

On the estimation in multi-server multi-core open queueing networks

SAULIUS MINKEVIČIUS

Institute of Data Science and Digital Technologies, Vilnius University
Akademijos St. 4, Vilnius, LT-04812
LITHUANIA

Abstract: The paper is devoted to the analysis of queueing systems in the context of the network and communications theory. We investigate the estimation in a multi-server multi-core open queueing networks and its applications to the theorems in heavy traffic conditions (fluid approximation, functional limit theorem, and law of the iterated logarithm) for a queue of jobs in a multi-server multi-core open queueing networks..

Key-Words: models of information systems, queueing theory, a multi-server multi-core open queueing networks, heavy traffic, queue length of jobs, fluid approximation, functional limit theorem, law of the iterated logarithm.

Received: January 21, 2020, Revised: August 12, 2020. Accepted: August 26, 2020. Published: September 8, 2020

1 Statement of the problem

The paper is devoted to the analysis of queueing systems in the context of the network and communications theory. We investigate the estimation in a multi-server multi-core open queueing networks and its applications to the theorems in heavy traffic conditions (fluid approximation, functional limit theorem, and law of the iterated logarithm) for a queue of jobs in a multi-server multi-core open queueing networks.

Now we shall survey the papers on a queue in heavy traffic conditions. In the paper of Chen, Xinyang and Yao [3], a semi-martingale reflecting the Brownian motion approximation is developed for the performance processes such as workload, queue, and sojourn time. In the paper of Massey and Srinivasan [7], the steady-state distribution of the queue process, using a tensor and Kronecker products, shows that it is of the matrix-geometric structure. J. Dai and W. Dai in [4] have proved that an appropriately normalized queue process converges in distribution to a d -dimensional mapping reflecting the Brownian motion under the heavy traffic conditions. Puhalskii [9] established moderate-deviation principles for the queue, virtual waiting time and sojourn processes. In [5], Yamada has showed that the normalized queue processes at the nodes converge in distribution to a reflected, multivariate diffusion process whose drift and diffusion coefficients are state dependent and nonsingular. In the article of Kushner and Martins [6], the authors study the problems of the pathwise average cost per unit time for controlled and uncontrolled open queueing networks in heavy traffic. In the paper of

Zhang Hanqin and Xu Guang-hui [13] strong approximations for an open queueing network in heavy traffic are proved. Peterson [8] has proved that, under heavy traffic conditions, the vector processes of total unfinished workloads converge to a multidimensional regulated Brownian motion. In the article of Reiman and Simon [11], the authors consider an open queueing network with multiple classes, priorities, "arbitrary" routing, and general service time distribution. Using a heavy traffic limit theorem for open queueing networks, Reiman [10] found the correct diffusion approximation for sojourn times in Jackson networks with a single-server station. As one can see, there are only several works designed to explore a queue in a more complicated than the classical single-server queue: tandem, multiphase queue, open queueing network (see the articles of Boxma [1, 2], Zhang Hanqin and Xu Guang-hui [13], Massey and Srinivasan [7], and Sakalauskas and Minkevičius [12]).

In this paper, we investigate a multi-server multi-core open queueing networks model in heavy traffic.

2 The network model

Let us consider a network of j stations, indexed by $j = 1, 2, \dots, J$ and the station j with c_j servers, where the server l has m core elements $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$. Next, $\{v_{jlm}, l \geq 1\}$, $j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ are sequences of service times, where $v_{jlm}(l) \geq 0$ is the service time for the l th job served in an m core element of the l th server of the node j of the multi-server multi-core open

queueing network. Denote $v_{jlm}(0) = 0$, $v_{jlm}(n) = \sum_{l=1}^n v_{jlm}(l)$, $n \geq 1$ and $x_{jlm}(t) = \sup\{n \geq 0 : v_{jlm}(n) \leq t\}$, where $x_{jlm} = \{x_{jlm}(t), t \geq 0\}$ is called a service process of the m core element of the l th server of the node j of the multi-server multi-core open queueing network, i.e. $x_{jlm}(t)$ counts the number of services completed during the server's busy time. We define $\mu_{jlm} = (E[v_{jlm}(l)])^{-1} > 0$, $\sigma_{jlm}^2 = D(v_{jlm}(l)) > 0$, $j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ and $\lambda_j = (E[u_j(l)])^{-1} > 0$, $a_j = D(u_j(l)) > 0$, $j = 1, 2, \dots, J$ with all of these terms assumed finite. In addition, let $\tilde{\tau}(t)$ be the total number of jobs, routed to the j th station of the network in the interval $[0, t]$, $\tau_j(t)$ be the total number of jobs after service departure from the j th station of the network in the interval $[0, t]$, and $\tilde{\tau}_{jlm}(t)$ be the total number of jobs routed to the m core element of the l th server of the node j of the network in the interval $[0, t]$. Next, let $\tau_{jlm}(t)$ be the total number of customers after service departure from the m core element of the l th server of the node j of the network in the interval $[0, t]$, and $\tau_{ijlm}(t)$ be the total number of jobs after service departure from the node i of the network and routed to the m core element of the l th server of the node j of the network in the interval $[0, t]$. In addition, let p_{ij} be a probability of the job after service at the i th station of the network, routed to the j th station of the network. Denote $p_{ijlm}^t = \frac{\tau_{ijlm}(t)}{\tau_{jlm}(t)}$ as a part of the total number of jobs which after service at the node i of the network are routed to the m core element of the l th server of the node j of the network in the interval $[0, t]$, $j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ and $t \geq 0$. Now we introduce the following process $Q_{jlm} = \{Q_{jlm}(t), t \geq 0\}$, where $Q_{jlm}(t)$ indicates the number of jobs waiting to be served by the m core element of the l th server of the j node of the network in the interval $[0, t]$, $j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ and $t \geq 0$. Assum, the arrival streams of jobs in the m core element of the l th server of the j node of the network be equal to $\frac{A_j(t)}{c_j \cdot c_{jl}}$, $j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ and $t \geq 0$.

3 Estimation of the queue length of jobs in a multi-server multi-core open queueing networks

Let us denote

$$w_{jlm}(t) = (x_{jlm}(t) - \tau_{jlm}(t)), \bar{w}_{jlm}(t) =$$

$$\sum_{i=1}^J x_{ijlm}(t) \cdot |p_{ij} - p_{ijlm}^t|, \hat{x}_{jlm}(t) = \frac{A_j(t)}{c_j \cdot c_{jl}} - x_{jlm}(t) \cdot p_j, p_j = 1 - \sum_{i=1}^J p_{ij}, p_{ijlm}^t = 1 - \sum_{i=1}^J p_{ijlm}^t, \hat{\beta}_{jlm} = \frac{\lambda_j}{c_j \cdot c_{jl}} - \mu_{jlm} \cdot p_j, \hat{\sigma}_{jlm}^2 = \frac{\lambda_j^3}{\sigma_j^2 \cdot \sigma_{jl}^2} + \mu_{jlm}^2 \cdot \sigma_{jlm} \cdot p_{ij}^2 > 0, \hat{p}_{ij} = \frac{1}{c_i} \cdot \frac{1}{c_{jl}} \cdot p_{ij}, i, j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl} \text{ and } t > 0.$$

We assume that the following conditions are fulfilled:

$$\frac{\lambda_j}{c_j \cdot c_{jl}} > \mu_{jlm} \cdot p_j, \quad (1)$$

$i, j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$. At first, we present the main estimation for the queue length of jobs in a multi-server multi-core open queueing network.

Lemma 3.1. *If $Q_{jlm}(0) = 0$, for $j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ and $t > 0$, then*

$$|Q_{jlm}(t) - \hat{x}_{jlm}(t)| \leq w_{jlm}(t) + \bar{w}_{jlm}(t).$$

Proof. By definition of the queue of jobs we obtained

$$\begin{aligned} Q_{jlm}(t) &= \tilde{\tau}_{jlm}(t) - \tau_{jlm}(t) = \sum_{i=1}^J \tau_{ijlm}(t) \\ &+ \frac{A_j(t)}{c_j \cdot c_{jl}} - \tau_{jlm}(t) = \sum_{i=1}^J \tau_{ijlm}(t) \cdot \frac{\tau_{ijlm}(t)}{\tau_{jlm}(t)} \\ &- \tau_{jlm}(t) + \frac{A_j(t)}{c_j \cdot c_{jl}} = \sum_{i=1}^J \tau_{jlm}(t) \cdot p_{ijlm}^t \\ &+ \frac{A_j(t)}{c_j \cdot c_{jl}} = \frac{A_j(t)}{c_j \cdot c_{jl}} - \tau_{jlm}(t) \cdot \left(1 - \sum_{i=1}^J p_{ijlm}^t\right) \\ &= \frac{A_j(t)}{c_j \cdot c_{jl}} - \tau_{jlm}(t) \cdot p_{jlm}^t. \end{aligned} \quad (2)$$

Hence it follows that

$$Q_{jlm}(t) = \frac{A_j(t)}{c_j \cdot c_{jl}} - \tau_{jlm}(t) \cdot p_{jlm}^t, \quad (3)$$

$j = 1, 2, \dots, J$, $l = 1, 2, \dots, c_j$, $m = 1, 2, \dots, c_{jl}$ and $t > 0$.

Besides, note that (see (3))

$$\begin{aligned}
 Q_{jlm}(t) &= \frac{A_j(t)}{c_j \cdot c_{jl}} - x_{jlm}(t) \cdot p_{jlm}^t \\
 &+ (x_{jlm}(t) - \tau_{jlm}(t)) \cdot p_{jlm}^t \leq \frac{A_j(t)}{c_j \cdot c_{jl}} \\
 &- x_{jlm}(t) \cdot p_{jlm}^t + x_{jlm}(t) - \tau_{jlm}(t) \cdot p_{jlm}^t \\
 &= \frac{A_j(t)}{c_j \cdot c_{jl}} - x_{jlm}(t) \cdot p_{jlm}^t + w_{jlm}(t) \\
 &= \frac{A_j(t)}{c_j \cdot c_{jl}} - x_{jlm}(t) \cdot p_j + x_{jlm}(t) \\
 &\cdot (p_j - p_{jlm}^t) + w_{jlm}(t) \leq \hat{x}_{jlm}(t) + w_{jlm}(t) \\
 &+ \sum_{i=1}^J x_{jlm}(t) \cdot |p_j - p_{ijlm}^t| = \hat{x}_{jlm}(t) \\
 &+ w_{jlm}(t) + \bar{w}_{jlm}(t), \quad j = 1, 2, \dots, J,
 \end{aligned} \tag{4}$$

$l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl}$ and $t > 0$.
So, we obtain

$$Q_{jlm}(t) \leq \hat{x}_{jlm}(t) + w_{jlm}(t) + \bar{w}_{jlm}(t), \tag{5}$$

$j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl}$
and $t > 0$.

Applying (3) we get that

$$\begin{aligned}
 Q_{jlm}(t) &\geq \frac{A_j(t)}{c_j \cdot c_{jl}} - x_{jlm}(t) \cdot p_{jlm}^t = \frac{A_j(t)}{c_j \cdot c_{jl}} \\
 &- x_{jlm}(t) \cdot p_j + x_{jlm}(t) \cdot (p_j - p_{jlm}^t) \\
 &\geq \hat{x}_{jlm}(t) - x_{jlm}(t) \cdot |p_j - p_{jlm}^t| \\
 &\geq \hat{x}_{jlm}(t) - \sum_{i=1}^J x_{jlm}(t) \cdot |p_j - p_{ijlm}^t| \\
 &= \hat{x}_{jlm}(t) - \bar{w}_{jlm}(t) \geq \hat{x}_{jlm}(t) - w_{jlm}(t) \\
 &- \bar{w}_{jlm}(t), \quad j = 1, 2, \dots, J, l = 1, 2, \dots, c_j,
 \end{aligned} \tag{6}$$

$m = 1, 2, \dots, c_{jl}$ and $t > 0$.

From (6) it follows that

$$Q_{jlm}(t) \geq \hat{x}_{jlm}(t) - w_{jlm}(t) - \bar{w}_{jlm}(t), \tag{7}$$

$j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl}$
and $t > 0$.

By combining (5) and (7) we can write

$$|Q_{jlm}(t) - \hat{x}_{jlm}(t)| \leq w_{jlm}(t) + \bar{w}_{jlm}(t), \tag{8}$$

$j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl}$
and $t > 0$.

The proof of the lemma is complete. \square

4 On applications of the main results

At first we present a Fluid Limit Theorem for the queue length of jobs in nodes of a multi-server multi-core open queueing network under overload heavy traffic conditions.

Theorem 4.1. *If conditions (1) are satisfied, then*

$$\left(\frac{Q_{1lm}(t)}{t}, \frac{Q_{2lm}(t)}{t}, \dots, \frac{Q_{jlm}(t)}{t} \right)$$

$$\xrightarrow[t \rightarrow \infty]{\Rightarrow} (\beta_{1lm}, \beta_{2lm}, \dots, \beta_{jlm}),$$

$j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl}$.

Next, we present a Functional Limit Theorem for the queue length of jobs in a multi-server multi-core open queueing network under overload heavy traffic conditions.

Theorem 4.2. *If conditions (1) are satisfied, then*

$$\left(\frac{Q_{1lm}(t) - \beta_{1lm} \cdot n \cdot t}{\sigma_{1lm} \cdot \sqrt{n}}; \frac{Q_{2lm}(t) - \beta_{2lm} \cdot n \cdot t}{\sigma_{2lm} \cdot \sqrt{n}}; \right.$$

$$\left. \dots; \frac{Q_{jlm}(t) - \beta_{jlm} \cdot n \cdot t}{\sigma_{jlm} \cdot \sqrt{n}} \right) \xrightarrow[n \rightarrow \infty]{\Rightarrow}$$

$$(z_{1lm}(t); z_{2lm}(t); \dots; z_{jlm}(t)), \text{ where}$$

$z_{jlm}(t), j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl}, 0 \leq t \leq 1$ are independent standard Wiener processes.

Finally, we present a Law of the Iterated Logarithm for the queue length of jobs in multi-server multi-core open queueing networks.

Theorem 4.3. *If conditions (1) are fulfilled, then*

$$P \left(\lim_{t \rightarrow \infty} \frac{Q_{jlm}(nt) - \hat{\beta}_{jlm} \cdot t}{\hat{\sigma}_{jlm} \cdot a(t)} = 1 \right)$$

$$= P \left(\lim_{t \rightarrow \infty} \frac{Q_{jlm}(nt) - \hat{\beta}_{jlm} \cdot t}{\hat{\sigma}_{jlm} \cdot a(t)} = -1 \right) = 1,$$

$j = 1, 2, \dots, J, l = 1, 2, \dots, c_j, m = 1, 2, \dots, c_{jl},$

$t > 0, a(t) = \sqrt{2t \ln \ln t}.$

Proof. The proof of Theorem 4.1, Theorem 4.2 and Theorem 4.3 is connected with the proof of Lemma 3.1, and we omit it. \square

References:

- [1] **O. Boxma.** On a tandem queueing model with identical service times at both counters. I. *Advances in Applied Probability.* 1979, 11(3), 616-643.
- [2] **O. Boxma.** On a tandem queueing model with identical service times at both counters. II. *Advances in Applied Probability.* 1979, 11(3), 644-659.
- [3] **H. Chen, X. Shen and D. Yao.** Brownian approximations of multiclass open-queueing networks. *Operations Research.* 2002, 50(6), 1032-1049.
- [4] **J. G. Dai and W. Dai.** A heavy traffic limit theorem for a class of open queueing networks with finite buffers. *Queueing Systems.* 1999, 32(1-3), 5-40.
- [5] **K. Yamada.** Diffusion approximation for open state-dependent queueing networks in the heavy traffic situation. *Annals of Applied Probability.* 1995, 5(4), 958-982.
- [6] **H. J. Kushner and L. F. Martins.** Limit theorems for pathwise average cost per unit time problems for controlled queues in heavy traffic. *Stochastics Reports.* 1993, 42(1), 25-51.
- [7] **W. A. Massey and R. Srinivasan.** Steady state analysis with heavy traffic limits for semi-open networks. *Stochastic models* (Ottawa, ON, 1998), 331-352, CMS Conf. Proc., 26, Amer. Math. Soc., Providence, RI, 2000.
- [8] **W. P. Peterson.** A heavy traffic limit theorem for networks of queues with multiple customer types. *Mathematics of Operations Research.* 1991, 16(1), 90-118.
- [9] **A. A. Puhalskii.** Moderate deviations for queues in critical loading. *Queueing Systems.* 1999, 31(3-4), 359-392.
- [10] **M. I. Reiman.** The heavy traffic diffusion approximation for sojourn times in Jackson networks. *Applied probability – computer science: the interface.* Vol. II, 3, 409-423, Birkhauser, Boston, MA, 1982.
- [11] **M. I. Reiman and B. Simon.** A network of priority queues in heavy traffic: one bottleneck station. *Queueing Systems.* 1990, 6(1), 33-58.
- [12] **L. L. Sakalauskas and S. Minkevičius.** On the law of the iterated logarithm in open queueing networks. *European Journal of Operational Research.* 2000, 120, 632-640.
- [13] **Zhang Hanqin and Xu Guang-hui.** Strong approximations for the open queueing network in heavy traffic. *Sci. China Ser. A.* 1992, 35(5), 521-535.

**Creative Commons Attribution License 4.0
(Attribution 4.0 International , CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US