Industrial Engineering
& Management Systems
Vol 15, No 2, June 2016, pp.131-142
ISSN 1598-7248 │ EISSN 2234-6473 │
http://dx.doi.org/10.7232/iems.2016.15.2.131

Queueing System Operating in Random Environment as a Model of a Cell Operation

Chesoong Kim*
Sangji University, Wonju, Korea

Alexander Dudin, Olga Dudina
Belarusian State University, Minsk, Belarus

Jiseung Kim
Kyungil University, Gyeongsan, Korea

(Received: March 3, 2016 / Revised: May 13, 2016 / Accepted: May 20, 2016)

ABSTRACT

We consider a multi-server queueing system without buffer and with two types of customers as a model of operation of a mobile network cell. Customers arrive at the system in the marked Markovian arrival flow. The service times of customers are exponentially distributed with parameters depending on the type of customer. A part of the available servers is reserved exclusively for service of first type customers. Customers who do not receive service upon arrival, can make repeated attempts. The system operation is influenced by random factors, leading to a change of the system parameters, including the total number of servers and the number of reserved servers. The behavior of the system is described by the multi-dimensional Markov chain. The generator of this Markov chain is constructed and the ergodicity condition is derived. Formulas for computation of the main performance measures of the system based on the stationary distribution of the Markov chain are derived. Numerical examples are presented.

Keywords: Multi-Server Queueing, Multi-Dimensional Markov Chain, Random Environment, Cell Operations

* Corresponding Author, E-mail: dowoo@sangji.ac.kr

1. INTRODUCTION

In mobile communication network, the entire coverage area is divided into cells. Indeed cell is a coverage area of a base station antenna. Nearby cells overlap, and an assembly of cells forms a network. Since users of a cellular network can move during communication sessions from one cell to another, the network has to perform a handover procedure from one base station to another without losing the connection. This procedure is called handover. So, in the cell the communication sessions generated in this cell and received from other cells (handover customers) should be simultaneously serviced. Different systems use different methods for processing the handover call. If the network does not give priority to the handover calls over the calls generated in a given cell, the probability of the current session interruption due to the movement of the subscriber will be equal to the probability of failure in the initialization of a new session. However, it is obvious that the current subscriber disconnection, for example, interruption of conversation, is much more irritant than the outgoing call drop. Therefore, to reduce the chance of losing the current communication sessions during the handover procedure, handover requests are given a priority. One of the possible ways of providing such a priority is a Guard Channel Concept, in which part of the communication channel is reserved exclusively for the maintenance of communication sessions that can be transferred into the cell from the outside. In other words, new sessions are blocked if the number of busy channels exceeds a certain threshold. The paper (Tran-Gia and Mandjes, 1997) considers a special case of this strategy, when only one channel is reserved for handover customer. In general, it is assumed that M channels are reserved, and the problem of finding such a number of channels $\bar{M}$, for which the system would be operating optimally according to a predetermined criterion is solved.

To solve such a type of problems, usually operation of a cell is modelled by a queueing system with heterogeneous customers and retrials, see, e.g., (Choi *et al*., 2008; Do, 2011; Zhou and Zhu, 2013; Kim *et al*., 2014). However, existing in the literature queueing models have some shortcomings which reduce their adequacy to real world systems. The shortcomings of the work (Choi *et al*., 2008) consists of assumptions that arrivals of handover and new customers are defined as the stationary Poisson processes and that the service time distribution is the same for different types of customers. In (Do, 2011), it is also assumed that arrivals of handover and new customers are defined as the stationary Poisson processes. In the papers (Do, 2011) and (Zhou and Zhu, 2013), it is assumed that the total intensity of retrials is constant, not dependent on the current number of retrying customers. As disadvantages of the model considered in (Kim *et al*., 2014) as well as in some other papers, the following two may be mentioned: (i) it is assumed that the blocked handover customers do not make retrials and permanently leave the system; (ii) customers are patient, they cannot leave the cell without getting service. The model considered in this paper is free of these disadvantages.

It is well known that quality of operation of the wireless communication networks may essentially depend on the weather conditions, noise in the transmission thread, including the natural and technogeneous ones, failures and breakdowns of equipment. Traffic intensity can be essentially influenced by time of the day or night, migration of users, breaking news, etc. All these factors may be random. Account of the influence of some random factors of characteristics of a service system, including telecommunication system, is usually performed via consideration of so called queues operating in the random environment (*RE*), for recent references see, e.g., (Kim *et al*., 2009; Kim *et al*., 2010; Wu *et al*., 2001; Cordeiro and Kharoufeh, 2012; Yang *et al*., 2013). The standard assumption in the existing literature about the queues operating in the *RE* is that the change of the state of the *RE* instantaneously causes the change of intensities of arrivals, service, retrials, impatience, etc. Queues where the change of the state of the *RE* possibly causes the change of the number of servers are more complicated for research. This is because the possible change, especially decrease of the number of available servers, due to the change of the state of the *RE*, creates difficulties in construction of the multi-dimensional Markov chain which describes behavior of the queueing system. In this paper, we analyse a queueing model of the cell operation where the number of available servers and the number of reserved servers may depend on some external random factors. This model suits, e.g., for description of situations when the dynamic redistribution of frequencies between the cells is possible, e.g., redistribution of available frequencies between the cells located in bedroom suburbs and city business districts during a day and night time. Similar to our model was recently analyzed in (Dudin *et al*., 2015). That model has another potential field of applications (modeling of cognitive radio systems). Primary customers have the preemptive priority over the cognitive customers. While in our model handover customers have the non-preemptive priority over the new customers. From the point of view of mathematical compexity of analysis, the model analyzed in (Dudin *et al*., 2015) is more simple because it was assumed in (Dudin *et al*., 2015) that the customers, service of which is terminated due to the reduction of the number of available servers caused by the change of the state of the *RE*, immediately leave the system. This guarantees that the generator of the multi-dimensional continuous time Markov chain describing behavior of the system has the three-block-diagonal structure. In our model, the generator can have more than three non-zero diagonals.

The paper is organized as follows. Mathematical model is described in section 2. In section 3, behavior of the considered model is described by the multi-dimensional continuous time Markov chain and the generator of this Markov chain is written down. Section 4 is addressed to derivation of condition of the stable operation of the system and computation of the stationary distribution of the Markov chain. Expressions for some key performance measures of the system are presented in section 5. Numerical examples are presented in section 6.

2. MATHEMATICAL MODEL

We consider a retrial multi-server queueing model without buffer and with two types of customers. The structure of the system is presented in Figure 1.

The system operation depends on the state of the *RE*. The *RE* is defined by the stochastic process $r_t,\ t \geq 0,$ which is an irreducible continuous time Markov chain with the finite state space $\{1, 2, \cdots, R\}$ and the infinitesimal generator G.

We assume that the number of servers depends on the state of the *RE*. So, under the fixed state $r,\ r = 1, \cdots, R,$ of the *RE*, the number of available servers is $N^{(r)}$. Without loss of generality, we assume that the states of the *RE*are enumerated in ascending order the number of servers, i.e. $N^{(1)} \leq N^{(2)} \leq \cdots \leq N^{(R)}$.

Arrival of customers is modelled by the marked Markovian arrival process (*MMAP*) and is described by the underlying process $\{r_t, \nu_t\},\ t \geq 0,$ where the process ν_t with a finite state space $\{0, 1, \cdots, W\}$ is defined as follows. Under the fixed state r of the *RE* the process ν_t

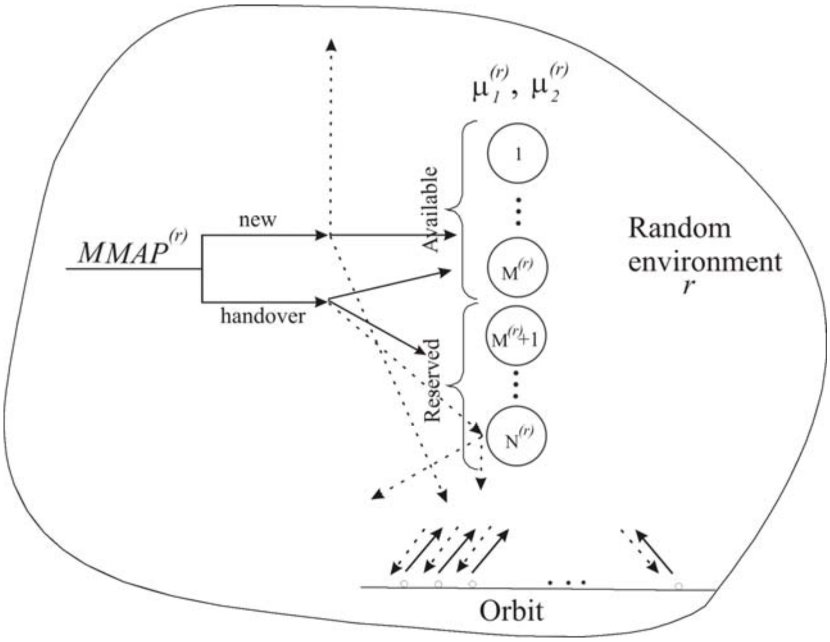

Figure 1. Structure of the system.

behaves as an irreducible continuous time Markov chain. The sojourn time of the chain $\nu_t, t \geq 0,$ in the state ν is exponentially distributed with the positive finite parameter $\lambda_\nu^{(r)}$, $\nu = 0, \cdots, W$, $r = 1, \cdots, R.$ When the sojourn time in the state ν expires, with probability $p_0^{(r)}(\nu, \nu')$ the process ν_t jumps to the state ν' without generation of customers, $\nu, \nu' = 0, \cdots, W, \nu \neq \nu', r = 1, \cdots, R.$ With probability $p_1^{(r)}(\nu, \nu')$ the process ν_t jumps to the state ν' (probably the same) with generation of type-1 customer, and with probability $p_2^{(r)}(\nu, \nu')$ the process ν_t jumps to the state ν' with generation of type-2 customer, $\nu, \nu' = 0, \cdots, W,$ $r = 1, \cdots, R.$

The behavior of the $MMAP$ under the fixed state r of the RE is completely characterized by the matrices $D_0^{(r)}$, $D_1^{(r)}$, and $D_2^{(r)}$ defined by the entries

$$(D_0^{(r)})_{\nu,\nu} = -\lambda_\nu^{(r)}, \nu = 0, \cdots, W,$$
$$(D_0^{(r)})_{\nu,\nu'} = \lambda_\nu^{(r)} p_0^{(r)}(\nu, \nu'),\ \nu, \nu' = 0, \cdots, W, \nu \neq \nu',$$
$$(D_l^{(r)})_{\nu,\nu'} = \lambda_\nu^{(r)} p_l^{(r)}(\nu, \nu'),\ \nu, \nu' = 0, \cdots, W,$$
$$l = 1, 2,\ r = 1, \cdots, R.$$

For more information about the $MMAP$, its parameters and features, see (He, 1996). Characteristics of the $MMAP$ influenced by the RE are given in (Dudin and Nazarov, 2015). For information about the $MMAP$ parameters fitting see, e.g., (Buchholz *et al*., 2010) and references therein.

We assume that during the epochs of transition of the RE the intensities of the process $\nu_t,\ t \geq 0,$ transitions are changed, but the state of this process does not change.

If during the arrival epoch of type-1 customer (priority customer, handover customer) there is a free server, the customer is admitted to the system and starts service. If at the arrival epoch of type-1 customer all servers are busy, then under the fixed state of the environment r this customer leaves the system with probability $q_1^{(r)}$, or with the complimentary probability $1 - q_1^{(r)}$, $r = 1, \cdots, R,$ goes into orbit and tries again later.

We assume that if, under the fixed state r, $r = 1, \cdots, R,$ of the RE, type-2 customer (non-priority customer, call that is generated within the cell) is admitted to the system only if the number of busy servers is less than the threshold $M^{(r)}$ where $N^{(r)} \geq M^{(r)} > 0.$ If during type-2 customer arrival epoch the number of busy servers is greater than or equal to $M^{(r)}$, this customer goes to orbit with probability $q_2^{(r)}$ and with the complimentary probability leaves the system, $r = 1, \cdots, R.$

Since the states of the RE are numbered in ascending order of the number of servers, it is reasonable to assume that the numbers of reserved servers also satisfy the inequality $M^{(1)} \leq M^{(2)} \leq \cdots \leq M^{(R)}.$

If the change of the state of the RE leads to a decrease of the number of available servers, we assume that the first of all the appropriate number of idle servers become unavailable. If this is not enough, the suitable number of servers providing service to type-2 customers is switched-off. If this is again not enough, the suitable number of servers providing service to type-1 customers is switched-off.

In the case of switching-off the servers (i.e., interruption and forced termination of customers service), under the fixed state of the environment r, each interrupted customer leaves the system with probability $p^{(r)}$, $r = 1, \cdots, R,$ independently of its type. With the complimentary probability, each interrupted customer goes to orbit.

Customers in orbit are indistinguishable, and under the fixed state $r,\ r=1,\cdots,R,$ of the RE, make repeated attempts to get service through exponentially distributed with the parameter $\alpha^{(r)},\ \alpha^{(r)}>0,$ time independently of other customers. If the attempt to get service is performed when the number of busy servers is greater than or equal to $M^{(r)}$, then the customer returns to orbit with probability $q_3^{(r)}$, and with the complimentary probability leaves the system, i.e., the customers from orbit can be nonpersistent.

Also, the customers from orbit may be impatient. We assume that, under the fixed state r of the RE, a customer leaves the orbit after an exponentially distributed with the parameter $\beta^{(r)},\ \beta^{(r)}>0,$ time since entering the orbit. If we assume that the customers from orbit are absolutely patient, we put $\beta^{(r)}=0,\ r=1,\cdots,R.$

The service time of type-l, $l=1,2,$ customer under the fixed state r of the RE has an exponential distribution with the parameter $\mu_l^{(r)},\ r=1,\cdots,R.$

3. PROCESS OF THE SYSTEM STATES

Let

- $i_t,\ i_t\ge 0,$ be the number of customers in orbit,
- $r_t,\ r_t=1,\cdots,R,$ be the state of the RE,
- $n_t,\ n_t=0,\cdots,N^{(r_t)},$ be the number of busy servers,
- $m_t, m_t=0,\cdots,\min\{n_t, M^{(r_t)}\},$ be the number of type-2 customers on service,
- $\nu_t, \nu_t=0,\cdots,W,$ be the state of the $MMAP$ underlying process during the epoch $t,\ t\ge 0.$

The process $\xi_t=\{i_t, r_t, n_t, m_t, \nu_t\},\ t\ge 0,$ is an irreducible continuous-time Markov chain.

Let us introduce the following notation:

- I is the identity matrix, O is a zero matrix of an appropriate dimension;
- $\mathbf{e}$ is a column vector consisting of ones, $\mathbf{0}$ is a row zero vector;
- $\bar{W}=W+1;$
- $\oplus$ and $\otimes$ are the symbols of Kroneker sum and product of matrices, see, e.g., (Graham, 1981);
- $K_r=(M^{(r)}+1)(N^{(r)}+1-M^{(r)}/2),\ r=1,\cdots,R;$
- $\mathrm{diag}\{A_1,\cdots,A_l\}$ is a block-diagonal matrix with the diagonal blocks $A_1,\cdots,A_l;$
- $C_n=\mathrm{diag}\{0,1,\cdots,n\}, \bar{C}_n=\mathrm{diag}\{n, n-1,\cdots,0\},$ $n=0,\ \cdots, M^{(R)};$
- $C_n^{(r)}=\mathrm{diag}\{n, n-1,\cdots, n-M^{(r)}+1, n-M^{(r)}\},$ $n=M^{(r)},\cdots,N^{(r)},\ r=1,\cdots,R;$
- $E_n^+,\ \bar{E}_n^+,\ n=0,\cdots,M^{(R)}-1,$ are the matrices of size $(n+1)\times(n+2),$ that are defined as follows

$$E_n^+=\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&0&\cdots&0&1\end{pmatrix},\ \bar{E}_n^+=\begin{pmatrix}1&0&\cdots&0&0\\0&1&\cdots&0&0\\ \vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&1&0\end{pmatrix};$$

- $E_n^-,\ \bar{E}_n^-,\ n=1,\cdots,M^{(R)},$ are the matrices of size $(n+1)\times n,$ that are defined as follows

$$E_n^-=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\\0&0&\cdots&0\end{pmatrix},\ \bar{E}_n^-=\begin{pmatrix}0&0&0&\cdots&0\\1&0&0&\cdots&0\\0&1&0&\cdots&0\\ \vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\end{pmatrix};$$

- $\tilde{I}_r,\ r=1,\cdots,R,$ are the diagonal matrix of size K_r with diagonal entries $\{\underbrace{0,\cdots,0,}_{(M^{(r)}+1)M^{(r)}/2}\ 1,\cdots,1\};$
- $\tilde{E}_r^-,\ r=1,\cdots,R,$ are the square matrices of size $M^{(r)}+1,$ that are defined as follows

$$\tilde{E}_r^-=\begin{pmatrix}0&0&0&\cdots&0&0\\1&0&0&\cdots&0&0\\0&1&0&\cdots&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&0&\cdots&1&0\end{pmatrix};$$

- $p(k,n)=\begin{cases}C_n^k(1-p^{(r)})^k(p^{(r)})^{n-k}, & k\le n,\\ 0, & k>n;\end{cases}$
- $\bar{N}=\max\{\max\{N^{(R)}-N^{(1)}, M^{(R)}-M^{(1)}\},1\}.$

Let us enumerate the state of the Markov chain ξ_t in the lexicographic order of the components (i, r, n, m, ν). We call the set of the states of the chain having the value (i, r) of two first components as macro-state (i, r).

Let A be the generator of the Markov chain ξ_t, $t\ge 0,$ that is formed by the blocks $A_{i,j}$, consisting of matrices $(A_{i,j})_{r,r'}$ that define the intensities of transitions of the Markov chain $\xi_t,\ t\ge 0,$ from macro-state (i, r) to macro-state $(j, r'),\ r, r'=1,\cdots,R.$ The diagonal entries of the matrix $A_{i,i}$ are negative and the modulus of each diagonal entry defines the intensity of leaving corresponding state of the Markov chain $\xi_t, t\ge 0.$

Theorem 1. The generator A has the following structure:

$$A=\begin{pmatrix}A_{0,0}&A_{0,1}&A_{0,2}&\ldots&A_{0,\bar{N}}&O&O&\ldots\\A_{1,0}&A_{1,1}&A_{1,2}&\ldots&A_{1,\bar{N}}&A_{1,\bar{N}+1}&O&\ldots\\O&A_{2,1}&A_{2,2}&\ldots&A_{2,\bar{N}}&A_{2,\bar{N}+1}&A_{2,\bar{N}+2}&\ldots\\ \vdots&\vdots&\vdots&\ldots&\vdots&\vdots&\vdots&\ddots\end{pmatrix}.$$

The non-zero blocks $A_{i,j},\ i,j\ge 0,$ have the following structure:

$$A_{i,i}=(A_{i,i})_{r,r'},\ r,r'=1,\cdots,R,$$

$$(A_{i,i})_{r,r}=\begin{pmatrix} L_r^{i,0} & B_0^{(r)} & O & \dots & O & O \\ F_1^{(r)} & L_r^{i,1} & B_1^{(r)} & \dots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & O & \dots & L_r^{i,N^{(r)}-1} & B_{N^{(r)}-1}^{(r)} \\ O & O & O & \dots & F_{N^{(r)}}^{(r)} & L_r^{i,N^{(r)}} \end{pmatrix} + q_2^{(r)}\tilde{I}_r\otimes D_2^{(r)}+I_{K_r}\otimes D_0^{(r)}+(G)_{r,r}I_{K_r\bar{W}}, i\ge 0,$$

$$(A_{i,i})_{r,r'}=(G)_{r,r'}Q_{r,r'}^{(0)}\otimes I_{\bar{W}},\ r'<r,\quad (A_{i,i})_{r,r'}=(G)_{r,r'}Q_{r,r'}^{+}\otimes I_{\bar{W}}, r'>r,$$

$$A_{i,i+k}=\begin{pmatrix} Z_{1,1}^{(k)} & O & O & \dots & O & O \\ Z_{2,1}^{(k)} & Z_{2,2}^{(k)} & O & \dots & O & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Z_{R-1,1}^{(k)} & Z_{R-1,2}^{(k)} & Z_{R-1,3}^{(k)} & \dots & Z_{R-1,R-1}^{(k)} & O \\ Z_{R,1}^{(k)} & Z_{R,2}^{(k)} & Z_{R,3}^{(k)} & \dots & Z_{R,R-1}^{(k)} & Z_{R,R}^{(k)} \end{pmatrix},\ i\ge 0,\quad A_{i,i-1}=\mathrm{diag}\{A_{i,i-1}^{(1)},\cdots,A_{i,i-1}^{(R)}\},\ i\ge 1,$$

where

- $$L_r^{i,n}=\begin{cases} -[\mu_2^{(r)}C_n+\mu_1^{(r)}\bar{C}_n+i(\alpha^{(r)}+\beta^{(r)})I_{n+1}]\otimes I_{\bar{W}}, & n<M^{(r)}, \\ -[\mu_2^{(r)}C_{M^{(r)}}+\mu_1^{(r)}C_n^{(r)}+i(q_3^{(r)}\alpha^{(r)}+\beta^{(r)})I_{M^{(r)}+1}]\otimes I_{\bar{W}}, & M^{(r)}\le n<N^{(r)}, \\ -[\mu_2^{(r)}C_{M^{(r)}}+\mu_1^{(r)}C_n^{(r)}+i(q_3^{(r)}\alpha^{(r)}+\beta^{(r)})I_{M^{(r)}+1}]\otimes I_{\bar{W}} & \\ \quad +q_1^{(r)}I_{M^{(r)}+1}\otimes D_1^{(r)}, & n=N^{(r)}, \end{cases}\quad i\ge 0;$$
- $$B_n^{(r)}=\begin{cases} E_n^{+}\otimes D_2^{(r)}+\bar{E}_n^{+}\otimes D_1^{(r)}, & n<M^{(r)}, \\ I_{M^{(r)}+1}\otimes D_1^{(r)}, & M^{(r)}\le n<N^{(r)}; \end{cases}$$
- $$F_n^{(r)}=\begin{cases} (\mu_2^{(r)}C_n\bar{E}_n^{-}+\mu_1^{(r)}\bar{C}_nE_n^{-})\otimes I_{\bar{W}}, & n\le M^{(r)}, \\ (\mu_2^{(r)}C_{M^{(r)}}\tilde{E}_r^{-}+\mu_1^{(r)}C_n^{(r)})\otimes I_{\bar{W}}, & M^{(r)}<n\le N^{(r)}; \end{cases}$$

- $Q_{r,r'}^{(k)},\ r=1,\cdots,R,\ r'=1,\cdots,r-1,$ are the block matrices of size $K_r\times K_{r'},$ consisting of non-zero blocks $(Q_{r,r'}^{(l)})_{n,n'},\ n=0,\cdots,N^{(r)},\ n'=0,\cdots,\min\{n,\cdots,N^{(r')}\},$ that are defined as:

$$(Q_{r,r'}^{(0)})_{n,n}=I_{n+1},\ n\le M^{(r')},\ (Q_{r,r'}^{(k)})_{n,n}=O_{n+1},\ n\le M^{(r')},\ k>0;$$

$(Q_{r,r'}^{(k)})_{n,n},\ n=M^{(r')}+1,\ \cdots,\ N^{(r')},$ are the matrices of size $(\min\{n,M^{(r)}\}+1)\times(M^{(r')}+1)$, that have the following form:

$$(Q_{r,r'}^{(0)})_{n,n}=\begin{pmatrix} I_{M^{(r')}+1} \\ O \end{pmatrix},\ (Q_{r,r'}^{(k)})_{n,n}=O, k>0;$$

$(Q_{r,r'}^{(k)})_{n,n'},\ n=M^{(r')}+1,\cdots,N^{(r')},\ n'=\min\{M^{(r')},n-M^{(r)}+M^{(r')}\},\cdots,n-1,$ are the matrices of size $(\min\{n,M^{(r)}\}+1)\times(M^{(r')}+1)$ with all zero entries except the entry $((Q_{r,r'}^{(k)})_{n,n'})_{n-n'+M^{(r')},M^{(r')}}=p(k,n-n')$;

$(Q_{r,r'}^{(k)})_{n,N^{(r')}},\ n=N^{(r')}+1,\cdots,N^{(r)},$ are the matrices of size $(\min\{M^{(r)},n\}+1)\times(M^{(r')}+1)$ with all zero entries except the entries $((Q_{r,r'}^{(k)})_{n,N^{(r')}})_{m,0}, m=0,\cdots,\min\{M_r,n-N^{(r')}\},$ $((Q_{r,r'}^{(k)})_{n,N^{(r')}})_{m,m-(n-N^{(r')})},\ m=n-N^{(r')}+1,\cdots,\min\{M^{(r)},M^{(r')}+n-N^{(r')}\},$ that are equal to $p(k,n-N^{(r')})$;

$(Q_{r,r'}^{(k)})_{n,n'},\ n=N^{(r')}+1,\cdots,N^{(r)},\ n'=\max\{M^{(r')},N^{(r')}-\max\{0,M^{(r)}-M^{(r')}-(n-N^{(r')})\}\},\cdots,N^{(r')}-1,$ are the matrices of size $(\min\{M^{(r')},n\}+1)\times(M^{(r')}+1)$ with all zero entries except the entry $((Q_{r,r'}^{(k)})_{n,n'})_{n-n'+M^{(r')},M^{(r')}}=p(k,n-n'),$

- $Q_{r,r'}^{+},\ r=1,\cdots,R-1,\ r'=r+1,\cdots,R,$ are the matrices of size $K_r\times K_{r'},$ that have the form $(\mathrm{diag}\{\Omega_0^{r,r'},\cdots,\Omega_{N^{(r)}}^{r,r'}\}\,|\,O)$

where

$$\Omega_n^{r,r'}=\begin{cases} I_{n+1}, & n\le M^{(r)}, \\ (I_{M^{(r)}+1}\,|\,O_{(M^{(r)}+1)\times(\min\{n-M^{(r)},M^{(r')}-M^{(r)}\})}), & n=M^{(r)}+1,\cdots,N^{(r)}; \end{cases}$$

- $Z_{r,r}^{(1)}=\mathrm{diag}\{\tilde{Z}_0^{(r)},\cdots,\tilde{Z}_{N^{(r)}}^{(r)}\},\ r=1,\cdots,R;$
- $Z_{r,r}^{(k)}=O, k>1,\ r=1,\cdots,R;$
- $Z_{r,r'}^{(k)}=(G)_{r,r'}Q_{r,r'}^{(k)}\otimes I_{\bar{W}},\ k\ge 1,\ r=1,\cdots,R,\ r'<r;$

- $\tilde{Z}_n^{(r)} = \begin{cases} O_{(n+1)\bar{W}}, & n < M^{(r)}, \\ (1-q_2^{(r)})I_{M^{(r)}+1} \otimes D_2^{(r)}, & M^{(r)} \le n < N^{(r)}, \\ (1-q_1^{(r)})I_{M^{(r)}+1} \otimes D_1^{(r)} + (1-q_2^{(r)})I_{M^{(r)}+1} \otimes D_2^{(r)}, & n = N^{(r)}; \end{cases}$

- $A_{i,i-1}^{(r)} = \begin{pmatrix} \tilde{L}_r^{i,0} & \tilde{B}_{i,0}^{(r)} & O & \dots & O & O \\ O & \tilde{L}_r^{i,1} & \tilde{B}_r^{i,1} & \dots & O & O \\ O & O & \tilde{L}_r^{i,2} & \dots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ O & O & O & \ddots & \tilde{L}_r^{i,N^{(r)}-1} & \tilde{B}_r^{i,N^{(r)}-1} \\ O & O & O & \dots & O & \tilde{L}_r^{i,N^{(r)}} \end{pmatrix}, i \ge 1;$

- $\tilde{B}_r^{i,n} = \begin{cases} i\alpha^{(r)} E_n^+ \otimes I_{\bar{W}}, & n < M^{(r)}, i \ge 1, \\ O_{(M^{(r)}+1)\bar{W}}, & n \ge M^{(r)}, i \ge 1; \end{cases}$

- $\tilde{L}_r^{i,n} = \begin{cases} i\beta^{(r)} I_{(n+1)\bar{W}}, & n < M^{(r)}, i \ge 1, \\ i(\beta^{(r)} + q_3^{(r)}\alpha^{(r)}) I_{(M^{(r)}+1)\bar{W}}, & M^{(r)} \le n \le N^{(r)}, i \ge 1. \end{cases}$

Proof. of Theorem 1 is implemented by analyzing all possible transitions of the Markov chain $\xi_t, t \ge 0,$ during an infinitesimal time interval and writing the intensities of these transitions in the block-matrix form.

4. ERGODICITY CONDITION AND THE STATIONARY PROBABILITIES OF THE SYSTEM STATES

To find the ergodicity condition of the system, we need the following assertion.

Lemma 1. The Markov chain ξ_t, $t \ge 0$, belongs to the class of the asymptotically quasi-Toeplitz continuous-time Markov chains, see (Klimenok and Dudin, 2006).

Proof. In order to prove that the Markov chain ξ_t, $t \ge 0$, belongs to the class of the asymptotically quasi-Toeplitz Markov chains, it is necessary to show the existence of matrices Y_k, $k = 0, \cdots, \bar{N}+1$, that are defined as

$$Y_0 = \lim_{t\to\infty} \bar{R}_i^{-1} A_{i,i-1},\ Y_1 = \lim_{t\to\infty} \bar{R}_i^{-1} A_{i,i} + I,\ Y_k = \lim_{t\to\infty} \bar{R}_i^{-1} A_{i,i+k-1},\ k = 2, \cdots, \bar{N}+1,$$

where the matrix $\bar{R}_i$ is the diagonal matrix with diagonal entries that are defined as the moduli of corresponding diagonal entries of the matrix $A_{i,i}$, $i \ge 0.$

It can be shown in a straight forward way that the explicit form of the matrices Y_k, $k = 0, \cdots, \bar{N}+1$, is given by:

$$Y_0 = \mathrm{diag}\{\tilde{\Omega}_1, \cdots, \tilde{\Omega}_R\},$$

$$Y_1 = \begin{pmatrix} \tilde{Q}_{1,1} & \tilde{Q}_{1,2} & \dots & \tilde{Q}_{1,R} \\ \tilde{Q}_{2,1} & \tilde{Q}_{2,2} & \dots & \tilde{Q}_{1,R} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{Q}_{R,1} & \tilde{Q}_{R,2} & \dots & \tilde{Q}_{R,R} \end{pmatrix},$$

$$Y_k = \begin{pmatrix} \tilde{Z}_{1,1}^{(k-1)} & O & O & \dots & O & O \\ \tilde{Z}_{2,1}^{(k-1)} & \tilde{Z}_{2,2}^{(k-1)} & O & \dots & O & O \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \tilde{Z}_{R-1,1}^{(k-1)} & \tilde{Z}_{R-1,2}^{(k-1)} & \tilde{Z}_{R-1,3}^{(k-1)} & \ddots & \tilde{Z}_{R-1,R-1}^{(k-1)} & O \\ \tilde{Z}_{R,1}^{(k-1)} & \tilde{Z}_{R,2}^{(k-1)} & \tilde{Z}_{R,3}^{(k-1)} & \dots & \tilde{Z}_{R,R-1}^{(k-1)} & \tilde{Z}_{R,R}^{(k-1)} \end{pmatrix},\ k = 2, \cdots, \bar{N}+1,$$

where

$$\tilde{Q}_{r,r'} = \begin{cases} R_r (A_{0,0})_{r,r'} + \delta_{r-r',0} \tilde{I}_r \otimes I_{\bar{W}}, & \text{if } q_3^{(r)} = 0 \text{ and } \beta^{(r)} = 0, \\ O, & \text{if } q_3^{(r)} \ne 0 \text{ or } \beta^{(r)} \ne 0, \end{cases} \quad r,\ r' = 1, \cdots, R,$$

$$\tilde{Z}_{r,r'}^{(k)} = \begin{cases} R_r Z_{r,r'}^{(k)}, & \text{if } q_3^{(r)} = 0 \text{ and } \beta^{(r)} = 0, \\ O, & \text{if } q_3^{(r)} \ne 0 \text{ or } \beta^{(r)} \ne 0, \end{cases} \quad r,\ r' = 1, \cdots, R,$$

$$\tilde{\Omega}_r = \begin{pmatrix} O & E_0^+ \otimes I_{\bar{W}} & O & \dots & O & \dots & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ O & O & O & \dots & E^+_{M^{(r)}-1} \otimes I_{\bar{W}} & \dots & O \\ O & O & O & \dots & O & \dots & O \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ O & O & O & \dots & O & \dots & O \end{pmatrix}, \text{ if } q_3^{(r)} = 0 \text{ and } \beta^{(r)} = 0,\ r = 1, \cdots, R,$$

$$\tilde{\Omega}_r = \begin{pmatrix} \frac{\beta^{(r)}}{\beta^{(r)}+\alpha^{(r)}} I_{\bar{W}} & \frac{\alpha^{(r)}}{\beta^{(r)}+\alpha^{(r)}} E_0^+ \otimes I_{\bar{W}} & \dots & O & O & O & \dots & O \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & \frac{\beta^{(r)}}{\beta^{(r)}+\alpha^{(r)}} I_{M^{(r)}\bar{W}} & \frac{\alpha^{(l)}}{\beta^{(r)}+\alpha^{(r)}} E^+_{M^{(r)}} \otimes I_{\bar{W}} & O & \dots & O \\ O & O & \dots & O & I_{(M^{(l)}+1)\bar{W}} & O & \dots & O \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & O & O & O & \dots & O \\ O & O & \dots & O & O & O & \dots & I_{(M^{(r)}+1)\bar{W}} \end{pmatrix},$$

if $q_3^{(r)} \neq 0$ or $\beta^{(r)} \neq 0$,

$R_r = \mathrm{diag}\{R_r^{(n)},\ n = 0, \cdots, N^{(r)}\}$,

$$R_r^{(n)} = \begin{cases} O_{(n+1)\bar{W}}, & n < M^{(r)}, \\ ((\mu_2^{(r)} C_{M^{(r)}} + \mu_1^{(r)} C_n^{(r)}) \otimes I_{\bar{W}} + \\ + I_{M^{(r)}+1} \otimes \Sigma_0^{(r)} - q_2^{(r)} I_{M^{(r)}+1} \otimes \Sigma_2^{(r)} - (G)_{r,r} I_{(M^{(r)}+1)\bar{W}})^{-1}, & M^{(r)} \leq n < N^{(r)}, \\ ((\mu_2^{(r)} C_{M^{(r)}} + \mu_1^{(r)} C_n^{(r)}) \otimes I_{\bar{W}} + I_{M^{(r)}+1} \otimes \Sigma_0^{(r)} - \\ - q_2^{(r)} I_{M^{(r)}+1} \otimes \Sigma_2^{(r)} - (G)_{r,r} I_{(M^{(r)}+1)\bar{W}} - q_1^{(r)} I_{M^{(r)}+1} \otimes \Sigma_1^{(r)})^{-1}, & n = N^{(r)}. \end{cases}$$

Here $\Sigma_0^{(r)}$, $\Sigma_1^{(r)}$ and $\Sigma_2^{(r)}$ are the diagonal matrices whose diagonal elements are determined by the corresponding diagonal elements of the matrices $-D_0^{(r)}$, $D_1^{(r)}$ and $D_2^{(r)}$, respectively. Thus, the lemma is proved.

Having proved that the Markov chain ξ_t, $t \geq 0$, belongs to the class of the asymptotically quasi-Toeplitz Markov chains, we can use the results previously obtained for this type of chains for derivation of conditions of the existence of the stationary mode of operation and computation of the stationary probabilities of the system states. As follows from (Klimenok and Dudin, 2006), a sufficient condition for the existence of the stationary distribution of the asymptotically quasi-Toeplitz Markov chain is the fulfillment of the inequality

$$\mathbf{y} Y_0 \mathbf{e} > \mathbf{y} \sum_{k=2}^{\bar{N}+1} (k-1) Y_k \mathbf{e} \tag{1}$$

where the row vector $\mathbf{y}$ is the unique solution to the system

$$\mathbf{y} \sum_{k=0}^{\bar{N}+1} Y_k = \mathbf{y}, \quad \mathbf{y}\mathbf{e} = 1. \tag{2}$$

In general, to check whether the Markov chain under study is ergodic or not, it is necessary to solve finite system of Eq. (2) on a computer, substitute the obtained solution to inequality (1) and verify the fulfillment of inequality (1). However, under certain conditions this steps can be avoided. Namely, it can be shown that if the customers from orbit are nonpersistent or impatient at least for one state of the *RE*, i.e. $\exists r,\ r \in \{1, \cdots, R\}$, that $q_3^{(r)} \neq 0$ or $\beta^{(r)} \neq 0$, then the Markov chain ξ_t is ergodic for any sets of other system parameters.

In what follows, we assume that the ergodicity condition is fulfilled. Then the following limits (stationary probabilities) exist:

$$p(i, r, n, m, \nu) = \lim_{t \to \infty} P\{i_t = i,\ r_t = r, n_t = n,\ m_t = m, \nu_t = \nu\},$$
$$i \geq 0,\ r = 1, \cdots, R,\ n = 0, \cdots, N^{(r)},$$
$$m = 0, \cdots, \min\{n, M^{(r)}\},\ \nu = 0, \cdots, W.$$

Let us form the row vectors $\mathbf{p}_i$ as follows:

$$\mathbf{p}(i, r, n, m) = (p(i, r, n, m, 0),\ p(i, r, n, m, 1), \cdots,$$
$$p(i, r, n, m, W)),\ m = 0, \cdots, \min\{n, M^{(r)}\},$$
$$\mathbf{p}(i, r, n) = (\mathbf{p}(i, r, n, 0), \mathbf{p}(i, r, n, 1), \cdots, \mathbf{p}(i, r, n, \min\{n, M^{(r)}\})),$$
$$n = 0, \cdots, N^{(r)},$$
$$\mathbf{p}(i, r) = (\mathbf{p}(i, r, 0),\ \mathbf{p}(i, r, 1), \cdots, \mathbf{p}(i, r, N^{(r)})),\ r = 1, \cdots, R,$$
$$\mathbf{p}_i = (\mathbf{p}(i, 1), \mathbf{p}(i, 2), \cdots, \mathbf{p}(i, R)),\ i \geq 0.$$

It is well known that the vectors $\mathbf{p}_i,\ i \ge 0,$ satisfy the system

$$(\mathbf{p}_0, \mathbf{p}_1, \cdots)A = \mathbf{0},\ (\mathbf{p}_0, \mathbf{p}_1, \cdots)\mathbf{e} = 1. \qquad (3)$$

System (3) has infinitely many equations and unknowns, and solution of this system is a quite difficult task. However, the numerically stable algorithm developed in (Klimenok and Dudin, 2006) can be successfully applied for computing the vectors $\mathbf{p}_i,\ i \ge 0.$ This algorithm does not assume solution of system (3). Instead, using the notion of the so called censored Markov chain, alternative to (3) system of equations for the vectors $\mathbf{p}_i,\ i \ge 0,$ is derived and solved.

The main steps of this algorithm are as follows.

- Calculate the matrix G as the minimal nonnegative solution of the matrix equation $G = \sum_{k=0}^{\bar{N}+1} Y_k G^k.$
- Calculate the matrices G_i by using the backward recursion
$$G_i = (-\sum_{n=i+1}^{i+\bar{N}+1} A_{i+1,n} G_{n-1} G_{n-2} \cdots G_{i+1})^{-1} A_{i+1,i},\ i \ge 0,$$
with the boundary condition $G_i = G,\ \ i \ge i_0$,

where the value i_0 is adaptively chosen as the minimal integer for which the norm of the matrix $G_{i_0-1} - G$ is less than some pre-assigned small value.

- Calculate the matrices $\bar{A}_{i,l},\ i \ge 0,\ \bar{N}+i \ge l \ge i,$ by the formula
$$\bar{A}_{i,l} = A_{i,l} + \sum_{n=l+1}^{\bar{N}+i} A_{i,n} G_{n-1} G_{n-2} \cdots G_l,\ l \ge i,\ i \ge 0.$$
- Calculate the matrices $F_l,\ l \ge 0,$ using the recursion
$$F_0 = I,\ F_l = \sum_{i=\max\{0,l-\bar{N}\}}^{l-1} F_i \bar{A}_{i,l} (-\bar{A}_{l,l})^{-1},\ l \ge 1.$$
- Calculate the vector $\mathbf{p}_0$ as the unique solution to the system of linear algebraic equations
$$\mathbf{p}_0(-\bar{A}_{0,0}) = \mathbf{0},\ \mathbf{p}_0 \sum_{i=0}^{\infty} F_i\,\mathbf{e} = 1.$$
- Calculate the vectors $\mathbf{p}_i,\ i \ge 1,$ by
$$\mathbf{p}_i = \mathbf{p}_0 F_i,\ \ i \ge 1.$$

Numerical stability of this algorithm stems from the fact that subtraction operation is not used in computations. All the involved matrices have non-negative entries.

5. PERFORMANCE MEASURES

After finding the stationary probabilities $\mathbf{p}_i,\ i \ge 0,$ it is possible to find the main performance measures of the system under study.

Distribution of the customers in orbit is defined as
$$\lim_{t\to\infty} P\{i_t = i\} = \mathbf{p}_i \mathbf{e},\ i \ge 0.$$

The average number of customers in orbit is calculated as

$$L_{orbit} = \sum_{i=1}^{\infty} i \mathbf{p}_i \mathbf{e}.$$

The mean number of customers in the system and orbit is

$$L = \sum_{i=0}^{\infty} \sum_{r=1}^{R} \sum_{n=0}^{N^{(r)}} (i+n) \mathbf{p}(i, r, n)\mathbf{e}.$$

The mean number of busy servers

$$N_{server} = \sum_{i=0}^{\infty} \sum_{r=1}^{R} \sum_{n=1}^{N^{(r)}} n \mathbf{p}(i, r, n)\mathbf{e}.$$

The average number of busy servers processing type-1 customers is defined as

$$N_{server}^{(1)} = \sum_{i=0}^{\infty} \sum_{r=1}^{R} \sum_{n=1}^{N^{(r)}} \sum_{m=0}^{\min\{n, M^{(r)}\}} (n-m) \mathbf{p}(i, r, n, m)\mathbf{e}.$$

The average number of busy servers processing type-2 customers is defined as

$$N_{server}^{(2)} = \sum_{i=0}^{\infty} \sum_{r=1}^{R} \sum_{n=1}^{N^{(r)}} \sum_{m=1}^{\min\{n, M^{(r)}\}} m \mathbf{p}(i, r, n, m)\mathbf{e} = N_{server} - N_{server}^{(1)}.$$

The intensity of output flow of type-1 customers is calculated as

$$\lambda_{out}^{(1)} = \sum_{i=0}^{\infty} \sum_{r=1}^{R} \sum_{n=1}^{N^{(r)}} \sum_{m=0}^{\min\{n, M^{(r)}\}} (n-m) \mu_1^{(r)} \mathbf{p}(i, r, n, m)\mathbf{e}.$$

The intensity of output flow of type-2 customers is calculated as

$$\lambda_{out}^{(2)} = \sum_{i=0}^{\infty} \sum_{r=1}^{R} \sum_{n=1}^{N^{(r)}} \sum_{m=1}^{\min\{n, M^{(r)}\}} m \mu_2^{(r)} \mathbf{p}(i, r, n, m)\mathbf{e}.$$

The intensity of output flow of customers is calculated as

$$\lambda_{out} = \lambda_{out}^{(1)} + \lambda_{out}^{(2)}.$$

Let us denote as λ_l the mean intensity of type-l customers arrival, $l = 1, 2.$ It can be shown that this intensity is defined by formula:

$$\lambda_l = \mathbf{q}\,\mathrm{diag}\{D_l^{(r)}, r = 1, \cdots, R\}\mathbf{e}$$

where the vector $\mathbf{q}$ is the unique solution of the following system:

$$\mathbf{q}(G \otimes I_{\bar{W}} + \mathrm{diag}\{D_0^{(r)} + D_1^{(r)} + D_2^{(r)}, r = 1, \cdots, R\}) = \mathbf{0},\ \ \mathbf{qe} = 1.$$

The probability that during an arbitrary type-1 customer arrival epoch all servers are busy and this customer leaves the system is defined as

$$P_1^{(loss-ent)} = \lambda_1^{-1}\sum_{i=0}^{\infty}\sum_{r=1}^{R} q_1^{(r)}\mathbf{p}(i, r, N^{(r)})(I_{M^{(r)}+1} \otimes D_1^{(r)})\mathbf{e}.$$

The probability that during an arbitrary type-1 customer arrival epoch all servers are busy and this customer goes to orbit is defined as

$$P_1^{(orb-ent)} = \lambda_1^{-1}\sum_{i=0}^{\infty}\sum_{r=1}^{R}(1-q_1^{(r)})\mathbf{p}(i, r, N^{(r)})(I_{M^{(r)}+1} \otimes D_1^{(r)})\mathbf{e}.$$

The loss probability of type-1 customer is defined as

$$P_1^{(loss)} = 1 - \frac{\lambda_{out}^{(1)}}{\lambda_1}.$$

The loss probability of type-1 customer caused by a decrease in the number of servers due to change the state of the RE is given by

$$P_1^{(loss-RE)} = \lambda_1^{-1}\sum_{i=0}^{\infty}\sum_{r=2}^{R}\sum_{r'=1}^{r-1}(G)_{r,r'}\sum_{n=N^{(r')}+1}^{N^{(r)}}\sum_{m=0}^{\min\{n-1,M^{(r)}\}}$$
$$p^{(r)}\max\{0, n-N^{(r')}-m\}\mathbf{p}(i, r, n, m)\mathbf{e}.$$

The probability that type-1 customer will go into orbit due to termination of its service caused by a decrease in the number of servers under the change of the state of the RE is given by

$$P_1^{(orb-RE)} = \lambda_1^{-1}\sum_{i=0}^{\infty}\sum_{r=2}^{R}\sum_{r'=1}^{r-1}(G)_{r,r'}\sum_{n=N^{(r')}+1}^{N^{(r)}}\sum_{m=0}^{\min\{n-1,M^{(r)}\}}$$
$$(1-p^{(r)})\max\{0, n-N^{(r')}-m\}\mathbf{p}(i, r, n, m)\mathbf{e}.$$

The loss probability of type-2 customer is

$$P_2^{(loss)} = 1 - \frac{\lambda_{out}^{(2)}}{\lambda_2 + \lambda_1(P_1^{(orb-RE)} + P_1^{(orb-ent)})}.$$

The loss probability of type-2 customer caused by a decrease in the number of servers due to the change of the state of the RE is defined as

$$P_2^{(loss-RE)} = (\lambda_2 + \lambda_1(P_1^{(orb-RE)} + P_1^{(orb-ent)}))^{-1}\sum_{i=0}^{\infty}\sum_{r=2}^{R}\sum_{r'=1}^{r-1}(G)_{r,r'}$$
$$\sum_{n=M^{(r')}+1}^{N^{(r)}}\sum_{m=1}^{\min\{n-1,M^{(r)}\}} p^{(r)} \times (\min\{m, \max\{n-N^{(r')}, 0\}\}$$
$$+\max\{0, m-\max\{n-N^{(r')}, 0\}-M^{(r')}\})\mathbf{p}(i, r, n, m)\mathbf{e}.$$

The probability that type-1 customer will go into orbit because the termination of its service caused by a decrease in the number of servers due to the change of the state of the RE is given by

$$P_2^{(orb-RE)} = (\lambda_2 + \lambda_1(P_1^{(orb-RE)} + P_1^{(orb-ent)}))^{-1}\sum_{i=0}^{\infty}\sum_{r=2}^{R}\sum_{r'=1}^{r-1}(G)_{r,r'}$$
$$\sum_{n=M^{(r')}+1}^{N^{(r)}}\sum_{m=1}^{\min\{n-1,M^{(r)}\}}(1-p^{(r)}) \times (\min\{m, \max\{n-N^{(r')}, 0\}\}$$
$$+\max\{0, m-\max\{n-N^{(r')}, 0\}-M^{(r')}\})\mathbf{p}(i, r, n, m)\mathbf{e}.$$

The loss probability of an arbitrary customer is calculated as

$$P^{(loss)} = 1 - \frac{\lambda_{out}}{\lambda_1 + \lambda_2}.$$

The loss probability of an arbitrary type-2 customer due to the business of more than $M^{(r)}-1$ servers upon its arrival is defined as

$$P_2^{(loss-ent)} = (\lambda_2 + \lambda_1(P_1^{(orb-RE)} + P_1^{(orb-ent)}))^{-1}$$
$$\sum_{i=0}^{\infty}\sum_{r=1}^{R}\sum_{n=M^{(r)}}^{N^{(r)}} q_2^{(r)}\mathbf{p}(i, r, n)(I_{M^{(r)}+1} \otimes D_2^{(r)})\mathbf{e}.$$

The probability of an arbitrary type-2 customer goes to orbit due to more than $M^{(r)}-1$ busy servers upon its arrival is defined as

$$P_2^{(orbit-ent)} = (\lambda_2 + \lambda_1(P_1^{(orb-RE)} + P_1^{(orb-ent)}))^{-1}$$
$$\sum_{i=0}^{\infty}\sum_{r=1}^{R}\sum_{n=M^{(r)}}^{N^{(r)}}(1-q_2^{(r)})\mathbf{p}(i, r, n)(I_{M^{(r)}+1} \otimes D_2^{(r)})\mathbf{e}.$$

The probability of loss of an arbitrary customer from orbit is

$$P^{(loss-from-orbit)} = P_2^{(loss)} - P_2^{(loss-ent)} - P_2^{(loss-RE)}.$$

The probability that an arbitrary type-2 customer from orbit makes an attempt to get service when the number of busy servers exceeds $M^{(r)}-1$ and returns to orbit is calculated as

$$P^{(return-to-orbit)} = \alpha^{-1}\sum_{i=1}^{\infty}\sum_{r=1}^{R}\sum_{n=M^{(r)}}^{N^{(r)}} i\alpha^{(r)}(1-q_3^{(r)})\mathbf{p}(i, r, n)\mathbf{e}$$

where $\alpha = \sum_{i=1}^{\infty}\sum_{r=1}^{R} i\alpha^{(r)}\mathbf{p}(i, r)\mathbf{e}$.

The probability that an arbitrary type-2 customer from orbit makes an attempt to get service when the number of busy servers exceeds $M^{(r)}-1$ and leaves the system is calculated as

$$P_1^{(loss-nonpersistent)} = \alpha^{-1}\sum_{i=1}^{\infty}\sum_{r=1}^{R}\sum_{n=M^{(r)}}^{N^{(r)}} i\alpha^{(r)}q_3^{(r)}\,\mathbf{p}(i, r, n)\mathbf{e}.$$

6. NUMERICAL EXAMPLES

Let us consider the following set of the system parameters. The number of the states of the RE is $R = 2$.

Under state 1 of the *RE*, the system has the following parameters characterizing the number of available servers, impatience of customers and intensities of their retrials and service:

$$N^{(1)} = 6,\ p^{(1)} = 0.4,\ q_1^{(1)} = 0.2,\ q_2^{(1)} = 0.6,\ q_3^{(1)} = 0.4,$$
$$\gamma^{(1)} = 0.2,\ \alpha^{(1)} = 0.5,\ \mu_1^{(1)} = 0.5,\ \mu_2^{(1)} = 0.8.$$

Under state 2 of the *RE*, the parameters are as follows:

$$N^{(2)} = 12,\ p^{(2)} = 0.2,\ q_1^{(2)} = 0.2,\ q_2^{(2)} = 0.3,\ q_3^{(2)} = 0.3,$$
$$\gamma^{(2)} = 0.1,\ \alpha^{(2)} = 0.6,\ \mu_1^{(2)} = 0.3,\ \mu_2^{(2)} = 0.6.$$

We assume that the arrival flow under state 1 of the *RE* is defined by the matrices

$$D_0^{(1)} = \begin{pmatrix} -0.6759 & 0 \\ 0 & -0.021941 \end{pmatrix},\ D_1^{(1)} = \begin{pmatrix} 0.26856 & 0.0018 \\ 0.004886 & 0.00389 \end{pmatrix},$$
$$D_2^{(1)} = \begin{pmatrix} 0.40284 & 0.0027 \\ 0.00733 & 0.005835 \end{pmatrix},$$

and under state 2 of the *RE* it is defined by the matrices

$$D_0^{(2)} = \begin{pmatrix} -2.7036 & 0.0 \\ 0.0 & -0.087766 \end{pmatrix}, D_1^{(2)} = \begin{pmatrix} 0.6714 & 0.0045 \\ 0.012216 & 0.009725 \end{pmatrix},$$
$$D_2^{(2)} = \begin{pmatrix} 2.0142 & 0.0135 \\ 0.03665 & 0.029175 \end{pmatrix}.$$

The average arrival intensity $\lambda^{(1)}$ of customers under state 1 of the *RE* is 0.5 (the average arrival rate of handover customers is $\lambda_1^{(1)} = 0.2$, and the average arrival rate of non-priority customers $\lambda_2^{(1)} = 0.3$). Under state 2 of the *RE*, the average arrival rate of customers $\lambda^{(2)}$ is 2 (the average arrival rate of priority customers $\lambda_1^{(2)} = 0.5$, and the average arrival rate of non-priority customers $\lambda_2^{(2)} = 1.5$).

Let the generator of the RE be

$$G = \begin{pmatrix} -0.06 & 0.06 \\ 0.006 & -0.006 \end{pmatrix}.$$

From the point of view of possible practical applications for design of a cell of mobile communication networks, the fixed above parameters of the system may be interpreted as follows. 91 percent of time the system operates in the normal mode and 9 percent of time the system operates in the congestion mode. The average arrival rate in the congestion mode is four times (2 versus 0.5) higher than in the normal mode. This may be caused by some peaks of business or driving activity in the vicinity of the modeled cell during certain periods of time. Taking existence of such peaks into account, the service provider may assign in advance to this cell radio frequencies sufficient for simultaneous service of 12 customers during the congestion mode versus 6 in the normal mode. The profit of the service provider is defined by the throughput of the cell. This throughput is defined as the intensity of the arriving flow of customers minus the intensity of the lost customers. Thus, to increase the throughput, it is necessary to reduce the probability of customers loss. From the perspective of the human psychology and image of the service provider, it is usually assumed that the probability of handover customer loss (sometimes referred to as the dropping probability) is more significant than the probability of the fresh customer loss (blocking probability). Therefore, these loss probabilities have to be computed and treated separately. Under the fixed state, r, of the *RE*, availability of the parameter (threshold) $M^{(r)}$ allows to vary the implicit priority given to thehandover customers. When $M^{(r)} = N^{(r)}$, thehandover customers do not have any priority. The dropping probability equals to the blocking probability. When $M^{(r)}$ decreases, i.e., a certain part of servers becomes available only for the handover (type-1) customers, it is clear that the dropping probability decreases at expense of the blocking probability increasing. To reach a reasonable trade-off between these two loss probabilities and to gain the best value of the throughput, the **quantitative** analysis of the behavior of the loss probabilities $P_1^{(loss)}$ of priority and $P_2^{(loss)}$ of non-priority customers as the functions of the thresholds $M^{(1)}$ and $M^{(2)}$ is required.To give such an analysis in our example, we vary the valueof the parameter $M^{(2)}$ from 1 to $N^{(2)}$, and the parameter $M^{(1)}$ from 1 to $\min\{N^{(1)}, M^{(2)}\}$. Figure 2, Figure 3 show the dependence of the loss probabilities $P_1^{(loss)}$ of priority and $P_2^{(loss)}$ of non-priority customers on the thresholds $M^{(1)}$ and $M^{(2)}$.

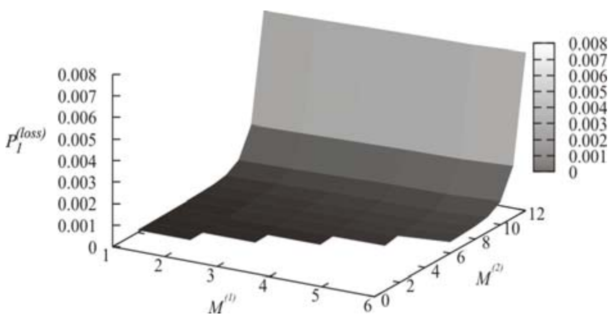

Figure 2. Dependence of $P_1^{(loss)}$ on the parameters $M^{(1)}$ and $M^{(2)}$.

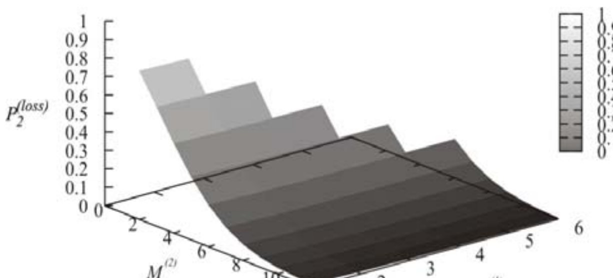

Figure 3. Dependence of $P_2^{(loss)}$ on the parameters $M^{(1)}$ and $M^{(2)}$.

These figures confirm our intuitive consideration that the increase of the parameters $M^{(r)}$ leads to the smaller dropping probability and the larger blocking probability. It is worth to note also that these figures show that the impact of the parameter $M^{(2)}$ is more essential (likely because the system is more congested under the state 2 of the *RE*) than the impact of the parameter $M^{(1)}$.

To analyse the influence of the thresholds $M^{(1)}$ and $M^{(2)}$ on the aggregate performance of the cell, various economical criteria can be used. Let us assume in this example that the quality of the system operation is defined by the following simple economical criterion:

$$J(M^{(1)}, M^{(2)}) = a_1\lambda_1 P_1^{(loss)} + a_2\lambda_2 P_2^{(loss)}$$

where a_l, $l = 1, 2,$ are the charges paid by the system for loss of type-*l* customer. It is worth noting that the problem of the suitable choice of cost coefficients (in our case, the coefficients a_1 and a_2) in the cost criterion always plays a crucial role in the successful implementation of optimization. We assume here that in our model the cost coefficients are obtained from experts in the telecommunication area. In this example, we fix $a_1 = 100$ and $a_2 = 10$.

The introduced criterion has the meaning of the lost revenues or the penalty of the system for the customers loss per unit of time. Therefore, our purpose is to define the values $M_*^{(1)}$ and $M_*^{(2)}$ for which the value $J(M^{(1)}, M^{(2)})$ of the criterion is minimal. Figure 4 presents the dependence of the cost criterion $J(M^{(1)}, M^{(2)})$ on parameters $M^{(1)}$ and $M^{(2)}$.

For the surface presented in this figure, the optimal values of the parameters $M^{(1)}$ and $M^{(2)}$ are $M_*^{(1)} = 5$ and $M_*^{(2)} = 11$, correspondingly, and the optimal value of the economical criterion is $J^*(M_*^{(1)}, M_*^{(2)}) = 0.24583$. If we do not use servers reservation in any state of the *RE*, i.e., we fix $M^{(1)} = N^{(1)} = 6$ and $M^{(2)} = N^{(2)} = 12$, we get $J(M^{(1)}, M^{(2)}) = 0.41308$. Thus, the optimal reservation of servers provides an essential economical profit. Therefore, the presented above results can be used by the specialists of mobile services providers to easy achieve better performance of the cell just by the means of the proper reservation of a few channels exclusively for service of handover customers. The best choice of the number of the reserved channels, under any fixed set of the system parameters, parameters characterizing the *RE* and cost coefficients, can be made based on the results of our analysis.

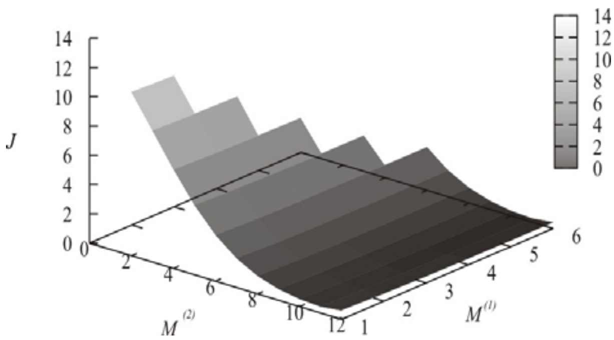

Figure 4. Dependence of $J(M^{(1)}, M^{(2)})$ on the parameters $M^{(1)}$ and $M^{(2)}$.

6. CONCLUSION

In this paper, a multi-server queueing model with two types of customers is considered. All the system parameters, including the number of servers, depend on the state of a random environment. The process of the system operation is considered, the ergodicity condition is derived, the main performance measures are calculated. The results can be effectively used to optimize the operation of the mobile network cell by means of the correct choice of admission control strategy of new and handover requests.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant No.2014R1A1A4A01007517).

REFERENCES

Buchholz, P., Kemper, P., and Kriege, J. (2010), Multiclass Markovian arrival processes and their parameter fitting, *Performance Evaluation*, **67**, 1092-1106.

Choi, B. D., Melikov, A., and Velibekov, A. (2008), A simple numerical approximation of joint probabilities of calls in service and calls in the retrial group in a picocell, *Applied Computational Mathematics*, **7**, 21-30.

Cordeiro, J. D. and Kharoufeh, J. P. (2012), The unreliable $M/M/1$ retrial queue in a random environment, *Stochastic Models*, **28**, 29-48.

Do, T.V. (2011), Solution for a retrial queueing problem in cellular networks with the Fractional Guard Channel policy, *Mathematical and Computer Modelling*, **53**, 2059-2066.

Dudin, A. N. and Nazarov, A. A. (2015), The $MMAP/M/R/0$ queueing system with reservation of servers operating in a random environment, *Problems of Information Transmission*, **51**, 274-283.

Dudin, A., Kim, C. S., Dudin, S., and Dudina, O. (2015), Priority Retrial Queueing Model Operating in Random Environment with Varying Number and Reservation of Servers, *Applied Mathematics and Com-*

putations, **269**, 674-690.

Graham, A. (1981), *Kronecker products and matrix calculus with applications*. Cichester: Ellis Horwood.

He, Q. M. (1996), Queues with marked calls, *Advances in Applied Probability*, **28**, 567-587.

Kim, C. S., Dudin, A. N., Klimenok, V. I., and Khramova, V. V. (2009), Erlang loss queueing system with batch arrivals operating in a random environment, *Computers and Operations Research*, **36**, 674-967.

Kim, C. S., Klimenok, V. I., and Dudin, A. N. (2014), Optimization of Guard Channel Policy in Cellular Mobile Networks with Account of Retrials, *Computers and Operation Research*, **43**, 181-190.

Kim, C. S., Klimenok, V., Mushko, V., and Dudin, A. (2010), The $BMAP/PH/N$ retrial queueing system operating in Markovian random environment, *Computers and Operations Research*, **37**, 1228-1237.

Klimenok, V. I. and Dudin, A.N. (2006), Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory, *Queueing Systems*, **54**, 245-259.

Tran-Gia, P. and Mandjes, M. (1997), Modeling of customer retrial phenomenon in cellular mobile networks, *IEEE Journal on Selected Areas in Communications*, **15**, 1406-1414.

Wu, J., Liu, Z., and Yang, G. (2001), Analysis of the finite source $MAP/PH/N$ retrial G-queue operating in a random environment, *Applied Mathematical Modelling*, **35**, 1184-1193.

Yang, G., Yao, L.-G., and Ouyang, Z.-S. (2013), The $MAP/PH/N$ retrial queue in a random environment. *Acta Mathematicae Applicatae Sinica*, **29**, 725-738.

Zhou, Z. and Zhu, Y. (2013), Optimization of the $(MAP_1, MAP_2)/(PH_1, PH_2)/N$ retrial queue model of wireless cellular networks with channel allocation, *Computers and Electrical Engineering*, **39**, 1637-1649.