Automatica 45 (2009) 1423–1430

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Optimization of queuing system via stochastic control[☆]

Boris M. Miller *

Monash University, School of Mathematical Sciences, Clayton, Victoria, 3800, Australia
Institute for Information Transmission Problems, RAS, Moscow, Russia

ARTICLE INFO

Article history:
Received 21 April 2008
Received in revised form
24 December 2008
Accepted 24 January 2009
Available online 28 February 2009

Keywords:
Queuing systems
Markov chains
Optimization
Stochastic control

ABSTRACT

The problem of access and service rate control in queuing systems as a general optimization problem for controlled Markov process with finite state space is considered. By using the dynamic programming approach we obtain the explicit form of the optimal control in the case of minimizing cost given as a mixture of an average queue length, number of lost jobs, and service resources. The problem is considered on a finite time interval in the case of nonstationary input flow. In this case we suggest the general procedure of the numerical solution which can be applied to a problems with constraints.

1. Introduction

Effective control of data flows is one of the most important problems of the Internet of the next generation, which conform both to new realities of super speed networks (according to the present concepts) with integral servicing. Even though there exist a number of empirical and engineering approach, the theoretical basis of data transmission control is still restricted by stationary methods of analysis. Moreover, the most existing approaches are based on asymptotic methods, which give only a qualitative description of network behavior, but do not correspond to real-time control and design of real network control algorithms, like various Active Queue Management (AQM) schemes (see Vasenin and Simonova (2005)).

Here we consider the nonstationary feedback type controls. These problems are inherent to queuing systems on finite horizon under restricted control resources. We underline that these problems are different from usually considered ones in infinite time interval (see, for example Hordijk and Spieksma (1989), Kelly, Maulloo, and Tan (1998), Low, Paganini, and Doyle (2002), Piunovskiy (2004), and Serfozo (1981)). First, in the case of infinite horizon one has to provide the stability, in other words the service rate must be greater than or equal to the rate of demands arriving. Another disadvantage concerns the necessity of stationary data which is not a case for real service systems. Moreover, the optimization of a stationary phase does not take into account the cost of transient phases and resources which are needed for their realization.

One of the most general problem statements is given in Hordijk and Spieksma (1989), where a queuing system can be controlled by restricting arrivals. Different settings of optimization problems related with stochastic networks are given in Kelly et al. (1998), where the approach to solution is also based on optimization techniques of convex mathematical programming. In Piunovskiy (2004) the problem of the input stream control arising in communication networks is also considered and reduced to a convex programming problem. It is worth to underline that in the problems with long-run average type criteria and stationary controls the optimal solution is very often has a threshold form.

Meanwhile, in the case of finite horizon the threshold type controls are inherent to settings with affine dependence on control action. In control of communication networks this result has been obtained, probably first in Bremaud (1979), where the problem of optimal thinning of a point process has been solved in the case of nonstationary input flow with deterministic intensity rate.

General approach to these problems is based on the martingale description of the process evolution (see Bremaud (1981), Elliott, Aggoun, and Moore (1995), and Liptser and Shiryaev (1979)). The existence of the optimal solution had been proved in Davis and Elliott (1977), and Wan and Davis (1979). In Elliott (1992), and Elliott et al. (1995) the general optimization problem for jump Markov process had been considered and the reduction to a problem with complete information had been proposed for a

☆ This work was supported in part by Russian Basic Research Foundation Grant 05-01-00508 and by Australian Research Council Grant DP0988685. The short version of this article had been presented in IFAC World Congress, July, 2008, Seoul, Korea. This paper was recommended for publication in revised form by Associate Editor George Yin, under the direction of Editor Ian R. Petersen.

* Corresponding address: Monash University, School of Mathematical Sciences, Clayton, Victoria, 3800, Australia. Tel.: +61 3 9905 5870; fax: +61 3 9905 9520.
E-mail address: boris.miller@sci.monash.edu.au.

0005-1098/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2009.01.011

wide class of optimal control problems. In Miller, Avrachenkov, Stepanyan, and Miller (2005) we extend this approach to a flow control with state–control-dependent rate.

In this article we consider some typical problems of a queuing systems control in nonstationary case. We extend the approach of Bremaud (1979) to a more wide class of the optimal control problems with complete information and prove the existence and characterization of the optimal control with the aid of dynamic programming approach. In this case the dynamic programming equation can be reduced to the system of ordinary differential equations. Then we apply these results to a problem of access and service rate control in the case of finite time horizon and finite buffer size.

We demonstrate that for a wide class of criteria the optimal control problem can be reduced to the solution of the system of ordinary differential equations. Moreover, the optimal control exists within the class of Markov strategies and therefore can be calculated in "program form" for each possible state of controlled Markov chain. The structure of the article is as follows. In the next section we provide some necessary results of the theory of controlled Markov chains. In Section 3 we apply these results to the simultaneous access and service rate control. In Section 4 we give some examples.

2. Controlled Markov chain

In this section we extend the approach of P. Bremaud (see Bremaud (1979)) to a more general class of controlled Markov chains based on their martingale description (see Aggoun and Elliott (2004), and Elliott et al. (1995)). This section presents a slight generalization of well-known result of Miller (1968) who considered the controls taking values in a finite set. It had been shown that the optimal control exists within the class of piecewise constant policies. Generally it is not always true for an arbitrary class of cost functions, however, for the most examples arising in network optimization it is still valid. Meanwhile, our Example 2 corresponds to the case when the optimal control does not piecewise constant (see formulae (15) and numerical results for the optimal service rate control on Fig. 4.

2.1. Martingale representation of controlled Markov chain

Assume that all processes are defined on a probability space $\{\Omega, \mathcal{F}, P\}$. Consider a process $\{X_t, t \in [0, T]\}$ which is a controlled jump Markov process with piecewise constant right-continuous trajectories. The state space of the process is the set of unit vectors $e_i \in \mathbb{R}^n$: $X_t \in S = \{e_1, \dots, e_n\}$, where e_i are vectors with unity as the ith element and zeros elsewhere (see Elliott et al. (1995, p. 17)).

Assumption 2.1. The matrix $A(t, u)$ with entries $a_{ij}(t, u)$ is a family of time-dependent generators, such that the probability column vector $p_t = (p_t^1, \dots, p_t^n)^*$, where $p_t^i = P(X_t = e^i)$ satisfies the Kolmogorov forward equation

$$\frac{dp_t}{dt} = A(t, u)p_t. \tag{1}$$

Here the control parameter $u \in U$, where U is some compact set in complete metric space and $A(t, u)$ is continuous on $[0, T] \times U$.

Introduce the following right-continuous sets of complete σ-algebras generated by X_t

$$\mathcal{F}_t^X = \sigma\{X_s : s \in [0, t]\}.$$

Assumption 2.2. We assume that the set $\underline{U}$ of admissible controls $\{u(\cdot)\}$ is the set of $\mathcal{F}_t^X$-predictable processes with values in U. This means that, if N_t is the number of the state changes, X_0^t is the series of states occurred from the origin at $t = 0$ until the current time $t \in [0, T]$, that is

$$X_0^t = \{(X_0, 0), (X_{\tau_1}, \tau_1), \dots, (X_{\tau_{N_t}}, \tau_{N_t})\}$$

is the set of states and jump times, then for $\tau_{N_t} < t \le \tau_{N_t+1}$ the control $u_t = u(t, X_0^t)$ is a function of X_0^t and the current time t (see Bremaud (1979, Sec. 2) and Elliott et al. (1995, chap. 12; p. 332)).

For each control function $u(\cdot) \in \underline{U}$ the process $\{X_t\}$ satisfies the following system of stochastic differential equations:

$$X_t = X_0 + \int_0^t A(s, u_s)X_{s-}ds + M_t, \tag{2}$$

where X_0 is the initial condition, and $M_t := \{M_t^1, \dots, M_t^n\}$ is a square integrable $(\mathcal{F}_t^X, P)$ martingales with the following quadratic variations [1](see Elliott et al. (1995, p. 340)):

$$\left\langle M \right\rangle_t = \int_0^t \text{diag } (A(s, u_s)X_s)\, ds - \int_0^t \left[A(s, u_s)(\text{diag } X_s) + (\text{diag } X_s)A^*(s, u_s)\right]ds, \tag{3}$$

where diag X denotes the matrix with diagonal entries $X^1, \dots, X^n$ and A^* denotes the transposed matrix of A.

Remark 2.3. In other words process $X(t) = \{X_t, t \in [0, T]\}$ is a solution of *controlled martingale problem* (2) and (3) for controlled Markov chain (see Elliott (1992, Sec. 5)). That is, an admissible $\mathcal{F}_t^X$ predictable control $u(\cdot)$ defines a probability measure $\mathcal{P}^u$ on the space of piecewise constant right-continuous functions taking values in S. This measure is unique and defines a process $X(t) = \{X_t \in S, t \in [0, T]\}$, satisfying conditions (2) and (3). The martingale in (2) and (3) depends on the control u, however we generally omit this dependence in our notations for the sake of simplicity.

2.2. Performance criterion

The optimization goal is to minimize some cost function of the Markov chain states and controls. This function could take into account the average queue length, which is related with the average time of service, or/and the price of rejected (thinned) demands, since they have to either repeatedly queue or choose another service center. Moreover, in the case of finite time horizon the final state of Markov chain is also very important, like in the case of congestion resolution. So we consider the following performance criterion

$$J[u(\cdot)] = \mathbf{E}\left\{\phi_0(X_T) + \int_0^T f_0(s, u_s, X_s)ds\right\} \to \min_{u(\cdot)}, \tag{4}$$

with

$$\phi_0(X) = \langle \phi_0, X \rangle, \qquad f_0(s, u, X) = \langle f_0(s, u), X \rangle,$$

where $\langle \cdot, \cdot \rangle$ is a sign of scalar product and

$$\phi_0 \in R^n,$$
$$f_0^*(s, u) = (f_0(s, u, e_1), \dots, f_0(s, u, e_n))$$

and each $f_0(s, \cdot, e_i)$ is a cost function when the Markov chain is in state e_i at time $s \in [0, T]$.

Assumption 2.4. (a) Each of functions $f_0(\cdot, \cdot, e_i)$ is continuous on $[0, T] \times U$ and bounded below;

(b) The set of vectors $\{A^i(t, U), f_0(t, U, e_i)\}$, where A^i is the ith column of A, is convex for any $i = 1, \dots, n$ and $t \in [0, T]$.

[1] $\left\langle M \right\rangle_t$ is the quadratic variation of a martingale M, and $\left\langle M, N \right\rangle_t$ is the mutual quadratic variation of martingales M_t and N_t.

2.3. Dynamic programming and optimal control

Define the value function

$$V(t,X)=\inf_{u(\cdot)} J[u(\cdot)|X_t=X], \quad (5)$$

where

$$J[u(\cdot)|X_t=X]=\mathbf{E}\left\{\phi_0(X_T)+\int_t^T f_0(s,u_s,X_s)\mathrm{d}s\,\middle|\, X_t=X\right\}. \quad (6)$$

Moreover, according to Assumption 2.4 the performance criterion (4) is bounded below, therefore, the infimum in (5) exists, and there exists a minimizing sequence of controls $\{u_k(\cdot)\}$. Since for each of controls $u_k(\cdot)$ function

$$J[u_k(\cdot)|X_t=X]=\langle\bar{\phi}_k(t),X\rangle$$

with continuous $\phi_k(t)$, then the function $V(t,X)$ admits the representation

$$V(t,X)=\lim_k\langle\bar{\phi}_k(t),X\rangle=\langle\bar{\phi}(t),X\rangle,$$

with some measurable vector-valued function $\bar{\phi}(t)=(\bar{\phi}^1(t),\dots,\bar{\phi}^n(t))^*\in R^n$.

Consider the following equation (*dynamic programming equation*) with respect to vector-valued function $\phi(t)$

$$\langle\phi'(t),X\rangle+\min_{u\in U}[\langle\phi(t),A(t,u)X\rangle+\langle f_0(t,u),X\rangle]=0, \quad (7)$$

with terminal condition

$$\phi(T)=\phi_0.$$

Remark 2.5. Later on it becomes clear that $\bar{\phi}(t)=\phi(t)$ provided that (7) has a unique solution.

Remark 2.6. The convexity condition of U (Assumption 2.4b) permits to write Eq. (7) in conventional form. Otherwise, we have to write it as follows

$$\langle\phi'(t),X\rangle+\min_{\mu(\mathrm{d}u)}\int_U[\langle\phi(t),A(t,u)X\rangle+\langle f_0(t,u),X\rangle]\,\mu(\mathrm{d}u)=0$$

with the minimum which have to be taken over all probability measures with support U (see, for example, the *Bellman equation* for *relaxed* control in Davis (1993, pp. 147-148, p. 163)). Even if this condition looks rather restrictive in practice it can be realized with the aid of the so-called *weak controls,* which provide the natural convexification of the right-hand side and criterion (see Lee and Marcus (1967, Thm. 4, Sec. 4.2)).

Since the function

$$H(\phi,t,u,X)=\langle\phi,A(t,u)X\rangle+\langle f_0(t,u),X\rangle,$$

is continuous with respect to (t,u) and affine with respect to ϕ, then for any $(t,X)\in[0,T]\times S$ function

$$\mathcal{H}(\phi,t,X)=\min_{u\in U}H(\phi,t,u,X)$$

is Lipschitz in ϕ with the constant $L=\max_{(t,u,X)}\|A(t,u)X\|$ and continuous in t for any $X\in S$.

Remark 2.7. Eq. (7) can be written as a system of ordinary differential equations

$$\frac{\mathrm{d}\phi^i(t)}{\mathrm{d}t}=-\mathcal{H}(\phi(t),t,e_i),\quad i=1,\dots,n \quad (8)$$

which can be obtained by substituting $X=e_i,\quad i=1,\dots,n$. Therefore, the right-hand side of Eq. (8) is Lipschitz in ϕ.

Next result follows from classic results of existence and uniqueness for solutions of ordinary differential equations (see Coddington and Levinson (1955)).

Proposition 1. *Let the Assumptions 2.1 and 2.4 hold. Then Eq. (7) has a unique solution on [0,T].*

The following characterization of the optimal control is the application of the results of Boel and Varaiya (1977, Thm. 5.6) to the case of controlled Markov chain. The proof is based on ideas of M.X.H. Davis (see Davis (1993, Thm. 42.8)).

Theorem 2.8. *Assume that:*

$\phi(t)$ is the solution of system (8) and there exists $u_0(t,X)\in U$ such that at each $(t,X)\in[0,T]\times S$ the value on the left-hand side of (7) and function $H(\phi(t),t,u,X)$ achieves the minimum at $u_0(t,X)$.

Then:

(1) *there exists $\hat{u}(t,X_0^t)$ in the class of $\mathcal{F}_t^X$ predictable controls which is the optimal control and $V(t,X)=\langle\phi(t),X\rangle=J[\hat{u}(\cdot)|X_t=X]$.*

(2) *this optimal control can be chosen as the Markov type control as follows*

$$\hat{u}(t,X_0^t)=u_0(t,X_{t-}).$$

Proof of Theorem 2.8. Let us introduce the notation $SD_{[0,T]}$ for the space of piecewise constant functions $X(t)\in D_{[0,T]}$ and such that $X(t)=\{X_t\in S,t\in[0,T]\}$.

(1) We consider the space of $\omega=X(t)\in SD_{[0,T]}$. For each $(t,\omega)=(t,X(t))\in[0,T]\times SD_{[0,T]}$ there is a $u_0\in U$ such that

$$\mathcal{H}(\phi(t),t,X(t))=H(\phi(t),t,u_0,X(t))=\min_{u\in U}H(\phi(t),t,u,X(t)).$$

According to Wan and Davis (1979, Theorem 4.2) there exists $\hat{u}(t,X_0^t)$ which belongs to the class of $\mathcal{F}_t^X$ predictable controls and such that for any $(t,X_t)\in[0,T]\times S$

$$\hat{u}(t,X_0^t)=\operatorname*{argmin}_{u\in U}H(\phi(t),t,u,X_t).$$

Moreover, according to the Theorem assumption this control can be chosen as the Markov type control

$$\hat{u}(t,X_0^t)=u_0(t,X_{t-}).$$

(2) Now we prove that $\hat{u}$ is the optimal control. Since $\hat{u}(\cdot)$ is a predictable control then for any initial condition $X_0\in S$ there exists a unique solution of the martingale problems (2) and (3). That is there exists a process $X^{\hat{u}}(t)\in SD_{[0,T]}$ satisfying the equation

$$\mathrm{d}X_t^{\hat{u}}=A(t,\hat{u}(t,(X^{\hat{u}})_0^t))X_{t-}^{\hat{u}}\mathrm{d}t+\mathrm{d}M_t^{\hat{u}}, \quad (9)$$

where $M^{\hat{u}}$ is a square integrable $\mathcal{F}_t^X$ martingale with quadratic variation (3).

Take some admissible control $u(s)$ and corresponding solution $X^u(\cdot)$ of the martingale problems (2) and (3) such that $X_t^u=X$. Then we apply Ito's formula to process $\langle\phi(t),X_t^u\rangle$, where $\phi(t)$ is the solution of Eq. (7), and add to both sides the expression

$$\int_t^T\langle f_0(s,u(s)),X_s^u\rangle\mathrm{d}s.$$

We have

$$\langle\phi(T),X_T^u\rangle-\langle\phi(t),X\rangle+\int_t^T\langle f_0(s,u(s)),X_s^u\rangle\mathrm{d}s$$

$$=\int_t^T[\langle\phi'(s),X_s^u\rangle+H(\phi(s),s,u(s),X_s^u)]\mathrm{d}s+\int_t^T\langle\phi(s),\mathrm{d}M^u(s)\rangle.$$

Take the expectation, since $\phi(s)$ is continuous deterministic function the expectation of the integral over martingale is equal to zero and the expectation of the first integral is nonnegative due to Eq. (7). Therefore,

$$J[u(\cdot)|X_t = X] = \mathbf{E}\left\{\langle\phi(T), X_T^u\rangle + \int_t^T \langle f_0(s, u(s)), X_s^u\rangle \mathrm{d}s\right\} \geq \langle\phi(t), X\rangle = V(t, X). \quad (10)$$

Note that the same calculations with the control $\hat{u}(t, X_0^t) = u_0(t, X_{t-})$ and Eq. (9) yields the equality

$$J[\hat{u}(\cdot)|X_t = X] = V(t, X),$$

which completes the proof. □

3. Access and service rate control model

We consider a queuing system that can be controlled by restricting arrivals and by changing of the service rate. We assume that the job flow constitutes a counting process with deterministic rate $\lambda(t) \geq 0$. It means that the arrival process is the process with independent increments (see Liptser and Shiryaev (1979, Thm. 18.9)) The number of jobs in the system is bounded by some constant $N < \infty$ and the service rate is $\mu \in [\underline{\mu}, \overline{\mu}]$, where $\underline{\mu} > 0$. Control $u(t) \in [0, 1]$ is a probability to accept the job at time $t \in [0, T]$. So the part of arriving jobs can be rejected and the performance criterion takes into account the number of rejected jobs and the average queue time for the accepted jobs. Our model is motivated by Hordijk and Spieksma (1989), where this problem is considered in infinite time horizon in the class of stationary controls of *threshold* or *thinning* types.

3.1. Controlled Markov chain model

Assume that the state X is a number of jobs in the system, so the number of states is $N+1$, and the corresponding state space S consists of unit vectors $\{e_0, \dots, e_N\}$.

Proposition 2. *$(N+1)\times(N+1)$ matrix $A(t, u, \mu)$ has a form*

$$A(t, u, \mu) = \begin{pmatrix} -\lambda(t)u & \mu & 0 & \dots & 0 & 0 & 0 \\ \lambda(t)u & -\mu-\lambda(t)u & \mu & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda(t)u & -\mu-\lambda(t)u & \mu \\ 0 & 0 & 0 & \dots & 0 & \lambda(t)u & -\mu \end{pmatrix}, \quad (11)$$

where control $(u, \mu) \in [0, 1] \times [\underline{\mu}, \overline{\mu}]$.

Proof of Proposition 2. Take some $\mathcal{F}_t^X$- predictable controls $u(t)$, $\mu(t)$. Let $N_t \in \{0, \dots, N\}$ is a current number of jobs in the system. This number changes due to the action of two flows: flow of arrivals and departure flow of outcome completed jobs. We assume that the arrival flow forms the counting process with deterministic intensity $\lambda(t) \geq 0$, so this is a *Poisson type process* and the number of demands N_t^a arriving to the system from origin at $t = 0$ until current time can be represented as follows (see Liptser and Shiryaev (1979, Sec. 18.2.4))

$$N_t^a = \int_0^t \lambda(s)\mathrm{d}s + M_t^a,$$

where M_t^a is a square integrable martingale with quadratic variation

$$\left\langle M^a \right\rangle_t = \int_0^t \lambda(s)\mathrm{d}s.$$

The departure flow is the counting process with state-dependent rate $\mu(t)I\{N_t > 0\}$, where $I\{\cdot\}$ is an indicator function. Therefore, departure flow N_t^d admits the representation

$$N_t^d = \int_0^t \mu(s)I\{N_s > 0\}\mathrm{d}s + M_t^d,$$

where M_t^d is a square integrable martingale with quadratic variation

$$\left\langle M^d \right\rangle_t = \int_0^t \mu(s)I\{N_s > 0\}\mathrm{d}s.$$

We suppose, that N_t^a and N_t^d are independent and do not have jumps at the same time, it means that the mutual quadratic variation $\left\langle M^a, M^d \right\rangle_t = 0$.

As shown by P. Bremaud (see in Bremaud (1979, Lemma 1)) the access control can be represented as a control of intensity of the arrival flow. Indeed, if $W(t)$ is an access control, that is it is a random variable, taking values in $\{0, 1\}$ and such that controlled arrival flow is equal to

$$N_t^{a,c} = \sum_{\tau \leq t} I\{N_\tau < N\}I\{W(\tau) = 1\}\Delta N_\tau^a,$$

where τ is the jump instant of N_t^a and $\Delta N_\tau^a = 1$. Then

$$E\{I\{W(t) = 1\}I\{N_t < N\}|\mathcal{F}_t^X\} = u(t)I\{N_t < N\} = u(t)I\{X_t \neq e_N\}, \quad (12)$$

and $u(t) \in [0, 1]$ is $\mathcal{F}_t^X$ predictable process.

Then,

$$\Delta N_t = \Delta N_t^{a,c} - \Delta N_t^d.$$

Taking into account the relation

$$I\{N_t = i\} = I\{X_t = e_i\}$$

one can write

$$\Delta X_t = A^+ X_{t-}\Delta N_t^{a,c} + A^- X_{t-}\Delta N_t^d,$$

where $(N+1)\times(N+1)$ matrices A^+, A^- have the following form

$$A^+ = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

$$A^- = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ 0 & 0 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & -1 \end{pmatrix}.$$

Then, by using the relation

$$X_t = X_0 + \sum_{\tau \leq t} \Delta X_\tau,$$

and taking into account that of $N_t^{a,c}$ and N_t^d are counting processes we obtain

$$X_t = X_0 + \sum_{\tau \leq t}[A^+ X_{\tau-}\Delta N_\tau^{a,c} + A^- X_{\tau-}\Delta N_\tau^d] = X_0 + \int_0^t A^+ X_{s-}\mathrm{d}N_s^{a,c} + \int_0^t A^- X_{s-}\mathrm{d}N_s^d.$$

Finally, we have to substitute the martingale representation of $N_t^{a,c}$ and N_t^d into the above equation and by taking the conditional expectation with respect to $\mathcal{F}_t^X$ we obtain that

$$X_t = X_0 + \int_0^t [A^+\lambda(s)u(s) + A^-\mu(s)]X_s \mathrm{d}s + M_t^{u,\mu}$$
$$= \int_0^t A(s, u(s), \mu(s))X_s \mathrm{d}s + M_t^{u,\mu},$$

where $M_t^{u,\mu}$ is a square integrable $\mathcal{F}_t^X$ martingale with quadratic variation

$$\left\langle M^{u,\mu}\right\rangle_t = \int_0^t [A^+X_sX_s^*(A^+)^*\lambda(s)u(s) + A^-X_sX_s^*(A^-)^*\mu(s)]\mathrm{d}s.$$

Routine calculation shows that this expression coincides with (3). □

Remark 3.1. Notice that the control set $U = [0, 1] \times [\underline{\mu}, \overline{\mu}]$ is convex and compact, therefore, the result of Theorem 2.8 can be applied with proper performance criteria.

3.2. The performance criterion

As we mentioned above the performance criterion takes into account the average time in queue, which can be estimated as follows

$$J_1 = \mathbf{E}\left\{\int_0^T \frac{N_s}{\mu(s)}\mathrm{d}s\right\} = \mathbf{E}\left\{\int_0^T \frac{\langle \mathbf{l}, X_s\rangle}{\mu(s)}\mathrm{d}s, \right\}$$

where

$$\mathbf{l}^* = (0, 1, 2, \dots, N) \in R^{N+1}.$$

Remark 3.2. If the service rate is a constant then J_1/T is the average value of waiting time (Little's formula).

Another criterion to be minimized is an average number of rejected jobs, which can be calculated with the aid of the integral representation given by Wong and Hajek (see Wong and Hajek (1985, p. 261, Lemma 3.2)) and by using the relation (12) as follows

$$J_2 = \mathbf{E}\{N_T^a - N_T^{a,c}\} = \mathbf{E}\left\{\int_0^T [1 - u(\tau)\langle \mathbf{1}, X_\tau\rangle]\lambda(\tau)\mathrm{d}\tau\right\},$$

where

$$\mathbf{1}^* = (1, 1, \dots, 1, 0) \in R^{N+1}.$$

Third criterion represents the service resources spent during the control interval

$$J_3 = \int_0^T \mu(\tau)\langle \mathbf{I}, X_\tau\rangle \mathrm{d}\tau,$$

where

$$\mathbf{I}^* = (0, 1, \dots, 1) \in R^{N+1},$$

Further we consider the performance criterion which is a mixture of J_1, J_2 and J_3, that is

$$J = k_1J_1 + k_2J_2 + k_3J_3, \tag{13}$$

where $k_i \geq 0$, $i = 1 \dots 3$.

3.3. Dynamic programming equation and optimal control

So we have to solve Eq. (7), where $A(t, u, \mu)$ is defined by (11), and cost function is defined by the function $f_0(t, u, \mu, X)$

$$f_0(t, u, \mu, X) = k_1\frac{\langle \mathbf{l}, X\rangle}{\mu} + k_2[1 - u\langle \mathbf{1}, X\rangle]\lambda(t) + k_3\mu\langle \mathbf{I}, X\rangle.$$

The Hamiltonian $H(\phi, t, u, \mu, X)$ is affine in u, i.e.

$$H(\phi, t, u, X) = H_0(\phi, t, \mu, X) + uH_1(\phi, t, X)$$

then the dynamic programming equation can be reduced to the following system of ordinary differential equations

$$0 = \frac{\mathrm{d}\phi^i(t)}{\mathrm{d}t} + \min_{\begin{pmatrix} u \in [0,1] \\ \mu \in [\underline{\mu}, \overline{\mu}]\end{pmatrix}} [H_0(\phi, t, \mu, e_i) + uH_1(\phi, t, e_i)]$$
$$= \frac{\mathrm{d}\phi^i(t)}{\mathrm{d}t} + \min_{u\in[0,1]}[\min_{\mu\in[\underline{\mu},\overline{\mu}]} H_0(\phi, t, \mu, e_i) + uH_1(\phi, t, e_i)]$$

for $i = 0, \dots, N$, where,

$$H_0(\phi, t, \mu, e_i) = \mu\langle\phi, A^-e_i\rangle + k_1\frac{\mathbf{l}_i}{\mu} + k_2\lambda(t) + k_3\mu\mathbf{I}_i,$$
$$H_1(\phi, t, e_i) = \lambda(t)[\langle\phi, A^+e_i\rangle - k_2\mathbf{1}_i].$$

Functions ϕ^i can be found from the system of equations

$$\frac{\mathrm{d}\phi^i(t)}{\mathrm{d}t} = -\min\{\min_\mu H_0(\phi, t, \mu, e_i), \min_\mu H_0(\phi(t), t, \mu, e_i)$$
$$+ H_1(\phi(t), t, e_i)\}, \quad \phi^i(T) = \phi_0^i, \tag{14}$$

and the optimal control $\mu(t, e_i)$ is equal

$$\mu(t, e_i) = \begin{cases} \sqrt{\frac{a}{b}} & \text{if } \sqrt{\frac{a}{b}} \in [\underline{\mu}, \overline{\mu}], \quad b > 0, \\ \underline{\mu} & \text{if } \sqrt{\frac{a}{b}} < \underline{\mu}, \quad b > 0, \\ \overline{\mu} & \text{if } \sqrt{\frac{a}{b}} > \overline{\mu}, \quad b > 0, \\ \overline{\mu} & \text{if } b \leq 0, \end{cases} \tag{15}$$

where

$$a(t, e_i) = k_1\mathbf{l}_i \geq 0, \quad b(t, e_i) = \langle\phi(t), A^-e_i\rangle + k_3\mathbf{I}_i.$$

The optimal control $u(t, e_i)$ is calculated with the aid of relation

$$u(t, e_i) = \begin{cases} 1, & \text{if } H_1(\phi(t), t, e_i) \leq 0, \\ 0, & \text{if } H_1(\phi(t), t, e_i) > 0. \end{cases} \tag{16}$$

Here,

$$H_0(\phi, t, \mu, e_i)$$
$$= \begin{cases} k_2\lambda(t) & \text{for } i = 0, \\ \mu(\phi^{i-1} - \phi^i) + \frac{k_1 i}{\mu} + k_2\lambda(t) + k_3\mu & \text{for } 0 < i < N, \\ \mu(\phi^{N-1} - \phi^N) + \frac{k_1 N}{\mu} + k_2\lambda(t) + k_3\mu & \text{for } i = N, \end{cases} \tag{17}$$

and

$$H_1(\phi, t, e_i) = \begin{cases} \lambda(t)(-\phi^0 + \phi^1 - k_2) & \text{for } i = 0, \\ \lambda(t)(-\phi^i + \phi^{i+1} - k_2) & \text{for } 0 < i < N, \\ 0 & \text{for } i = N. \end{cases} \tag{18}$$

Remark 3.3. Notice that for each state $e_i \in S$ the optimal control can be chosen as Borelean measurable function which coincides with (15) and (16) almost everywhere in [0,T]. Moreover, the number of the state changes is finite almost surely and therefore, the composition $u(t, X_t)$ is progressively measurable one, and there exists a predictable version of this control. Full details can be found in Wan and Davis (1979).