On Stationary Regime for Queueing Networks with Controlled Input

EUGENE LEBEDEV, HANNA LIVINSKA
Applied Statistics Department
Taras Shevchenko National University of Kyiv
Volodymyrska str. 64, Kyiv, 01601
UKRAINE
stat@unicyb.kiev.ua, livinskaav@gmail.com

Abstract: - In this paper stationary properties of queueing network with exponentially distributed service times are investigated provided that an input flow is controlled by a Markov chain. We consider two cases. In the case, where we have one node only, a generating function of a stationary distribution is obtained. The form of the generating function is a matrix version of the Takacs formula. In the second case, a network with $r>1$ service nodes is considered. For a multivariate service process the condition of a stationary regime existence and a correlation matrix are found.

Key-Words: - Multi-channel queueing networks, controlled input, stationary regime

1 Introduction

There is a great potential for using queueing network models in the performance analysis of computer and communication networks. Their structure is determined by the probabilistic characteristics of input flows and data processing algorithms. It is necessary to find their characteristics to optimize them, and to develop the corresponding control algorithms. This is particularly important at the stage of network design since it is necessary to obtain estimates of basic network performance measures.

For many mathematical models describing probabilistic structure of real systems, there has been obtained using queueing theory large number of analytical and numerical results characterizing the performance of these systems. But classical queueing models do not assume any external intervention in a working process.

In reality there are many systems with the primary feature of their control capability during operation, since in this case a s ignificant improvement in characteristics can be achieved, for example, reducing the queues length, increasing throughput or decreasing the operational expenses. The criteria of quality control, depending on t he objectives of the study, are the various characteristics of the queueing models: system performance, the average number of customers in the system, the average queue length, the average sojourn time of customers in the system, the average waiting time, the probability of loss, a system load factor or the average idle time of devices etc., as well as various economic indicators related to these characteristics.

Development and research of analysis techniques for queueing networks with controlled input flow are actual directions of development of the queueing networks theory. In particular, special class of stochastic processes, named switching processes, was introduced by V.V. Anisimov ([1], [2], [3]). This class is a suitable tool for modelling of the control procedure of local model characteristics. Another control technique is connected with MAP (Markovian arrival process) and BMAP (batch Markovian arrival process) inputs (see, for example, [6], [15], [16], etc.). Models with a controlled input flow were studied also in [4], [8], [10], [11], [12], [17], [18], etc. Problems of analysis of queueing networks with input rates, dependent on time or on the state of the model are of considerable interest. Works [1], [2], [3], [13], [14], etc. are devoted to elaboration and development of methods for studying of this class of queueing systems and networks.

The main model we consider is a queueing network consisting of r service nodes. Each network node is a queueing system and it consists of infinite number of servers. Therefore if a customer arrives at such a system, then it begins processing immediately. Input flow arriving at the network is controlled by a Markov process. We define a service process in the network as an r-dimensional stochastic process $Q(t)=\left(Q_1(t),...,Q_r(t)\right)'$, $t\geq 0$, where $Q_i(t)$, $i=1,2,...,r$, is the number of customers at the i-th node at time t, symbol $(\cdot)'$ stands for a transposition operator.

In this paper we study a stationary regime for such multi-channel networks in two cases. In Section 2 we consider one-dimensional case, where the network has the only service node. In this case a generating function of the stationary distribution for the process $Q(t)$, $t \geq 0$, is obtained. In Section 3 the network with $r > 1$ service nodes is studied. Formulas for the first and second moments of service process in stationary regime are found.

2 Case of Controlled System with One Service Node

Firstly, let us consider the case of a single node.

Let $\eta(t)$ be a homogeneous continuous time Markov chain with a finite set of states $S = \{1,2,...,r\}$ and $A = \left\|\lambda_{ij}\right\|_1^r$ be its infinitesimal matrix. It is assumed that the instants of customers arrivals to the system are the same as jump instants t_n, $n = 1,2,...$, of $\eta(t)$. A customer arrived to the system immediately begins to be served anywhere on a free server. The service time is distributed exponentially with parameter μ. In accordance with notations introduced above, let us denote the number of busy servers in the system at instant t by $Q(t)$. For the process $Q(t)$, $t \geq 0$, the following result takes place.

Theorem 1. *If the controlling Markov chain $\eta(t)$ is ergodic, then for any $i \in S$ and $j = 1,2,...$*

$$\lim_{t\to\infty} E\left\{z^{Q(t)} / \eta(0) = i, Q(0) = j\right\} = 1 + \sum_{n=1}^{\infty} (z-1)^n \pi' \times$$
$$\times \prod_{k=1}^{n} \left[\left(\Delta(\lambda) + A\right)\left(k\mu I - A\right)^{-1} \right] \bar{1}, \; |z| \leq 1, \quad (1)$$

where $\pi' = (\pi_1, ..., \pi_r)$ is an ergodic distribution for $\eta(t)$; $\Delta(\lambda) = \left\|\lambda_i \delta_{ij}\right\|_1^r$ is a diagonal matrix (δ_{ij} is the Kronecker delta), $I = \left\|\delta_{ij}\right\|_1^r$ is an identity matrix, $\lambda' = (\lambda_1, ..., \lambda_r)$, $\lambda_i = \sum_{j\neq i} \lambda_{ij}$, $i = 1,2,...,r$, $\bar{1}$ is an r-dimensional vector-column with unit entries.

Proof. Let

$$\Phi_i(t,z) = E\left\{z^{Q(t)} / \eta(0) = i, Q(0) = 0\right\}, \quad (2)$$

$i = 1,2,...,r$, $|z| \leq 1$, be a generating function of the customers number in the system at time $t \geq 0$ if at the initial instant the system is empty and the controlling chain is in the state i. Then

$$E\left\{z^{Q(t)} / \eta(0) = i, Q(0) = j\right\} =$$
$$= \left[1 - (1-z)e^{-\mu t}\right]^j \Phi_i(t,z)$$

and

$$\lim_{t\to\infty} E\left\{z^{Q(t)} / \eta(0) = i, Q(0) = j\right\} = \lim_{t\to\infty} \Phi_i(t,z). \quad (3)$$

Since the functions $\Phi_i(t,z)$, $i = 1,2,...,r$, satisfy the equation system

$$\Phi_i(t,z) = e^{-\lambda_i t} +$$
$$+ \sum_{j\neq i} \int_0^t e^{-\lambda_i u} \lambda_{ij} \Phi_j(t-u,z) \left[1 - (1-z)e^{-\mu(t-u)}\right] du, \quad (4)$$

the existence of limits (3) follows from Markov renewal theory ([9], sect. 12.6).

Let us introduce

$$\varphi_i(s,z) = \int_0^\infty e^{-st} \Phi_i(t,z)\, dt, \; \mathrm{Re}(s) > 0, \quad (5)$$

Laplace transform for the functions $\Phi_i(t,z)$, $t \geq 0$, $i = 1,2,...,r$. Then the system (4) can be written as

$$(\lambda_i + s)\varphi_i(s,z) = 1 +$$
$$+ \sum_{j\neq i} \lambda_{ij} \varphi_j(s,z) + (z-1) \sum_{j\neq i} \lambda_{ij} \varphi_j(s+\mu, z),$$

or

$$(sI - A)\varphi(s,z) = \bar{1} +$$
$$+ (z-1)\left(\Delta(\lambda) + A\right)\varphi(s+\mu, z), \quad (6)$$

where $\varphi'(s,z) = \left(\varphi_1(s,z), ..., \varphi_r(s,z)\right)$. For $\mathrm{Re}(s) > 0$ we find from (6):

$$\varphi(s,z) = (sI - A)^{-1} \times$$
$$\times \left\{\bar{1} + (z-1)\left(\Delta(\lambda) + A\right)\varphi(s+\mu, z)\right\}. \quad (7)$$

Substituting consecutively $s + k\mu$, $k = 1,2,...,n-1$, instead of s in (7), we obtain:

$$\varphi(s,z) = (sI - A)^{-1} \times$$
$$\times \left\{\bar{1} + (z-1)\left(\Delta(\lambda) + A\right)\left((\mu+s)I - A\right)^{-1}\bar{1} + ... + (z-1)^{n-1} \times\right.$$
$$\times \prod_{k=1}^{n-1} \left[\left(\Delta(\lambda) + A\right)\left((k\mu+s)I - A\right)^{-1}\right]\bar{1} + (z-1)^{n-1} \times$$
$$\times \prod_{k=1}^{n-1} \left[\left(\Delta(\lambda) + A\right)\left((k\mu+s)I - A\right)^{-1}\right] \times$$
$$\left. \times \left(\Delta(\lambda) + A\right)\varphi(s+n\mu, z)\right\} \quad (8)$$

for $n = 1,2,...$.

It is not difficult to show that for any $|z| \leq 1$

$$\sum_{n=0}^{\infty} |z-1|^n \left\| \prod_{k=1}^{n-1} \left[\left(\Delta(\lambda) + A\right)\left((k\mu+s)I - A\right)^{-1}\right] \right\| < \infty,$$

where $\|\cdot\|$ is Euclidean matrix norm. Therefore, based on [7] (sect. 5.6) we find from (8):

$$\varphi(s,z) = (sI - A)^{-1} \times \tag{9}$$
$$\times\left\{ I + \sum_{n=0}^{\infty}(z-1)^n \prod_{k=1}^{n}\left[(\Delta(\lambda)+A)((k\mu+s)I - A)^{-1} \right] \right\}\bar{1}.$$

Since the controlling chain $\eta(t)$ is ergodic, then

$$\lim_{s\to 0} s(sI - A)^{-1} = \Pi, \tag{10}$$

where Π is a matrix whose rows coincide with the ergodic distribution $\pi = (\pi_1,...,\pi_r)$.

From (9), (10) we find:

$$\lim_{s\to 0} s\varphi(s,z) = \bar{1} +$$
$$+\sum_{n=0}^{\infty}(z-1)^n \Pi \prod_{k=1}^{n}\left[(\Delta(\lambda)+A)(k\mu I - A)^{-1} \right]\bar{1}.$$

The theorem is proved.

Equation (1) is a matrix version of the Takacs formula ([20], p. 159). Notice also, that formula (9) contains the Laplace transform of the probability characteristics for the service process in the transient regime.

3 Case of Network with Many Service Nodes

Let us now consider a queueing network consisting of $r > 1$ nodes. At servicing nodes a common input flow of customers arrives. This flow is controlled by a Markov chain $\eta(t)$ according to the following algorithm. As before, the instants of customers arrivals are the same as jump moments t_n, $n = 1,2,...$, of the chain $\eta(t)$. If the chain $\eta(t)$ jumps into state i at the instant t_n, the customer numbered 1 arrives for service into the i-th node. There it occupies a free server for the time distributed exponentially with parameter μ_i. After service in the i-th node the customer arrives to the j-th node with probability p_{ij}, $j = 1,2,...,r$, and leaves the network with probability $1 - \sum_{j=1}^{r} p_{ij}$. $P = \|p_{ij}\|_1^r$ is a switching matrix of the network.

As before, denote the number of occupied servers in the i-th node at instant $t \ge 0$ by $Q_i(t)$, $i = 1,2,...,r$, $Q(t)' = (Q_1(t),...,Q_r(t))$. If there exists a stationary regime for the queueing network, we will denote the number of occupied servers in the i-th node in the stationary regime by Q_i, $i = 1,2,...,r$, $Q' = (Q_1,...,Q_r)$.

Let $a' = (a_1,...,a_r) = (\lambda_1\pi_1,...,\lambda_r\pi_r)$, a_i^{-1} be the mean time between two consecutive hits to the state i of the controlling chain $\eta(t)$ provided that it is ergodic.

For functions $A_{ij}(t) = E\{Q_j(t) / \eta(0) = i, Q_1(0) = 0,...,Q_r(0) = 0\}$, $i, j = 1,2,...,r$, we introduce their Laplace transforms:

$$a_{ij}(s) = \int_0^{\infty} e^{-st} A_{ij}(t)dt,\ \mathrm{Re}(s) > 0,\ A(s) = \|a_{ij}(s)\|_1^r.$$

The following result holds.

Theorem 2. *If the controlling Markov chain $\eta(t)$ is ergodic and the spectral radius of the switching matrix P is strictly less than 1, then a stationary regime for the service process $Q(t)$, $t \ge 0$, exists, and*

1) $EQ_i = \theta_i / \mu_i$, $i = 1,2,...,r$, *(11)*

$\theta' = (\theta_1,...,\theta_r)$ *is the solution of the balance equation* $\theta' = a' + \theta' P$*;*

2) $EQ_i(Q_j - \delta_{ij}) =$

$$= \sum_{m=1}^{r} a_m \int_0^{\infty}\left[A_{mi}(t)p_{mj}(t) + A_{mj}(t)p_{mi}(t) \right]dt, \tag{12}$$

where $P(t) = \|p_{ij}(t)\|_1^r = \exp(Bt)$, $B = \Delta(\mu)(P - I)$, $\mu' = (\mu_1,...,\mu_r)$, $\Delta(\mu) = \|\mu\delta_{ij}\|_1^r$.

The proof needs an auxiliary result.

Lemma 1. *The matrix* $A(s)$, $\mathrm{Re}(s) > 0$, *can be represented in the explicit form via the model parameters as follows:*

$$A(s) = (sI - A)^{-1}(\Delta(\lambda) + A)(sI - B)^{-1}. \tag{13}$$

Proof. Let us introduce generating functions in the same way as in (2):

$$\Phi_i(t, z_1,...,z_r) = \Phi_i(t,z) =$$
$$= E\{z_1^{Q_1(t)} \cdot ... \cdot z_r^{Q_r(t)} / \eta(0) = i, Q_1(0) = 0,...,Q_r(0) = 0\},$$

$t \ge 0$, $|z| \le 1$, $z = (z_1,...,z_r)$.

Functions $\Phi_i(t,z)$, $i = 1,2,...,r$, satisfy the following equation system of Markov renewal type:

$$\Phi_i(t,z) = e^{-\lambda_i t} + \tag{14}$$
$$+\sum_{m\neq i}\int_0^t e^{-\lambda_i u}\lambda_{im}\Phi_m(t-u,z)\left[1 - \sum_{j=1}^{r} p_{mj}(t-u)(1-z_j)\right]du,$$
$$i = 1,2,...,r.$$

From (14) we obtain Markov renewal equations for functions $A_{ij}(t)$, $i, j = 1,2,...,r$:

$$A_{ij}(t) = \sum_{m\neq i}\int_0^t e^{-\lambda_i u}\lambda_{im}\left[A_{mj}(t-u) + p_{mj}(t-u) \right]du. \tag{15}$$

System (15) can be solved in terms of Laplace transforms.

Let

$$\tilde{p}_{ij}(s)=\int_0^\infty e^{-st}p_{ij}(t)dt,\ \mathrm{Re}(s)\ge 0,$$

$$\widetilde{P}(s)=\left\|\tilde{p}_{ij}(s)\right\|_1^r,$$

be the Laplace transform for transient probabilities $p_{ij}(t)$ of Markov chain which is defined by the following infinitesimal $(r+1)\times(r+1)$ matrix:

$$\widetilde{B}=\begin{pmatrix} B & \begin{matrix}\mu_1 p_{1,r+1}\\ \vdots \\ \mu_r p_{r,r+1}\end{matrix} \\ 0 & \dots 0\end{pmatrix}.$$

The state with a number $r+1$ is an absorbing one.

Using Laplace transforms, system (15) can be represented in the form:

$$(sI-A)A(s)=(\Delta(\lambda)+A)\widetilde{P}(s). \qquad (16)$$

From the Kolmogorov differential equations for $p_{ij}(t)$ we obtain:

$$\widetilde{P}(s)=(sI-B)^{-1}.$$

Substitution of this formula in (16) gives (13).

Proof of Theorem 2.

Subsequently, point 1 of theorem 2 is conclusion of (13) and Tauberian theorems for Laplace transforms. Indeed, we can write

$$\lim_{s\downarrow 0} sA(s)=\begin{pmatrix} EQ_1 & EQ_2 & \dots & EQ_r\\ \dots & \dots & \dots & \dots \\ EQ_1 & EQ_2 & \dots & EQ_r\end{pmatrix}=$$

$$=\Pi(\Delta(\lambda)+A)(I-P)^{-1}\Delta^{-1}(\mu)=$$

$$=\Pi\Delta(\lambda)(I-P)^{-1}\Delta^{-1}(\mu)=$$

$$=\begin{pmatrix} \theta_1/\mu_1 & \theta_2/\mu_2 & \dots & \theta_r/\mu_r\\ \dots & \dots & \dots & \dots \\ \theta_1/\mu_1 & \theta_2/\mu_2 & \dots & \theta_r/\mu_r\end{pmatrix}.$$

We now proceed to prove point 2 of theorem 2. Let us denote second moments of the process $Q(t)'=(Q_1(t),...,Q_r(t))$ by $D_{ij}^{\ m}(t)$:

$$D_{ij}^{\ m}(t)=$$

$$=E\left\{Q_i(t)\left(Q_j(t)-\delta_{ij}\right)/\eta(0)=m, Q_1(0)=0,...,Q_r(0)=0\right\}.$$

From (14) it follows that the functions $D_{ij}^{\ m}(t)$, $m,i,j=1,2,...,r$, $t\ge 0$, satisfy the system of Markov renewal equations:

$$D_{ij}^{\ m}(t)=$$

$$=\sum_{k\ne m}\int_0^t e^{-\lambda_m u}\lambda_{mk}\left[D_{ij}^{\ k}(t-u)+A_{ki}(t-u)p_{kj}(t-u)+\right.$$

$$\left.+A_{kj}(t-u)p_{ki}(t-u)\right]du, \qquad (17)$$

$$m,i,j=1,2,...,r,\ t\ge 0.$$

In order to analyse system (17), we introduce the Laplace transforms of $D_{ij}^{\ m}(t)$:

$$\widetilde{D}_{ij}^{\ m}(s)=\int_0^\infty e^{-st}D_{ij}^{\ m}(t)dt,$$

$$d_{ij}^{\ m}(s)=\int_0^\infty e^{-st}A_{mi}(t)p_{mj}(t)dt,\ \mathrm{Re}(s)>0,$$

$$\widetilde{D}_{ij}(s)=\left(\widetilde{D}_{ij}^{\ 1}(s),...,\widetilde{D}_{ij}^{\ r}(s)\right)',$$

$$d_{ij}(s)=\left(d_{ij}^{\ 1}(s),...,d_{ij}^{\ r}(s)\right)',\ m,i,j=1,2,...,r.$$

Note that functions $A_{ij}(t)$ are defined by Laplace transforms (13), therefore $d_{ij}^{\ m}(s)$ are known.

Now, (17) can be written in a vector-matrix form in terms of Laplace transforms:

$$\Delta(\lambda+s)\widetilde{D}_{ij}(s)=$$

$$=(\Delta(\lambda)+A)\widetilde{D}_{ij}(s)+(\Delta(\lambda)+A)\left(d_{ij}(s)+d_{ji}(s)\right),$$

$$i,j=1,2,...,r.$$

Taking into account (10) and Tauberian theorems for Laplace transforms, we find:

$$\lim_{s\downarrow 0} s\widetilde{D}_{ij}(s)=\begin{pmatrix} EQ_i(Q_j-\delta_{ij})\\ \vdots \\ EQ_i(Q_j-\delta_{ij})\end{pmatrix}=$$

$$=\Pi(\Delta(\lambda)+A)\left(d_{ij}(0)+d_{ji}(0)\right)=$$

$$=\begin{pmatrix} \pi_1\lambda_1 & \pi_2\lambda_2 & \dots & \pi_r\lambda_r\\ \dots & \dots & \dots & \dots \\ \pi_1\lambda_1 & \pi_2\lambda_2 & \dots & \pi_r\lambda_r\end{pmatrix}\left(d_{ij}(0)+d_{ji}(0)\right).$$

The last expression is another form of (13). Theorem 2 is proved.

4 Conclusion

Comparison of (11), (12) with the results from [11] shows that the service process for the network with controlled input has similar limit properties to the service process in the $[GI\,|\,M\,|\,\infty]^r$-network. Recurrent input flow at the i-th node is formed by a sequence of return instants of controlling chain $\eta(t)$ into the i-th state.

The obtained results allow to solve modeling and optimization problem for complex systems such as mobile networks, distributed computer networks, automobile traffic flows, call-centers.

Note that in the case when the service rates μ_i depend on n (the number of series) and

$$\lim_{n\to\infty} n\mu_i(n) = \mu_i \neq 0,\ i = 1,2,...,r,$$

the networks operates in heavy traffic regime. With this condition it is possible to approximate the service process by a G aussian process with its characteristics written via the network parameters. Such an approximation of some multi-channel stochastic networks was considered, for example, in [13], [14]. In these papers for multi-channel queueing networks in heavy traffic regime we develop asymptotic methods of investigation which are based on the approximation of jump-type service processes of calls by continuous Gaussian processes.

References:

[1] Anisimov, V., Averaging in Markov Models with Fast Markov Switches and Applications to Queueing Models, *Annals of Operation Research*, Vol.112, No.1, 2002, pp. 63-82.

[2] Anisimov, V., *Switching Processes in Queueing Models*, Wiley-ISTE, 2013.

[3] Anisimov, V.V., Lebedev, E.A., *Stochastic Service networks. Markov models*, Kyiv–Lybid, 1992 (in Russian).

[4] Bäuerle, N., Optimal Control of Queueing Networks: An Approach via Fluid Models, *Advances in Applied Probability*, Vol.34, No.2, 2002, pp. 313-328.

[5] Bocharov, P.P., Pechinkin, A.V., *Queueing Theory,* RUDN-Moscow, 1995 (in Russian).

[6] Breuer, L., The Periodic BMAP/PH/c Queue, *Queueing Syst,* No.38(1), 2001, pp. 67-76.

[7] Horn, R.A., Johnson, Ch.R., *Matrix Analysis*, Cambr. Univ. Press, 1985.

[8] Korolyuk, V.S., Korolyuk, V.V., *Stochastic Models of Systems*, Kluwer Acad. Press, Dordrecht, 1999.

[9] Kovalenko, I.N., Kuznetsov, N.Yu., Shurenkov, V.M. *Stochastic Processes*, Kyiv, 1983 (in Russian).

[10] Kushner, H.J., *Heavy Traffic Analysis of Controlled Queueing and Communication Network*, Springen-Verlag, New-York, 2001.

[11] Lebedev, E.A., Service Networks with Multi-channel Nodes and Recurrent Input Flow, *Cybernetics and System Analysis*, No.4, 2001, pp. 133-141.

[12] Lebedev, E., Makushenko, I., On Profit Maximization and Risk Minimization in Networks of Semi-Markov Type, *Cybernetics and System Analysis*, No.2, 2007, pp. 65-79.

[13] Lebedev, E., Livinska, G., Gaussian Approximation of Multi-channel Networks in Heavy Traffic, *Communications in Computer and Information Science*, No.356, 2013, pp. 122-130.

[14] Livinska, H.V., Lebedev, E.O., Conditions of Gaussian Non-Markov Approximation for Multi-channel Networks, *Proceedings of the ECMS-2015*, 2015, pp. 642-649.

[15] Moiseev, A., Nazarov, A., Queueing Network MAP−(GI/∞)^K with High-rate Arrivals, *European Journal of Operational Research,* Vol.254, No.2., 2016, pp. 161-168.

[16] Nazarov, A.A., Moiseeva, S.P., *Method of Assimptotic Analysis in Queueing Theory*, NTL–Tomsk, 2006 (in Russian).

[17] Rykov, V.V., Controlled Queueing Systems, *Results of Science and Technology: Prob. Theory, Math. Stat., Theor. Cybernetics,* Vol.12, 1975, p. 43-153 (in Russian).

[18] Rykov, V.V., Efrosinin, D.V., Numerical Analysis of Optimal Control Policies for Queueing Systems with Heterogeneous Servers, *Queueing Systems,* Vol.46. 2004. pp. 389-407.

[19] Sean Meyn, *Control Techniques for Complex Networks*, Cambridge Univ. Press, 2008.

[20] Takasc, L., *Introduction to the Theory of Queues*, Oxford Univ. Press, 1962.