MULTI-SERVER QUEUEING MODEL *MAP/PH/N* WITH BROADCASTING SERVICE IN UNRELIABLE SERVERS

A. Dudin, B. Sun

Belarusian State University
Minsk, Belarus
dudin@bsu.by

We consider a multi-server queueing model with broadcasting service discipline which assumes that the customer who sees several free servers upon arrival is served by all these servers. Errors can occur during the service that leads to incorrect service. The key performance measures of the system including probability of correct delivering of a customer are calculated. Effect of correlation in the arrival process is numerically illustrated.

Keywords: Broadcasting, queueing analysis, buffers, reliability, communication system traffic, Markov processes.

1. INTRODUCTION

Multi-server queueing systems model many communication networks and have got a lot of attention in literature since the pioneering works by A.K. Erlang in the early 1900th. The standard assumption in analysis of multi-server queues is that each customer is served by one server. In this paper, we investigate the case when the customer gets a service from all servers that are idle at the customer arrival epoch. Such a service discipline creates some redundancy, but it can help to decrease the average delivering time of a first copy of the broadcasted customer, see [1, 2, 3]. In this paper we focus on the model with unreliable servers. Servers do not physically broken, but just can provide a wrong service. It is intuitively clear that the possible parallel processing in several servers increases probability that a customer (at least one of its copies) will get correct service. Here we illustrate that a customer sojourn time and probability of successful (correct) service are influenced by correlation in the arrival process. We assume that the input flow is described by the *MAP* (Markov Arrival Process) which well suits for modeling the correlated bursty traffic in the modern telecommunication networks. The distribution of service time and time till an error occurrence in a server is of *PH* (phase type). *PH* distribution is the well recognized descriptor of the service process in the multi-server queues which still allows to get analytically tractable results.

The rest of the paper is organized as follows. In section 2, the model is described and stationary distribution of the number of customers in this model is analyzed. Sojourn time distribution in the system and probability of successful service are calculated in section 3. Section 4 contains the numerical results.

2. MATHEMATICAL MODEL AND DISTRIBUTION OF THE NUMBER OF CUSTOMERS IN A SYSTEM

We consider an N-server queueing system. The customers arrive to the system according to a *MAP* (*MarkovianArrivalProcess*). The notion of the *MAP* and its detailed description is given by M. Neuts, V. Ramaswami and D. Lucantoni, see, e.g., paper [4] were the currently standard in literature notation for the *MAP* is introduced.

We denote the directing process of the *MAP* by $\nu_t, t \geq 0$. The process $\nu_t, t \geq 0$, is an irreducible continuous time Markov chain with state space $\{0, 1, \dots, W\}$. Intensities of transitions of this Markov chain, which are accompanied by generation of k customers, $k = 0, 1$, are combined into the matrices $D_k, k = 0, 1$. The matrix $D(1) = D_0 + D_1$ is the generator of the process $\nu_t, t \geq 0$. The average arrival rate λ is defined by $\lambda = \boldsymbol{\theta} D_1 \boldsymbol{e}$ where $\boldsymbol{\theta}$ is the invariant vector of the stationary distribution of the Markov chain $\nu_t, t \geq 0$. The vector $\boldsymbol{\theta}$ is the unique solution to the system $\boldsymbol{\theta} D(1) = \boldsymbol{0}, \boldsymbol{\theta e} = 1$. Here e is the column-vector of appropriate size consisting of 1's and $\boldsymbol{0}$ is the row-vector of appropriate size consisting of 0's.

The servers are assumed to be identical and independent of each other. The service process in each server is assumed to be of phase (PH) type. The PH type service process is described by a continuous time Markov process $\eta_t, t \geq 0$. The state of this process at the epoch of a service start is defined according to a probabilistic row-vector $\beta = (\beta_1, ..., \beta_M)$. Further, transitions of the process $\eta_t, t \geq 0$, are defined by a matrix S of dimension M. We set $S_0 = -Se$. The average service time b_1 is given by $b_1 = \beta(-S)^{-1}\boldsymbol{e}$. For more details on PH type distributions see the book [5].

If the arriving customer meets several free servers upon arrival, the customer is copied and all free servers start, independently of others, the service of the copies of this customer. We assume that sojourn time of the customer in the system finishes at the earliest epoch of one of its copies service completion. However, all other copies of this customer continue the service in the corresponding servers but they are deleted after the service completion. If all the servers are busy upon arrival, the customer is placed into the buffer of an infinite capacity and then it will be picked up from the queue according to the $FIFO$ (First In - First Out) discipline.

Errors can happen during a service. Time till occurrence of an error has PH distribution, which is described by the row vector $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_R)$ and a subgenerator matrix Γ. The average intensity of errors φ is given by $\varphi^{-1} = \boldsymbol{\gamma}(-\Gamma)^{-1}\boldsymbol{e}$.

For use in the sequel, we introduce the following denotations.

- I is an identity matrix. O is zero square matrix. If the dimension of the matrix is not clear from context, it will be indicated by the suffix. E.g., $I_{\bar{W}}$ is the identity matrix of dimension $\bar{W}$, $\bar{W} = W + 1$.
- $\otimes$ ($\oplus$) is the symbol of Kronecker product (sum) of the matrices.
- $\beta^{\otimes l} \stackrel{def}{=} \underbrace{\beta \otimes \ldots \otimes \beta}_{l}, \quad S^{\oplus l} \stackrel{def}{=} \underbrace{S \oplus \ldots \oplus S}_{l}, \quad l \geq 1, \quad \beta^{\otimes 0} \stackrel{def}{=} I_M, S^{\oplus 0} \stackrel{def}{=} O.$
- $S_0^{\oplus l} \stackrel{def}{=} \sum\limits_{m=0}^{l-1} I_{M^m} \otimes S_0 \otimes I_{M^{l-m-1}}, \quad l \geq 1.$

- $\mathcal{M}_r = I_{\bar{W}} \otimes S_0^{\oplus r}, r = \overline{1,N}, \mathcal{M}'_N = I_{\bar{W}} \otimes (S_0\beta)^{\oplus N}, \mathcal{N}_r = D_0 \oplus S^{\oplus r}, r = \overline{0,N}$.
- $\mathcal{D}_r = D_1 \otimes I_{M^{N-r}} \otimes \beta^{\otimes r}, r = \overline{0,N}, \mathcal{K}_i = -\mathcal{M}_{i+1}\mathcal{N}_i^{-1}, i = \overline{0,N-1}$.

Let i_t be the number of customers in the system, $i_t \geq 0$, $m_t^{(j)}$ be the state of the directing process of the service on the j-th busy server, $m_t^{(j)} = \overline{1,M}$, $j = \overline{1, \min\{i_t, N\}}$ (we assume here that the busy servers are numerated in order of their occupying, i.e. the server, which begins the service, is appointed the maximal number among all busy servers; servers that start the service at the same epoch get sequential numbers with equal probabilities; when some server finishes the service, the servers are correspondingly enumerated), ν_t be the state of the underlying process of the MAP, $\nu_t = \overline{0,W}$, at the epoch $t, t \geq 0$.

Consider the multi-dimensional process $\xi_t = (i_t, \nu_t, m_t^{(1)}, \dots, m_t^{(\min\{i_t,N\})})$, $t \geq 0$. It is easy to see that this process is a continuous time irreducible regular Markov chain.

Let Q be the generator of the Markov chain $\xi_t, t \geq 0$, with blocks $Q_{i,j}$ corresponding to the transitions of the component $i_t, t \geq 0$, from the state i into the state j, $i, j \geq 0$.

Lemma 1. *The generator Q of the Markov chain ξ_t, $t \geq 0$, is given by*

$$Q = \begin{pmatrix} \mathcal{N}_0 & O & \dots & O & \mathcal{D}_N & O & O & \dots \\ \mathcal{M}_1 & \mathcal{N}_1 & \dots & O & \mathcal{D}_{N-1} & O & O & \dots \\ O & \mathcal{M}_2 & \dots & O & \mathcal{D}_{N-2} & O & O & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ O & O & \dots & \mathcal{N}_{N-1} & \mathcal{D}_1 & O & O & \dots \\ O & O & \dots & \mathcal{M}_N & \mathcal{N}_N & \mathcal{D}_0 & O & \dots \\ O & O & \dots & O & \mathcal{M}'_N & \mathcal{N}_N & \mathcal{D}_0 & \dots \\ O & O & \dots & O & O & \mathcal{M}'_N & \mathcal{N}_N & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Proof of the lemma is implemented by means of analysis of the probabilities of the Markov chain $\xi_t, t \geq 0$, transitions during an infinitesimal time interval.

Denote the stationary probabilities of this process by

$$p(i, \nu, m^{(1)}, \dots, m^{(\min\{i,N\})}) =$$
$$= \lim_{t\to\infty} P\{\, i_t = i,\ \nu_t = \nu,\ m_t^{(k)} = m^{(k)},\ k = \overline{1, \min\{i,N\}}\},\ i \geq 0, \nu = \overline{0,W}, m^{(j)} = \overline{1,M}.$$

It can be shown that the limits exist if the stability condition $\rho = \frac{\lambda}{N} b_1 < 1$ holds good. In what follows we assume that this condition is fulfilled.

Let us enumerate the states of the Markov chain ξ_t in the lexicographic order and form the row-vectors $\boldsymbol{p}_i$ of the stationary probabilities $p\,(i, \nu, m^{(1)}, \dots, m^{(\min\{i,N\})})$, corresponding to the state i of the first component of the chain, $i \geq 0$. The dimensionality of these vectors is equal to $K_i = \bar{W}M^i$ for $i = \overline{0,N}$ and $K = \bar{W}M^N$ for $i > N$.

Theorem 1. *Vectors $\boldsymbol{p}_i$, $i \geq 0$, are computed as follows:*

$$\boldsymbol{p}_i = \boldsymbol{p}_N \mathcal{B}_i, i = \overline{0,N},\ \boldsymbol{p}_N = \boldsymbol{p}_{N+1}\mathcal{M}'_N\mathcal{A}_N^{-1},\ \boldsymbol{p}_i = \boldsymbol{p}_{N+1}\mathcal{R}^{i-N-1},\quad i \geq N+1,$$

where $\mathcal{B}_i = \mathcal{K}_{N-1} \times \cdots \times \mathcal{K}_i, i = \overline{0, N-1}, \mathcal{B}_N = I,\ \mathcal{A}_N = -\left(\mathcal{N}_N + \sum_{l=0}^{N-1} \mathcal{B}_l \mathcal{D}_{N-l}\right)$, *the matrix* $\mathcal{R}$ *is the minimal non-negative solution to the system* $\mathcal{R}^2\mathcal{M}'_N + \mathcal{R}\mathcal{N}_N + \mathcal{D}_0 = O$, *and the vector* $\mathbf{p}_{N+1}$ *is the unique solution to the system of linear algebraic equations*

$$\mathbf{p}_{N+1}\left[\mathcal{N}_N + \mathcal{R}\mathcal{M}'_N + \mathcal{M}'_N\mathcal{A}_N^{-1}\mathcal{D}_0\right] = \mathbf{0},\ \mathbf{p}_{N+1}\left[(I-\mathcal{R})^{-1}\mathbf{e} + \mathcal{M}'_N\mathcal{A}_N^{-1}\sum_{l=0}^{N}\mathcal{B}_l\mathbf{e}\right] = 1.$$

Proof is based on careful use of specifics of the generator Q in the first $N+1$ block columns with combination with technique by M. Neuts [5].

3. DISTRIBUTION OF THE SOJOURN TIME AND PROBABILITY OF SUCCESSFUL SERVICE

Let $V(x)$ be the distribution function of the actual sojourn time of an arbitrary customer in the system under study and $v(u)$ be the corresponding Laplace-Stieltjes transform: $v(u) = \int_0^\infty e^{-ux}dV(x),\ \ Re\ u > 0$. The sojourn time of a customer in the system is the time since the customer arrival to the system till the service completion of the earliest copy of this customer.

Lemma 2. *Let* $\xi_k, k = \overline{1,m}$, *be* m *independent identically distributed random variables having PH distribution defined by the irreducible representation* (β, S).

Then a random $\xi = \min_{k=\overline{1,m}} \xi_k$ *has PH distribution defined by the irreducible representation* $(\beta^{\otimes m}, S^{\oplus m})$.

Theorem 2. *Laplace-Stieltjes transform* $v(u)$ *of the sojourn time distribution is calculated by*

$$v(u) = \lambda^{-1}\Bigg[\sum_{i=0}^{N-1}\mathbf{p}_i(D_1 \otimes I_{M^i})\mathbf{e}\gamma_{N-i}(u) +$$

$$\left[\mathbf{p}_N + \mathbf{p}_{N+1}[I - \mathcal{R}\mathcal{Z}(u)]^{-1}\mathcal{Z}(u)\mathbf{e}\right]\mathcal{Z}(u)(D_1 \otimes I_{M^N})\mathbf{e}\gamma_1(u)\Bigg]$$

where $\mathcal{Z}(u) = I_{\bar{W}} \otimes ((uI - S^{\oplus N})^{-1}(S_0\beta)^{\oplus N}),\ \gamma_i(u) = \beta^{\otimes i}(uI - S^{\oplus i})^{-1}(-S^{\oplus i})\mathbf{e},\ i = \overline{1,N}$.

Proof of theorem is implemented by means of generalization of the known method of catastrophes by H. Kesten, J.Th. Runnenberg and G. van Dantzig to the matrix case.

Corollary 1. *The mean sojourn time* V *in the system is computed by*

$$V = \lambda^{-1}\Bigg[\sum_{i=0}^{N-1}\mathbf{p}_i(D_1 \otimes I_{M^i})\mathbf{e}b_1^{(N-i)} + \mathbf{p}_N(\mathcal{S}^{(1)} + b_1^{(1)}I)(D_1 \otimes I_{M^N})\mathbf{e} +$$

$$\mathbf{p}_{N+1}[I - \mathcal{R}\mathcal{S}^{(0)}]^{-1}\left[\mathcal{R}\mathcal{S}^{(1)}[I-\mathcal{R}\mathcal{S}^{(0)}]^{-1}(\mathcal{S}^{(0)})^2 + \mathcal{S}^{(0)}\mathcal{S}^{(1)} + \mathcal{S}^{(1)}\mathcal{S}^{(0)} + \mathcal{S}^{(0)}\mathcal{S}^{(0)}b_1^{(1)}\right](D_1 \otimes I_{M^N})\mathbf{e}\Bigg]$$

where $S^{(0)} = I_{\bar{W}} \otimes ((-S^{\oplus N})^{-1}(S_0\beta)^{\oplus N})$, $S^{(1)} = I_{\bar{W}} \otimes ((-S^{\oplus N})^{-1})S^{(0)}$, *the value* $b_1^{(i)}$ *as expectation of distribution of the minimum of i independent identically distributed random variables having PH distribution with irreducible representation* (β, S) *is given by* $b_1^{(i)} = \beta^{\otimes i}(-S^{\oplus i})^{-1}e$, $i = \overline{1,N}$.

It is easy to compute probability q that an error does not occur during the service of a customer: $q = (\gamma \otimes \beta)(-(\Gamma \oplus S)^{-1})(e_R \otimes S_0)$. Then formula for probability $P_{success}$ that an arbitrary customer will be served correctly follows from the law of total probability as

$$P_{success} = \lambda^{-1}\left[\sum_{i=0}^{N-1}(1-(1-q)^{N-i})p_i(D_1 \otimes I_{M^i})e + q\left[p_N + p_{N+1}[I - \mathcal{R}]^{-1}\right](D_1 \otimes I_{M^N})e\right].$$

4. NUMERICAL ILLUSTRATIONS

In this section we show an influence of correlation in the arrival process. We assume that $N = 5$, PH service process and breakdowns processes in the first experiment are defined by the vectors $\beta = (1,0)$, $\gamma = (1,0)$ and sub-generators

$$S = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}, \Gamma = \begin{pmatrix} -1 & 1 \\ 0 & -0.5 \end{pmatrix}.$$

Average intensities of the service and breakdowns processes are equal to 1 and 0.33333, respectively. Squared coefficients of variation for these processes are equal to 0.5 and 0.55555.

In the second experiment, sub-generator S is of form $S = \begin{pmatrix} -6 & 3 \\ 3 & -9 \end{pmatrix}$. Average intensity of the service is equal to 3.75. Squared coefficient of the service time variation is equal to 0.875.

We consider six different MAPs having the same fundamental rate $\lambda = 5$ and different variation coefficient and correlation coefficients.

The MAP, which is coded as MAP_0, is characterized by the matrices

$$D_0 = \begin{bmatrix} -2.002637 & 0.793175 & 1.209462 \\ 6.473711 & -347.245854 & 340.772143 \\ 6.473711 & 6.473711 & -1350.141152 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 501.447648 & 835.746082 & 0 \end{bmatrix}.$$

This MAP is the IPP (interrupted Poisson process). It has the correlation coefficient $c_{cor} = 0$, and the variation $c_{var} = 4$.

The MAP, which is coded as MAP_1, is characterized by the matrices

$$D_0 = \begin{bmatrix} -3 & 2.5 \\ 2.5 & -4 \end{bmatrix}, D = \begin{bmatrix} 0.5 & 0 \\ 0 & 1.5 \end{bmatrix}.$$

This MAP has the correlation coefficient $c_{cor} = 0.0052$, and the variation $c_{var} = 1.0869$.

The MAP, which is coded as MAP_2, is characterized by the matrices

$$D_0 = \begin{bmatrix} -13.33463 & 0.588578 & 0.617293 \\ 0.692663 & -2.446574 & 0.422942 \\ 0.682252 & 0.414363 & -1.635426 \end{bmatrix}, D = \begin{bmatrix} 11.54694 & 0.363141 & 0.21867 \\ 0.384249 & 0.865869 & 0.08085 \\ 0.285172 & 0.04255 & 0.21109 \end{bmatrix}.$$

This MAP has the correlation coefficient $c_{cor} = 0.1$, and the variation $c_{var} = 4$.

The MAP, which is coded as MAP_3, is characterized by the matrices

$$D_0 = \begin{bmatrix} -15.732675 & 0.606178 & 0.592394 \\ 0.517816 & -2.289674 & 0.467885 \\ 0.597058 & 0.565264 & -1.959664 \end{bmatrix}, D = \begin{bmatrix} 14.1502 & 0.302098 & 0.08181 \\ 0.10707 & 1.03228 & 0.16463 \\ 0.08583 & 0.197946 & 0.51357 \end{bmatrix}.$$

This MAP has the correlation coefficient $c_{cor} = 0.2$, and the variation $c_{var} = 4$.

The MAP, which is coded as MAP_4, is characterized by the matrices

$$D_0 = \begin{bmatrix} -25.539839 & 0.393329 & 0.361199 \\ 0.14515 & -2.2322 & 0.200007 \\ 0.295961 & 0.387445 & -1.752618 \end{bmatrix}, D = \begin{bmatrix} 24.24212 & 0.466868 & 0.076323 \\ 0.034097 & 1.666864 & 0.186082 \\ 0.009046 & 0.255508 & 0.804658 \end{bmatrix}.$$

This MAP has the correlation coefficient $c_{cor} = 0.3$, and the variation $c_{var} = 4$.

The MAP, which is coded as MAP_5, is the stationary Poisson process. It has the correlation coefficient $c_{cor} = 0$, and the variation $c_{var} = 1$.

Figures 1 and 2 show dependence of the probability $P_{success}$ on the service rate μ and on the rate φ of errors occurrence for the arrival processes with the same arrival rate but different correlation. One may conclude that ignorance of effect of correlation can lead to the wrong performance evaluation of the model under study.

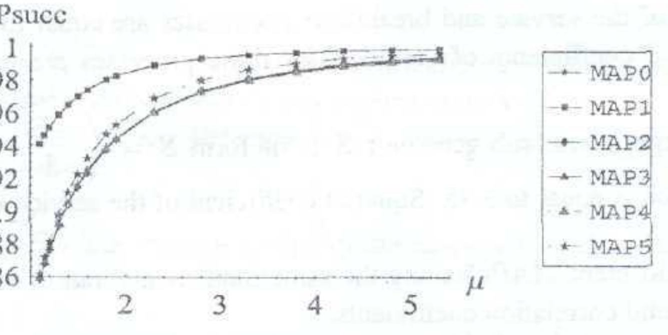

Fig. 1. Dependence of the probability $P_{success}$ on the service rate μ for different correlation in arrival process

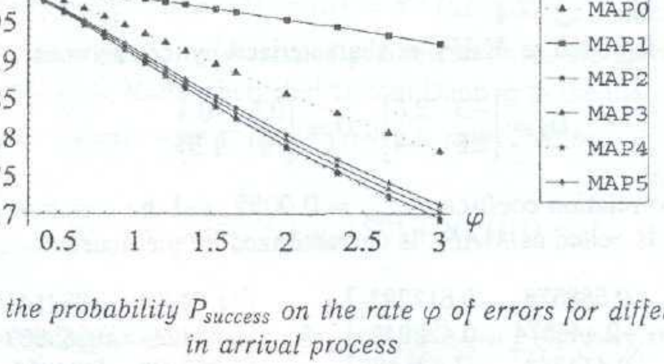

Fig. 2. Dependence of the probability $P_{success}$ on the rate φ of errors for different correlation in arrival process

REFERENCES

1. *Lee M.H., Dudin A.N., Klimenok V.I.* The SM/M/N queueing system with broadcasting service // Mathematical Problems in Engineering. 2006. V. 2006. Article ID 98171. 18 p.
2. *Lee M.H., Dudin A., Klimenok V., Kim C. J.* Multi-server queueing model with broadcasting service // IEEE Communications Letters. 2007. V. 11. № 6. P. 546–548.
3. *Kim C.S., Lee M.H., Dudin A.N., Klimenok V.I.* Multi-server queueing systems with cooperation of the servers // Annals of Operations Research. 2008. V. 162. P. 57–68.
4. *Lucantoni D.M.* New results on the single server queue with a batch Markovian arrival process // Communications in Statistics-Stochastic Models. 1991. V. 7. № 1. P. 1–46.
5. M.F. Neuts, *Matrix-Geometric Solutions in Stochastic Models – An Algorithmic Approach*. Johns Hopkins University Press, 1981.