Asymptotic analysis of some controlled finite-source queueing systems

By V. V. Anisimov and J. Sztrik

KIEV STATE UNIVERSITY
FACULTY OF CYBERNETICS
VLADIMIRSKAJA 64.
KIEV—17, USSR

UNIVERSITY OF DEBRECEN
DEPARTMENT OF MATHEMATICS
DEBRECEN, PF. 12., 4010
HUNGARY

1. Introduction

Stochastic processes are powerful tools for the investigation of behavior of complex systems. Results of queueing theory are effectively used in solving problems encountered in practical applications. Different methods and approaches have been developed in order to that the involved models should be mathematically tractable. Since there exists an overwhelming body of literature devoted to the study of queues the interested reader is referred to, among others, Franken et. al. [4], Gnedenko and Kovalenko [5], Gnedenko [6], Gnedenko and König [7], Jaiswal [8], Kleinrock [9], Koroljuk and Turbin [11], Kovalenko [12], König and Stoyan [13], Lavenberg [14], Lifsic and Malc [15], Takács [16], White et. al. [17].

It is also well-known, that a great majority of problems can be treated by the help of Semi-Markov Processes (SMP). In many situations the distribution of time until the SMP gets out of a subset of its state space is of great practical importance. Recently, however, due to the rapid development of technical devices there are cases where the exit from a given subset is a "rare" event, that is, it occurs with a small probability. Thus, it is natural to investigate the asymptotic behaviour of sojourn time in a given subset, provided that the propability of exit from it tends to zero.

The purpose of the present paper is three-fold. On the one hand, without proofs we give a brief survey of preliminary results mainly due to Koroljuk and Anisimov. On the other hand, we deal with an asymptotic analysis of some controlled finite-source queueing systems under the assumption of "fast" service. We show that the time to the first system failure converges in distribution, under appropriate norming, to an exponentially distributed random variable. Finally, applications of these models in the field of reliability theory and computer performance are considered.

2. Preliminary results

Let us begin with the non-asymptotic case (see Koroljuk [11]).

(i) Let $(\xi(t),\ t\geqq 0)$ be a Semi-Markov Process with state space $\{0, 1, \dots, r\}$ given by the embedded Markov chain $(X_n,\ n\geqq 0)$ and by transition matrix $\|p(i,j)\|$, $i,j=\overline{0,r}$. Furthermore, let $\tau(i,j)$ be mutually independent random variables denoting the time spent in state i, given that the next state is j, $i,j=\overline{0,r}$.

Let $\Omega(k)$ denote the sojourn time of $\xi(t)$ in subset $\{1, \dots, r\}$ started in state k, that is

$$\Omega(k) = \inf\{t\colon t > 0,\ \xi(t) = 0/\xi(0) = k,\ k \neq 0\}.$$

For $\Omega(k)$ we have the following stochastic relations

$$\Omega(k) = \begin{cases} \tau(k,0) & \text{with probability (w.p.)} \quad p(k,0), \\ \tau(k,j)+\Omega(j) & \text{w.p.} \quad p(k,j), \quad j=\overline{1,r}. \end{cases} \tag{1}$$

Let us introduce some notations:

$$\varphi(u,k,j) = \mathbf{E}\exp\big(iu\tau(k,j)\big), \quad \psi(u,k) = \mathbf{E}\exp\big(iu\Omega(k)\big),$$

$$\underline{\varphi}(u) = \begin{pmatrix} p(1,0)\varphi(u,1,0) \\ \vdots \\ p(r,0)\varphi(u,r,0) \end{pmatrix}, \quad \underline{\psi}(u) = \begin{pmatrix} \psi(u,1) \\ \vdots \\ \psi(u,r) \end{pmatrix},$$

$$\Phi(u) = \|p(k,j)\varphi(u,k,j)\|, \quad k,j=\overline{1,r}.$$

When passing in (1) to characteristic functions we obtain

$$\underline{\psi}(u) = \Phi\underline{\psi}(u)+\underline{\varphi}(u). \tag{2}$$

Supposing that for any $j\in\{1, \dots, r\}$ there exists a sequence of transitions with positive probabilities leading to $\{0\}$, that is $\{1, \dots, r\}$ is not closed and does not contain any closed subset, from (2) we get

$$\underline{\psi}(u) = \big(E-\Phi(u)\big)^{-1}\underline{\varphi}(u).$$

In particular, for the mean sojourn times we have

$$\underline{M} = (E-P)^{-1}\underline{m},$$

where $\underline{M}$ and $\underline{m}$ are column vectors with components $\mathbf{E}\Omega(k)$, $m_k=\sum_{j=0}^{r} p(k,j)\mathbf{E}\tau(k,j)$ respectively, $k=\overline{1,r}$.

(ii) Suppose that $X_n=X_n(\varepsilon)$, $p(k,j)=p_\varepsilon(k,j)$ and $\tau(k,j)=\tau_\varepsilon(k,j)$, that is, $(\xi(t),\ t\geqq 0)$ depends on some small parameter ε, such that $p_\varepsilon(k,0)\to 0$ as $\varepsilon\to 0$. Therefore, it is natural to investigate the limit distribution of $\Omega(k)=\Omega_\varepsilon(k)$ as $\varepsilon\to 0$.

Assume that the following conditions are satisfied:

1. $p_\varepsilon(k,j)\to p_0(k,j)$, $k,j=\overline{1,r}$ and matrix $P_0=\|p_0(k,j)\|$, $k,j=\overline{1,r}$ corresponds to a single essential class.

2. $p_\varepsilon(k,0)/\varepsilon \to b_k < \infty$, $k=\overline{1,r}$, $\Sigma b_k \neq 0$.

3. There exists a normalizing factor β_ε such that

a) $\varphi_\varepsilon(\beta_\varepsilon u, k, j) = 1+\varepsilon a_{kj}(u)+o(\varepsilon), \quad k, j = \overline{1, r},$

b) $\varphi_\varepsilon(\beta_\varepsilon u, k, 0) \to \varrho_k(u), \quad k = \overline{1, r}.$

Theorem 1 (Koroljuk [10], [11]). If conditions (1)—(3) are satisfied, then independently of the initial state $j; j=\overline{1, r}$ the distribution of $\beta_\varepsilon \Omega_\varepsilon(j)$ converges weakly to a distribution with characteristic function

$$\Big(\sum_{k=1}^{r} \pi_k b_k \varrho_k(u)\Big)\Big/\Big(\sum_{i=1}^{r} \pi_i b_i - \sum_{i,j=1}^{r} \pi_i p_0(i,j) a_{ij}(u)\Big),$$

where $\{\pi_k, \ k=\overline{1, r}\}$ is the stationary distribution for the chain with matrix P_0.

Corollary 1. If the random variables $\tau_\varepsilon(k,j)$ do not depend on ε, that is $\tau_\varepsilon(k,j)=\tau_0(k,j), \ k,j=\overline{0, r}$ and $\mathbf{E}\tau_0(k,j)=m_{kj}<\infty$, furthermore conditions (1), (2) are satisfied, then for any $j, j=\overline{1, r}$ we have

$$\mathbf{P}\{\varepsilon\Omega_\varepsilon(j) < x\} \to 1-\exp\left\{-\frac{b}{m}x\right\}, \quad x > 0$$

where

$$m = \sum_{k=1}^{r} \pi_k p_0(k,j) m_{kj}, \quad b = \sum_{k=1}^{r} \pi_k b_k.$$

(iii) Sometimes, however, there are practical situations (for example systems with "fast" service, or highly reliable systems) when the set $\{1, ..., r\}$ in the limit may split into several essential classes and, possibly, inessential states. To assert the correspondig theorem we need the notion of s-set, introduced by Anisimov (see [1]).

Let $(X_\varepsilon(k), k \geqq 0)$ be a Markov chain with state space $\{0, 1, ..., r\}$ and let $\|p_\varepsilon(i,j)\|$ denote its transition matrix, $i,j=\overline{0, r}$.

Furthermore, let $\langle\alpha\rangle$ be a subset from $\{1, ..., r\}$. Set

$$V_\varepsilon(i, \langle\alpha\rangle) = \min\{k\colon \ k > 0, X_\varepsilon(k) \notin \langle\alpha\rangle / X_\varepsilon(0) = i \in \langle\alpha\rangle\},$$

$q_\varepsilon(i,j, \langle\alpha\rangle)=\mathbf{P}\{X_\varepsilon(l)=j,$ for at least one $l, \ l<V_\varepsilon(i, \langle\alpha\rangle\}$ i.e. $q_\varepsilon(i,j, \langle\alpha\rangle)$ is the probability of a visit to j up the time when the chain exits from $\langle\alpha\rangle$, given that the initial state was $i, \ i,j \in \langle\alpha\rangle$.

Definition. A set of states

$$\langle\alpha\rangle = \{i_1, ..., i_l\}$$

is called an s-set (communicating set) if for any $i,j \in \langle\alpha\rangle \ q_\varepsilon(i,j, \langle\alpha\rangle) \to 1$ as $\varepsilon \to 0$.

Practically, it means that initiated from any state the chain visits each state asymptotically infinitely many times before leaving. (The simplest example for an s-set is a set which in the limit forms a simple essential class.)

Let $(\xi_\varepsilon(t), t \geqq 0)$ be a SMP with state space $\{0, 1, ..., r\}$ given by the embedded Markov chain $(X_\varepsilon(n), \ n \geqq 0)$, the transition matrix $\|p_\varepsilon(k,j)\|, \ k,j=\overline{0, r}$ and the random variables $\tau_\varepsilon(k,j)$. Assume that the subset $\{1, ..., r\}$ forms an s-set.

Set

$$g_\varepsilon = \sum_{k=1}^{r} \pi_k p_\varepsilon(k, 0)$$

where $\{\pi_k,\ k=\overline{1, r}\}$ is the stationary distribution for the chain with transition matrix

$$\|p_\varepsilon(i, j)/(1-p_\varepsilon(i, 0)),\quad i, j = \overline{1, r}.$$

Furthermore, suppose that

$$\pi_\varepsilon(k) p_\varepsilon(k, 0)/g_\varepsilon \to b_k,\quad k = \overline{1, r},$$

and there exists a normalizing factor β_ε such that

a) $\varphi_\varepsilon(\beta_\varepsilon u, k, j) = 1+g_\varepsilon a_{kj}(u)+o(g_\varepsilon),\quad k, j = \overline{1, r},$

b) $\varphi_\varepsilon(\beta_\varepsilon u, k, 0) \to \varrho_k(u),\quad k = \overline{1, r}.$

Theorem 2 (Anisimov [2], [3]). If the above conditions are satisfied, then independently of the initial state $j, j=\overline{1, r}$ the distribution of $\beta_\varepsilon \Omega_\varepsilon(j)$ converges weakly to a distribution with characteristic function

$$\Big(\sum_{k=1}^{r} b_k \varrho_k(u)\Big)\Big/\Big(1-\sum_{k,j=1}^{r} \pi_0(k) p_0(k, j) a_{kj}(u)\Big),$$

where

$$\pi_0(k) = \lim_{\varepsilon\to 0} \pi_\varepsilon(k),\quad p_0(k, j) = \lim_{\varepsilon\to 0} p_\varepsilon(k, j).\quad k, j = \overline{1, r}.$$

The most crucial part of applying Theorem 2 to particular situations is finding the normalizing factor β_ε.

In the following an example is given on which our further considerations are based.

Example (see Anisimov et. al. [3] pp. 151).

Let $(X_\varepsilon(k),\ k\geqq 0)$ be a Markov chain with state space

$$E = \{(i, q),\ i = \overline{1, r},\ q = \overline{0, m+1}\}$$

defined by the transition matrix $\|p_\varepsilon[(i, q), (j, z)]\|$ satisfying the following conditions:

1. $p_\varepsilon[(i, 0), (j, 0)] \to p_{ij},\quad i, j = \overline{1, r},$

and the matrix $\|p_{ij}\|$, $i, j=\overline{1, r}$ is irreducible,

2. $p_\varepsilon[(i, q), (j, q+1)] = \varepsilon\alpha_{ij}^{(q)}+o(\varepsilon),\quad i, j = \overline{1, r},\quad q = \overline{0, m},$

3. $p_\varepsilon[(i, q), (j, q)] \to 0,\quad i, j = \overline{1, r},\quad q \geqq 1,$

4. $p_\varepsilon[(i, q), (j, z)] \equiv 0,\quad i, j = \overline{1, r},\quad |z-q| \geqq 2.$

In the sequel the set of states $\{(i, q),\ i=\overline{1, r}\}$ is called the q-th level of the chain, $q=\overline{0, m+1}$.

It is easy to see that conditions (1)—(4) have the following meaning. Level 0 in the limit forms an essential class, the transition probability from the q-th level to the $q+1$-th level is of order ε, on the q-th level the transition probability tends to zero, finally, the summarized one-step transition probability from the q-th level to a lower level tends to 1.

Let us single out the subset of states

$$\langle\alpha\rangle = \{(i, q),\ i = \overline{1, r},\ q = \overline{0, m}\}.$$

Denote by $\pi_\varepsilon(i, q)$ the stationary distribution of $X_\varepsilon(k)$ and by $g_\varepsilon(\langle\alpha\rangle)$ the steady state probability of exit from $\langle\alpha\rangle$, that is

$$g_\varepsilon(\langle\alpha\rangle) = \sum_{i=1}^{r} \pi_\varepsilon(i, m) \sum_{j=1}^{r} p_\varepsilon[(i, m), (j, m+1)].$$

Let

$$P = \|p_{ij}\|,\quad i, j = \overline{1, r},\quad A^{(q)} = \|\alpha_{ij}^{(q)}\|,\quad i, j = \overline{1, r},\quad q = \overline{0, m},$$

$\{\pi_k,\ k = \overline{1, r}\}$ the stationary distribution for the chain with matrix P,

$$\bar\pi_\varepsilon^{(q)} = (\pi_\varepsilon^{(q)}(i, q),\ i = \overline{1, m}),\quad \bar\pi = (\pi_1, \ldots, \pi_r)$$

row-vectors.

Conditions (1)—(4) enable us to compute the main terms of the asymptotic expression for $\bar\pi_\varepsilon^{(q)}$ and $g_\varepsilon(\langle\alpha\rangle)$, namely, we obtain

$$\begin{aligned} \bar\pi_\varepsilon^{(q)} &= \varepsilon^q \bar\pi A^{(0)} A^{(1)} \ldots A^{(q-1)} + o(\varepsilon^q),\quad q \geqq 1, \\ g_\varepsilon(\langle\alpha\rangle) &= \varepsilon^{m+1} \bar\pi A^{(0)} \ldots A^{(m)} \underline{1} + o(\varepsilon^{m+1}), \end{aligned} \tag{3}$$

where $\underline{1} = (1, \ldots, 1)^T$.

Now, making use of Theorem 2 and formula (3) we get the following asymptotic result. (See Anisimov [2], [3].)

Let $(\xi_\varepsilon(t),\ t \geqq 0)$ be a SMP given by the embedded Markov chain $(X_\varepsilon(k),\ k \geqq 0)$ satisfying conditions (1)—(4).

Let the times $\tau_\varepsilon((l, s), (j, z))$ transition time from state (l, s) to state (j, z) fulfill the condition

$$\mathbf{E} \exp\{i\theta\beta_\varepsilon \tau_\varepsilon((l, s), (j, z))\} = 1 + a_{lj}(s, z, \theta)\varepsilon^{m+1} + o(\varepsilon^{m+1}),$$

where β_ε is a normalizing factor.

Denote by $\Omega_\varepsilon(j, s)$ the instant at which the SMP reaches the $q+1$-th level for the first time, provided $\xi_\varepsilon(0) = (j, z)$, $s \leqq m$.

Corollary 2. If the above conditions are satisfied, then

$$\lim_{\varepsilon \to 0} \mathbf{E}\{iu\beta_\varepsilon \Omega(j, s)\} = (1 - A(\theta))^{-1},$$

where

$$A(\theta) = \Big(\sum_{k, j=1}^{r} \pi_k p_{kj} a_{kj}(0, 0, \theta)\Big) / (\bar\pi A^{(0)} \ldots A^{(m)} \underline{1}).$$

In particular, if $a_{ij}(s, z, \theta)=i\theta m_{ij}(s, z)$, then the limit is an exponentially distributed random variable with parameter

$$(\bar{\pi}A^{(0)}\dots A^{(m)}\underline{1})/(\sum_{k,j=1}^{r}\pi_k p_{kj} m_{kj}(0, 0)).$$

3. The mathematical models

In this section we show how the above results for sojourn time problems can be applied to the asymptotic analysis of controlled finite-source queueing systems under the assumption of fast service.

3.1. System $\langle N/M_u/M_u/n\rangle$.

Let the requests emanate from a finite source of size N which are served by one of n $(n\leqq N)$ servers at a service facility according to a First-In-First-Out (FIFO) discipline. If there is no idle server, then a waiting line is formed and the customers are delayed. Suppose that the system is evolving in a random environment governed by an irreducible, aperiodic Markov chain $(x(t),\ t\geqq 0)$ with state space $\{1, \dots, r\}$ and transition density matrix

$$\{a_{ij},\ i,j=\overline{1,r},\quad a_{ii}=\sum_{j\neq i}a_{ij}\}.$$

Whenever the environmental process is in state i and there are s, $s=\overline{0, N-1}$ customers at the service facility, each request is assumed to stay in the source for a random time having exponential distribution with parameter $\lambda(i, s)$. Furthermore, the service time of each customer is supposed to be an exponentially distributed random variable with parameter $\mu_\varepsilon(i, s)$, $i=\overline{1,r}$, $s=\overline{1,N}$.

All random variables involved here are assumed to be independent of each other.

Let us consider the system under the assumption of fast service, that is, $\mu_\varepsilon(i, s)\to\infty$ as $\varepsilon\to 0$. For simplicity let $\mu_\varepsilon(i, s)=\mu(i, s)/\varepsilon$. Denote by $y_\varepsilon(t)$ the number of customers staying at the service facility at time t, and let

$$Z_\varepsilon(t) = (X(t), y_\varepsilon(t)).$$

Clearly, $(Z(t),\ t\geqq 0)$ is a two-dimensional Markov process with state space

$$E = \{(i, s),\ i=\overline{1,r},\ s=\overline{0,N}\}.$$

Let $\Omega_\varepsilon(k, q\colon m)$ denote the instant at which $y_\varepsilon(t)=m+1$ for the first time, provided that the initial state of $Z_\varepsilon(t)$ was (k, q), $k=\overline{1,r}$, $q=\overline{0,m}$, $m=\overline{1,N-1}$. That is,

$$\Omega_\varepsilon(k, q\colon m) = \inf\{t\colon t>0,\ y_\varepsilon(t)=m+1/Z_\varepsilon(0)=(k, q)\}$$

which is termed, in the sequel, as the time to the first system failure. It is easy to see, that $\Omega_\varepsilon(k, q\colon m)$ is the first exit time of $Z_\varepsilon(t)$ from the subset

$$\langle\alpha\rangle = \{(i, s),\ i=\overline{1,r},\ s=\overline{0,m}\},$$

provided that $Z_\varepsilon(0)=(k, q)$.

We are interested in the limiting distribution of the random variable $\Omega_\varepsilon(k, q\colon m)$ as $\varepsilon\to 0$.

It can easily be verified that the following transitions occur in an arbitrary time interval $(t, t+h)$

$$(i, s)\xrightarrow{h}\begin{cases}(j, s) & \text{w.p.}\quad a_{ij}h+o(h), \quad i\neq j,\\ (i, s+1) & \text{w.p.}\quad (N-s)\lambda(i, s)+o(h), \quad 0\leqq s<N,\\ (i, s-1) & \text{w.p.}\quad (S_n\mu(i, s)/\varepsilon)h+o(h), \quad 1\leqq s\leqq N,\\ (i, s) & \text{w.p.}\quad 1-h[(N-s)\lambda(i, s)+S_n\mu(i, s)/\varepsilon+a_{ii}]+o(h),\end{cases}$$

$$S_n \doteq \min(s, n)$$

The sojourn time of $Z_\varepsilon(t)$ in state (i, s) is exponentially distributed with parameter

$$(N-s)\lambda(i, s)+S_n\mu(i, s)/\varepsilon+a_{ii}.$$

Thus, the transition probabilities for the embedded Markov chain are

$$p_\varepsilon[(i, s), (j, s)] = a_{ij}[(N-s)\lambda(i, s)+S_n\mu(i, s)/\varepsilon+a_{ii}]^{-1}, \quad 1\leqq s\leqq N,$$

$$p_\varepsilon[(i, 0), (j, 0)] = a_{ij}[N\lambda(i, 0)+a_{ii}]^{-1}, \quad s=0,$$

$$p_\varepsilon[(i, s), (i, s+1)] = (N-s)\lambda(i, s)[(N-s)\lambda(i, s)+S_n\mu(i, s)/\varepsilon+a_{ii}]^{-1}, \quad 1\leqq s\leqq N,$$

$$p_\varepsilon[(i, 0), (i, 1)] = N\lambda(i, 0)[N\lambda(i, 0)+a_{ii}]^{-1}, \quad s=0,$$

$$p_\varepsilon[(i, s), (i, s-1)] = \frac{S_n\mu(i, s)}{\varepsilon}[(N-s)\lambda(i, s)+S_n\mu(i, s)/\varepsilon+a_{ii}]^{-1}, \quad 1\leqq s\leqq N.$$

As $\varepsilon\to 0$, this implies

$$p_\varepsilon[(i, s), (j, s)] = o(1), \quad s\geqq 1,$$

$$p_\varepsilon[(i, s), (i, s+1)] = \frac{\varepsilon(N-s)\lambda(i, s)}{S_n\mu(i, s)}(1+o(1)), \quad 1\leqq s\leqq N,$$

$$p_\varepsilon[(i, s), (i, s-1)] \to 1, \quad 1\leqq s\leqq N,$$

$$p_\varepsilon[(i, 0), (j, 0)] = a_{ij}/(N\lambda(i, 0)+a_{ii}), \quad i, j=\overline{1, r},$$

$$p_\varepsilon[(i, 0), (i, 1)] = N\lambda(i, 0)/(N\lambda(i, 0)+a_{ii}), \quad i=\overline{1, r}.$$

This agrees with the conditions (1)—(4) of Example, but here the zero level is the set $\{(i, 0), (i, 1), i=\overline{1, p}\}$ while the q-th level is $\{(i, q+1), i=\overline{1, r}\}$. Since the level 0 in the limit forms an essential class, the probabilities $\pi_0(i, q)$, $i=\overline{1, r}$, $q=\overline{0, 1}$ satisfy the following system of equations

$$\pi_0(j, 0) = \sum_{i\neq j}\pi_0(i, 0)a_{ij}/(N\lambda(i, 0)+a_{ii})+\pi_0(j, 1), \tag{4}$$

$$\pi_0(j, 1) = \pi_0(j, 0)N\lambda(j, 0)/(N\lambda(j, 0)+a_{jj}). \tag{5}$$

Substituting eq. (5) to eq. (4) we get

$$\pi_0(j, 0)\frac{a_{jj}}{N\lambda(j, 0)+a_{jj}} = \sum_{i\neq j}\pi_0(i, 0)\frac{a_{ij}}{N\lambda(i, 0)+a_{ii}}. \tag{6}$$

Denote by π_k, $k=\overline{1, r}$ the stationary distribution of the governing Markov chain $(x(t), t\geqq 0)$.

Since

$$\pi_j a_{jj} = \sum_{i\neq j}\pi_i a_{ij}, \quad j=\overline{1, r},$$

from (6) we have

$$\pi_0(i, 0) = B\pi_i[N\lambda(i, 0)+a_{ii}],$$

$$\pi_0(i, 1) = B\pi_i N\lambda(i, 0), \quad i=\overline{1, r},$$

where

$$B = \Big[\sum_{k=1}^{r}\pi_k\big(a_{kk}+2N\lambda(k, 0)\big)\Big]^{-1}.$$

By using formula (3), it is easy to obtain that

$$\pi_\varepsilon(i, q) = \varepsilon^{q-1}B\pi_i N\lambda(i, 0)\prod_{s=1}^{q-1}\frac{(N-s)\lambda(i, s)}{S_n\mu(i, s)}(1+o(1)),$$

$$\prod_{s=1}^{0} = 1, \quad q\geqq 1,$$

and

$$g_\varepsilon(\langle\alpha\rangle) = \varepsilon^m NB\sum_{i=1}^{r}\pi_i\lambda(i, 0)\prod_{s=1}^{m}\frac{(N-s)\lambda(i, s)}{S_n\mu(i, s)}(1+o(1)).$$

Taking into consideration the exponentiality of $\tau_\varepsilon(i, s)$ for a fixed u we have

$$\mathbf{E}\exp\{i\varepsilon^m u\tau_\varepsilon(j, 0)\} = 1+\varepsilon^m\frac{iu}{a_{jj}+N\lambda(j, 0)}(1+o(1)),$$

$$\mathbf{E}\exp\{i\varepsilon^m u\tau_\varepsilon(j, s)\} = 1+o(\varepsilon^m), \quad s>0.$$

Notice, that $\beta_\varepsilon=\varepsilon^m$. Therefore, by the help of Corollary 2 we immediately get the following theorem.

Theorem 3. For the system $\langle N/M_u/M_u/n\rangle$ under the above conditions, for any $k=\overline{1, r}$, $q\leqq m$ the distribution of the normalized random variable $\varepsilon^m\Omega_\varepsilon(k, q\colon m)$ converges weakly to an exponentially distributed random variable with parameter

$$\Lambda = \sum_{i=1}^{r}\pi_i N\lambda(i, 0)\prod_{s=1}^{m}\frac{(N-s)\lambda(i, s)}{S_n\mu(i, s)}.$$

Consequently, the distribution of time to the first system failure can be approximated by

$$\mathbf{P}\big(\Omega_\varepsilon(k, q\colon m)>t\big)\approx\exp(-\varepsilon^m\Lambda t).$$

In particular, for the system $\langle N/M/M/n\rangle$ the parameter Λ can be written as

$$\Lambda = \lambda\left(\frac{\lambda}{\mu}\right)^m N \prod_{s=1}^{m} \frac{(N-s)}{S_n}.$$

Furthermore, for the random variable $\varepsilon^{n+m}\Omega_\varepsilon(k, q\colon n+m)$, Λ is given by

$$\Lambda = \lambda\left(\frac{\lambda}{\mu}\right)^{n+m} \frac{N}{n!n^m} \prod_{s=1}^{n+m} (N-s). \tag{7}$$

It is well-known, that if $N\to\infty$ and $\lambda\to 0$ such that $N\lambda\to\lambda'$, then the stationary characteristics of the system $\langle N/M/M/n\rangle$ coincide with the corresponding characteristics of the system $M/M/n$ with Poisson arrivals with parameter λ' and exponentially distributed service times with parameter μ/ε. (See [8].) In fact, applying (7) as $N\to\infty$, $\lambda\to 0$, $N\lambda\to\lambda'$ we easily get

$$\Lambda = \frac{1}{n!n^m}\lambda'\left(\frac{\lambda'}{\mu}\right)^{n+m}$$

which agrees with the result of Anisimov [3] pp. 157.

3.2. The system $\langle N/\vec{M}_u/\vec{M}_u/1\rangle$.

Let us consider problem 3.1. with the following modifications. The requests are stochastically different, unit k is characterized by arrival rate $\lambda_k(i, s)$ and service rate $\mu_k(i, s)/\varepsilon$, provided that the underlying Markov chain is in state i and there are s customers at the service facility consisting of one server.

We are interested in the limiting distribution of $\Omega_\varepsilon(m)$ under the assumption of fast service, that is, as $\varepsilon\to 0$.

Let

$$Z_\varepsilon(t) = \{X(t),\ y_\varepsilon(t)\colon \gamma_1(t), \dots, \gamma_{y_\varepsilon(t)}(t)\}$$

be a multi-dimensional Markov process with state space

$$E = \{(i, s\colon k_1, \dots, k_s);\ i = \overline{1, r},\ s = \overline{0, N},\ (k_1, \dots, k_s)\in V_N^s,\ k_0 = 0\},$$

where

$X(t)$ is the governing Markov chain,

$y_\varepsilon(t)$ is the number of customers staying at the service facility at time t,

$\gamma_1(t), \dots, \gamma_{y_\varepsilon(t)}(t)$ are the indices of requests staying at the service facility at time t, ordered lexicographically,

V_N^s is the set of all variations of order k of integers $1, \dots, N$.

Let us single out the subset of states

$$\langle\alpha\rangle = \{(i, q\colon k_1, \dots, k_q),\ i = \overline{1, r},\ q = \overline{0, m},\ (k_1, \dots, k_q)\in V_N^q\}.$$

Notice, that $\Omega_\varepsilon(m)$ is the first exit time of $Z_\varepsilon(t)$ from $\langle\alpha\rangle$. On the analogy of 3.1, it is not difficult to verify, that the transition probabilities of the embedded Markov chain as $\varepsilon \to 0$, are

$$p_\varepsilon[(i, 0\colon 0), (j, 0\colon 0)] = a_{ij}/[\sum_{l=1}^{N} \lambda_l(i, 0) + a_{ii}], \quad i = \overline{1, r}.$$

$$p_\varepsilon[(i, 0\colon 0), (i, 1\colon k)] = \lambda_k(i, 0)/[\Sigma\lambda_l(i, 0) + a_{ii}], \quad i = \overline{1, r}, \quad k = \overline{1, N},$$

$$p_\varepsilon[(i, s\colon k_1, \ldots, k_s), (j, s\colon k_1, \ldots, k_s)] = o(1), \quad s \geqq 1,$$

$$p_\varepsilon[(i, s\colon k_1, \ldots, k_s), (i, s+1\colon k_1, \ldots, k_{s+1})] = \frac{\varepsilon\lambda_{k_{s+1}}(i, s)}{\mu_{k_1}(i, s)}(1+o(1)),$$

$$p_\varepsilon[(i, s\colon k_1, \ldots, k_s), (i, s-1, k_2, \ldots, k_s)] \to 1, \quad 1 \leqq s \leqq N,$$

$$(i, 0\colon k_2, \ldots, k_1) = (i, 0\colon 0).$$

Now, we can obtain that

$$\pi_\varepsilon(i, q\colon k_1, \ldots, k_q) = \varepsilon^{q-1} B\pi_i \frac{\lambda_{k_1}(i, 0)\cdot\lambda_{k_2}(i, 1)\ldots\lambda_{k_q}(i, q-1)}{\mu_{k_1}(i, 1)\cdot\mu_{k_1}(i, 2)\ldots\mu_{k_1}(i, q-1)}(1+o(1)),$$

$$\pi_\varepsilon(i, q) = \sum_{(k_1, \ldots, k_q)} \pi_\varepsilon(i, q\colon k_1, \ldots, k_q),$$

and

$$g_\varepsilon(\langle\alpha\rangle) = \varepsilon^m B \sum_{i=1}^{r} \pi_i \sum_{(k_1, \ldots, k_{m+1})} \frac{\lambda_{k_1}(i, 0)\ldots\lambda_{k_{m+1}}(i, m)}{\mu_{k_1}(i, 1)\ldots\mu_{k_1}(i, m)}(1+o(1)),$$

where

$$B = [\sum_{i=1}^{r} \pi_i(a_{ii} + 2\sum_{l=1}^{N} \lambda_l(i, 0))]^{-1}.$$

Making use of Corollary 2 we are ready to get the following theorem.

Theorem 4. For the system $\langle N/\vec{M}_u/\vec{M}_u/1\rangle$ under the the above conditions, independently of the initial state, the distribution of the normalized random variable $\varepsilon^m \Omega_\varepsilon(m)$ converges weakly to an exponentially distributed random variable with parameter

$$\Lambda = \sum_{i=1}^{r} \pi_i \sum_{(k_1, \ldots, k_{m+1})} \frac{\lambda_{k_1}(i, 0)\ldots\lambda_{k_{m+1}}(i, m)}{\mu_{k_1}(i, 1)\ldots\mu_{k_1}(i, m)}.$$

4. Applications

In this section we show how the above results can be applied to problems arised in the field of reliability theory and computer performance.

4.1. Consider a repairable system operating in a random environment, which consists of N elements and one repairman. The expected life time of the k-th element is assumed to be an exponentialy distributed random variable with failure rate λ_k. When an element fails, it enters the repair facility and will be immediately repaired unless the repairman is busy, otherwise it will wait in a queue in the order of its arrival. The required repair time of the k-th element is supposed to be exponentially distributed random variable with parameter μ_k, $k = \overline{1, N}$. Furthermore, we assume that the failure and repair intensities depend on the state of the environmental process and the number of failed elements. Namely, whenever the environmental chain is in state i and there are s elements at the repair facility, the rate at which the remaining duration of life times decreases is $a(i, s)$ and the rate at which the remaining duration of repair time decreases is $b(i, s)/\varepsilon$.

The involved random variables are supposed to be independent of each other.

The system is said to be failed if the number of failed elements is $m+1$. Therefore, the instant $\Omega_\varepsilon(m)$ at which the queue length reaches the level $m+1$ for the first time is of great practical importance. Hence, the problem is to find the asymptotic distribution of the random variable under the assumption of fast repair, that is, as $\varepsilon \to 0$. This assumption is justified, since usually the average repair time is many times smaller than the average failure-free oparation time.

Clearly, this problem corresponds to the system $\langle N/\vec{M}_u/\vec{M}_u/1\rangle$, thus the distribution of $\Omega_\varepsilon(m)$ can be approximated by

$$\mathbf{P}(\Omega_\varepsilon(m) > t) \approx \exp(-\varepsilon^m \Lambda t),$$

where

$$\Lambda = \sum_{i=1}^{r} \pi_i \sum_{(k_1, \dots, k_{m+1})} \frac{\lambda_{k_1} a(i, 0) \dots \lambda_{k_{m+1}} a(i, m)}{\mu_{k_1} b(i, 1) \dots \mu_{k_1} b(i, m)}.$$

4.2. Let us consider a CP-terminal system consisting of N terminals coupled to one Central Processor Unit (CPU). The system is operating in a random environment which influences the service rates at the terminals and at the CPU. At the terminals the think times are exponentially distributed random variables with parameter λ_k for terminal k, $k=\overline{1, N}$. The processing times for jobs at the CPU are exponentially distributed random variables with mean ε/μ_k, for the job from terminal k, $k=\overline{1, N}$, where ε is a small parameter. The service discipline at the CPU is FIFO. Whenever the environmental process is in state i the rate at which the remaining duration of think times, processing time decreases is $a(i)$, $b(i)$ respectively.

The think and processing times are supposed to be independent of each other. Let us assume that the average CPU times are many times smaller than the average think times, that is, $\varepsilon \approx 0$.

We are interested in the distribution of the instant $\Omega_\varepsilon(m)$ at which the number of jobs at the CPU reaches the level $m+1$, $1<m<N$.

It is easy to see, that applying Theorem 4 we get the following approximation

$$\mathbf{P}(\Omega_\varepsilon(m) > t) \approx \exp\{-\varepsilon^m \Lambda t\}$$

where

$$\Lambda = \sum_{i=1}^{r} \pi_i \sum_{(k_1, \dots, k_{m+1})} \frac{\lambda_{k_1} a(0) \dots \lambda_{k_{m+1}} a(m)}{\mu_{k_1} b(1) \dots \mu_{k_1} b(m)}.$$

Remark. We must admit that for terminal systems this characteristic is a less effective performance measure but sometimes it is useful to know.

Abstract

This paper is concerned with an asymptotic analysis of some controlled finite-source queueing systems under the assumption of fast service. Firstly, a brief summary of preliminary results related the asymptotic behavior of SMP is given. Secondly, models of queueing systems with fast service is treated. It is shown, that the time to the first system failure converges in distribution, under appropiate norming, to an exponentially distributed random variable. Finally, applications of the systems in the field of reliability theory and computer performance are considered.

Keywords: SMP, sojourn time, fast service, time to the first system failure, weak convergence.

References

[1] Anisimov V. V., Limit theorems for sums of random variables in an array of sequence defined on a subset of states of a Markov chain up to the exit time, Theor. Probability and Math. Statist. 1974 No 4, 15—22.
[2] Anisimov V. V., Asymptotic methods in analysing the behavior of stochastic systems, Mecniereba, Tbilisi, 1984 (in Russian).
[3] Anisimov V. V.—Zakusilo O. K.—Doncsenko V. S., Elements of queueing theory and asymptotic analysis of systems, Visa Skola, Kiev, 1987 (in Russian).
[4] Franken P.—König D.—Arndt U.—Schmidt V., Queues and Point Processes, Akademie-Verlag, Berlin, 1981.
[5] Gnedenko B. V. and Kovalenko I. N., Introduction to Queueing Theory, Nauka, Moscow, 1966 (in Russian).
[6] Gnedenko B. V. (Editor), Mathematical problems in reliability theory, Radio i Svjaz, 1983 (in Russian).
[7] Gnedenko B. V. and König D., Handbuch der Bedienungstheorie, Vol. I., II., Akademie-Verlag, Berlin, 1983, 84.
[8] Jaiswal N. K., Priority Queues, Academic Press, New York, 1968.
[9] Kleinrock L., Queueing Systems, Vol. II. Computer applications, Wiley-Interscience, New York, 1976.
[10] Koroljuk V. S., Asymptotic behavior of the sojourn time of Semi-Markov Process in a subset of the states, Ukrain. Mat. Z. 21 (1969) 842—845 (in Russian).
[11] Koroljuk V. S. and Turbin A. F., Semi-Markov Processes and their applications, Naukova Dumka, Kiev, 1976 (in Russian).
[12] Kovalenko I. N., Studies in reliability analysis of complex systems, Naukova Dumka, Kiev, 1975 (in Russian).
[13] König D. and Stoyan D., Methoden der Bedienungstheorie, Academie-Verlag, Berlin, 1976.
[14] Lavenberg S. S. (Editor), Computer Performance Modeling Handbook, Academic Press, New York, 1983.
[15] Lifsic A. L. and Malc E. A., Simulation of queueing systems, Sovetskoe Radio, Moscow, 1978 (in Russian).
[16] Takács L., Introduction to the Theory of Queues, Oxford University Press, New York, 1962.
[17] White J. A.—Schmidt J. W.—Bennett G. K., Analysis of Queueing Systems, Academic Press, New York, 1975.

Received April 5, 1988