PLISKA
STUDIA MATHEMATICA
BULGARICA

ПЛИСКА
БЪЛГАРСКИ
МАТЕМАТИЧЕСКИ
СТУДИИ

Pliska Stud. Math. Bulgar. **20** (2011), 53–62

PLISKA
STUDIA MATHEMATICA
BULGARICA

SINGLE-SERVER QUEUEING SYSTEM WITH MARKOV-MODULATED ARRIVALS AND SERVICE TIMES

Mitko Dimitrov

Markov-modulated queueing systems are those in which the input process or service mechanism is influenced by an underlying Markov chain. Several models for such systems have been investigated .

In this paper we present heavy traffic analysis of single queueing system with Poisson arrival process whose arrival rate is a function of the state of Markov chain and service times depend on the state of the same Markov chain at the epoch of arrivals.

1. Introduction

Markov process is natural random environment and is extensively used for modeling the arrival rate and service times in queueing theory and many other applications. The presence of this Markov process is attributed to stochastic variations or heterogeneity of arrivals and service and actually it is a component of the queueing system considered.

Significant effort is currently being devoted to the development of integrated communication systems, which can support a wide range applications including voice, video and data.

There are many exact analytical results, computational techniques and approximations of the basic characteristics of Markov-modulated queueing model

2000 *Mathematics Subject Classification:* 60K25.
Key words: Markov-modulated queues, waiting time, heavy traffic.

[1], [2], [3], [4], [5]. The asymptotic behavior of performance characteristics of such queues are derived in some limit cases as heavy traffic. Heavy traffic analysis of Markov-modulated queueing models is suggested by Burman and Smith [2] where an $M/G/1$ type queue with Markov-modulated arrivals was investigated. Genadi Falin and Anatoli Falin [3] suggested another approach to the same problem. It is based on certain "semi-explicit" formula for the stationary distribution of the virtual waiting time and its mean value in heavy traffic assumptions. We applied this approach to single queueing system with arrival rate and service time depending on the state of Markov chain at the arrival epoch.

2. Some preliminary results

Let us consider an irreducible continuous time Markov chain $Z(t)$ with finite state space $S = \{0, 1, \dots, K\}$, transition rates α_{nm}, $\alpha_n = -\alpha_{nn}$ and stationary distribution $\pi = (\pi_0, \pi_1, \dots, \pi_K)$. The infinitesimal matrix of the process is denoted as Q.

The linear equations

$$Q(x_0, x_1, \dots, x_K)^T = (b_0, b_1, \dots, b_K)^T$$

has a solution if the vector $(b_0, b_1, \dots, b_K)$ is orthogonal to the stationary vector π. Now let $f(n)$ be a nonnegative function defined on the state space of $Z(t)$,$H_n(x)$ is the probability distribution with two moments h_{n1}, h_{n2}, $n = 0, 1, \dots, K$ and $f = \sum\limits_{n=0}^{K} \pi_n f(n) h_{n,1}$.

Then the linear algebraic equations

$$\sum_{m=0}^{K} \alpha_{nm} x_m = f - f(n) h_{n1}, 0 \le n \le K \tag{1}$$

with unknowns $x_0, x_1, \dots, x_K$ always has a solution.

Since the Markov chain $Z(t)$ is finite and irreducible, it is also ergodic chain, and $\pi_j > 0$ for $j = 0, 1, 2, \dots, K$. Then the stationary distribution $\pi = (\pi_0, \pi_1, \dots, \pi_K)$ can be found from the Kolmogorov equations $\pi Q = 0$ and normalization condition $\sum\limits_{n=0}^{K} \pi_n = 1$. In coordinate form Kolmogorov equations are $\pi_0 \alpha_{0j} + \sum\limits_{i=0}^{K} \pi_i \alpha_{ij} = 0, 0 \le j \le K$.

If we consider the equations

$$\pi_0 \alpha_{0j} + \sum_{i=0}^{K} \pi_i \alpha_{ij} = 0, \quad 0 \leq j \leq K,$$

then $\pi_0(\alpha_{01}, \alpha_{02}, \ldots, \alpha_{0K}) = -(\pi_1, \ldots, \pi_K)A$

where $A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1K} \\ \vdots & & & \\ \alpha_{K1} & \alpha_{K2} & \ldots & \alpha_{KK} \end{pmatrix}$, and $\det A \neq 0$.

Then $(\alpha_{01}, \ldots, \alpha_{0K})A^{-1} = -\dfrac{(\pi_1, \ldots, \pi_K)}{\pi_0}$

Now, if we define matrix R as $R = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & A^{-1} & \\ 0 & & & \end{pmatrix}$, then matrix QR is equal to

$$\begin{pmatrix} 0 & \dfrac{-\pi_1}{\pi_0} & \dfrac{-\pi_2}{\pi_0} & \ldots & \dfrac{-\pi_K}{\pi_0} \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & 1 \end{pmatrix} \tag{2}$$

and for any vector $\overline{y} = (y_0, y_1, \ldots, y_K)$ we have $\overline{y}QR = \overline{y} - \dfrac{y_0}{\pi_0}\pi$.

It is obvious that vector

$$x = (x_0, x_1, \ldots, x_K)^T = R(f - f(0)h_{01}, f - f(1)h_{01}, \ldots, f - f(K)h_{K1})$$

is a solution to the linear equations (1).

3. An M/G/1 queueing system with Markov modulated arrivals and service times

We make the following assumptions

1. The modulating Markov chain $Z(t)$ is exactly as in the Section 2.
2. During a time-interval in which $Z(t) = n$, customers arrive in a Poisson

process at a rate $\lambda f(n)(0 < f(n) < \infty)$ and then service times have a distribution H_n. Given the states of the Markov chain, the service times are conditionally independent and also independent of the arrival process.

3. There is a single server and the queue discipline is first come, first served.

We denote by $W(t)$ the virtual waiting time. The system is ergodic if the inequality $\rho = \sum\limits_{n=0}^{K} \lambda f(n) h_{n1} \pi_n < 1$ holds. From now on we assume that $\rho < 1$ and system is in steady state. The presence of the Markov chain $Z(t)$ implies that we need to designate the waiting time process as $\{W, Z\} = \{W(t), Z(t), t \geq 0\}$. Given the state of $Z(t) = n$, the probability that a customer arrives during $(t, t+dt]$ and make a service demand not exceeding x is given by $\lambda f(n) H_j(x) dt + 0(dt)$. As in the standard model it is seen that (W, Z) is a time-homogeneous Markov process on the state space $R_+ \mathrm{x} S$. To derive integro-differential equations for $F(x,t) = P\{W(t) < x, Z(t) = n\}$ we process as on the standard model but note the presence of the modeling chain $Z(t)$. Thus considering the process over the intervals $(0,t]$,$(t, t+dt]$ we find that

$$F_n(x, t+dt) = F_n(x+dt, t)(1 - \lambda f(n)dt + \alpha_{nn} dt) +$$

$$+ \int_0^\infty F_n(x + dt - \nu, t) \lambda f(n) dH_n(\nu) + \sum_{m \neq n} \alpha_{mn} F_m(x,t) dt.$$

This leads to

$$F'_n(x) = (\lambda f(n) + \alpha_n) F_n(x) + \lambda f(n) \int_0^x F_n(x-y) dH_n(y) + \sum_{m \neq n} \alpha_{mn} F_m(x).$$

Let

$$\begin{aligned} P_{n0} &= P\{W(t) = 0, Z(t) = n\} \\ \varphi_n(s) &= E\{e^{-sW(t)}; Z(t) = n\} = P_{n0} + \int_{0^+}^\infty e^{-sx} dF_n(x) \\ h_n(s) &= \int_0^\infty e^{-sx} dH_n(x). \end{aligned}$$

In terms of Laplace transform we obtain for $\varphi_n(s)$, $n = 0, 1, \dots, K$ the next equations

$$\sum_{m=0}^{K} \varphi_m(s)\alpha_{mn} = [\lambda f(n)(1 - h_n(s) - s)]\, \varphi_n(s) + sP_{n0}. \tag{3}$$

From $\sum\limits_{n=0}^{K} \alpha_{mn} = 0$ and summing both sides of (3) we get

$$\sum_{n=0}^{K} P_{n0} = \sum_{n=0}^{K} \left[1 - \lambda f(n)\frac{1 - h_n(s)}{s}\right] \varphi_n(s). \tag{4}$$

If $s \to 0$ the sum $\sum\limits_{n=0}^{K} \left[1 - \lambda f(n)\dfrac{1 - h_n(s)}{s}\right] \varphi_n(s)$ tends to

$$\sum_{n=0}^{K} [1 - \lambda f(n)h_{n,1}]\, \pi_n = 1 - \sum_{n=0}^{K} \lambda f(n)h_{n1}\pi_n = 1 - \rho$$

and $\sum\limits_{n=0}^{K} P_{n0} = 1 - \rho$. Taking the first derivative at the point $s = 0$ from the both sides of (4) we have

$$\sum_{n=0}^{K} (1 - \lambda f(n)h_{n1})\, W_n = \sum_{n=0}^{K} \frac{\lambda f(n)h_{n2}}{2}\pi_n. \tag{5}$$

Now multiplying both sides of (3) by x_n and summing with respect to n we get

$$\sum_{n=0}^{K}\sum_{m=0}^{K} \varphi_m(s)\alpha_{mn}x_n = \sum_{n=0}^{K} [\lambda f(n)\,[1 - h_n(s)] - s]\, \varphi_n(s)x_n + s\sum_{n=0}^{K} P_{n0}x_n,$$

or

$$\sum_{m=0}^{K} \varphi_m(s)(f - f(m)h_{m1}) = \sum_{n=0}^{K} x_n\, [\lambda f(n)(1 - h_n s) - s]\, \varphi_n(s) + s\sum_{n=0}^{K} P_{n0}x_n.$$

Differentiating this equation with respect to s at the point $s = 0$ we have

$$-\sum_{m=0}^{K} W_m\, (f - f(m)h_{m1}) = \sum_{n=0}^{K} x_n\pi_n(\lambda f(n)h_{n1} - 1) + \sum_{n=0}^{K} P_{n0}x_n. \tag{6}$$

Multiplying both sides of (6) with λ and summing (5) and (6) we get the equation

$$\sum_{m=0}^{K} W_m(1-\rho) = \sum_{n=0}^{K} \frac{\lambda f(n) h_{n2}}{2}\pi_n + \sum_{n=0}^{K} \lambda x_n \pi_n (\lambda f(n) h_{n1} - 1) + \sum_{n=0}^{K} \lambda P_{n0} x_n.$$

Thus we obtain the main result.

Theorem 1. *The mean virtual waiting time is given by the formula*

$$W = \frac{\sum\limits_{n=0}^{K} \frac{\lambda f(n)\pi_n h_{n2}}{2} + \sum\limits_{n=0}^{K} \lambda x_n \pi_n \left(\lambda f(n) h_{n1} - 1\right) + \sum\limits_{n=0}^{K} \lambda P_{n0} x_n}{1-\rho}$$

where the variables x_n are a solution of the linear algebraic equations (1), *and the probabilities P_{n0} satisfy equation $\sum\limits_{n=0}^{K} P_{n0} = 1-\rho$.*

Our goal in this paper is to investigate the queue in heavy traffic condition.

If $\rho = \lambda \sum\limits_{n=0}^{K} f(n) h_{n1} \pi_n \to 1$ then all probabilities P_{n0}, $n = 0, 1, \ldots, K$ tends to 0 because $\sum\limits_{n=0}^{K} P_{n0} = 1 - \rho \to 0$ as $\rho \to 1$.

Thus we prove the next theorem.

Theorem 2. *If $\rho \to 1$ (under heavy traffic)*

$$(1-\rho)W = \sum_{n=0}^{K} \frac{x_n f(n)\pi_n h_{n2}}{2f} + \sum_{n=0}^{K} \frac{x_n \pi_n (f(n) h_{n1} - f)}{f^2}$$

From $P_{n0} > 0$, $n = 0, 1, \ldots, K$, $\sum\limits_{n=0}^{K} P_{n0} = 1-\rho$ follows that:

$$\sum_{n=0}^{K} P_{n0} x_n \in (1-\rho)(x_*, x^*), \;\; x_* = \min_n \{x_n\}, \quad \text{and} \quad x^* = \max_n \{x_n\},$$

and the next theorem is proved.

Theorem 3. *Let* $x_* = \min_n\{x_n\}$, $x^* = \max_n\{x_n\}$.
Then

$$\lambda x_* \le W - \frac{\sum\limits_{n=0}^{K} \dfrac{\lambda x_n f(n)\pi_n h_{n2}}{2} + \sum\limits_{n=0}^{K} \lambda x_n \pi_n (\lambda f(n) h_{n1} - 1)}{1-\rho} \le \lambda x^*$$

We shall consider stationary distribution of the process $(W(t), Z(t))$ *as* $\rho \to 1$.

Theorem 4.
a) random variables $W(t)$ *and* $Z(t)$ *are asymptotically independent;*
b) the random variable $(1-\rho)W(t)$ *is asymptotically exponential with the mean*

$$\frac{\sum\limits_{n=0}^{K} \pi_n h_{n2} f(n)}{2f} + \frac{\sum\limits_{n=0}^{K} \pi_n (h_{n1} f(n) - f) x_n}{f^2}.$$

P r o o f. Let us consider several notations as $\varphi(s) = (\varphi_0(s), \varphi_1(s), \ldots, \varphi_K(s))$, $\overline{P}_0 = (P_{00}, \ldots, P_{K0})$ and diagonal matrices $F = \operatorname{diag}(f(0), f(1), \ldots, f(K))$, $\tilde{H}(s) = \operatorname{diag}(1 - h_0(s), 1 - h_1(s), \ldots, 1 - h_K(s))$, $\tilde{H}_1 = \operatorname{diag}(h_{01}, h_{11}, h_{21}, \ldots, h_{K1})$ and $\tilde{H}_2 = \operatorname{diag}(h_{02}, h_{12}, \ldots, h_{K2})$, $E = \operatorname{diag}(1, 1, \ldots, 1)$.

In this notations we can rewrite equation (3) as

$$\bar{\varphi}(s)Q = \bar{\varphi}(s)\left[\lambda\tilde{H}(s)F - sE\right] + s\bar{P}_0 \tag{7}$$

Multiplying from the right both sides of (7) with $\bar{e}$ and $\bar{P}_0\bar{e} = 1 - \rho \equiv \varepsilon$ we get

$$\bar{\varphi}(s)\left[\lambda\tilde{H}(s)F - sE\right]\bar{e}^T + \varepsilon s = 0 \tag{8}$$

From equation (2) multiplying again both sides of (7) from the right side by the matrix R we have

$$\bar{\varphi}(s)QR = \bar{\varphi}(s)\left[\lambda\tilde{H}(s)F - sE\right]R + s\bar{P}_0 R$$

or

$$\bar{\varphi}(s) = \varphi_0(s)\frac{\pi}{\pi_0} + \bar{\varphi}(s)\left[\lambda\tilde{H}(s)F - sE\right]R + s\bar{P}_0 R \tag{9}$$

Now, let us replace vector $\bar{\varphi}(s)$ from (9) to the right side of (9), and after some algebraic transformation we get the equation

$$\bar{\varphi}(s) = \varphi_0(s)\frac{\pi}{\pi_0}\left[E + [\lambda\tilde{H}(s)F - sE]R\right] +$$

$$+s^2\bar{\varphi}(s)\left[\left[\lambda\frac{\tilde{H}(s)}{s}F - E\right]R\right]^2 + s\bar{P}_0R\left[E + [\lambda\tilde{H}(s)F - sE]R\right] = \tag{10}$$

$$= \varphi_0(s)\frac{\pi}{\pi_0}[E + \lambda\tilde{H}(s)FR - sR] + s\bar{Y}(s)$$

where

$$\bar{Y}(s) = s\bar{\varphi}(s)\left[\left[\lambda\frac{\tilde{H}(s)}{s}F - E\right]R\right]^2 + \bar{P}_0R\left[E + [\lambda\tilde{H}(s)F - sE]R\right]$$

Substituting $\bar{\varphi}(s) = \varphi_0(s)\frac{\pi}{\pi_0}[E + \lambda\tilde{H}(s)FR - sR] + s\bar{Y}(s)$ into (8) we,ll find an expression for $\bar{\varphi}_0(s)$.

Really

$$\left[\varphi_0(s)\frac{\pi}{\pi_0}[E + \lambda\tilde{H}(s)FR - sR] + s\bar{Y}(s)\right] . \left[\lambda\tilde{H}(s)F - sE\right]\bar{e}^T + \varepsilon s =$$

$$= \frac{\varphi_0(s)}{\pi_0}\left[\pi\left[\lambda\tilde{H}(s)F - sE\right]\bar{e}^T + \pi[\lambda\tilde{H}(s)F - sE]R[\lambda\tilde{H}(s)F - sE]\bar{e^T}\right] +$$

$$+sY(s)[\lambda\tilde{H}(s)F - sE]\bar{e}^T + \varepsilon s = \frac{\varphi_0(s)}{\pi_0}\left(B_1(s) + B_2(s)\right) + A(s) = 0$$

where

$$\begin{aligned} A(s) &= \varepsilon s + s\bar{Y}(s)[\lambda\tilde{H}(s)F - sE]\bar{e}^T, \\ B_1(s) &= \pi[\lambda\tilde{H}(s)F - sE]\bar{e}^T \\ B_2(s) &= \pi[\lambda\tilde{H}(s)F - sE]R[\lambda\tilde{H}(s)F - sE]\bar{e}^T \end{aligned}$$

Now, replace s by εs and let λ increase to $1/f$ as $\varepsilon \to 0$. Finally we obtain

$$A(\varepsilon s) = \varepsilon^2 s + \varphi^2 s^2\bar{Y}(\varepsilon s)\left[\lambda\frac{\tilde{H}(\varepsilon s)F}{\varepsilon s} - E\right]\bar{e}^T = \varepsilon^2 s + o(\varepsilon^2),$$

because $\bar{Y}(\varepsilon s) \to 0$ as $\varepsilon \to 0$.

$$B_1(\varepsilon s) = \pi[\lambda \tilde{H}(\varepsilon s)F - \lambda \tilde{H}_1 F \varepsilon s + \lambda \tilde{H}_1 F \varepsilon s - \varepsilon s E]\bar{e}^T =$$

$$= -\varepsilon s \left(\pi \bar{e}^T - \lambda \pi \tilde{H}_1 F \bar{e}^T\right) - \lambda \pi \frac{\tilde{H}_1 \varepsilon s - \tilde{H}(\varepsilon s)}{\varepsilon^2 s^2} \varepsilon^2 s^2 F \bar{e}^T =$$

$$= -\varepsilon s(1-\rho) - \frac{\lambda \pi \tilde{H}_2 F \bar{e}^T}{2} \varepsilon^2 s^2 + o(\varepsilon^2) = -\varepsilon^2 s - \frac{\pi \tilde{H}_2 F \bar{e}^T}{2f} \varepsilon^2 s^2 + o(\varepsilon^2),$$

$$B_2(\varepsilon s) = \pi \varepsilon^2 s^2 \left[\frac{\lambda \tilde{H}(\varepsilon s)}{\varepsilon s} F - E\right] R[\frac{\lambda \tilde{H}(\varepsilon s)}{\varepsilon s} F - E]\bar{e}^T =$$

$$= \pi \varepsilon^2 s^2 [\lambda \tilde{H}_1 F - E] R [\lambda \tilde{H}_1 F - E] \bar{e}^T + o(\varepsilon^2) =$$

$$= -\pi \frac{\varepsilon^2 s^2}{f^2} [\tilde{H}_1 F - f] R [f \bar{e}^T - \tilde{H}_1 F \bar{e}^T] + o(\varepsilon^2)$$

because $\dfrac{\tilde{H}(\varepsilon s)}{\varepsilon s} \to \tilde{H}_1$, $\dfrac{\tilde{H}_1 \varepsilon s - \tilde{H}(\varepsilon s)}{\varepsilon^2 s^2} \to \dfrac{\tilde{H}_2}{2}$.

Finally we get

$$\varphi_0(\varepsilon s) = \pi_0 \frac{\varepsilon^2 s + o(\varepsilon^2)}{\varepsilon^2 s + \dfrac{\pi \tilde{H}_2 F \bar{e}^T \varepsilon^2 s^2}{2f} + \dfrac{\pi}{f^2}\left[H_1 E - fE\right] R \left[f \bar{e}^T - \tilde{H}_1 F \bar{e}^T\right] \varepsilon^2 s^2 + o(\varepsilon^2)}$$

and

$$\lim_{\varepsilon \to 0} \bar{\varphi}(\varepsilon s) = \lim_{\varepsilon \to 0} \frac{\varphi_0(\varepsilon s)\pi}{\pi_0} = \frac{\pi}{1 + M.s},$$

where

$$M = \frac{\sum\limits_{n=0}^{K} \pi_n f(n) h_{n2}}{2f} + \frac{1}{f^2} \sum_{n=0}^{K} \pi_n (h_{n1} f(n) - f) x_n,$$

since

$$R[fE - \tilde{H}_1 F]\bar{e}^T = \bar{x} = (x_0, x_1, \ldots, x_K). \qquad \square$$

REFERENCES

[1] F. Baccelli, A. M. Makowski. Martingale relations for the M/G/1 queue with Markov -modulated Poisson input. *Stochastic Processes Appl.*, **38** (1991), V. 2, 99–133.

[2] D. Y. Burman, D. R. Smith. An asymptotic analysis of a queueing system with Markov-modulated arrivals. *Operations Research*, **34** (1986), No 1, 105–119.

[3] G. I. Falin, A. I. Falin. Heavy traffic analysis of M/G/1type queueing systems with Markov-modulated arrivals. *TOP*, **7** (1999), No 2, 279–291.

[4] W. Fisher, K. Hellstern. The Markov-modulated poisson process (MMPP) cook-book. *Performance Evaluation*, **18** (1992), 149–171.

[5] N. U. Prabhu,Y. Zhu. Markov modulated queueing systems. *Queueing systems Theory Appl.*, **5** (1989), 215–245.

Mitko Dimitrov
University of National and World Economy
Student town, Hristo Botev
8, December Str.
1700 Sofia, Bulgaria
e-mail: `mcdimitrov@abv.bg`