韓國經營科學會誌
第20卷 第2號
1995年 8月

휴가가 존재하는 M/M/2 대기 시스템의 한계치를 이용한 제어정책

이 효 성*

The Threshold Policy in the M/M/2 Queue with Server Vacations

Hyo-Seong Lee*

Abstract

In this study, a threshold policy is considered for the M/M/2 queueing system with server vacations. The probability generating function for the number of customers present in the system is derived using an embedded Markov chain approach. Then, assuming a linear cost structure, an efficient procedure to find an optimal threshold policy is presented. The Laplace-Stieltjes transform for the waiting time of an arbitrary customer under a "FIFO" discipline is also derived.

1. 서론

대기 시스템을 경제적으로 운용하기 위해서는 대기 시스템을 효율적으로 제어하기 위한 정책이 필요하다. 대기 시스템의 제어에 대한 연구는 주어진 비용구조하에서 시스템을 가장 경제적으로 운용할 수 있는 정책을 도출함을 목적으로 하며 서비스 속도나 서어버의 서비스 시점, 고객의 도착률 등을 조절함으로써 대기 시스템의 제어가 이루어진다[4,11]. 특히 서어버의 유휴기간이 끝나고 서비스가 시작될 때 준비비용(setup cost)이 발생하는 경우에는 빈번한 준비비용의 발생을 피하기 위하여 대기중인 고객의 수가 특정 한계치를 넘어섰을 때 서비스를 시작하는 제어정책을 사용하게 된다. 이런 형태의 제어정책을 한계치를 이용한 제어정책(threshold control policy)이라 하며, 이에 대해서는 Yadin[12], Heyman[5, 6], Bell[3] 등의 연구를 기점으로 그동안 많은 연구가 수행되어져 왔

* 경희대학교 공과대학 산업공학과

다. 최근에는 M/G/1 서어버 휴가모형과 관련하여 한계치를 이용한 대기 시스템의 제어정책이 활발히 연구되어 지고 있다. Kella[9]는 M/G/1 서어버 휴가 시스템에 대하여 한계치를 이용한 최적 제어정책을 연구하였으며 이효성과 Srinivasan[10]은 $M^X/G/1$ 서어버 복수 휴가 시스템에 대하여 한계치를 이용한 최적 제어정책을 연구하였다. 이효성[2]은 후속 연구로서 상기 시스템의 안정상태확률을 효율적으로 구하기 위한 알고리듬을 개발하였다. 이순석[1]은 $M^X/G/1$ 서어버 단일 휴가 시스템 및 복수 휴가 시스템에 대하여 한계치를 이용한 제어정책을 부가 변수법(supplementary variable technique)을 이용하여 분석하였으며 이호우 등[7]은 한계치를 이용한 제어정책이 사용되는 $M^X/G/1$ 서어버 복수 휴가 시스템의 운용특성을 연구하였다.

앞에서 언급된 연구는 모두 서어버의 수가 한명인 단일 서어버 시스템에 국한된 연구이다. 한계치를 이용한 복수 서어버 시스템의 제어정책 연구는 매우 희귀하며 Igaki[8]의 연구 이외에는 찾아보기 힘든 것으로 보인다. Igaki는 다음과 같은 제어정책이 사용되는 M/M/2 서어버 복수 휴가 모형과 서어버 단일 휴가 모형을 분석하였다. 서어버의 가동기간(busy period)이 끝나 시스템 내에 고객이 존재하지 않게 되면 두명의 서어버 중 한명은 즉시 휴가를 떠나고 다른 한명의 서어버는 시스템에 남아 새로 도착하는 고객의 서비스를 담당한다. 휴가를 떠난 서어버가 휴가에서 돌아왔을 때 시스템에 존재하는 고객의 수가 한계치 N을 넘으면, 휴가에서 돌아온 서어버도 즉시 서비스에 투입된다. 서어버가 첫번째 휴가에서 돌아왔을 때 시스템에 존재하는 고객의 수가 N 미만이면 이때 취할 수 있는 서어버의 행태(behavior)에 따라 다음과 같은 두가지 종류의 모형이 고려될 수 있다. 첫째는 서어버가 휴가에서 돌아왔을 때 시스템에 존재하는 고객의 수가 N이상이 될때까지 서어버가 반복해서 휴가를 떠나는 서어버 복수 휴가 모형이다. 둘째는 첫번째 휴가에서 돌아온 서어버는 더이상 휴가를 떠나지 않고 고객의 수가 N이 될 때까지 시스템에 계속 남아 있는 서어버 단일 휴가 모형이다. 위의 두가지 모형에 대해 Igaki는 안정상태확률을 구하는 식을 유도하였으며 유도된 식의 추계적 분해성질(decomposition property)을 규명하였다.

설명된 바와 같이 Igaki의 연구에서는 가동기간이 끝난 후 한명의 서어버 만이 휴가를 떠나는 것으로 가정하였다. 그러나 현실상황에서는 두명의 서어버가 모두 휴가를 떠나고 휴가에서 돌아온 서어버가 동시에 서비스를 시작하는 경우도 흔히 존재한다. 이러한 사실을 감안하여 본 연구에서는 두명의 서어버가 모두 휴가를 떠나고 대기중인 고객의 수가 미리 정해진 한계치 N을 넘어섰을 때 두명의 서어버가 동시에 서비스를 시작하는 M/M/2 대기시스템의 제어정책을 분석하고자 한다. 분석을 위하여 본 연구에서는 embedded Markov chain 기법이 이용되며, 고객 수에 대한 pgf(probability generating function) 및 선입선출(FIFO)방식 하에서의 고객의 대기시간에 대한 LST(Laplace-Stieltjes transform)가 유도된다. 또한 유도된 식의 특성을 이용하여 선형비용구조를 가정한 후 단위기간당 기대비용을 최소화해주는 최적 제어정책을 구하게 된다.

2. 수학적 모형

본 연구는 다음과 같은 대기 시스템에 대하여 한계치를 이용한 제어정책을 분석한다. 서비스를 받기 위한 고객은 도착률이 λ인 Poisson 과정에 따라 도착하고 한명의 고객이 서비스를 받는데 소요되는 시간은 서비스률이 μ인 지수분포를 따른다. 서어버의 수는 2명이고 대기 시스템의 용량은 무한하다고 가정한다. 서비스는 선입선출방식에 따라 이루어지며 서비스가 일단 시작되면 두명의 서어버는 대기 시스템에 고객이 존재하지 않을 때까지 서비스를 계속한다(exhaustive service discipline). 서비스가 끝나 대기 시스템에 고객이 존재하지 않으면 서어버의 가동기간(busy period)이 끝나고 유휴기간(idle period)이 시작된다. 유휴기간은 도착한 고객의 수가 제어한계치 N 이상이 되면 끝나게 되고 유휴기간이 끝나는 대로 2명의 서어버는 서비스를 시작하게 된다. 이러한 시스템이 안정상태에 도달하기 위한 조건은 $\rho=\frac{\lambda}{2\mu}<1$ 이며 본 연구에서는 이 조건이 충족된다고 가정한다. 한계치를 이용한 이와 같은 제어모형은 유휴기간 중에 취할 수 있는 서어버의 행태(behavior)에 따라 다음과 같이 세가지 형태로 분류될 수 있다. 첫째, 유휴기간 동안 서어버가 시스템을 떠나지 않고 고객을 기다리는 경우(서어버의 휴가가 존재하지 않는 모형) 둘째, 가동기간이 끝나면 서어버는 V 시간이 소요되는 휴가를 반복해서 떠나며, 서어버가 휴가에서 돌아온 순간 대기중인 고객의 수가 제어 한계치를 최초로 넘어서면 서비스를 시작하는 모형(서어버 복수 휴가 모형) 셋째, 서어버는 휴가를 한번 만 떠나며 서어버가 휴가에서 돌아온 순간 대기중인 고객의 수가 제어 한계치에 미달되면 제어 한계치에 도달될 때까지 서어버가 시스템에서 기다리는 모형(서어버 단일 휴가 모형)으로 분류된다.

서어버의 휴가기간중 두명의 서어버는 부차적인 일을 같이 수행하는 경우가 많다. 따라서 본 연구에서는 두명의 서어버가 휴가기간을 독립적으로 취하지 않고 동시에 휴가를 떠나서 동시에 휴가로부터 돌아온다고 가정한다. 이와 같은 상황 하에서 본 연구에서는 위 세가지 모형에 모두 적용 가능한 일반식을 유도하고자 하며 유도된 일반식을 이용하여 시스템의 중요한 성능 분석치를 구해 보고자 한다. 본 연구에서는 또한, 가동기간이 시작되기 전 일정한 기간의 준비시간(setup time)이 필요한 경우의 분석도 아울러 수행한다.

본 연구에서 분석하고자 하는 대기 시스템은 고객의 도착이 집단으로 이루어지지 않고 고객의 서비스도 집단으로 이루어지지 않는다. 따라서 시스템에 도착하는 고객이 본 시스템의 상태(도착 시점에서의 시스템의 상태)와 서비스를 마치고 시스템을 떠나는 고객이 본 시스템의 상태(서비스 종료시점에서의 시스템의 상태)는 동일한 확률분포를 갖는다. 또한 고객의 도착이 Poisson 과정을 따르므로 PASTA (Poisson Arrivals See Time Averages) 성질에 의해 시스템에 도착하는 고객은 시스템을 안정상태 하에서 보게 된다. 따라서 시스템을 떠나는 고객도 시스템을 안정상태 하에서 보게 되며 시스템에 존재하는 고객수에 대한 안정상태 확률을 구하기 위해서는 서비스 종료시점에서의 고객수의 확률분포를 구하면 된다.

ξ_n을 n번째 서비스가 종료된 직후의 시스템 내의 고객수라 정의하자. 그러면 서비스 종료 시점에서의 고객수가 j일 안정상태확률은 다음과 같이 표현된다.

$$\pi_j = \lim_{n\to\infty} \Pr\{\xi_n = j\}.$$

π_j를 구하기 위하여 다음 기호를 도입한다.

I_N = 제어 한계치로 N이 사용될 경우 가동기간이 시작되는 시점에서의 시스템 내의 고객수,

$I_N(z)$ = I_N의 확률발생함수(pgf),

S = 고객 한명이 서비스를 받는데 소요되는 시간,

U = 두명의 서어버가 모두 서비스를 수행하고 있을 때 고객 한명의 서비스가 끝날 때까지 소요되는 시간,

$S(t)$, $U(t)$=S와 U의 분포함수

a_n = S 동안 도착한 고객의 수가 n일 확률,

b_n = U 동안 도착한 고객의 수가 n일 확률.

$S(t)$는 평균이 $\frac{1}{\mu}$인 지수분포를 따르고 $U(t)$는 평균이 $\frac{2}{\mu}$인 지수분포를 따름을 확인할 수 있다. 또한 π_j를 구하기 위해 필요한 a_n, b_n의 값은 다음과 같이 계산한다.

$$a_n = \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^n}{n!}\, dS(t) = \left(\frac{\lambda}{\lambda+\mu}\right)^n\left(\frac{\mu}{\lambda+\mu}\right), \tag{1.a}$$

$$b_n = \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^n}{n!}\, dU(t) = \left(\frac{\lambda}{\lambda+2\mu}\right)^n\left(\frac{2\mu}{\lambda+2\mu}\right). \tag{1.b}$$

만일 제어 한계치의 값이 $N \ge 2$이면 ξ_{n+1}에 대한 다음 관계식이 성립함을 보일 수 있다.

$j \ge 1$:

$$\Pr\{\xi_{n+1}=j/\xi_n=i\} = \begin{cases} \sum_{k=N}^{j+1} \Pr\{I_N=k\} b_{j-k+1}, & i=0 \\ (1-a_0)b_{j-1}, & i=1 \\ b_{j-i+1}, & 2\le i\le j+1 \\ 0 & i\ge j+2 \end{cases} \tag{2.a}$$

$j=0$:

$$\Pr\{\xi_{n+1}=j/\xi_n=i\} = \begin{cases} a_0, & i=1 \\ 0, & i\ne 1 \end{cases} \tag{2.b}$$

식 (2)로부터 ξ_{n+1}은 ξ_n에만 의존하고 $\xi_m(m<n)$에는 의존하지 않음을 알 수 있다. 따라서 $\{\xi_n, n=1, 2,\cdots\}$은 Markov chain을 이루며, Markov chain의 안정상태확률 π_j는 다음과 같은 식으로 표현된다.

$$\pi_j = \begin{cases} \pi_0 \sum_{k=N}^{j+1} Pr\{I_N=k\} b_{j-k+1} + \pi_1(1-a_0)b_{j-1} \\ + \sum_{i=2}^{j+1} \pi_i b_{j-i+1}, & j\ge 1 \\ \pi_1 a_0, & j=0 \end{cases} \tag{3}$$

식(1)을 이용하면 π_j의 pgf $\Pi(z)$는 다음과 같이 표현된다.

$$\begin{aligned}\Pi(z) &= \sum_{j=0}^{\infty} \pi_j z^j \\ &= \pi_1 a_0 + \pi_0 \sum_{j=1}^{\infty}\sum_{k=N}^{j+1} Pr\{I_N=k\} b_{j-k+1} z^j + \\ &\quad \pi_1 \sum_{j=1}^{\infty}(1-a_0)b_{j-1}z^j + \sum_{j=1}^{\infty}\sum_{i=2}^{j+1} \pi_i b_{j-i+1} z^j \\ &= \pi_1\left(\frac{\mu}{\lambda+\mu}\right) + \left(\frac{2\mu}{\lambda+2\mu}\right)\pi_0 \sum_{k=N}^{\infty}\sum_{j=k-1}^{\infty} \\ &\quad Pr\{I_N=k\}\left(\frac{\lambda}{\lambda+2\mu}\right)^{j-k+1} z^j \\ &\quad + \pi_1\left(\frac{\mu}{\lambda+\mu}\right)\left(\frac{2\mu}{\lambda+2\mu}\right)\sum_{j=1}^{\infty}\left(\frac{\lambda}{\lambda+2\mu}\right)^{j-1} z^j \\ &\quad + \left(\frac{2\mu}{\lambda+2\mu}\right)\sum_{i=2}^{\infty}\pi_i \sum_{j=i-1}^{\infty}\left(\frac{\lambda}{\lambda+2\mu}\right)^{j-i+1} z^j.\end{aligned}$$

$\alpha=\left(\frac{\lambda}{\lambda+2\mu}\right)$, $\beta=\left(\frac{\lambda}{\lambda+\mu}\right)$라 놓고 정리하면 위 식은 다음과 같다.

$$\Pi(z)=(1-\beta)\ \pi_1+\frac{(1-\alpha)\pi_0\ I_N(z)}{z(1-\alpha z)}+\frac{(1-\alpha)\beta\pi_1 z}{1-\alpha z}+\frac{(1-\alpha)}{z(1-\alpha z)}\{\Pi(z)-\pi_0-\pi_1 z\}. \quad (4)$$

식(4)를 Π(z)에 대해서 풀은 후 $\pi_0=\pi_1 a_0$의 관계를 이용하여 정리하면 다음과 같다.

$$\Pi(z)=\frac{\alpha z^2-\alpha z+\alpha-1+(1-\alpha)I_N(z)}{-\alpha z^2+z+\alpha-1}\ \pi_0. \quad (5)$$

식(5)는 $N\geq 2$라는 가정 하에서 유도된 결과이다. 그러나 $N=1$일 경우(제어가 존재하지 않는 경우)에도 동일한 방법으로 유도하면 Π(z)는 식(5)와 같음을 보일 수 있으며 따라서 식(5)는 모든 N의 값에서 성립한다. 식 (5)에 $\lim_{z\to 1}\Pi(z)=1$의 관계를 이용하면 π_0는 다음과 같이 구해진다.

$$\pi_0=\frac{1-2\alpha}{\alpha+(1-\alpha)E(I_N)}=\frac{1-\rho}{E(I_N)+\rho}. \quad (6)$$

따라서 시스템 내의 고객 수에 대한 pgf는 다음과 같다.

$$\Pi(z)=\frac{\alpha z^2-\alpha z+\alpha-1+(1-\alpha)I_N(z)}{-\alpha z^2+z+\alpha-1}\cdot\frac{1-\rho}{E(I_N)+\rho}. \quad (7)$$

$i_N^{(1)}=E(I_N)$, $i_N^{(2)}=[E(I_N(I_N-1)]$으로 놓고 식(7)을 이용하여 시스템 내의 기대 고객 수를 구하면 다음과 같다.

$$L_N=\frac{\rho}{1-\rho}+\frac{\rho+\frac{1}{2}i_N^{(2)}}{\rho+i_N^{(1)}}. \quad (8)$$

식 (8)에 Little의 공식을 적용하면 임의의 고객의 시스템 내에서의 기대 체류 시간은 $T_N=\frac{L_N}{\lambda}$으로 주어진다. 식(7)과 (8)로 부터 $I_N(z)$만 구할 수 있으면 시스템 내에 존재하는 고객수에 대한 pgf 및 기대 고객 수를 구할 수 있음을 알 수 있다. 세가지 모형에 대해 $I_N(z)$의 값은 다음과 같이 구해진다.

(1) 서어버의 휴가가 존재하지 않는 모형

이 모형은 가동기간이 끝나도 서어버가 시스템을 떠나지 않고 고객을 기다리는 경우의 모형이다. 서어버는 도착한 고객의 총 수가 N이 되는 순간 서비스를 시작한다. 따라서 이 경우의 $I_N(z)$ 및 $i_N^{(1)}$, $i_N^{(2)}$는 다음과 같다.

$$I_N(z)=z^N,\ i_N^{(1)}=N,\ i_N^{(2)}=N(N-1). \quad (9)$$

$N=1$일 경우는 제어가 존재하지 않는 모형, 즉 $M/M/2$ 모형으로 귀착되며, 이 경우 $I_N(z)=z$를 식(7)에 대입하여 정리하면 $\Pi(z)=\frac{(1-\rho)(1+\rho z)}{(1+\rho)(1-\rho z)}$를 얻게 된다.

(2) 서어버 복수 휴가 모형

서어버 복수 휴가 모형에서는 서비스가 끝나 대기행렬 내에 고객이 존재하지 않으면 두 명의 서어버는 휴가를 떠난다. 휴가기간의 길이는 일반분포 $V(t)$를 따른다고 가정하며, 휴가에서 돌아올 때마다 서어버는 대기 시스템에서 대기 중인 고객의 총 수가 N을 넘어섰는지 조사한다. 만일 대기 중인 고객의 수가 N 미만이면 두명의 서어버는 또다시 휴가를 떠나고 N이상이면 두명의 서어버는 즉시 서비스를 시작하게 된다. 이 경우의 $I_N(z)$는 $M^X/G/1$시스템의 연구에서 이순석[1]에 의해 다음과 같이 유도되었다.

$$I_N(z) = 1 + \frac{V^*(\lambda-\lambda z) - 1}{1 - r_0} \sum_{j=0}^{N-1} \phi_j z^j. \qquad (10)$$

여기에서 r_k는 한번의 서어버 휴가기간 동안 k명의 고객이 도착할 확률이고, $V^*(\theta)$는 $V(t)$의 LST이다. 또한 ϕ_j는 다음과 같이 정의된다.

$$\phi_j = \begin{cases} 1, & j = 0 \\ \sum_{i=1}^{j} \frac{r_i}{1 - r_0} \phi_{j-i}, & j > 0 \end{cases}$$

식(10)로 부터 I_N의 1차, 2차 적률을 구하면 다음과 같다.

$$\mathbf{i}_N^{(1)} = \mathbf{i}_{N-1}^{(1)} + \frac{\lambda \mathbf{E}(\mathbf{V})}{1 - \mathbf{r}_0} \phi_{N-1},$$

$$\mathbf{i}_N^{(2)} = \mathbf{i}_{N-1}^{(2)} + \frac{\lambda^2 \mathbf{E}(\mathbf{V}^2) - 2\lambda \mathbf{E}(\mathbf{V})(\mathbf{N}-1)}{1 - \mathbf{r}_0} \phi_{N-1} \quad (11)$$

(3) 서어버 단일 휴가 모형

서어버 단일 휴가 모형에서는 가동기간이 끝난 후 서어버는 단 한번만의 휴가를 갖는다. 휴가기간의 길이는 일반분포 $V(t)$를 따르며, 서어버가 휴가에서 돌아왔을 때 시스템에 대기중인 고객의 수가 N이상이면 즉시 서비스를 시작하고 N미만이면 고객의 수가 N에 이를 때까지 서어버는 시스템을 떠나지 않고 기다린다. 이 경우의 $I_N(z)$도 $M^X/G/1$ 시스템의 연구에서 이순석[1]에 의해 유도되었으며 그 형태는 다음과 같다.

$$I_N(z) = V^*(\lambda - \lambda z) + (z-1) \sum_{j=0}^{N-1} \psi_j z^j, \qquad (12)$$

여기서 $\psi_j = \sum_{i=0}^{j} r_j$.

식(12)로 부터 I_N의 1차, 2차 적률을 구하면 다음과 같다.

$$i_N^{(1)} = i_{N-1}^{(1)} + \psi_{N-1}.$$

$$i_N^{(2)} = i_{N-1}^{(2)} + 2(N-1)\psi_{N-1}. \qquad (13)$$

(4) 준비시간이 존재하는 경우의 모형

유휴기간이 끝나고 서어버의 서비스가 시작되기 위해서는 준비시간(setup time)이 필요할 경우가 있다. 이 경우 준비시간이 확률분포 $H(t)$를 따른다면 $I_N(z)$는 다음과 같이 표현된다.

$$I_N(z) = \tilde{I}_N(z) H^*(\lambda - \lambda z). \qquad (14)$$

식(14)에서 $H^*(\theta)$는 $H(t)$의 LST이고 $\tilde{I}_N(z)$는 유휴기간이 끝난 시점에서의 고객수에 대한 pgf로서 식(9),(10),(12)로부터 구해질 수 있다.

3. 최적 제어 정책

본 절에서는 2절의 결과를 이용하여 선형비용구조 하에서 단위기간당 발생하는 기대비용을 최소화하는 제어정책을 구해보고자 한다. 본 연구에서 가정하는 비용구조는 다음과 같다. 가동기간이 시작될 때 준비비용(setup cost) K가 발생하고 고객 한명이 단위기간 동안 시스템에서 체류함에 따라 대기비용(waiting cost) h가 발생한다. 이러한 비용구조 하에서 단위기간당 발생하는 기대비용은 단위기간당 발생하는 기대 대기비용 $h \cdot L_N$과 단위기간당 발생하는 기대 준비비용의 합으로 표시된다. 단위기간당 발생하는 기대준비비용은 K를 사이클의 기대길이 X_N으로 나눈 값이 된다.

다. 사이클의 기대길이 X_N은 다음과 같이 구하여 질 수 있다. 유휴기간의 기대길이를 A, 가동기간의 기대길이를 B라고 하면 유휴기간의 기대길이는 $A = \frac{i_N^{(1)}}{\lambda}$이고, 한 사이클 중 고객이 존재하지 않는 기대시간은 $\frac{1}{\lambda}$이 된다. 이러한 사실과 재생과정에서의 정리(regenerative process theorem)으로 부터 시스템 내에 고객이 존재하지 않을 안정상태확률은 다음과 같이 표현된다.

$$\pi_0 = \frac{\text{한 사이클 동안 고객이 존재하지 않는 기대시간}}{\text{한 사이클의 기대길이}}$$

$$= \frac{1}{i_N^{(1)} + \lambda B} \tag{15}$$

식(6),(15)로 부터 가동기간의 기대길이와 사이클의 기대길이는 각각 다음과 같이 얻어진다.

$$B = \frac{\rho(1 + i_N^{(1)})}{\lambda(1-\rho)}$$

$$X_N = A + B = \frac{\rho + i_N^{(1)}}{\lambda(1-\rho)} \tag{16}$$

이상의 결과로 부터 N-정책을 사용할 경우 단위기간당 발생하는 기대비용 C_N은 다음과 같이 주어진다.

$$C_N = \frac{K}{X_N} + hL_N$$

$$= \frac{K\lambda(1-\rho)}{\rho + i_N^{(1)}} + h\left\{\frac{\rho}{1-\rho} + \frac{\rho + \frac{1}{2} i_N^{(2)}}{\rho + i_N^{(1)}}\right\}$$

$$= \frac{\theta + \frac{h}{2} i_N^{(2)}}{\rho + i_N^{(1)}} + \frac{\rho h}{1-\rho}, \tag{17}$$

여기서 $\theta = K\lambda(1-\rho) + h\rho$.

식 (17)로 부터 C_N은 N에 대하여 단조증가하거나 unimodal한 함수임을 보일 수 있다. 이에대한 증명은 $M^X/G/1$시스템에서의 증명방법[10]과 동일하므로 생략한다. 따라서 단위기간당 기대비용을 최소화 해주는 최적제어 한계치 N^*는 다음과 같이 구하여 진다.

$$N^* = \min\{N \in I / C_N < C_{N+1}\}, \tag{18}$$

여기서 I는 자연수 집합을 의미한다.

이로부터 최적제어 한계치는 $C_N < C_{N+1}$을 최초로 만족하는 N에 도달할 때까지 $N = 1$로 부터 C_N의 값을 순차적으로 계산해 나가면 구할 수 있음을 알 수 있다. C_N을 구성하는 $i_N^{(1)}$, $i_N^{(2)}$의 순환적 구조로 부터 이와같은 순차적 탐색방법은 매우 효율적인 방법임이 입증된다.

4. 대기시간의 분포

앞에서 유도된 바와 같이 한 사이클의 기대길이는 $\frac{\rho + i_N^{(1)}}{\lambda(1-\rho)}$, 유휴기간의 기대길이는 $\frac{i_N^{(1)}}{\lambda}$, 가동기간의 기대길이는 $\frac{\rho(1 + i_N^{(1)})}{\lambda(1-\rho)}$이다. 따라서 한 사이클 내에서 유휴기간이 차지하는 시간적 비율은 $\frac{(1-\rho)\, i_N^{(1)}}{\rho + i_N^{(1)}}$이고 가동기간이 차지하는 시간적 비율은 $\frac{\rho(1 + i_N^{(1)})}{\rho + i_N^{(1)}}$이다. 유휴기간 동안 도착한 임의의 고객이 대기하는 시간(서비스 시간을 제외한)의 LST를 $W_1^*(\theta)$, 가동기간 중에 도착한 임의의 고객이 대기하는 시간의 LST를 $W_2^*(\theta)$라 하면 임의의 고객의 대기시간의 LST, $W^*(\theta)$는 다음과 같다.

$$W^*(\theta) = \frac{(1-\rho)\, i_N^{(1)}}{\rho + i_N^{(1)}} W_1^*(\theta) + \frac{\rho(1 + i_N^{(1)})}{\rho + i_N^{(1)}} W_2^*(\theta). \tag{19}$$

따라서 $W_1^*(\theta)$ 및 $W_2^*(\theta)$를 구하면 임의의 고객의 대기시간의 LST를 구할 수 있으며, 이들 값은 다음과 같이 구하여 질 수 있다.

유휴기간 중 도착한 임의의 고객이 도착시점에서 본 시스템 내의 고객수를 Q_1, 가동기간 중 도착한 임의의 고객이 도착시점에서 본 시스템 내의 고객수를 Q_2라 하고 이들의 pgf를 각각 $Q_1(z)$, $Q_2(z)$라 하면 다음식이 성립된다.

$$\Pi(z)=\frac{(1-\rho)\,i_N^{(1)}}{\rho+i_N^{(1)}}\,Q_1(z)+\frac{\rho(1+i_N^{(1)})}{\rho+i_N^{(1)}}\,Q_2(z). \tag{20}$$

식 (20)의 $Q_1(z)$는 이산적 시간을 갖는 재생과정(discrete time renewal process)에서의 backward recurrence time 식을 이용하면 다음과 같이 표현된다.

$$Q_1(z)=\frac{1-I_N(z)}{i_N^{(1)}(1-z)} \tag{21}$$

$Q_1(z)$가 구하여졌으므로 식 (7), (20)을 이용하여 이제 $Q_2(z)$를 구할 수 있다. 가동기간 중 도착한 임의의 고객은 도착시점에서 항상 한명 이상의 고객이 시스템에 존재함을 보며($Q_2\geq 1$), 이 고객이 서비스를 받기 전 대기하여야만 하는 대기시간은 Q_2-1명이 서비스를 끝내는데 소요되는 시간과 같으므로 $W_2^*(\theta)$는 다음과 같이 표현된다.

$$W_2^*(\theta)=\frac{Q_2(U^*(\theta))}{U^*(\theta)}, \tag{22}$$

여기서 $U^*(\theta)=\dfrac{2\mu}{\theta+2\mu}$이며 $U(t)$의 LST 이다.

$W_1^*(\theta)$는 $M^X/G/1$ 시스템에서의 연구결과 [1]를 이용하여 구하여질 수 있으며, 유휴기간 동안의 서어버의 행태에 따라 각기 구하는 방법이 달라진다. 특히 서어버의 휴가가 존재하는 경우에는 $W_1^*(\theta)$의 유도는 복잡하다. 여기에서는 서어버의 휴가가 존재하지 않는 모형에서의 $W_1^*(\theta)$를 구하여 보도록 한다. 서어버의 휴가가 존재하지 않는 모형에서 유휴기간 중 도착한 임의의 고객이 도착시점에서 n명의 고객을 볼 확률은 $\dfrac{1}{N}$이다 ($n=0,1,\cdots,N-1$).

n이 0이나 1일 경우에는 이 고객은 가동기간이 시작되는 즉시 서비스를 받을 수 있으며 따라서 서비스를 받기전 대기해야만하는 시간은 $N-n-1$명의 고객이 도착하는데 소요되는 시간과 같다. n이 2이상인 경우에는 이 고객이 대기해야만 하는 시간은 $(N-n-1)$명의 고객이 도착하는데 소요되는 시간과 먼저 도착한 n명의 고객 중 $n-1$명의 고객이 서비스 를 끝내는데 소요되는 시간(서어버의 수가 2이므로)의 합이다. $(n-1)$명의 고객이 서비스를 끝마치는 데 소요되는 시간의 LST는 $\{U^*(\theta)\}^{n-1}$이므로 $W_1^*(\theta)$는 다음과 같이 표현된다.

$$\begin{aligned}W_1^*(\theta)&=\frac{1}{N}\Big[\Big(\frac{\lambda}{\theta+\lambda}\Big)^{N-1}+\Big(\frac{\lambda}{\theta+\lambda}\Big)^{N-2}\\&\quad+\sum_{n=2}^{N-1}\Big(\frac{\lambda}{\theta+\lambda}\Big)^{N-n-1}\{U^*(\theta)\}^{n-1}\Big]\\&=\frac{1}{N}\Big[\Big(\frac{\lambda}{\theta+\lambda}\Big)^{N-1}\\&\quad+\frac{\Big(\frac{\lambda}{\theta+\lambda}\Big)^{n-1}-\{U^*(\theta)\}^{n-1}}{\Big(\frac{\lambda}{\theta+\lambda}\Big)-\{U^*(\theta)\}}\Big].\end{aligned} \tag{23}$$

휴가가 존재하지 않는 모형에서는 $I_N(z)=z^N$이므로 $Q_1(z)=\dfrac{1-z^N}{N(1-z)}$으로 주어지고 식 (7)과 (20)을 이용하면 $Q_2(z)$는 다음과 같다.

$$Q_2(z)=\frac{(1-\rho)z}{(1+N)}\,\frac{(2-z-z^N)}{(1-z)(1-\rho z)}. \tag{24}$$

식(22)로 부터 $W_2^*(\theta)$는 다음과 같이 주어진다.

$$W_2^*(\theta) =$$

$$= \frac{(1-\rho)}{(1+N)} \frac{\{2 - U^*(\theta) - (U^*(\theta))^N\}}{\{1 - U^*(\theta)\}\{1 - \rho U^*(\theta)\}} . \quad (25)$$

고객의 대기시간의 LST, $W^*(\theta)$는 식(19), (23),(25)로 부터 다음과 같이 표현된다.

$$W^*(\theta) =$$

$$\frac{1-\rho}{N+\rho} \left\{ \left(\frac{\lambda}{\theta+\lambda}\right)^{N-1} + \frac{\left(\frac{\lambda}{\theta+\lambda}\right)^{N-1} - (U^*(\theta))^{N-1}}{\left(\frac{\lambda}{\theta+\lambda}\right) - U^*(\theta)} \right\}$$

$$+ \frac{\rho(1-\rho)}{N+\rho} \left\{ \frac{2 - U^*(\theta) - (U^*(\theta))^N}{(1 - U^*(\theta))(1 - \rho U^*(\theta))} \right\} \quad (26)$$

식 (26)은 제어가 존재하지 않는 경우(N=1) 에도 성립함을 쉽게 확인할 수 있다.

5. 결 론

본 연구에서는 한계치를 이용한 M/M/2 서어버 휴가 시스템의 제어정책이 분석되었다. 서어버의 휴가가 존재하지 않는 모형, 서어버 단일 휴가 모형, 서어버 복수 휴가 모형 등 세 가지 모형이 고려되었으며 이들 세가지 모형에서 모두 성립되는 고객 수의 분포에 대한 pgf가 유도되었다. 또한 선형비용구조 하에서 단위 기간당 발생하는 기대비용을 최소화 해주는 최적제어정책을 찾기 위한 방법이 제시되었으며 선입선출방식 하에서의 고객의 대기시간의 LST를 구하기 위한 방법이 제시되었다. 이러한 결과를 얻기 위하여 본 연구에서는 embeded Markov chain기법이 이용되었다. 그러나 부가변수법을 이용하여도 본 연구에서 얻은 결과와 동일한 결과가 유도될 수 있으리라 판단된다. 본 연구와 관련된 향후 연구로는 고객수의 분포에 대한 확률적 분해 성질 규명과 그에 대한 확률적 해석을 들 수 있겠다.

참 고 문 헌

[1] 이순석, 「Threshold와 휴가가 있는 집단 대기행렬의 운영 특성」, 박사학위논문, 성균관대학교 산업공학과, 1993.

[2] 이효성, "제어운영정책 하에 있는 집단으로 도착하는 서어버 휴가모형의 안정상태 확률," 「한국경영과학회지」, 제16권, 제2호 (1991). pp. 36~48.

[3] Bell, C. E., "*Characterization and Computation of Optimal Policies for Operating an M/G/1 Queueing System with Removable Server*," Opns. Res., Vol. 19 (1971), pp. 208~218.

[4] Grabill, T. D., D. Gross and N. Magazine, "*A Classified Bibliography of Research on Optimal Design and Control of Queues*," Opns. Res., Vol. 25(1977), pp. 219~232.

[5] Heyman, D. P., "*Optimal Operating Policies for M/G/1 Queueing Systems*," Opns. Res., Vol. 16(1968), pp. 362~382.

[6] Heyman, D. P., "*The T-Policy for the M/G/1 Queue*." Mgmt. Sci., Vol. 23 (1977), pp. 775~778.

[7] Lee, H. W., H. S. Lee and K. C. Chae, "*Operating Charactristics of $M^X/G/1$ Queue with N-Policy*." *Queueing Systems*. Vol. 15(1994), pp. 387~399.

[8] Igaki, N., "*Exponential Two Server Queue with N-Policy and General*

Vacations," Queueing Systems, Vol. 10 (1992), pp. 279~294.

[9] Kella, O., "*The Threshold Policy in the M/G/1 Queue with Server Vacations*," Naval Research Logistics, Vol. 36(1989), pp. 111~123.

[10] Lee, H. S. and M. M. Srinivasan, "*Control Policies for the $M^X/G/1$ Queueing System*," Mgmt. Sci., Vol. 35(1989), pp. 708~721.

[11] Teghem, J., "*Control of the Service Process in a Queueing System*," Euro. J. of Oper. Res., Vol. 23(1986), pp. 141~158.

[12] Yadin, M. and P. Naor, "*Queueing Systems with a Removable Service Station*," Oper. Res., Quart., Vol. 14 (1963), pp. 393~405.