Chapter-9

A FINITE CAPACITY M/G/1 QUEUE WITH VACATION

9.1 INTRODUCTION

In the ordinary M/G/1 queue, a poisson stream of customers with i.i.d. service times of arbitrary distribution enters a single server facility with continuous service. When the system capacity is finite with room for atmost N customers in the system, i.e. in queue or in service, the system will be designated by M/G/1/N.

Queueing systems with server vacations arise naturally as models of many diverse fields such as computer, communication and production systems. Under specified conditions, the server after finishing the customer in service, discontinues service for an independent vacation period. When a vacation period ends, either customers are present in the queue or the queue is empty. If customers are present, service is resumed. Otherwise, a new vacation period, having the same distribution as the previous one and independent of it, begins. Customers are served in the order of their arrival. Different models are distinguished by the rules which determine when service stops and a vacation begins. For an M/G/1 vacation system with exhaustive service, i.i.d vacations are performed.
whenever the queue is empty. In a single service discipline a vacation period begins after every service completion, or after any vacation period if the queue is empty. In a Bernoulli schedule discipline the server begins a vacation if the queue is empty. If at a service completion epoch the queue is not empty, the service is resumed with probability $p$ and with probability $1-p$ a vacation commences.

An extensive literature on single server vacation systems has been developed in recent years. Notable among them are Courtois (1980), Scholl and Kleinrock (1983), Fuhrman and Cooper (1985), Keilson and Servi (1986a), Levy and Yechiali (1975), Heymann (1977), Ramachandran Nair (1987) etc. For a survey of Queueing vacation models one may refer to Doshi (1986). While various server vacation policies have been considered for the M/G/1 system with infinite capacity, very few papers deal with the finite system. Finite capacity systems have been studied previously in Hoskstad (1977), Miller (1975), Levenberg (1975), Courtois (1980) and Loris-Teghem (1988). Lee (1984) studied an M/G/1 queue with finite waiting space and server vacation where he considered two types of service discipline: viz. 1) exhaustive service discipline, and 2) limited service discipline ie., the server will begin a vacation if either the queue has been emptied or $M$ customers have been served during a visit. In all these studies the steady state behaviour of the system is of major concern.
In this chapter we consider a rather large class of vacation policies which contains, in particular, the policies described above. A precise description of the model is given below. The model is analysed, using Markov renewal theory, in section 9.2. The time dependent queue size distribution is obtained in section 9.3 and virtual waiting time distribution is attempted in section 9.4.

We consider an M/G/1/N queueing system where an arriving customer finding N customers present in the system may not enter the system and is lost. Customers arrive according to a Poisson process of rate \( \mu \). The successive service times are i.i.d random variables with distribution function \( G(\cdot) \) and they are also independent of interarrival times. The queue discipline is FIFO. The server goes for vacation either when the queue becomes empty or after serving a random number of customers. ie, if the queue is empty after a service completion then the server begins a vacation period for a duration having a general distribution. At the end of a vacation period service begins if at least one customer is present in the queue. Otherwise, one or more additional vacations are repeated until at least one customer is present (multiple vacation). If \( k \) units were served continuously since his arrival to the system after the last vacation, the server may go for vacation with
probability $p_k$ and resumes his service with probability $1-p_k$. We have $p_1 < p_2 < \ldots < p_M = 1$, where $M$ is the maximum number of customers served in a busy period. The vacation times are i.i.d random variables with a common probability distribution function $H(.)$ and they are assumed to be independent of the interarrival times and service times. Without loss of generality we assume that $N < M$.

One can model many service systems using the present model. In certain cases it often happens that the service rate decreases with the increasing number of units served. In that case the service time of customers also increases and thereby results in increased cost to the system. Hence there will be an optimal number of units the server can be allowed to serve continuously so as to make the system most efficient and least expensive.

The following notations are used in the sequel:

The lower case letters denote the probability density functions (assuming that they exist).

$$
\mu_n(x) = \frac{e^{-\mu x} (\mu x)^n}{n!}, \ n = 0,1,2, \ldots, N-1
$$

$$
\mu_N(x) = \sum_{n=N}^{\infty} \frac{e^{-\mu x} (\mu x)^n}{n!}
$$
* denotes the convolution

\[ f^{*n}(.) = n\text{-fold convolution of } f(.) \text{ with itself (} f^{*0}(.) \equiv 1 \]  

\[ N^0 = \{0, 1, 2, \ldots, \} \]  

\[ E = \{0, 1, 2, \ldots, N\} \]  

\[ b(x) = \sum_{m=0}^{\infty} h^{*m}(x) \]  

\[ \langle r \rangle = \max(o, r). \]

9.2 ANALYSIS OF THE MODEL

Let us denote by \( X(t) \) the number of customers present in the system at time \( t \). Then the process

\[ X = \{X(t), \ t \geq 0\} \]

is a semi-regenerative process with state space \( E = \{0, 1, 2, \ldots, N\} \) and the following embedded Markov renewal process.

Let \( 0 = T_0, T_1, T_2, \ldots \) be the instants of successive busy period terminations, and \( X_n \) be the number of customers left behind at the termination of the \( n^{\text{th}} \) busy period. Then \( (X, T) = \{(X_n, T_n); \ n \in N^0\} \) is a time homogeneous Markov renewal process. The associated semi-Markov kernel over the set \( E \) is

\[ Q = \{Q(i, j, t) : i, j \in E, \ t \geq 0\} \]
where \( Q(i,j,t) = \Pr \{ X_{n+1} = j, T_{n+1} - T_n \leq t \mid X_n = i \} \) is given by

\[
Q(i,j,t) = \begin{cases} 
   \int_0^t \sum_{k=\langle j-1 \rangle}^{M-i+j} \mu_k(u)(h^*(i+k-j))(u)p_{i+k-j}du, & i \neq 0, j \in E; \\
   \int_0^t e^{-\mu_b(u)} \sum_{k=0}^{M} \mu_j+k(t-u)(h^*(k-j))(t-u)p_{k-j}du, & i = 0, j \in E
\end{cases}
\tag{9.2.1}
\]

Let \( Q(t) = [Q(i,j,t)]_{i,j \in E} \)

Define for \( n \in \mathbb{N}^0 \)

\[
Q^n(i,j,t) = \Pr \{ X_n = j, T_n \leq t \mid X_0 = i \} , i,j \in E, t \geq 0.
\]

Then

\[
Q^0(i,j,t) = \begin{cases} 
   1 & \text{if } i = j \text{ for all } t \geq 0, \\
   0 & \text{if } i \neq j
\end{cases}
\]

and we have, for \( n \geq 0 \), the recurrence relation

\[
Q^{n+1}(i,k,t) = \sum_{j \in E} \int_0^t Q(i,j,du)Q^n(j,k,t-u)
\]
Then the Markov renewal functions are given by

\[ R(i,j,t) = \sum_{n=0}^{\infty} Q^n(i,j,t), \quad i,j \in E, \quad t \geq 0 \quad (9.2.2) \]

Denote \( R(t) = \{ R(i,j,t) \}_{i,j \in E} \). It is the Markov renewal kernel corresponding to \( Q(t) \).

Here we note that, since the state space \( E \) is finite we can compute the Markov renewal kernel by the relation

\[ \hat{R}_\alpha = (I - \hat{Q}_\alpha)^{-1} \quad (9.2.3) \]

where \( \hat{Q}_\alpha = \{ \hat{Q}_\alpha(i,j) \}_{i,j \in E} \) and \( \hat{R}_\alpha = \{ \hat{R}_\alpha(i,j) \}_{i,j \in E} \),

\[ \hat{Q}_\alpha(i,j) = \int_0^\infty e^{-\alpha t} Q(i,j,dt), \]

\[ \hat{R}_\alpha(i,j) = \int_0^\infty e^{-\alpha t} R(i,j,dt). \]

9.3 THE STATE SPACE AND TRANSIENT SYSTEM SIZE PROBABILITIES

Our basic concern is with the process \( X(t) \), the number of customers in the system at time \( t \). Consider the trivariate stochastic process \( \mathbf{X}(t) = (X(t), Y(t), Z(t)) \), where

\[ X(t) = \text{number of customers in the system} \]
Y(t) = \begin{cases} 
0 & \text{if a vacation is in progress at time } t \\
1 & \text{if a service is in progress at time } t 
\end{cases}

and Z(t) = \begin{cases} 
\text{number of customers served up to } t \text{ since} \\
\text{the commencement of the current busy} \\
\text{period (termination of the last vacation} \\
\text{period), if a service is in progress at} \\
\text{time } t \\
0, \text{if a vacation is in progress at time } t.
\end{cases}

This process may be discussed on its state space

\[ S = \left\{ (i,j,k); 0 \leq i \leq N, j = 0,1, 0 \leq k \leq M-1 \right\}. \]

We assume that at time \( t = 0 \), the server just completes a busy period and enters a vacation period so that the initial state of the process is

\[ W(0) = (X(0), Y(0), Z(0)) = (a,0,0) \text{ for some } a \in E \]

For each \( a,i \in E \), \( j \in \{0,1\} \), \( k \in \{0,1,2,\ldots,M-1\} \), \( t \geq 0 \) define

\[ P(t,a,i,j,k) = \Pr \left[ W(t) = (i,j,k) \mid W(0) = (a,0,0) \right] \]

Then we have
Theorem 9.3.1. For any \( a, i \in E, j \in \{0,1\}, 0 \leq k \leq M-1 \) and \( t \geq 0 \),

\[
P(t, a, i, j, k) = \sum_{l \in E} \int_0^t R(a, l, du) K(t-u, l, i, j, k)\]

where \( K(t, l, i, j, k) = \text{Pr}\{W(t)=(i,j,k), T_1 > t | \mathcal{Y}(0) = (l,0,0)\} \)

are as given below.

\[
K(t, l, i, j, k) = \begin{cases} 
\int_0^t e^{-\mu u} b(u) \mu_i(t-u)[1-H(t-u)]du \\
& \text{for } l=0, 0 \leq i \leq N, j=0, k=0;
\end{cases}
\]

\[
\int_0^t e^{-\mu u} b(u) \int_0^t h(v-u) \mu_{i+k}(t-u)[G^k(t-v)-G^*(k+1)(t-v)]dv du \\
& (1-p_k)dv du \\
& \text{for } l=0, 1 \leq i \leq N, j=1, 0 \leq k \leq M-1; 
\]

\[
[l-H(t)] \mu_{i-l}(t) \text{ for } 1 \leq l \leq N, l \leq i \leq N, j=0, k=0; 
\]

\[
\mu_{i+k-l}(t) \int_0^t h(u) [G^k(t-u)-G^*(k+1)(t-u)](1-p_k)du \\
& \text{for } 1 \leq l \leq N, l \leq i \leq N, j=1, 1 \leq k \leq M-1 
\]

\[
\mu_{i-l}(t) \int_0^t h(u)[1-G(t-u)]du \\
& \text{for } 1 \leq l \leq N, l \leq i \leq N, j=1, k=0 
\]

0 otherwise
Proof:

\[ P(t,a,i,j,k) = Pr \{ W(t) = (i,j,k), T_1 > t \mid W(0) = (a,0,0) \} \]
\[ + Pr \{ W(t) = (i,j,k), T_1 < t \mid W(0) = (a,0,0) \} \]
\[ = K(t,a,i,j,k) + \sum_{l \in E_0} \int_{R(a,l,du)} P(t-u,l,i,j,k) \]

which is a Markov renewal equation (see for eg. Cinlar (1975b)) and its solution is given by

\[ P(t,a,i,j,k) = \sum_{l \in E_0} \int_{R(a,l,du)} K(t-u,l,i,j,k) \]

Since there are only finitely many states, this solution is unique.

Now the expressions for \( K(t,l,i,j,k) \) can be seen to be as given in the statement of the theorem by considering different possible values of \( l,i,j,k \).

\[ Q.E.D. \]

9.4 VIRTUAL WAITING TIME DISTRIBUTION

Let \( \eta(t) \) be the virtual waiting time of a customer in the queue. Here we obtain the probability distribution of \( \eta(t) \) conditioned on the state of the system at time \( t \).
Denote $\varphi_t(x/(i,j,k)) = Pr \{ \eta(t) \leq x \mid \omega(t) = (i,j,k), \omega(0) = (a,0,0) \}$

Then we consider the following cases separately.

Case (i): $i=0, j=0, k=0$,

$$\varphi_t(x/(i,j,k)) = \int_0^t R(a,0,du) \mu_0(t-u) \int_u^t b(v-u) \left[ H(t+x-v) - H(t-v) \right] dv \ du$$

Case (ii): $1 \leq i \leq N, j=0, k=0$

$$\varphi_t(x/(i,j,k)) = \int_0^t R(a,0,du) \mu_i(t-u) \int_u^t b(w-u) \int_w^{t+x} h(v-w) \left( \sum_{k=0}^i (G^i H^k)(t+x-v) \right) dv \ dw \ du$$

$$+ \int_1^t \int_0^t R(a,0,du) \mu_{i-l}(t-u) \int_u^{t+x} h(v-u) \left( \sum_{n=0}^{i-l} (G^i H^n)(t+x-v) \right) dv \ du$$

and finally,

Case (iii): $1 \leq i \leq N, j=l, 0 \leq k \leq M-1$

$$\varphi_t(x/(i,j,t)) = \int_0^{t+j} \int_1^{t+k} R(a,l,du) \mu_{i+k-l}(t-u) \int_u^{t+x} h(v-u) \left( \sum_{n=1}^{i+k} (G^{i+k} H^n)(t+x-v) \right) dv \ du$$
\[ + \int_0^t R(a, O, du) \mu_{i+k}(t-u) \int_{u}^{t} b(w-u) \int_{u}^{t+x} h(v-w) \]

\[ \sum_{n=1}^{i+k} (G^*(k+i) \ast H^*n)(t+x-v)dv \, du. \]

**Remark:**

In this model if we put \( M=1 \) and \( p_1=1 \) we get single service discipline and if we put \( M=\infty \) and \( p_1=p \) \((i=1, 2, \ldots)\) we have the Bernoulli schedule discipline and when \( p=0 \) it becomes the exhaustive service discipline.
CONCLUDING REMARKS

In this thesis, (s,S) inventory systems with non-identically distributed interarrival demand times and random lead times, state dependent demands, varying ordering levels and perishable commodities with exponential life times have been studied. The queueing system of the type $E^k/G^a,b/1$ with server vacations, service systems with single and batch services, queueing system with phase type arrival and service processes and finite capacity $M/G/1$ queue when server going for vacation after serving a random number of customers are also analysed.

In inventory theory, one can extend the present study to the case of multi-item, multi-echelon problems. The study of perishable inventory problem when the commodities have a general life time distribution would be a quite interesting problem. The analogy between the queueing systems and inventory systems could be exploited in solving certain models.

Consider an inventory system with more than one ordering levels and more than one server. Assume that some of the servers take vacation when the inventory is less than a prescribed quantity. Here again one can investigate the transient as well as steady state solutions.
The techniques used to derive the time dependent solutions may be of special interest to any stochastic system having regenerative or semi-regenerative structure. The most important problem one can think of is to develop an algorithm to compute the given transient solutions numerically. To the application point of view this is a quite worthwhile work. For developing the algorithm, possibly one can effectively use some fast transform techniques (see, Elliott and Rao (1982)) because here we cannot use the usual procedure of Laplace transforms.

In vacation models, one important result is the stochastic decomposition property of the system size or waiting time. One can think of extending this to the transient case. The distribution of virtual waiting time may be used for the decomposition property of the waiting time, since it can be defined as the unfinished work (see Kleinrock (1975)).