A TANDEM SYSTEM $BMAP/G/1 \to \bullet/M/N/0$ WITH NON-ELASTIC TRAFFIC

V.I. Klimenok, O.S. Taramin

Belarusian State University,
4, Nezavisimosti Ave., 220030 Minsk, Belarus
{klimenok@bsu.by,taramin@mail.ru}

Introduction. Tandem queueing systems are widely used in capacity planning, performance evaluation and optimization of computer and communication systems, service centers, manufacturing systems, etc. In this presentation, we deal with a tandem queue under the assumption that customers arrive according to the $BMAP$. Such an arrival process includes many input flows considered previously in literature. Generally speaking, the $BMAP$ is correlated, so it

is ideal to model correlated and bursty traffic in modern telecommunication networks. Tandem queueing systems with the $BMAP$ input were investigated in literature only for single server phases of a tandem. In this presentation, we consider the model with multi-server second phase under additional difficulty: with a fixed probability, service of a customer requires a fixed number of servers at the second stage.

Mathematical model and analysis. We consider tandem queue which consists of two stages. The first stage is represented by the $BMAP/G/1$ queue. The input flow is described by the $BMAP$. The $BMAP$ is defined by the directing process ν_t, $t \geqslant 0$, which is an irreducible continuous time Markov chain with state space $\{0, \ldots, W\}$. The successive service times of customers at the first stage are independent random variables with general distribution. After receiving the service at the first stage a customer proceeds to the second stage which is represented by N independent identical servers. The service time by a server has exponential distribution. Customers are non-elastic, i.e., with a fixed probability the customer requires a fixed number of servers to start the service at the first stage. No queue is allowed between the first and the second stage. In case a customer completes the service at the first stage and does not see the required number of free servers at the second stage, we assume that the customer leaves the system forever with the fixed probability. With supplementary probability, the customer waits until the required number of servers at the second stage become free and then occupies these servers immediately. The waiting period is accompanied by blocking the first stage server operation. Such an assumption allows to consider in unified way models with losses and blocking which are usually distinguished in the literature.

Behavior of the described tandem queueing model is described by the non-Markovian process $\xi_t = \{i_t, r_t, \nu_t\}$, $t \geqslant 0$, where i_i is the number of customers at the first stage (not counting the blocked customer); r_t is the number of busy servers at the second stage $r_t = \overline{0, N}$; ν_t is the state of the arrival directing process ν_t at the epoch t, $i_t \geqslant 0$, $\nu_t = \overline{0, W}$. Analysis of this process starts of preliminary investigation of the multi-dimensional Markov chain $\xi_n = \{i_n, r_n, \nu_n\}$, $n \geqslant 1$, where i_n is the number of customers at the first stage at the epoch $t_n + 0$ of the nth service completion at the first stage; r_n is the number of busy servers at the second stage at the epoch $t_n - 0$; ν_n is the state of the arrival directing process ν_t at the epoch t_n. Criterion of ergodicity of this Markov chain is obtained in analytically tractable form and its stationary state distribution is computed. Based on the theory of Markov renewal processes, stationary distribution of the non-Markovian process $\xi_t = \{i_t, r_t, \nu_t\}$, $t \geqslant 0$, at arbitrary time is computed. Based on the method of collective marks, virtual and real sojourn times in the system are analyzed. Effectiveness of the proposed algorithms is numerically demonstrated. Impact of correlation in the arrival process is clarified.