PROCEEDINGS OF THE YEREVAN STATE UNIVERSITY

Physical and Mathematical Sciences 2009, № 3, p. 26–31

Mathematics

ON THE NUMBER OF SERVED CUSTOMERS IN $BMAP(t)|G|\infty$ MODEL

A. R. KHAJKHNIDI*

Chair of Probability Theory and Mathematical Statistics, YSU

In the present paper the $BMAP(t)|G|\infty$ model with non-stationary input stream of customers is considered. The distribution of customers' number served during time t and its mean value are obtained. Estimates for characteristics of $MAP(t)|G|\infty$ and $M^x(t)|G|\infty$ models are obtained as well.

***Keywords*:** queuing model, MAP stream, infinite number of servers.

Introduction. The models with infinite number of servers are applied for analyzing and estimating characteristics in communication and telecom-mutation networks, computer networks, network databases, multimedia, distance learning systems, provider systems, etc. [1].

Application of the $M|M|\infty$ model in communication networks allowed forecasting load of telephone lines and number of channels, needed to provide service quality [2].

Application of the $M|G|\infty$ model is substantiated by the fact that its stationary distribution is invariant with respect to the service time distribution [3, 4] and the Poisson property of the output stream. The model was being used to estimate the channel load in package commutation networks [5], to forecast losses in queuing systems [6], to determine the optimal level of reservation [7], etc. However, the assumptions on stationarity and Poisson property of the input stream were hindering further application of the model in computer networks. Possible mistakes while using such models to estimate characteristics of computer and mobile networks, to forecast the time and quantity of the channel peak load are mentioned in [8, 9]. The cause of those mistakes is the non-stationary and non-Poisson character of the network stream. The network traffic differs from the Poisson one and is characterized by non-ordinarity, non-stationarity, long-term dependence and after-action [10, 11]. As a rule, it is pulsing, and the customers arrive in groups of different lengths [11].

To describe the network traffic one uses *BMAP* (Batch Markovian Arrival Process) models [12]. *BMAP* stream properties and its applications are presented in [13].

Papers [14–16] are devoted to the study of the $BMAP|G|\infty$ models. The models with phase type input stream $PH|G|\infty$ are considered in [17, 18]. Papers

* E-mail: hajhnidi@yahoo.com

[8, 9, 19, 20] are devoted to the study of $M(t)|G|\infty$ model with non-stationary Poisson input stream. Here there are also presented examples and arguments justifying the use of models with non-stationary input streams for telecommunication, mobile network and production system analysis.

Models with phase type input stream and service $PH(t)|PH(t)|\infty$ are considered in [21], where the system load and the distribution of number of busy servers are studied with the help of numerical methods. The distribution of number of busy servers for the model $BMAP(t)|G|\infty$ is obtained in [14].

In literature on infinite linear queues there are mostly studied distributions of the number of busy servers in steady-state and/or non-steady-state modes, its moments and distribution of the model's busy period. However, as a performance measure for practical systems one should use the number of served customers during the period of the system's functioning [5, 9].

For systems with non-Markovian and non-stationary input streams the study of this characteristic encounters some difficulties. Here one should mention the paper [22], where the models $M|G|\infty$ and $M^x|G|\infty$ with stationary input streams are considered.

In the present paper the distribution of number of customers served during time *t* and the joint distribution of numbers of customers served and being under service by the moment *t* for the $BMAP(t)|G|\infty$ model are considered.

Model Description. The model consists of unlimited number of identical servers; the input stream of customers is non-stationary *BMAP* stream. Service time is a random variable γ, not depending on input process, on model state, having arbitrary distribution $G(t)=P(\gamma\le t)$ and finite average $\bar{\gamma}$. The input *BMAP* stream is given by basic Markovian process (*MP*) [13] $J(t)$ (phase process (*PP*)) with finite set $E=\{0,1,...,m\}$ of states and the sequence of characteristic matrices $\{D_k(t),\, k\ge 0\}$ of $m\times m$ size. $D_0(t)$ is a non-singular matrix with negative diagonal and non-negative extra-diagonal elements, the column-wise sum of which is less or equal to zero. $D_0(t)$ manages *PP* $J(t)$ transitions, that are not accompanied with customer generation. Non-negative matrices $D_k(t),\, k\ge 1$, manage *PP* $J(t)$ transitions, that are accompanied with generation of groups of $k\ge 1$ customers. Assume that *PP* $J(t)$ is a non-expandable *MP* with generating matrix $D(t)$, with set E of states and with vector of stationary distribution $\pi(t)=\{\pi_0(t),\pi_1(t),...,\pi_m(t)\}$, $t\ge 0$. Here $D(t)$ is the matrix of size $m\times m$:

$$D(t)=\sum_{k=0}^{\infty} D_k(t),\ \ (D(t)\ne D_0(t)),\ \ t\ge 0,\ \ \pi(t)D(t)=0,\ \ D(t)e=0,\ \ \pi(t)e=1,$$

where e is a unit vector column. *BMAP* stream customer arrival rate $\lambda(t)$ at moment t is

$$\lambda(t)=\pi(t)\sum_{k=1}^{\infty} kD_k(t)e,\ \ t\ge 0.$$

Transitions probabilities $P(x,t)$ *PP* $J(t)$ during time $[x,t)$ are found from Kolmogorov equations:

$$\frac{dP(u)}{du} = P(u)D(u), u \in [x,t)$$

and may be presented in matrix-exponential form $P(x,t) = e^{\int_x^t D(u)du}$.

Model Analysis. Let the states of the system (number of served customers and $PP\ J(u)$) are observed during the time interval $[u,t]$. Denote by $N(u,t)$ the number of customers arrived at moment u, $0 \le u \le t$, and served until moment t, $N(t) = N(0,t)$. Take $N(0) = 0$. Let $Q_{k,i,j}(u,t)$ be the probability that k custommers are served by moment t, and $PP\ J(u)$ is in the phase $j \in E$ under condition that at initial moment $u \le t$ the system was empty, and $PP\ J(u)$ was in phase $i \in E$, $Q_{k,i,j}(u,t) = P(N(u,t) = k,\ J(u) = j \mid N(0,t) = 0,\ J(0) = i)$. $Q_k(u,t)$ is a $m \times m$ matrix with elements $Q_k(s,t,i,j)$. Assume also $Q(u,t) = (Q_k(u,t),\ k \ge 0)$, $Q(t) = Q(0,t)$.

$b_k(n,t-u)$ is the probability that $n \le k$ customers that had arrived at moment u from a group of k customers are served during time $t-u$. As is shown in [14], the equality holds

$$b_k(n,t-u) = \binom{k}{n} G(t-u)^n (1-G(t-u))^{k-n}, \quad 0 \le k \le n.$$

$R_n(u,t)$ is the probability that the service of n customers arrived at moment u will be finished during time $t-u$. $R_n(0,t) = R_n(t) = (R_{n;ij}(t))_{i,j\in\{1,2,\ldots,m\}}$.

$$R_n(u,t) = \sum_{k=n}^{\infty} D_k(u) b_k(n,t-u) = \sum_{k=n}^{\infty} D_k(u) \binom{k}{n} G(t-u)^n (1-G(t-u))^{k-n}.$$

The probabilities $Q_k(u,t), k \ge 0$, may be estimated from the system of differential equations [15]

$$\frac{dQ_k(u,t)}{du} = \sum_{n=0}^{k} Q_n(u,t) R_{k-n}(u,t) \quad (1)$$

with initial condition $Q_0(u,0) = I, Q_k(u,0) = 0,\ k \ge 1$, where I and 0 are the unit and null matrices.

Let $\tilde{Q}(u,t,z)$ be the generating function (GF) of number of customers served by moment t:

$$\tilde{Q}(u,t,z) = \sum_{n=0}^{\infty} z^n Q_n(u,t), |z| \le 1.$$

Then from (1) for $\tilde{Q}(u,t,z)$ we get

$$\frac{\partial \tilde{Q}(u,t,z)}{\partial u} = \tilde{Q}(u,t,z) \tilde{R}(u,t,z), |z| \le 1, \quad (2)$$

with initial condition $Q(u,0,z) = I$. Here

$$\tilde{R}(u,t,z) = \sum_{n=0}^{\infty} z^n R_n(u,t) = \sum_{n=0}^{\infty} D_n(u)(1-G(t-u)+zG(t-u))^n.$$

Solution (2) is presented in matrix-exponential form

$$\tilde{Q}(u,t,z) = e^{\int_u^t \tilde{R}(x,t,z)sx}, \quad 0 \le u \le t, \quad |z| \le 1. \tag{3}$$

Hence for $u = 0$ $\tilde{Q}(0,t,z) = \tilde{Q}(t,z) = e^{\int_0^t \tilde{R}(x,t,z)sx}$.

Let $q(t) = (q_1(t), q_2(t),..., q_m(t))$ be the average number of customers served in the system during time t. If for all i and t the condition of model's stability $\lambda_{ij}(t) = \sum_{n=1}^{\infty}\sum_{j=1}^{m} nD_{n;ij}(t) < \infty, \ i \in E$, holds, i.e. the rate for all phases and moments of time is limited, then for the average number of customers served during time t we get

$$q(t) = \int_0^t P(0,u)\sum_{n=1}^{\infty} nD_n(u)P(u,t)G(t-u)du. \tag{4}$$

If the phase process has stationary distribution π, and the *BMAP* starts from balanced state, the average number of customers served during time t we get

$$q_\pi(t) = \pi q(t)1_m = \int_0^t \pi\sum_{n=1}^{\infty} nD_n(u)1_m G(t-u)du = \int_0^t \lambda(u)G(t-u)du. \tag{5}$$

Let us consider the joint distribution of number of customers served during time t and the number of busy servers at moment t. Let $Q_{r,m}(u,t)$ be the probability that by moment t r customers are already served in the system, and m servers are busy, $r, m \ge 0$.

Probabilities $Q_{r,m}(u,t), \ r,m \ge 0$, may be estimated from the system of differential equations:

$$\frac{\partial Q_{r,m}(u,t)}{\partial u} = \sum_{k=0}^{m}\sum_{n=0}^{r} Q_{r-n,\, m-k}(u,t)b_k(n,t-u) \tag{6}$$

with initial condition $Q_{0,0}(u,0) = I, \ Q_{r,m}(u,t) = 0, \ \ r,m \ge 0.$

Let $\tilde{Q}(u,t,z,y)$ be the joint GF of the number of customers served and the number of busy servers by moment t:

$$\tilde{Q}(u,t,z,y) = \sum_{n,m=0}^{\infty}\sum_{r=0}^{\infty} z^r y^m Q_{r,m}(u,t), \quad |z| \le 1, \quad |y| \le 1.$$

From (6) for $\tilde{Q}(u,t,z,y)$ follows

$$\frac{\partial \tilde{Q}(u,t,z,y)}{\partial u} = \tilde{Q}(u,t,z,y)\tilde{R}(u,t,z,y), \quad |z| \le 1, \quad |y| \le 1 \tag{7}$$

with initial condition $Q(u,0,z,y) = I$. Here

$$\tilde{R}(u,t,z,y) = \Phi(u, \alpha(u,t,z,y)), \Phi(u,z) = \sum_{n=0}^{\infty} z^n D_n(u), \alpha(u,t,z,y) = y\bar{G}(t-u) + zG(t-u),$$

$$\tilde{R}(u,t,z,y) = \sum_{r=0}^{\infty}\sum_{m=0}^{\infty} z^r y^m b_r(m,u,t) = \sum_{r=0}^{\infty}\sum_{m=0}^{\infty} z^r y^m D_{k+n}(u)\binom{k+n}{n} G(t-u)^n \bar{G}(t-u)^k.$$

Particular Cases.

1. Let the input stream be a *MAP* process with characteristic matrices $C(u)$ and $D(u)$. Then in the model considered above we have $D_0(u)=C(u)$, $D_1(u)=D(u)$, $D_k(u)=0$, $k>1$. Hence for $R_n(u,t)$, $n=0,1$, and $\tilde{R}(u,t,z)$ we get:

$$R_0(u,t)=C(u)+D(u)(1-G(t-u)),\quad R_1(u,t)=D(u)G(t-u),$$

$$\tilde{R}(u,t,z)=C(u)+D(u)[1-G(t-u)+zG(t-u)].$$

Substituting the values $R_n(u,t)$, $n=0,1$, and $\tilde{R}(u,t,z)$ in (3), (4) and (5), for $q(t)$ and $\tilde{Q}(u,t,z)$ we get

$$\tilde{Q}(u,t,z)=e^{\int_u^t \tilde{R}(x,t,z)dx}=e^{\int_u^t \{C(x)+D(x)[1-G(t-x)(1-z)]\}dx}.$$

$$q(t)=\int_0^t P(0,u)D(u)P(u,t)G(t-u)du,\quad q_\pi(t)=\int_0^t \lambda(u)G(t-u)du,\quad \lambda(u)=\pi D(u)1_m.$$

In the case, when $D_0(u)=C$, $D_1(u)=D$, $D_k(u)=0$, $k>1$, i.e. the parameters of input stream do not depend on time, we get

$$\tilde{Q}(u,t,z)=e^{\int_u^t \{C(x)+D(x)[1-G(t-x)(1-z)]\}dx}=e^{C(t-u)+D\int_u^t[1-G(t-x)(1-z)]dx}.$$

$$q(t)=\int_0^t P(0,u)DP(u,t)G(t-u)du,\quad q_\pi(t)=\lambda\int_0^t G(u)du,\quad \lambda=\pi D1_m.$$

2. Let $D_k(u)=\beta_k\lambda(u)$, $k\ge 0$, i.e. the customers arrive into the system in groups of size β_k at moments, given by the Poisson process with parameter $\lambda(u)$. Then if the condition of the model stability is satisfied, for $q(t)$ and $\tilde{Q}(u,t,z)$ we get:

$$\tilde{Q}(u,t,z)=e^{\int_u^t \tilde{R}(x,t,z)dx}\; e^{\int_u^t \lambda(x)\Phi(1-G(t-x)+zG(t-x))dx},\quad 0\le u\le t,\quad |z|\le 1;$$

$$\tilde{R}(u,t,z)=\lambda(u)\Phi(1-G(t-u)+zG(t-u)),\quad q(t)=\bar{\beta}\int_0^t \lambda(u)G(t-u)du,$$

where $\Phi(z)=\sum_{n=1}^{\infty}\beta_n z^n$ is the GF of sizes of arriving customers' groups.

Hence, if $\lambda(u)=\lambda$ for all $u\ge 0$, then we get the result well-known for the model $M^x|G|\infty$ [22], which states that the average number of customers served and input stream intensity are correspondingly equal to

$$q_\pi(t)=\bar{\beta}\lambda\int_0^t G(u)du,\quad \lambda_\beta=\bar{\beta}\lambda.$$

The obtained results may be applied for estimating characteristics, as well as searching optimal strategies for managing resources of wide class of systems, whereas the model $BMAP(t)|G|\infty$ may be used as models of these systems.

Received 29.04.2009

REFERENCES

1. **Tijms H.C.** A First Course in Stochastic Models. John Wily & Sons, 2003, 480 p.
2. **Erlang A.K.** The Post Office Electrical Engineers' Journal, 1918, v. 10, p. 189–197.
3. **Sevastyanov B.A.** Theory of Probability and its Applications, 1957, v. 2, № 1, p. 106–116 (in Russian).
4. **Gnedenko B.V., Kovalenko I.N.** Introduction of Theory of Queues. M.: Nauka, 1987 (in Russian).
5. **Kerobyan Kh.V.** Resource Sharing Performance Models for Estimating Internet Service Characteristic in View of Restrictions on Buffer Capacity. Int. NATO Conf. DF for SM, Kluwer Academic Pb., 2003.
6. **Grigorian T.A.** The Development and Study of Internet-Service Proecting Models for Enter to Datebuses. PhD Thesis, Ye., 2004 (in Russian).
7. The Relabily of Technical Systems. (Eds. I.A. Ushakova). Radio i Sviaz, 1985, 608p. (in Russian).
8. **Leung K., Massey W.A. and Whitt W.** Journal on Selected Areas in Communications, 1994, v. 12, p. 1353–1364.
9. **Massey W.A.** Telecommunication Systems, 2002, v. 21, № 2–4, p. 173–204.
10. **Neuts M.F.** J. Appl. Probab., 1979, v. 16, p. 764–779.
11. **Lucantoni D.M.** Commun. Statist. Stochastic Models, 1991, v. 7, № 1, p. 1–46.
12. **Paxson V., Floyd S**. Wide-area Traffic: The Failure of Poisson Modeling. Proc. Of the ACM94, p. 257–268.
13. **Park K., Kim G., Crovella M.** On the Effect of Traffic Self-Si,ilarity on Network Performance, Proc. SPIE Int'l. Conf. Perf. and Control., 1997, p. 296–310.
14. **Eick S., Massey W.A. and Whitt W.** Management Science, 1993, v. 39, p. 241–252.
15. **Eick S., Massey W.A. and Whitt W.** Operations Research, 1993, v. 41, p. 400–408.
16. **Keilson J., Servi L.D.** Networks of Non-homogenous $M|G|\infty$ Systems. GTE Laboratories. Waltham, MA, 1989.
17. **Breuer L.** Spatial Queues. Ph.D. Thesis, University of Trier, Germany, 2000.
18. **Masuyama H**. Studies on Algorithmic Analysis of Queues with Batch Markovian Arrival Streams. PhD Thesis, Kyoto University, 2003, 146 p.
19. **Baum D., Kalashnikov V.** Queueing Systems, 2004, v. 46, p. 231–247.
20. **Ramaswami V. and Neuts M.F.** Journal of Applied Probability, 1980, № 17, p. 498–514.
21. **Barry L. Nelson, Michael R. Taaffe** Informs Journal on Computing, 2004, v. 16, № 3, p. 266–274.
22. **Matveev V.F., Ushakov V.G.** Systems of Queues. M.: MGU, 1984, 239 p. (in Russian).

Proc. of the Yerevan State Univ. Phys. and Mathem. Sci., 2009, № 3, p. 26–31.

$\boldsymbol{BMAP(t)|G|\infty}$ մոդելում սպասարկված հարցումների թվի մասին

Աշխատանքում դիտարկված է հարցումների ոչ ստա-ցիոնար մուտքային հոսքով մոդելը: Ստացված են t ժամանակում սպա-սարկված հարցումների թվի բաշխումը և նրա միջինը: Ստացված են նաև մոդելների բնութագրերի համար գնահատա-կաններ:

О числе обслуженных запросов в модели $\boldsymbol{BMAP(t)|G|\infty}$

В работе рассмотрена модель $BMAP(t)|G|\infty$ с нестационарным входящим потоком запросов. Найдено распределение числа обслуженных за время t запросов. Получены также оценки для характеристик моделей $MAP(t)|G|\infty$, $M^x(t)|G|\infty$.