Journal of the Korean Institute of Industrial Engineers
Vol. 26, No. 2, pp. 81-87, June 2000.

N-정책과 T-정책이 적용되는 M/G/1 시스템의 분석†

허 선 · 이훈규 · 김종수

한양대학교 산업공학과

An Analysis of M/G/1 System with N and T-Policy

Sun Hur · Hun-Gyu Lee · Jong-Soo Kim

As for M/G/1 queueing system, we use various control policies, with which we can optimize the system. Up to now the most widely adopted policies are N-Policy, T-Policy, D-Policy, and so on.
The existing researches are largely concerned to find an optimal operation condition or to optimize the system under single policy in M/G/1 system. There are, however, few literatures dealing with multiple control policies at once to enhance the flexibility of the model. In this study, we consider M/G/1 system adopting N-Policy and T-Policy simultaneously. If one of two conditions is satisfied, then, the server starts the service. We call this Min(N,T)-Policy. We find the probability distribution of the number of customers and mean waiting time in steady state and derive a cost function. Next, we seek the N^*, optimal threshold under various N values. Finally, we reveal the characteristics of cost function.

1. 서 론

대기시스템의 분석에 있어서 운용상황에 변화를 주어 어떠한 목적을 달성하기 위하여 제어정책을 사용하게 된다. 예를 들어 시스템의 운용비용을 최소화하기 위하여 서버의 바쁜기간이 시작되는 조건을 준다든가, 또는 시스템을 떠났다가 돌아오는 시간 등을 조절하는 것이다. 대표적인 제어정책에는 N-정책, T-정책, D-정책 등이 있다. N-정책이란 시스템내의 고객수가 N명이 되어야 유휴중인 서버가 서비스를 시작하는 제어정책이고, T-정책이란 서비스할 고객이 없으면 서버는 시스템을 떠나서 고정된 시간 T시간 후에 돌아와 고객이 존재하면 서비스를 개시하고 서비스할 고객이 없으면 다시 T시간의 휴가를 떠나는 제어정책이다. 그리고 D-정책이란 도착하는 고객들의 서비스시간의 총합, 즉 총 로드가 임계치 D를 넘어야 유휴중인 서버가 서비스를 시작하는 제어정책이다. 위의 제어정책은 한 가지의 변수만을 고려하므로 단일제어정책이라 한다.

다중제어정책이란 N-정책, D-정책, T-정책 등의 단일제어정책의 조합으로 만들어지는 제어정책이다. 즉, 위의 단일제어정책 세 가지 중 두 가지를 조합하여 두 정책 중 하나가 먼저 만족되면 바쁜 기간을 시작하거나(Min정책), 두 정책이 모두 만족되어야 바쁜 기간이 시작되도록 하는 정책(Max정책)이 있을 수 있다. 이 밖에 위의 세 가지의 정책을 조합하여 세 가지의 조건 중 어느 하나가 먼저 만족하면 바쁜 기간이 시작되는 Min(N, D, T)-정책과 두 가지의 조건이 만족하면 바쁜 기간이 시작되는 Med(N, D, T)-정책, 그리고 세 가지의 조건이 모두 만족해야 바쁜 기간이 시작되는 Max(N, D, T)-정책이 있다.

제어정책이 있는 대기 시스템에 관한 연구는 Yadin과 Naor (1963)가 N-정책에 관한 연구를 한 이후 Heyman (1968), Sobel (1969), Bell (1971) 등에 의해 확장되었고, Minh (1992), Medhi와 Templeton (1992) 등은 준비기간을 고려한 M/G/1 대기 시스템에서의 N-정책에 관하여 분석하였고, Lee와 Park (1997)은 시스템의 고객수가 N명이 되기 전에 미리 setup을 시작하는 조기준비정책하의 N-정책에 대하여 연구하였다. 또한, Hur와 Paik (1999)은 준비기간과 N-정책을 갖는 M/G/1 대기시스템에서 고객의 도착률이 시스템의 상태에 따라 달라지는 경우 단위 시간당 대기비용을 최소화하는 최적의 제어정책에 관하여 연구하였다. T-정책이 적용되는 시스템은 Heyman(1977)에 의하여 최초로 소개가 되었는데 여기서 고정된 휴가의 길이 T를 확률변수로 만들면 복수휴가형 모형과 같다.

D-정책은 Balachandran(1973)에 의하여 최초로 연구가 시작되

† 이 논문은 1999년도 두뇌한국21사업 핵심분야에 의하여 지원되었음.

였고 Balachandran과 Tijms (1975)는 선형비용구조를 가질 때 D-정책이 N-정책보다 상위에 있음을 보였다. 위의 연구들은 N, T, D 등의 단일 변수하에서 운영되는 대기 시스템에서 최적의 비용 또는 최적의 운영조건을 찾아내는 데 중점을 둔 논문들이다.

다중제어정책에 관한 논문으로, Kella(1989)는 고정된 휴가의 길이 T를 확률변수 V로 만들어서 Max(N, T) - 정책을 보다 일반화한 Max (N, V) - 정책에서 단위 시간당 평균비용을 최소화하는 N의 결정에 대하여 연구를 하였고, Rhee (1997)는 다중 제어정책에서 가상확률밀도함수를 이용하여 바쁜 기간의 분포를 연구하였다. 또한, Gakis 등 (1995)은 단일제어정책과 이변수 제어정책에서 바쁜 기간과 평균길이에 대하여 연구하였다.

그러나 지금까지는 비용함수를 고려한 최적의 다중제어정책을 고려한 연구가 거의 없다. 이 연구에서는 N-정책과 T-정책을 동시에 고려한 Min(N, T)-정책을 적용하는 M/G/1 대기모형의 여러 성능치들을 구하고 재가동비용, 고객유지비용, 고객검사비용 등이 포함된 총 비용함수를 도입하여 이를 최적화하는 N 값을 구하는 것을 목적으로 한다.

이 Min(N, T) 정책의 적용사례로서 다음과 같은 불량품 수선 문제를 들 수가 있다. 어떤 라인에서 생산되는 제품들 중 포아송과정에 의해 발생되는 불량품은 용량이 N인 저장장소에 보관된다. 기술자는 이 불량품을 수리하기도 하지만 또다른 어떤 일(예: 서류작업)도 한다. 이 일은 고정된 시간 T가 걸린다. 이 일이 끝나면 기술자는 저장장소를 살펴보고 불량품이 발생한 것이 있으면 이를 수리하고 불량품이 없으면 다시 서류작업을 계속한다. 그러나 서류작업 중이라도 불량품의 개수가 N을 넘어서면 서류작업을 중단하고 불량품의 수리를 한다.

2. 기호와 모델의 설명

이 연구는 안정상태에 있는 M/G/1 대기행렬모형을 연구대상으로 한다. 고객은 도착률이 λ인 푸아송과정으로 시스템에 도착한다. 고객의 서비스시간은 서로 독립이고, 평균이 $1/\mu$인 동일한 확률분포를 따른다. 서비스 규칙은 FIFO를 따른다. 시스템과 고객 모집단의 크기는 무한대이며, 기타 사항들은 일반적인 M/G/1 대기모형의 가정에 준한다.

이 연구에서 사용될 기호는 다음과 같다.

(i) $A(t)$: 시간 t까지 시스템에 도착하는 고객수

$$P[A(t)=n]=\frac{(\lambda t)^n}{n!}e^{-\lambda t},\quad n=0,1,2,\cdots$$

(ii) S_n : n 번째 고객의 시스템 도착 시간

$$P[A(t)=n]=P[S_n\le t]-P[S_{n+1}\le t]$$

(iii) $F_n(t)$: n 번째 고객의 도착시점을 나타내는 분포함수

$$F_n(t)=\int_0^t\frac{\lambda(\lambda t)^{n-1}}{(n-1)!}e^{-\lambda t}dt=P[A(t)\ge n],\quad t\ge 0$$

$$P[A(t)=n]=F_n(t)-F_{n+1}(t),\quad n=0,1,2,\cdots$$

(iv) S : 서비스시간을 나타내는 확률변수

이 모형은 제어정책이 있는 대기행렬 시스템에서 대기하는 고객수가 $N(\ge 1)$명이 되면 유휴중인 서버가 서비스를 시작하는 N-정책과, 서비스할 고객이 없으면 서버가 시스템을 떠났다가 T(상수)시간 후에 돌아와 대기공간을 조사하여 고객이 존재하면 서비스를 시작하고 고객이 없으면 다시 T기간의 휴가를 떠나는 T-정책 중 먼저 한 조건이 만족되면 서비스를 시작하는 제어정책을 사용한다. 이러한 제어정책을 Min (N, T)-정책이라고 부르기로 한다.

Min(N,T)-정책에서 서버가 서비스를 제공하기 시작하는 것은 다음 두 가지 경우로 나누어진다. 첫째, N-정책에 의하여 바쁜 기간이 시작되는 경우이다. N-정책에 의하여 바쁜 기간이 시작되려면, 서버가 휴가를 떠난 시점을 0이라고 했을 때 $[0, T]$기간 동안에 고객의 도착이 N명 이상이어야 한다. 반면 T-정책에 의하여 바쁜 기간이 시작되려면 $[0, T]$기간 동안에 고객의 도착이 1명 이상 $N-1$명 이하이어야 한다.

바쁜 기간이 시작되면 서버는 시스템을 setup하게 되는데, 이 연구에서는 setup시간은 무시하고 비용만 고려한다. 또한, 서버는 T시간 동안의 휴가를 마치고 돌아오면 휴가를 다시 떠날지의 여부를 결정하기 위하여 시스템내의 고객수를 검사하게 되는데 매 검사순간마다 검사비용이 발생한다. 또한, 고객이 시스템에 단위 시간 동안 머물게 되면 고객유지비용이 발생한다.

3. 고객수 PGF와 대기시간

이 절에서는 고객수의 PGF를 유도하고 이를 이용하여 평균 고객수와 평균 대기시간을 구한다.

3.1 고객수 PGF

N-정책으로 바쁜 기간이 시작되려면 2절에서 지적한 바와 같이 서버가 휴가를 떠난 시점을 0이라고 했을 때 (구간 $[0, T]$ 동안 고객이 도착하지 않았으면 서버는 다시 휴가를 떠나고 시점 T를 다시 0으로 둔다) 구간 $[0, T]$ 동안에 실제로 고객의 도착이 있었다는 조건하에 고객의 도착이 N명 이상이어야 한다. 따라서 N-정책에 의해 바쁜 기간이 시작될 확률 P_N은 다음 식 (1)과 같다.

$$P_N=\Pr(T\ \text{동안}\ N\text{명 이상 도착}\mid \text{고객의 도착이 있음})$$

$$= \sum_{n=N}^{\infty} [F_n(T) - F_{n+1}(T)]/F_1(T)$$
$$= \frac{1}{F_1(T)} F_N(T) \tag{1}$$

T-정책에 의해 바쁜 기간이 시작되려면 N-정책에 의한 경우와 유사하게 $[0, T]$ 동안에 도착한 고객의 수가 1명 이상 $N-1$명 이하이어야 한다. 이 확률을 P_T라 하면 다음 식 (2)와 같다.

$$P_T = \Pr(T \text{ 동안 1명 이상 } N-1\text{명 이하 도착} \mid \text{고객의 도착이 있음})$$
$$= \sum_{n=1}^{N-1} [F_n(T) - F_{n+1}(T)]/F_1(T)$$
$$= \frac{1}{F_1(T)} [F_1(T) - F_N(T)] \tag{2}$$

위에서 구한 P_N과 P_T는 서로 배반사상이고 바쁜 기간이 시작될 수 있는 모든 경우를 포함하고 있으므로 그 합은 1 이 됨을 알 수 있다.

Min(N, T)-정책에서 바쁜 기간 시작시의 고객수를 N_B, 이의 PGF를 $N_B(z)$라 하자. $N_B(z)$는 식(1)과 (2)를 이용하여 구하면 다음과 같다.

$$N_B(z) = \frac{1}{F_1(T)} \left\{ \sum_{n=1}^{N-1} z^n [F_n(T) - F_{n+1}(T)] + z^N F_N(T) \right\} \tag{3}$$

따라서 바쁜 기간 시작시 평균 고객수는 다음과 같다.

$$\frac{d}{dz} N_B(z)|_{z=1} = E(N_B) = \frac{1}{F_1(T)} \sum_{n=1}^{N} F_n(T) \tag{4}$$

식 (4)를 살펴보면, Min(N, T)-정책에서 바쁜 기간 시작시의 평균 고객수는 휴가가 없는 일반적인 M/G/1 대기행렬모형의 바쁜 기간 시작시의 평균 고객수 1의 $\frac{1}{F_1(T)} \sum_{n=1}^{N} F_n(T)$배가 됨을 알 수 있는데 이는 제어정책에 의하여 바쁜 기간 시작시의 평균 고객수가 증가하기 때문이다.

일반적으로 완전서비스를 갖는 휴가형 대기행렬시스템에서 고객수 PGF는 확률적 분해성질에 의해 다음과 같이 분해가 된다(Fuhrmann and Cooper, 1985).

$$P_{vac}(z) = P_{M/G/1}(z) \cdot \frac{1 - N_B(z)}{E(N_B)(1-z)} \tag{5}$$

이 연구에서 사용된 Min(N, T)-정책하의 M/G/1 모형도 휴가형 대기행렬모형에 해당되며 따라서 고객수 PGF는 식 (5)를 이용하여 식 (6)과 같이 구할 수 있다.

$$P_{Min(N,T)}(z) = P_{M/G/1}(z) \cdot \frac{1 - N_B(z)}{E(N_B)(1-z)}$$
$$= \frac{(1-\rho) S^*(\lambda - \lambda z)}{S^*(\lambda - \lambda z) - z} \times \frac{F_1(T) - \left[\sum_{n=1}^{N-1} z^n (F_n(T) - F_{n+1}(T)) + z^N F_N(T) \right]}{\sum_{n=1}^{N} F_n(T)} \tag{6}$$

위 식에서 $T \to \infty$일 경우 고객수 PGF는 N-정책에서의 PGF와 같고, $N \to \infty$일 경우 T-정책에서의 고객수 PGF와 일치하는 것을 확인할 수 있다. 또한 $T \to \infty$이고 N을 1로 두면 일반적인 M/G/1 대기행렬모형과 동일하게 된다(이호우, 1998).

3.2 평균 고객수

식 (6)으로부터 안정상태에서의 평균 고객수 $L_{Min(N,T)}$는 다음 식 (7)과 같이 구할 수 있다.

$$L_{Min(N,T)} = \rho + \frac{\lambda^2 E(S^2)}{2(1-\rho)} + \frac{\sum_{n=1}^{N-1} n(n-1)(F_n(T) - F_{n+1}(T)) + N(N-1)F_N(T)}{2 \sum_{n=1}^{N} F_n(T)} \tag{7}$$

위 식 (7)을 살펴보면 Min(N, T)-정책에서의 평균 고객수는 휴가가 없는 일반적인 M/G/1 대기행렬에서의 평균 고객수보다 우변의 세 번째 항만큼 증가하게 된다는 것을 알 수 있는데, 이는 제어정책에 의하여 바쁜 기간의 시작시점이 조절되기 때문에 증가하는 양이다.

3.3 평균 대기시간

Little의 정리를 이용하여 평균 대기시간 $W_{q,Min(N,T)}$를 구하면 식 (8)과 같다.

$$W_{q,Min(N,T)} = \frac{1}{\rho}(L_{Min(N,T)} - \rho) = \frac{\lambda E(S^2)}{2(1-\rho)} + \frac{\sum_{n=1}^{N-1} n(n-1)[F_n(T) - F_{n+1}(T)] + N(N-1)F_N(T)}{2\lambda \sum_{n=1}^{N} F_n(T)} \tag{8}$$

식 (8)은 휴가가 없는 일반적인 M/G/1 대기행렬 보다 우변의 두 번째 항만큼 크게 나타나는데, 이는 바쁜 기간이 첫 고객의 도착 즉시 시작하는 것이 아니라 N명의 고객이 차거나 매 T 시간이 경과한 후에 시작되기 때문이다.

4. 사이클 길이

이 절에서는 바쁜 기간의 길이와 유휴기간의 길이를 유도하고

이를 이용하여 전체 사이클 길이를 구한다.

4.1 바쁜 기간의 길이

Min(N,T)-정책에서 바쁜 기간의 길이 $B_{Min(N,T)}$는 N_B명으로 시작하는 바쁜 기간의 길이가 되므로 M/G/1의 바쁜 길이 (B)의 N_B차 중합이다(이호우, 1998). 즉,

$$B_{Min(N,T)} = B_1 + B_2 + \cdots + B_{N_B} \tag{9}$$

따라서 바쁜 기간의 LST $B^*_{Min(N,T)}(\theta)$는 식 (10)과 같이 유도할 수 있다.

$$\begin{aligned} B^*_{Min(N,T)}(\theta) &= N_B(B^*(\theta)) \\ &= \frac{1}{F_1}(T)\left[\sum_{n=1}^{N-1}(B^*(\theta))^n(F_n(T) - F_{n+1}(T)) + B^*(\theta)^N F_N(T)\right] \end{aligned} \tag{10}$$

식 (10)을 θ에 대하여 미분하고, 0을 대입하면 식 (11)과 같이 바쁜 기간의 평균길이가 구해진다.

$$E(B_{Min(N,T)}) = \frac{\sum_{n=1}^{N} F_n(T)}{F_1(T)} \cdot \frac{E(S)}{1-\rho} \tag{11}$$

위 식을 보면 Min(N, T)-정책에서 바쁜 기간의 평균 길이는 휴가가 없는 일반적인 M/G/1 대기행렬의 평균 바쁜 길이 $\frac{E(S)}{1-\rho}$의 $\frac{\sum_{n=1}^{N} F_n(T)}{F_1(T)}$ 배가됨을 알 수 있는데, 이는 바쁜 기간 시작시의 고객수가 M/G/1 대기행렬의 고객수인 1 명보다 N_B 배 크기 때문이다.

4.2 유휴기간의 길이

유휴기간이란 서버가 서비스를 제공하지 않는 시간간격을 말하는데 이는 여러 개의 연속적인 휴가로 이루어질 수 있고 하나의 휴가로 이루어질 수도 있다.

4.2.1 T-정책으로 유휴기간이 종료되는 경우

T-정책으로 바쁜 기간이 시작되는 경우를 고려하자. 이 때 유휴기간의 길이를 I_T라 한다. 처음 $m(m=0,1,2,\cdots)$개의 T시간 동안에 고객의 도착이 이루어지지 않았고, $(m+1)$번째 T시간 동안에 고객이 도착한다고 하자. 이 경우 $(m+1)$번째 T에서 T-정책에 의해 유휴기간이 종료되려면 T시간이 경과하는 동안에 도착한 고객의 수가 1명 이상 $N-1$명 이하여야 하고 그 확률은 식 (12)와 같다.

$$\begin{aligned} \sum_{n=1}^{N-1} \Pr(N(T)=n) &= \sum_{n=1}^{N-1}(F_n(T) - F_{n+1}(T)) \\ &= F_1(T) - F_N(T) \end{aligned} \tag{12}$$

또한, 유휴기간의 길이가 $(m+1)T$가 되려면 처음 m개의 T시간 동안에 고객의 도착이 없어야 하며 그 확률은 $(e^{-\lambda T})^m$가 된다. 따라서, 유휴기간의 길이가 $(m+1)T$가 될 확률은 식 (13)과 같다.

$$\begin{aligned} \Pr(I_T &= (m+1)T) \\ &= e^{-m\lambda T}(F_1(T) - F_N(T)), \quad m=0,1,2,\cdots \end{aligned} \tag{13}$$

식 (12)와 (13)을 이용하여 T-정책에 의해 유휴기간이 종료되는 경우의 평균길이를 구하면 식 (14)와 같다.

$$\begin{aligned} E(I_T) &= \sum_{m=0}^{\infty}(m+1)T \cdot e^{-m\lambda T}(F_1(T) - F_N(T)) \\ &= \frac{T}{(1-e^{-\lambda T})^2}(F_1(T) - F_N(T)) \end{aligned} \tag{14}$$

이 값은 일반적인 T-정책에서보다 길이가 더 짧은데 그 이유는 T시간이 경과하기 전에 N-정책에 의하여 바쁜 기간이 시작되는 경우가 확률 P_N으로 존재하기 때문이다.

4.2.2 N-정책으로 유휴기간이 종료되는 경우

T-정책으로 시작하는 경우와 동일하게 처음 m개의 T시간 동안에 고객의 도착이 없었고, 다음 $(m+1)$번째 T시간 동안에 고객이 도착했다고 가정하면, 이 경우 N-정책에 의해 유휴기간이 종료되려면 $(m+1)$번째 T시간 동안에 도착한 고객의 수가 N명 이상이 되어야 한다. 여기서 $(m+1)$번째 T에서 N번째 고객의 도착시점을 S_N, N-정책으로 유휴기간이 끝날 경우의 유휴기간의 길이를 I_N이라 하면 식(15)와 같다.

$$I_N = mT + S_N, \quad (S_N < T) \tag{15}$$

그러면 유휴기간의 길이가 $mT+t$ $(0 \le t < T)$가 될 확률은 식 (16)과 같다.

$$\begin{aligned} &\Pr(I_N \le mT+t) = e^{-m\lambda T} \cdot \Pr(S_N \le t,\ A(T) \ge N) \\ &= e^{-m\lambda T} \cdot \sum_{n=N}^{\infty} \Pr(S_N \le t | A(T)=n) \cdot \Pr(A(T)=n), \\ &0 \le t < T \end{aligned} \tag{16}$$

위 식에서 $[S_N | A(T)=n]$의 분포는 $B(N, n-N+1)$ 분포를 따르고 평균이 $\frac{NT}{n+1}$이다(이호우, 1998). N-정책으로 바쁜 기간이 시작될 경우의 평균 유휴기간의 길이는 식 (17)과 같다.

$$E(I_N)=\sum_{m=0}^{\infty}\sum_{n=N}^{\infty}e^{-m\lambda T}\cdot\left[mT+\frac{NT}{n+1}\right]\cdot\frac{e^{-\lambda T}(\lambda T)^n}{n!}$$
$$=\frac{\lambda TF_N(T)(1-F_1(T))+NF_{N+1}(T)F_1(T)}{\lambda(F_1(T))^2} \quad (17)$$

4.2.3 총 유휴기간의 평균 길이

위에서 구한 N-정책으로 유휴기간이 종료되는 경우와 T-정책으로 유휴기간이 종료되는 경우는 서로 배반사상이므로 총 유휴기간의 평균 길이는 앞에서 구한 값들의 합으로 구성된다. 따라서, Min(N,T)-정책에서 유휴기간의 길이 $I_{Min(N,T)}$는 다음 식 (18)과 같다.

$$E(I_{Min(N,T)})=E(I_T)+E(I_N)=\frac{\sum_{n=1}^{N}F_n(T)}{F_1(T)}\cdot\frac{1}{\lambda} \quad (18)$$

식 (18)을 보면 Min(N,T)-정책에서의 유휴기간의 평균 길이는 휴가가 없는 일반적인 M/G/1모형에서의 유휴기간 길이 $1/\lambda$보다 $\sum_{n=1}^{N}F_n(T)/F_1(T)$배가 됨을 알 수 있다.

4.3 사이클의 평균길이

사이클의 평균 길이는 유휴기간의 평균길이와 바쁜 기간의 평균 길이의 합이므로, 식 (11)과 (18)의 결과를 이용하여 사이클의 길이 $C_{Min(N,T)}$를 식 (19)와 같이 구할 수 있다.

$$E(C_{Min(N,T)})=E(I_{Min(N,T)})+E(B_{Min(N,T)})$$
$$=\frac{\sum_{n=1}^{N}F_n(T)}{F_1(T)}\cdot\frac{1}{\lambda}+\frac{\sum_{n=1}^{N}F_n(T)}{F_1(T)}\cdot\frac{E(S)}{1-\rho}$$
$$=\frac{1}{\lambda(1-\rho)}\cdot\frac{\sum_{n=1}^{N}F_n(T)}{F_1(T)} \quad (19)$$

식 (19)에서 T를 무한대로 보낼 경우, 사이클의 평균 길이는 N-정책에서의 평균 길이와 동일하게 되고, N을 무한대로 보낼 경우 T-정책에서와 평균 길이와 일치함을 알 수 있다.

5. 비용함수의 유도

5.1 기호정의

비용함수를 유도하기 위해 사용될 기호는 다음과 같다.

(i) C_{su} : Min(N, T)-정책에서 유휴중인 서버가 서비스를 시작할 경우 발생하는 재가동 비용

(ii) C_{sc} : Min(N,T)-정책에서 서버가 휴가에서 돌아와서 고객수를 파악하는데 드는 비용

(iii) C_h : 고객 1명이 시스템 내에서 단위 시간을 보내는 데 발생하는 유지비용

5.2 시스템 재가동비용(setup cost)

시스템 재가동은 매 사이클마다 1회씩 발생하므로 단위 시간당 평균 재가동비용은 C_{su}를 사이클의 길이로 나눈 값이므로 다음과 같다.

• 단위 시간당 평균 시스템 재가동비용

$$=\frac{\lambda(1-\rho)F_1(T)}{\sum_{n=1}^{N}F_n(T)}\cdot C_{su} \quad (20)$$

5.3 고객수 검사비용(scanning cost)

고객수 검사비용은 휴가의 횟수에 관련된 비용으로 N-정책에 의해 바쁜 기간이 시작되면 휴가의 수보다 한번이 적은 횟수의 고객수 검사비용이 발생하고, T-정책으로 바쁜 기간이 시작되면 휴가의 수만큼 검사비용이 발생한다. 한 사이클에서 고객수 검사횟수를 M이라 하면 검사횟수가 k일 확률은 다음과 같다.

$$\Pr(M=k)$$
$$=\begin{cases}F_N(T), & k=0\\(e^{-\lambda T})^kF_N(T)+(e^{-\lambda T})^{k-1}(F_1(T)-F_N(T)), & k\ge 1\end{cases} \quad (21)$$

따라서 평균 고객수 검사횟수는 다음과 같다.

$$E(M)=\frac{1-F_N(T)}{F_1(T)} \quad (22)$$

단위 시간당 고객수 검사비용은 사이클 당 평균 고객수 검사비용을 사이클의 길이로 나눈 값이므로 다음과 같다.

• 단위 시간당 평균 고객수 검사비용

$$=\frac{E(M)\cdot C_{sc}}{E(C_{Min(N,T)})}$$
$$=\frac{\lambda(1-\rho)(1-F_N(T))}{\sum_{n=1}^{N}F_n(T)}\cdot C_{sc} \quad (23)$$

위 식에서 T가 무한대로 가면 고객수 파악 비용은 N-정책에서와 같이 0이 되고, N이 무한대가 되면 T-정책에서와 같이 $\frac{1-\rho}{T}\cdot C_{sc}$가 된다.

5.4 고객유지비용

단위 시간당 평균 고객유지비용이란 고객이 시스템에 단위 시간을 머물면서 발생하는 비용이다. 이는 평균 고객수에 단위 시간당 고객유지비용을 곱하여 계산한다.

- 단위 시간당 고객유지비용

$$= L_{Min(N,T)} \cdot C_h$$

$$= \left[\rho + \frac{\lambda^2 E(S^2)}{2(1-\rho)} + \frac{\sum_{n=1}^{N-1} n(n-1)(F_n(T) - F_{n+1}(T)) + N(N-1)F_N(T)}{2\sum_{n=1}^{N} F_n(T)} \right] \cdot C_h \tag{24}$$

5.5 단위 시간당 총 평균 비용

단위 시간당 총 평균 비용 $TC(N)$은 위에서 구한 세 가지 비용요소의 합으로 구성된다.

- $TC(N)$ = 식(20) + (23) + (24)

$$= \frac{\lambda(1-\rho)(1-F_N(T))}{\sum_{n=1}^{N} F_n(T)} \cdot C_{sc} + \frac{\lambda(1-\rho)F_1(T)}{\sum_{n=1}^{N} F_n(T)} \cdot C_{su} + \left[\rho + \frac{\lambda^2 E(S^2)}{2(1-\rho)} + \frac{\sum_{n=1}^{N-1} n(n-1)(F_n(T) - F_{n+1}(T)) + N(N-1)F_N(T)}{2\sum_{n=1}^{N} F_N(T)} \right] C_h \tag{25}$$

5.6 최적의 N^*결정

주어진 T 값에 대하여 단위 시간당 총 평균 비용을 최소화하기 위한 임계치 N^*를 최적고객수라 하자. 최적고객수를 결정하기 위해서는 비용함수의 성격을 알아야 하는데, 관련된 모수들이 많아 해석적으로 비용함수의 모양을 파악하는 것이 매우 어렵다. 이 연구에서는 광범위한 수치분석을 통해서 비용함수의 성격을 다음과 같이 파악하였다. 이 수치분석에서는 서비스가 *Erlang*(2,4)인 경우의 총 비용함수의 그래프만을 제시하였으나 다른 분포의 서비스 시간의 경우도 비슷한 결과를 얻었다.

5.6.1 평균 비용함수의 성격

총 평균 비용함수는 평균 고객수 검사비용, 시스템 재가동비용, 고객유지비용의 합으로 구성된다. 이 중 평균 고객수 검사비용과 고객유지비용 함수는 N에 대한 증가함수이므로 이들의 합도 증가함수가 되고 시스템 재가동비용함수는 N이 증가함에 따라 $\frac{(1-\rho)F_1(t)}{T} \cdot C_{su}$에 점점 근접해 가는 convex 형태의 감소함수이다. 따라서, 총 평균 비용함수는 두 개의 증가함수와 하나의 감소함수의 합으로 구성이 되므로 다음과 같은 두 가지의 형태가 나타난다.

① 증가함수의 영향이 큰 경우
(고객수 검사비용과 고객유지비용의 영향이 큰 경우)

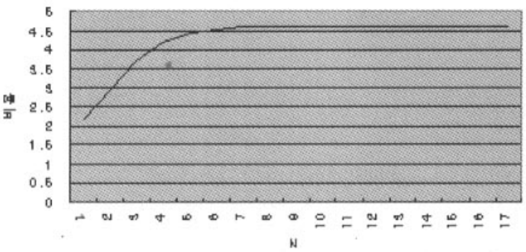

그림 1. $\lambda=1, \rho=0.5, T=2, C_{sc}=10, C_{su}=C_h=1.$

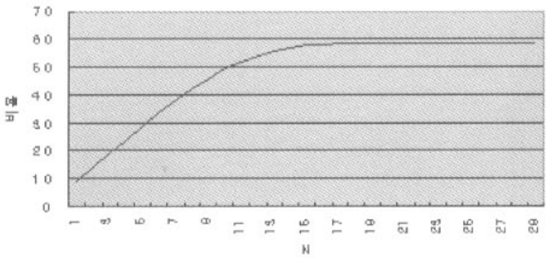

그림 2. $\lambda=1, \rho=0.5, T=10, C_{sc}=C_{su}=1, C_h=10.$

<그림 1>과 <그림 2>는 각각 고객수 검사비용과 고객유지비용이 큰 경우를 나타내고 있다. 이 경우 증가함수인 평균 고객수 검사비용함수나 평균 고객유지비용 함수의 영향력이 평균 시스템 재가동비용보다 크다. 즉 N이 증가함에 따라 평균 고객수 검사비용과 고객유지비용에 의하여 증가하는 양이 평균 시스템 재가동비용에 의하여 감소하는 양보다 많은 경우가 이에 해당된다. 이 경우는 총 비용함수가 처음부터 증가를 한다. 따라서 최적고객수는 1이 된다.

② 감소함수의 영향이 큰 경우
(시스템 재가동비용의 영향이 큰 경우)

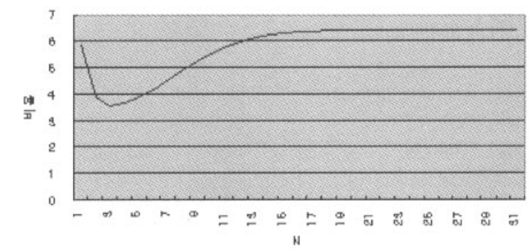

그림 3. $\lambda=1, \rho=0.5, T=10, C_{sc}=C_h=1, C_{su}=10.$

<그림 3>은 평균 시스템 재가동 비용의 영향력이 큰 경우 중 초기에는 시스템 재가동 비용의 영향을 많이 받아 전체적인 비용이 줄어들다가 나중에는 평균 고객수검사비용과 고객유지비용의 영향력이 커져서 다시 총 비용이 증가하는 경우이다. 이것은 평균 시스템 재가동 비용함수가 convex 형태의 감소함수이고 양의 값에 수렴하는 형태이기 때문이다.

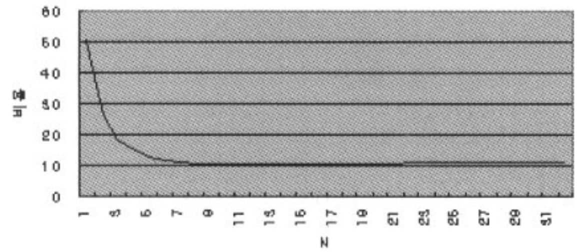

그림 4. $\lambda=1,\ \rho=0.5,\ T=10,\ C_{sc}=C_h=1,\ C_{su}=100.$

위 <그림 4>의 경우는 평균 시스템 재가동비용이 상대적으로 매우 커서 비용함수가 일정한 값에 수렴할 때까지 그 영향력이 지속되는 경우이다. 하지만, 이는 매우 극단적인 경우로 현실적으로는 나타나지 않는다고 볼 수 있다.

5.6.2 최적고객수 결정

위의 수치분석에 의한 그래프에서 보는 바와 같이 최소 비용점은 단 한 곳에서만 나타남을 알 수 있다. 따라서, Min(N,T)-정책에서 고정된 길이의 휴가기간 T가 주어진 경우, 단위 시간당 평균 비용을 최소화할 수 있는 최적의 고객수는 다음과 같은 식을 이용하여 구할 수 있다.

$$N^* = \mathrm{Min}\{\, m \mid TC(m+1) - TC(m) \geq 0,\ m=1,2,3,\cdots\} \quad (26)$$

즉, m을 1부터 시작하여 점점 증가시켜 가면서 $TC(m+1)-TC(m)$ 값을 계산하여 그 값이 0과 같거나 0보다 커지는 최초의 점에서 최적고객수 N^*을 구할 수 있다.

6. 결론 및 향후 연구 과제

이 연구에서는 M/G/1 대기행렬의 제어정책으로 지금까지 연구된 바가 없는 Min(N, T)-정책을 사용하였는데 이는 기존의 N-정책이나 T-정책보다 시스템의 유연성을 보다 향상시킨 모형이다. 이 연구에서는 안정상태에서 시스템의 각종 성능들을 분석하였다. 고객수 PGF를 유도하고 바쁜 기간의 길이와 유휴기간의 길이를 구하여 사이클의 길이를 유도하였다. 제 5 절에서는 시스템을 운용하는 데 발생하는 비용을 고객수 검사비용, 시스템 재가동비용, 고객유지비용으로 나누어 단위 시간당 평균 비용함수를 구하고 비용함수를 이용하여 최적의 고객수를 찾는 방법을 제시하였다. 이 방법을 이용하면 시스템을 운용하는 비용을 절감할 수 있는 효과를 가져올 수 있다.

추후 연구과제로는 N-정책과 T-정책, 그리고 Max(N, T)-정책과 비용함수를 비교하여 비용이 최소가 되는 영역을 밝혀 내는 것이 필요하다. 또한 휴가의 길이를 확률변수로 확장을 하는 연구도 필요하다.

참고문헌

이호우 (1998), 대기행렬이론, 시그마프레스.

Balachandran, K. R. and Tijms, H. (1975), On the D-policy for the M/G/1 Queue, *Management Science*, 21, 1073-1076.

Balachandran, K. R. (1973), Control Policies for a single Server Systems, *Management Science*, 19, 1013-1018.

Bell, C. E. (1971), Characterization and Computation of Optimal Policies for Operating an M/G/1 Queueing System with Removable Server, *Operations Research*, 19, 208-218.

Fuhrmann, S. W. and Cooper, R. (1985), Stochastic Decompositions in the M/G/1 Queue with Generalized Vacations, *Operations Research*, 33(5), 1117-1129.

Gakis, K. G., Rhee, H. K. and Sivazlian, B. D. (1995), Distributions and First Moments of the Busy and Idle Periods in Controllable M/G/1 Queueing Models with Simple and Dyadic Policies, *Stoch. Analysis and Appl.*, 13(1), 47-81,.

Heyman, D. P. (1968), Optimal Operating Policies for M/G/1 Queueing Systems, *Operations Research*, 16, 362-382.

Heyman, D. P. (1977), T-policy for the M/G/1 queue, *Management Science*, 23, 775-778.

Hur, S. and Paik, S. J. (1999), The effect of different arrival rates on the N-policy of M/G/1 with server setup, *Applied Mathematical Modelling*, 23, 289-299.

Kella, O. (1989), The Threshold Policy in the M/G/1 Queue with server Vacations, *Naval Research Logistics*, 36, 111-123.

Lee, H. W. and Park, J. O. (1997), Optimal Strategy in N-Policy System with Early Setup, *Journal of Operations Research Society*, 48, 306-313.

Medhi, J. and Templeton, J. G. C. (1992), A Poisson Input Queue under N-Policy and with a General Start up Time, *Computers and Operations Research*, 19(1), 34-41.

Minh, D. L. (1992), Transient Solutions for some Exhaustive M/G/1 Queues with Generalized Independent Vacations, *European Journal of Operational Research*., 36, 197-201.

Rhee, H. K. (1997), Development of a New Methodology to find the Expected Busy Periods for Controllable M/G/1 Queueing Models Operating under the Multi-variable Operating Policies, *대한산업공학회지*, 23, 729-739.

Sobel, M. J. (1969), Optimal Average-Cost Policy for a Queue with Start-up and Shut-down Costs, *Operations Research*, 17, 145-162.

Yadin, M. and Naor, P. (1963), Queueing Systems with a Removable Service Station, *Operations Research Quarterly*, 14, 393-405.