{A(\varepsilon s) / \varepsilon^2 s}{(B_1(\varepsilon s) + B_2(\varepsilon s)) / \varepsilon^2 s} = \pi_0 \frac{1}{1 + Ns} ,$$

where $N = \frac{1}{2\rho^0} \pi \sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k^0(n) v_{k,n} + \frac{-\pi}{\rho^0} (\sum_{k=1}^{K} \sum_{n=1}^{\infty} n D_k^0(n) h_k - \rho^0 I) R^0 \ (\sum_{k=1}^{K} \sum_{n=1}^{\infty} n D_k^0(n) h_k e - \rho^0 e)$. Using (6) we can transform the parameter N to the form

$$N = \frac{1}{2\rho^0} \pi \sum_{k=1}^{K} \sum_{n=1}^{\infty} D_k^0(n) v_{k,n} + \frac{\pi}{\rho^0} (\sum_{k=1}^{K} \sum_{n=1}^{\infty} n D_k^0(n) h_k - \rho^0 I) a.$$

Finally, from (16) and (17) we get that there exists

$$\lim_{\varepsilon \to 0} \varphi(\varepsilon s) = \lim_{\varepsilon \to 0} \frac{\varphi_0(\varepsilon s)}{\pi_0} \pi = \pi \frac{1}{1 + Ns}.$$

Thus the theorem is proved.

4. Conclusion

We studied the virtual waiting time of a queueing model introduced in [1]. We used the analytical approach introduced in [6] and derived that the scaled virtual waiting time converges to an exponentialy distributed random variable in a heavy traffic. This limit random variable is independent of the state of the environment.

REFERENCES

1. Massuyama, H. & Takine, T. (2003) Analysis of computation of the joint queue length distribution in a FIFO single-server queue with multiple batch Markovian arrival streams. *Stochastic Models*. 19. pp. 349–381.
2. Lucantoni, D.M., Meier-Hellstern, K.S. & Neuts, M.F. (1990) A single server queue with server vacations and a class of non-renewal arrival process. *Advances in Applied Probability*. 22. pp. 676–705. DOI: 10.2307/1427464
3. Lucantoni, D.M. (1991) New results on the single server queue with a batch Markovian arrival process. *Stochastic Models.* 7. pp. 1–46. DOI: 10.1080/15326349108807174
4. He, Q.M. (1996) Queues with Marked customers. *Advances in Applied Probability*. 28. pp. 567–587. DOI: 10.1017/S000186780004862X
5. Takine, T. & Hasegawa, T. (1994) The workload in the $MAP|G|1$ queue with state dependent services: Its applications to a queue with preemptive resume priority. *Stochastic Models.* 10. pp. 183–204. DOI: 10.1080/15326349408807292
6. Falin, G.I. & Falin, A.I. (1999) Heavy traffic analysis of $M|G|1$ type queueing systems with Markov modulated arrivals. *TOP*. 7(2). pp. 279–291. DOI: 10.1007/BF02564727

Received: December 21, 2018

Dimitrov M.Tz. (2019) HEAVY TRAFFIC ANALYSIS OF A QUEUE WITH BATCH MMAP. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitelnaja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 47. pp. 24–29

DOI: 10.17223/19988605/47/3

Димитров М.Ц. АСИМПТОТИЧЕСКИЙ АНАЛИЗ ОДНОЛИНЕЙНОЙ СИСТЕМЫ ОБСЛУЖИВАНИЯ С ВХОДЯЩИМ МАРКОВСКИМ МАРКИРОВАНЫМ ПОТОКОМ. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика.* 2019. № 47. С. 24–29

В работе рассматривается однолинейная система массового облуживания с дисциплиной обслуживания в порядке поступления, на вход которой поступает групповой марковский маркированый поток. Времена обслуживания требований зависят от номера поступающего потока. Исследован двумерный случайный процес, первая компонента которого является номером состояния сопроводящей марковской цепи, вторая – виртуальное время ожидания. Для его исследования предложен метод асимптотического анализа при условии высокой интенсивности входящего потока. Показано, что распределение состояния марковской цепи и виртуального времени ожидания независимы в стационарном режиме. Получено асимптотическое распределение виртуального времени ожидания в стационарном режиме в условиях большой нагрузки.

Ключевые слова: однолинейная система обслуживания; FIFO; марковский маркированый входящий поток; виртуальное время ожидания; высокая интенсивность входящего потока.

DIMITROV Mitko Tzanov (Candidate of Physics and Mathematics, Associate Professor, University of National and World Economy, Sofia, Bulgaria).
E-mail: mcdimitrov@abv.bg