Journal of the Society of Korea Industrial and Systems Engineering
Vol, 25, No. 5, pp.31－38, October 2002.

전용서버가 있는 이계층고객 M/M/2 대기모형

정재호 · 허 선

한양대학교 산업공학과

M/M/2 system with two customer classes and exclusive server

Jae-Ho Jung · Sun Hur

Dept. of Industrial Engineering, Hanyang University

In this paper, we model a two-server queueing system with priority, to which we put a restriction on the number of servers for each customer class. customers are divided into two different classes. Class 1 customers have non-preemptive priority over class 2 customers. They are served by both servers when available but class 2 customers are served only by a designated server. We use a method of generating function depending on the state of servers. We find the generating function of the number of customers in queue, server utilization, mean queue length and mean waiting time for each class of customers.

Keywords : priority, customer class, queueing system, exclusive server, multi-server

1. 서 론

1.1 연구 배경 및 목적

고객의 도착이 포아송 과정(Poisson process)을 따르고, 서비스 시간이 지수분포를 따르며 복수의 서버를 갖는 M/M/c 대기행렬모형에서는 현실 모형을 단순화하기 위해 여러 가지 가정을 둔다. 즉, 도착하는 고객들의 유형에는 차이가 없으며, 각 서버의 서비스율은 같다고 가정한다. 그러나, 현실 시스템에서는 도착하는 고객의 유형이 다른 경우가 많고 고객 유형에 따라 그 도착률이 다를 수 있고, 각 서버의 서비스율도 다를 수 있다. 예를 들어, 다품종 생산 시스템에서 여러 종류의 제품을 생산할 수 있고, 기계에 따라 생산 속도가 다를 수 있다. 또한, 고객의 유형에 따라 서비스를 받을 수 있는 서버의 수가 제한될 수도 있다. 예를 들어, 통신 시스템에서 real-time 처리 traffic(class 1)과 비 real-time 처리 traffic(class 2)으로 유형을 나누고 real- time traffic에 우선권을 준다. class 1은 가급적 서버 1(대용량 서버)로 서비스를 받으나, 서버 1이 바쁘면 서버 2(저용량 서버)도 사용한다. class 2는 서버 2에서만 서비스를 받는다. 이와 같이 보다 현실적인 모형으로 접근하기 위해서는 M/M/c 대기행렬시스템에 대한 변형된 모형의 연구가 필요하다.

M/M/c 대기행렬시스템의 변형된 모형 중에서 순서적 서버 사냥 모형이 있다. 일반적인 M/M/c 대기행렬모형에서는 유휴한 서버 중에서 아무나 선택하여 서비스를 받는 반면, 순서적 서버 사냥 모형에서는 도착하는 고객이 유휴한 서버 중에서 미리 정해진 순서에 따라 하나의 서버를 선택하여 서비스를 받는다. 본 연구에서는 이러한 순서적 서버 사냥의 확장된 경우로, 도착하는 고객이 순서적으로 하나의 서버를 선택하되, 고객의 유형에 따라 서비스를 받을 수 있는 서버의 개수가 제한되어 있는 시스템에 대하여 분석하고자 한다.

한편, 여러 가지 서비스 규칙 중 고객 집단이 여러 class로 나누어져 있을 때 class 별로 우선순위를 정하여 우선순위가 높은 class의 고객이 우선 순위가 낮은 class의 고객에 대하여 우선권을 가지고 서비스를 받는 우선순위 서비스 규칙을 적용하여 모형을 분석하고자 한다.

기존 연구를 살펴보면, Cooper(1990)에서는 여러 형태의 순서적 서버 사냥모형에 관한 분석 기법이 제시되어

있다[2]. Kella와 Yechiali(1985)에서는 비축출형 우선순위를 가지는 M/M/c 대기행렬모형에 대한 분석을 하였다[4]. 이 논문에서는 복수 휴가를 가지는 M/G/1 대기행렬모형과 M/M/c 대기행렬모형이 확률적으로 동일함을 이용하여 대기고객수 PGF를 유도하였다. Wagner (1997)에서는 비축출형 우선순위 서비스를 적용하는 시스템 용량이 유한한 복수 서버 모형을 분석하였다[5]. Choi et al.(2001)에서는 impatient 성향을 가지는 상위 클래스 고객이 하위 클래스 고객에 대하여 nonpre- emptive priority를 가지는 이계층 M/M/1 대기 모형을 분석하였다[3].

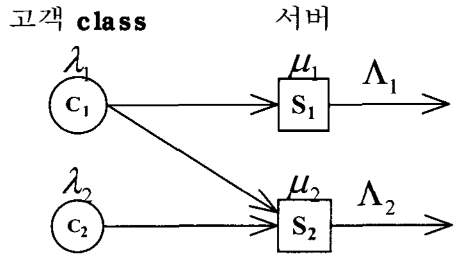
본 논문에서는 순서적 서버 사냥 모형을 적용하되, 고객 집단을 두 가지 class로 나누어, class 1의 고객이 class 2의 고객에 대하여 우선순위를 가질 때, 이 고객의 유형에 따라 서비스를 받는 서버의 수에 제한이 있는 M/M/2 대기행렬모형에 대하여 분석한다. 이 시스템에서 안정상태에서의 고객 class 별 대기고객수의 GF (probability generating function)를 유도하고, 각 서버의 서버 이용률, 고객 class 별 평균대기고객수와 평균대기시간 등을 유도하는 것을 본 연구의 목적으로 한다.

2. 모형 설명

본 연구는 두 가지 유형의 class의 고객들을 서비스하는 서버의 수에 제한이 있는 M/M/2 대기행렬시스템을 연구 대상으로 한다. 본 연구에 사용될 가정은 다음과 같다.

(1) 시스템에 도착하는 고객의 class는 두 가지이다.
(2) 각 class의 고객은 도착률이 λ_i $(i=1,2)$이고 서로 독립인 포아송과정으로 시스템에 도착한다.
(3) 각 서버는 서로 독립적으로 운영되며, 서비스 시간들은 각각 평균이 $1/\mu_i$ $(i=1,2)$인 지수분포를 따른다.
(4) 각 class의 고객이 서비스를 받을 수 있는 서버의 수에 제한을 둔다. 즉, class 1의 고객은 두 서버 중에서 아무나 하나의 서버를 선택하여 서비스를 받을 수 있는 반면, class 2의 고객은 서버 2에서만 서비스를 받을 수 있다.
(5) 두 서버가 모두 유휴할 때 도착한 class 1의 고객은 서버 1을 선택하여 서비스를 받는다.
(6) 두 class의 고객이 대기열에 함께 있을 때에는 class 1의 고객이 먼저 서비스를 받으며, class 1의 고객이 대기열에 없고 서버 2가 유휴하면 비로소 class 2의 고객이 서비스를 받는다. 하지만, 서버 1이 바쁘고, 서버 2에서 class 2의 고객이 서비스를 받고 있을 때 도착한 class 1의 고객은 두 서버 중 서비스가 먼저 끝나는 서버를 선택하여 서비스를 받게 된다. 즉, class 2의 고객이 먼저 서비스를 받고 있을 때에는 class 1의 고객이 class 2의 고객을 축출하지 않고 서비스가 끝날 때까지 기다린다.
(7) 시스템의 용량과 고객의 모집단의 크기는 무한하다.
(8) 기타 사항들은 일반적인 대기행렬모형의 가정에 준한다[1].

다음 <그림 2-1>를 통하여 모형의 전반적인 형태와 서비스 규칙에 대하여 설명하고자 한다. 고객의 class는 두 가지이고, 서버가 두 대인 대기행렬시스템이다. class 1의 고객은 도착률이 λ_1이고, 서버 1과 서버 2 중에서 유휴한 서버를 선택하여 서비스를 받는다. 반면에, class 2의 고객은 도착률이 λ_2이고, 서버 2에게만 서비스를 받는다.



<그림 2-1> class 별 서버의 수에 제한이 있는 M/M/2 대기행렬모형

3. 대기고객수 GF

3.1 전이율 다이어그램

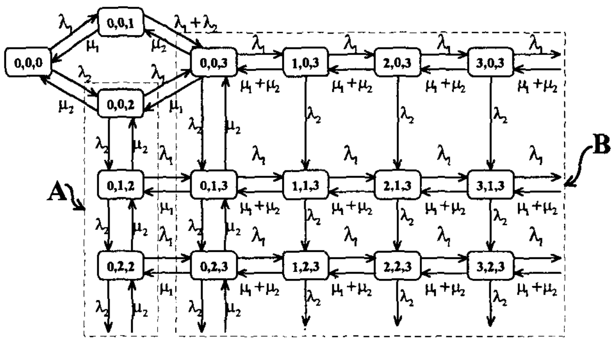
본 모형에 대한 전이율 다이어그램을 나타내기 위해 2절에서 정의한 기호들을 바탕으로 시스템의 상태를 아래와 같이 서버의 상태와 대기고객수로 정의한다.

상태 정의 : $(\ i,\ j,\ s\ )$
여기서,
$i=$ 대기공간에 있는 class 1의 고객수, $i=0,1,2,\ldots$
$j=$ 대기공간에 있는 class 2의 고객수, $j=0,1,2,\ldots$
$s=$ 서버의 상태

$$=\begin{cases} 0, & \text{두 서버가 모두 유휴한 경우} \\ 1, & \text{서버 1만 바쁜 경우} \\ 2, & \text{서버 2만 바쁜 경우} \\ 3, & \text{두 서버가 모두 바쁜 경우} \end{cases}$$

상태 정의를 바탕으로 전이율 다이어그램(transition diagram)을 <그림 3-1>과 같이 나타낼 수 있다.

<그림 3-1>에서와 같이 각 상태를 두 부분으로 나누어 GF를 유도하게 된다. 그림의 A 부분은 서버 2만 바쁠 때 class 2의 대기고객수를 나타내며, B 부분은 두 서버가 모두 바쁠 때 각 class의 대기고객수를 나타내는 부분이다. 이 두 부분에 포함되지 않는 상태들 (0,0,0), (0,0,1)의 확률은 4장에서 구할 서버 이용률로부터 유도할 수 있다.



<그림 3-1> 전이율 다이어그램(transition diagram)

3.2 안정 상태 방정식

$P_{i,j,s}$을 안정 상태에서 서버의 상태가 s이고, class 1의 대기고객수가 i명, class 2의 대기고객수가 j명일 확률이라고 하자. 즉,

$$P_{i,j,s} = \Pr[X_1 = i, X_2 = j, S = s],\quad i,j = 0,1,2,\ldots,\quad s = 0,1,2,3$$

<그림 3-1>의 전이율 다이어그램으로부터 전체 안정 상태 방정식들을 유도하면 다음 식 (3-1)~(3-8)과 같다.

$$(\lambda_1 + \lambda_2)P_{0,0,0} = \mu_1 P_{0,0,1} + \mu_2 P_{0,0,2} \qquad (3\text{-}1)$$

$$(\mu_1 + \lambda_1 + \lambda_2)P_{0,0,1} = \lambda_1 P_{0,0,0} + \mu_2 P_{0,0,3} \qquad (3\text{-}2)$$

$$(\mu_2 + \lambda_1 + \lambda_2)P_{0,0,2} = \lambda_2 P_{0,0,0} + \mu_1 P_{0,0,3} + \mu_2 P_{0,1,2} \qquad (3\text{-}3)$$

$$(\mu_1 + \mu_2 + \lambda_1 + \lambda_2)P_{0,0,3} = (\lambda_1 + \lambda_2)P_{0,0,1} + \lambda_1 P_{0,0,2} + (\mu_1 + \mu_2)P_{1,0,3} + \mu_2 P_{0,1,3} \qquad (3\text{-}4)$$

$$(\lambda_1 + \lambda_2 + \mu_2)P_{0,j,2} = \lambda_2 P_{0,j-1,2} + \mu_1 P_{0,j,3} + \mu_2 P_{0,j+1,2},\quad j \ge 1 \qquad (3\text{-}5)$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)P_{0,j,3} = \lambda_1 P_{0,j,2} + \lambda_2 P_{0,j-1,3} + (\mu_1 + \mu_2)P_{1,j,3} + \mu_2 P_{0,j+1,3},\quad j \ge 1 \qquad (3\text{-}6)$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)P_{i,0,3} = \lambda_1 P_{i-1,0,3} + (\mu_1 + \mu_2)P_{i+1,0,3},\quad i \ge 1 \qquad (3\text{-}7)$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)P_{i,j,3} = \lambda_1 P_{i-1,j,3} + \lambda_2 P_{i,j-1,3} + (\mu_1 + \mu_2)P_{i+1,j,3},\quad i \ge 1,\ j \ge 1 \qquad (3\text{-}8)$$

여기서 구하고자 하는 값들은 각 대기고객수 확률인 $P_{i,j,s}$ 값들이지만, 위 식들로부터 직접 구하기 어려우므로 GF(generating function)를 이용하여 대기고객수 GF를 유도한다.

3.3 대기고객수 GF

3.2 절에서 구한 안정상태 방정식들 (3-1)~(3-8)을 이용하여 대기고객수 GF를 유도하는 과정을 단계별로 제시하였다.

단계 1) 필요한 GF를 정의한다.

$$P(z,w,3) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} P_{i,j,3}\, z^i w^j \qquad (3\text{-}9)$$

$$P(0,w,s) = \sum_{j=0}^{\infty} P_{0,j,s}\, w^j,\quad s = 2,3 \qquad (3\text{-}10)$$

여기서, $P(z,w,3)$은 두 서버가 바쁠 때 각 class의 대기고객수, 즉, A 부분의 GF를 나타내며, $P(0,w,2)$는 서버 2만 바쁠 때 class 2의 대기고객수 GF를, $P(0,w,3)$은 두 서버가 모두 바쁠 때 class 2의 대기고객수 GF를 나타낸다.

단계 2) 식 (3-5)~(3-8)에 각각 z^i, w^j를 곱한 후, i, j의 모든 범위에 대하여 양변을 각각 더한다.

$$(\lambda_1 + \lambda_2 + \mu_2)\sum_{j=1}^{\infty} P_{0,j,2} w^j = \lambda_2 \sum_{j=1}^{\infty} P_{0,j,3} w^j + \mu_1 \sum_{j=1}^{\infty} P_{0,j-1,3} w^j + \mu_2 \sum_{j=1}^{\infty} P_{0,j+1,2} w^j \qquad (3\text{-}11)$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\sum_{j=1}^{\infty} P_{0,j,3} w^j = \lambda_1 \sum_{j=1}^{\infty} P_{0,j,2} w^j + \lambda_2 \sum_{j=1}^{\infty} P_{0,j-1,3} w^j + (\mu_1 + \mu_2)\sum_{j=1}^{\infty} P_{1,j,3} w^j + \mu_2 \sum_{j=1}^{\infty} P_{0,j+1,3} w^j \qquad (3\text{-}12)$$

$$(\lambda_1 + \lambda_2 + \mu_1 + \mu_2)\sum_{i=1}^{\infty} P_{i,0,3} z^i = \lambda_1 \sum_{i=1}^{\infty} P_{i-1,0,3} z^i + (\mu_1 + \mu_2)\sum_{i=1}^{\infty} P_{i+1,0,3} z^i \qquad (3\text{-}13)$$

$$(\lambda_1+\lambda_2+\mu_1+\mu_2)\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{i,j,3}z^iw^j$$
$$=\lambda_1\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{i-1,j,3}z^iw^j+\lambda_2\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{i,j-1,3}z^iw^j$$
$$+(\mu_1+\mu_2)\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}P_{i+1,j,3}z^iw^j \qquad (3\text{-}14)$$

단계 3) 식 (3-1)~(3-4)와 식 (3-12)~(3-14)의 양변을 각각 더하여 GF $P(z,w,3)$, $P(0,w,2)$, $P(0,w,3)$의 식으로 정리하여 식 (3-15)와 같이 유도한다.

$$(\lambda_1+\lambda_2+\mu_1+\mu_2-\lambda_1z-\lambda_2w-(\mu_1+\mu_2)/z)P(z,w,3)$$
$$=\lambda_1P(0,w,2)+(\mu_2/w-(\mu_1+\mu_2)/z)P(0,w,3)$$
$$+(\lambda_1+\lambda_2+\mu_1+\mu_2)P_{0,0,3}-\lambda_1P_{0,0,2}$$
$$-(\mu_1+\mu_2)P_{1,0,3}-\mu_2P_{0,1,3}-(\mu_2/w)P_{0,0,3} \qquad (3\text{-}15)$$

단계 4) 식 (3-11)을 $P(0,w,3)$에 관하여 정리하면 식 (3-16)과 같다.

$$P(0,w,3)=(1/\mu_1)[(\lambda_1+\lambda_2-\lambda_2w-\mu_2/w)P(0,w,2)$$
$$-(\lambda_1+\lambda_2+\mu_2-\mu_2/w)P_{0,0,2}+\mu_1P_{0,0,3}$$
$$+\mu_2P_{0,1,2}] \qquad (3\text{-}16)$$

단계 5) 식 (3-15)을 식 (3-16)에 대입하여 정리하면 식 (3-17)과 같다.

$$(\lambda_1-\lambda_1z+\lambda_2-\lambda_2w+\mu_1+\mu_2-(\mu_1+\mu_2)/z)P(z,w,3)$$
$$=P(0,w,2)[\lambda_1+1/\mu_1(\mu_2/w-(\mu_1+\mu_2)/z)$$
$$\cdot(\lambda_1+\lambda_2-\lambda_2w+\mu_2-\mu_2/w)]$$
$$+P_{0,0,1}[\lambda_1+\lambda_2-(\lambda_1+\lambda_2+\mu_1)/w$$
$$-(1/w)(\mu_2/w-\mu_1+\mu_2/z)]$$
$$+P_{0,0,0}[[\mu_2/w-(\mu_1+\mu_2)/z][\lambda_1+\lambda_2(1-w)]/(w\mu_1)$$
$$+\lambda_1/w] \qquad (3\text{-}17)$$

단계 6) 식 (3-17)의 좌변에서 $P(z,w,3)$의 계수가 0이 되는 z의 값을 $f(w)$라 둔다. 즉,

$$\lambda_1-\lambda_1z+\lambda_2-\lambda_2w+\mu_1+\mu_2-(\mu_1+\mu_2)/z=0$$

에서 식 (3-18)과 같이 유도된다.

$$z=f(w)$$
$$=1/(2\lambda_1)[\lambda_1+\lambda_2(1-w)+\mu_1+\mu_2$$
$$-\sqrt{(\lambda_1+\lambda_2-\lambda_2w+\mu_1+\mu_2)^2-4\lambda_1(\mu_1+\mu_2)}] \qquad (3\text{-}18)$$

단계 7) 식 (3-18)을 식 (3-17)에 대입하면 $P(0,w,2)$에 관한 식을 유도할 수 있다.

$$P(0,w,2)=$$
$$\left[\lambda_1+\frac{1}{\mu_1}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)(\lambda_1+\lambda_2-\lambda_2w+\mu_2-\frac{\mu_2}{f(w)})\right]^{-1}$$
$$\cdot\left[P_{0,0,1}\left\{\lambda_1+\lambda_2-\frac{\lambda_1+\lambda_2+\mu_2}{w}-\frac{1}{w}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)\right\}\right.$$
$$\left.+P_{0,0,0}\left\{\frac{\lambda_1}{w}+\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)\frac{\lambda_1+\lambda_2(1-w)}{w\mu_1}\right\}\right] \qquad (3\text{-}19)$$

단계 8) 식 (3-19)을 식 (3-17)에 대입하면 $P(z,w,3)$에 관한 식을 얻는다.

$$P(z,w,3)=$$
$$[\lambda_1-\lambda_1w+\lambda_2-\lambda_2w+\mu_1+\mu_2-\frac{\mu_1+\mu_2}{z}]^{-1}\cdot$$
$$\left\{\frac{\lambda_1+\frac{1}{\mu_1}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{z}\right)(\lambda_1+\lambda_2-\lambda_2w+\mu_2-\frac{\mu_2}{w})}{\lambda_1+\frac{1}{\mu_1}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)(\lambda_1+\lambda_2-\lambda_2w+\mu_2-\frac{\mu_2}{w})}\right.$$
$$\cdot\left[P_{0,0,1}\left(\lambda_1+\lambda_2-\frac{\lambda_1+\lambda_2+\mu_2}{w}-\frac{1}{w}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)\right)\right.$$
$$\left.+P_{0,0,0}\left(\frac{\lambda_1}{w}+\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)\frac{\lambda_1+\lambda_2(1-w)}{w\mu_1}\right)\right]$$
$$+P_{0,0,1}\left[\lambda_1+\lambda_2-\frac{\lambda_1+\lambda_2+\mu_1}{w}-\frac{1}{w}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{z}\right)\right]$$
$$\left.+P_{0,0,0}\left[\frac{\lambda_1}{w}+\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{z}\right)\frac{\lambda_1+\lambda_2(1-w)}{w\mu_1}\right]\right\} \qquad (3\text{-}20)$$

단계 9) 식 (3-20)을 $P_{0,0,0}$, $P_{0,0,1}$에 관하여 정리하면 식 (3-21)과 같다.

$$P(z,w,3)=$$
$$\frac{(1/z-1/f(w))(\mu_1+\mu_2)\frac{w-1}{w}}{(\lambda_1+\lambda_2-\lambda_1z-\lambda_2w+\mu_1+\mu_2+(\mu_1+\mu_2)/z)K(w)}$$
$$\cdot[P_{0,0,1}\{(\lambda_1+\lambda_2)\lambda_1/\mu_1$$
$$+(1-(\lambda_1+\lambda_2)(w-1)/\mu_1)(\lambda_2-\mu_2/w)\}$$
$$+P_{0,0,0}\cdot(\lambda_1\mu_2)/(\mu_1w)]$$

여기서,

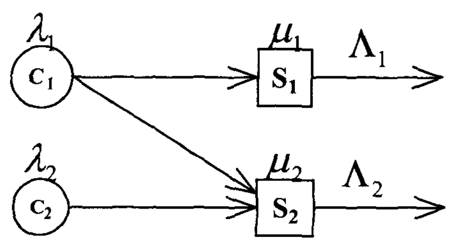
$$K(w)=\lambda_1+\frac{1}{\mu_1}(\mu_2/w-\mu_1+\mu_2/f(w))\cdot(\lambda_1+\lambda_2+\mu_2-\lambda_2 w-\frac{\mu_2}{w}) \tag{3-21}$$

식 (3-19)와 식 (3-21)과 같이 본 모형에 필요한 GF $P(0,w,2)$와 $P(z,w,3)$을 구하였다. 여기서, 미지수인 $P_{0,0,1}$와 $P_{0,0,0}$은 4장에서 구하게 될 시스템 이용률로부터 얻을 수 있다.

4. 시스템 이용률

4.1 시스템 이용률 관계식

먼저 <그림 4-1>에서와 같이 고객 class와 서버와의 관계를 이용하여 각 서버에서의 고객 이탈률 Λ_1, Λ_2에 관한 관계식을 유도한다.



<그림 4-1> 고객 class와 서버와의 관계

$$\Lambda_1=\lambda_1\cdot\frac{\mu_1}{\mu_1+\mu_2}\Pr(\text{서버 }1,2\text{ 모두 }busy)+\lambda_1\Pr(\text{서버 }1\ idle) \tag{4-1}$$

$$\Lambda_2=\lambda_2+\lambda_1\cdot\frac{\mu_2}{\mu_1+\mu_2}\Pr(\text{서버 }1,2\text{ 모두 }busy)+\lambda_1\Pr(\text{서버 }1\text{ 만 }busy) \tag{4-2}$$

여기서, $\Lambda_1+\Lambda_2=\lambda_1+\lambda_2$임을 확인할 수 있다.

각 서버의 이용률을 ρ_1, ρ_2라 하면 식 (4-1)에서 Pr(서버 1 *idle*)은 식 (4-3)과 같이 나타낼 수 있다.

$$\Pr(\text{서버 }1\ idle)=1-\rho_1 \tag{4-3}$$

식 (4-3)을 식 (4-1)에 대입하고, 식 (4-1)과 식 (4-2)를 각각 μ_1과 μ_2로 나누어주면 ρ_1, ρ_2는 각각 식 (4-4)와 식 (4-5)와 같이 표현된다.

$$\rho_1=\frac{\Lambda_1}{\mu_1}=\frac{\lambda_1}{\mu_1+\mu_2}\Pr(\text{서버 }1,2\text{ 모두 }busy)+\frac{\lambda_1}{\mu_1}(1-\rho_1) \tag{4-4}$$

$$\rho_2=\frac{\Lambda_2}{\mu_2}=\frac{\lambda_2}{\mu_2}+\frac{\lambda_1}{\mu_1+\mu_2}\Pr(\text{서버 }1,2\text{ 모두 }busy)+\frac{\lambda_1}{\mu_2}\Pr(\text{서버 }1\text{만 }busy) \tag{4-5}$$

식 (4-4), (4-5)를 이용하여 Pr(서버 1,2 모두 *busy*), Pr(서버 1만 *busy*)를 식 (4-6), (4-7)과 같이 얻을 수 있다.

$$\Pr(\text{서버 }1,2\text{ 모두 }busy)=\frac{\mu_1+\mu_2}{\lambda_1}(1+\frac{\lambda_1}{\mu_1})\rho_1-\frac{\mu_1+\mu_2}{\mu_1} \tag{4-6}$$

$$\Pr(\text{서버 }1\text{만 }busy)=\frac{\rho_2-(1+\lambda_1/\mu_1)\rho_1-\lambda_2/\mu_2+\lambda_1/\mu_1}{\lambda_1/\mu_2} \tag{4-7}$$

식 (4-6), (4-7)로부터 나머지 관련 확률들을 구할 수 있다.

$$\Pr(\text{서버 }2\text{만 }busy)=\rho_2-\Pr(\text{서버 }1,2\text{ 모두 }busy)=\rho_2-\frac{\mu_1+\mu_2}{\lambda_1}(1+\frac{\lambda_1}{\mu_1})\rho_1-\frac{\mu_1+\mu_2}{\mu_1} \tag{4-8}$$

$$\Pr(\text{서버 }1,2\text{ 모두 }idle)=(\frac{\mu_1+\mu_2}{\lambda_1}+\frac{\mu_2}{\mu_1}+\frac{\mu_1}{\mu_2})\rho_1-\frac{\lambda_1+\lambda_2}{\mu_2}-\frac{\mu_2}{\mu_1} \tag{4-9}$$

여기서, Pr(서버 1만 *busy*), Pr(서버 1,2 모두 *idle*)은 각각 $P_{0,0,1}$, $P_{0,0,0}$와 같고 이는 3장에서 구한 대기고객수 GF 식들에서의 미지수들이므로, 이들을 알면 $P(0,w,2)$와 $P(z,w,3)$을 구할 수 있게 된다.

시스템에 입력되는 고객들은 모두 시스템을 나가게 되므로, 다음의 관계식이 성립한다.

$$\Lambda_1+\Lambda_2=\lambda_1+\lambda_2 \tag{4-10}$$

여기서, $\Lambda_1=\mu_1\rho_1$, $\Lambda_2=\mu_2\rho_2$이므로 식 (4-10)으로부터 ρ_1과 ρ_2의 관계식을 유도하여 ρ_2에 관하여 정리하면 식 (4-11)과 같다.

$$\rho_2 = -\frac{\mu_1}{\mu_2}\rho_1 + \frac{\lambda_1+\lambda_2}{\mu_2} \tag{4-11}$$

식 (4-11)을 식 (4-7), (4-9)에 대입하면 $P_{0,0,1}$, $P_{0,0,0}$를 ρ_1에 관한 관계식으로 표현할 수 있다.

$$P_{0,0,1} = -\left(\frac{1}{\lambda_1}+\frac{\mu_2}{\lambda_2}+\frac{\mu_2}{\mu_1}\right)\rho_1 + 1 + \frac{\mu_2}{\mu_1} \tag{4-12}$$

$$P_{0,0,0} = \left(\frac{\mu_1+\mu_2}{\lambda_1}+\frac{\mu_2}{\mu_1}+\frac{\mu_1}{\mu_2}\right)\rho_1 - \frac{\lambda_1+\lambda_2}{\mu_2} - \frac{\mu_2}{\mu_1} \tag{4-13}$$

4.2 시스템 이용률

이 절에서는 3장에서 구한 대기고객수 GF와 4.1 절에서 얻은 식들로부터 시스템 이용률 ρ_1과 ρ_2를 유도한다. 먼저 3 장에서 구한 GF $P(z,w,3)$의 식 (3-21)에 w 대신 1을 대입하면 식 (4-14)를 얻는다.

$$P(z,1,3) = \frac{\left(\frac{1}{z}-1\right)(\mu_1+\mu_2)}{\lambda_1-\lambda_1 z+\mu_1+\mu_2-\frac{\mu_1+\mu_2}{z}} \cdot \frac{P_{0,0,1}\left\{\frac{(\lambda_1+\lambda_2)\lambda_1}{\mu_1}+\lambda_2-\mu_2\right\}+P_{0,0,0}\frac{\lambda_1\mu_2}{\mu_1}}{\frac{\lambda_1\mu_2+(\lambda_2-\mu_2)(\mu_1+\mu_2)}{\mu_1+\mu_2-\lambda_1}\cdot\frac{\lambda_1}{\mu_1}+\lambda_2-\mu_2} \tag{4-14}$$

식 (4-14)에 z 대신 1을 대입하면 $\frac{0}{0}$ 형태가 되므로 로피탈의 정리를 이용하여 풀면 식 (4-15)를 얻을 수 있다.

$$P(1,1,3) = \frac{\mu_1+\mu_2}{\mu_1+\mu_2-\lambda_1} \cdot \frac{P_{0,0,1}\left\{\frac{(\lambda_1+\lambda_2)\lambda_1}{\mu_1}+\lambda_2-\mu_2\right\}+P_{0,0,0}\frac{\lambda_1\mu_2}{\mu_1}}{\frac{\lambda_1\mu_2+(\lambda_2-\mu_2)(\mu_1+\mu_2)}{\mu_1+\mu_2-\lambda_1}\cdot\frac{\lambda_1}{\mu_1}+\lambda_2-\mu_2} \tag{4-15}$$

그리고, 식 (4-15)는 Pr(서버 1,2 모두 *busy*)의 확률과 같으므로 식 (4-15)에 식 (4-12), (4-13)을 대입하여 ρ_1에 관하여 풀면 식 (4-16)을 얻을 수 있다.

$$\rho_1 = \frac{\frac{A_3}{\mu_1}-A_2}{A_1+A_3\cdot\frac{1}{\lambda_1}\left(1+\frac{\lambda_1}{\mu_1}\right)}$$

여기서,

$$\begin{aligned} A_1 =& -\left(\frac{1}{\lambda_1}-\frac{\mu_2}{\lambda_1}-\frac{\mu_2}{\mu_1}\right)\left\{\frac{(\lambda_1+\lambda_2)\lambda_1}{\mu_1}+\lambda_2-\mu_2\right\} \\ &+(\mu_1+\mu_2)\mu_2+\frac{\mu_2^2\lambda_1}{\mu_1}+\lambda_1\mu_1 \\ A_2 =& \left(1+\frac{\mu_2}{\mu_1}\right)\left\{\frac{(\lambda_1+\lambda_2)\lambda_1}{\mu_1}+\lambda_2-\mu_2\right\} \\ &-\left(\lambda_1+\lambda_2+\frac{\mu_2^2\lambda_1}{\mu_1}\right) \\ A_3 =& \{\lambda_1\mu_2+(\lambda_2-\mu_2)(\mu_1+\mu_2)\}\cdot\frac{\lambda_1}{\mu_1} \\ &+(\mu_1+\mu_2-\lambda_1)(\lambda_2-\mu_2) \end{aligned} \tag{4-16}$$

그리고, 식 (4-11)에 식 (4-16)을 대입하면 ρ_2를 구할 수 있다.

$$\rho_2 = -\frac{\mu_1}{\mu_2}\cdot\frac{A_3/\mu_1-A_2}{A_1+A_3\cdot(1/\lambda_1)(1+\lambda_1/\mu_1)}+\frac{\lambda_1+\lambda_2}{\mu_2} \tag{4-17}$$

5. 평균대기고객수와 평균대기시간

5.1 class 1의 평균대기고객수와 평균대기시간

class 1의 평균대기고객수 L_1를 구하기 위해서는 식 (4-14)를 z에 관하여 1 차 미분하여 z 대신 1을 대입하면 된다. 식 (4-14)에 z 대신 1을 대입해보면 0/0 형태를 띄므로, z에 관하여 1차 미분한 뒤 로피탈의 정리를 두 번 사용하여 z 대신 1을 대입하면 class 1의 평균대기고객수를 식 (5-1)과 같이 구할 수 있다.

$$\begin{aligned} L_1 =& \frac{d}{dz}P(z,1,3)|_{z=1} \\ =& -\lambda_1(\mu_1+\mu_2)\left[P_{0,0,1}\left\{\frac{(\lambda_1+\lambda_2)\lambda_1}{\mu_1}+\lambda_2-\mu_2\right\}\right. \\ &\left.+P_{0,0,0}\frac{\lambda_1\mu_2}{\mu_1}\right]/(\mu_1+\mu_2-\lambda_1)^2 \\ &/\left[\frac{\{\lambda_1\mu_2+(\lambda_2-\mu_2)(\mu_1+\mu_2)\}\lambda_1}{(\mu_1+\mu_2-\lambda_1)\mu_1}+(\lambda_2-\mu_2)\right] \end{aligned} \tag{5-1}$$

식 (5-1)에서 $P_{0,0,1}$과 $P_{0,0,0}$ 대신 각각 식 (4-12)와 식 (4-13)을 대입하면 class 1의 평균대기고객수를 구할 수 있다.

다음으로, class 1의 평균대기시간 W_1은 Little의 법칙을 이용하여 식 (5-2)와 같이 유도할 수 있다.

$$W_1 = \frac{L_1}{\lambda_1} \tag{5-2}$$

5.2 class 2의 평균대기고객수와 평균대기시간

class 2의 평균대기고객수 L_2를 구하기 위해서는 3장에서 유도한 대기고객수 GF $P(0,w,2)$와 $P(z,w,3)$을 모두 이용해야 한다.

먼저 $P(0,w,2)$의 식 (3-19)을 w에 관하여 미분해야 한다. 미분을 쉽게 하기 위해 식 (3-19)를 식 (5-3)과 같이 표현한다.

$$P(0,w,2) = -\frac{B_2(w)}{B_1(w)}P_{0,0,1} - \frac{B_3(w)}{B_1(w)}P_{0,0,0}$$

여기서,

$$\begin{aligned}
&B_1(w) \\
&= \lambda_1 + \frac{1}{\mu_1}\left(\frac{\mu_2}{w} - \frac{\mu_1+\mu_2}{f(w)}\right)\left(\lambda_1+\lambda_2+\mu_2-\lambda_2-\frac{\mu_2}{w}\right) \\
&B_2(w) = \lambda_1+\lambda_2 - \frac{\lambda_1+\lambda_2+\mu_1}{w} - \frac{1}{w}\left(\frac{\mu_2}{w} - \frac{\mu_1+\mu_2}{f(w)}\right) \\
&B_3(w) = \frac{\lambda_1}{w} + \frac{1}{\mu_1 w}\left[\frac{\mu_2}{w} - \frac{\mu_1+\mu_2}{f(w)}\right][\lambda_1+\lambda_2(1-w)]
\end{aligned} \tag{5-3}$$

식 (5-3)을 1차 미분하여 로피탈의 정리를 2번 이용하면 $P(0,w,2)$의 미분값이 식 (5-4)와 같이 얻어진다.

$$\begin{aligned}
&\frac{d}{dw}P(0,w,2)|_{w=1} \\
&= -\frac{B_2^{(2)}(1)B_1^{(1)}(1) - B_2^{(1)}(1)B_1^{(2)}(1)}{2[B_1^{(1)}(1)]^2}P_{0,0,1} \\
&\quad -\frac{B_3^{(2)}(1)B_1^{(1)}(1) - B_3^{(1)}(1)B_1^{(2)}(1)}{2[B_1^{(1)}(1)]^2}P_{0,0,0}
\end{aligned}$$

여기서,

$$\begin{aligned}
&B_1^{(2)}(1) = \left[-\mu_2 + \frac{(\mu_1+\mu_2)\lambda_2}{\mu_1+\mu_2-\lambda_1}\right]\frac{\lambda_1}{\mu_1} - (\mu_2-\lambda_2) \\
&B_1^{(2)}(1) = 2\left[\mu_2 + \frac{(\mu_1+\mu_2)\lambda_2^2\lambda_1}{(\mu_1+\mu_2-\lambda_1)^3}\right]\frac{\lambda_1}{\mu_1} + 2\mu_2 \\
&\qquad + 2\left[-\mu_2 + \frac{(\mu_1+\mu_2)\lambda_2}{\mu_1+\mu_2-\lambda_1}\right]\frac{\mu_2-\lambda_2}{\mu_1} \\
&B_2^{(1)}(1) = \frac{\lambda_1(\mu_1+\mu_2-\lambda_1-\lambda_2)}{\mu_1+\mu_2-\lambda_1} \\
&B_2^{(2)}(1) = -2(\lambda_1+\lambda_2+2\mu_2) \\
&\qquad + \frac{2\lambda_2(\mu_1+\mu_2)}{\mu_1+\mu_2-\lambda_1}\left(1 + \frac{\lambda_1\lambda_2}{(\mu_1+\mu_2-\lambda_1)^2}\right) \\
&B_3^{(1)}(1) = \lambda_2 - \frac{\lambda_1\mu_2}{\mu_1} + \frac{(\mu_1+\mu_2)\lambda_1\lambda_2}{\mu_1(\mu_1+\mu_2-\lambda_1)} \\
&B_3^{(2)}(1) \\
&= 2(2\lambda_1+\lambda_2)\left(1+\frac{2\mu_2}{\mu_1}\right) \\
&\quad + \frac{2(\mu_1+\mu_2)\lambda_2}{\mu_1(\mu_1+\mu_2-\lambda_1)}\left(\frac{\lambda_1^2\lambda_2}{(\mu_1+\mu_2-\lambda_1)^2} + \lambda_1 + \lambda_2\right)
\end{aligned} \tag{5-4}$$

다음으로, $P(z,w,3)$을 w에 관하여 미분하기 위해 식 (3-21)에 z 대신 1을 대입하면 식 (5-5)와 같다.

$$\begin{aligned}
&P(1,w,3) \\
&= \frac{(1-1/f(w))(\mu_1+\mu_2)(w-1)/w}{(\lambda_2-\lambda_2 w)K(w)} \\
&\quad \cdot \left[P_{0,0,1}\left\{(\lambda_1+\lambda_2)\frac{\lambda_1}{\mu_1} + \left(1 - \frac{(\lambda_1+\lambda_2)(w-1)}{w}\right)\right.\right. \\
&\quad \left.\left. \cdot\left(\lambda_2 - \frac{\lambda_2}{w}\right)\right\} + P_{0,0,0}\frac{\lambda_1\mu_2}{\mu_1 w}\right]
\end{aligned}$$

여기서,

$$\begin{aligned}
K(w) = {}& \lambda_1 + \frac{1}{\mu_1}\left(\frac{\mu_2}{w} - \frac{\mu_1+\mu_2}{f(w)}\right) \\
& \cdot\left(\lambda_1+\lambda_2+\mu_2-\lambda_2 w - \frac{\mu_2}{w}\right)
\end{aligned} \tag{5-5}$$

미분을 쉽게 하기 위해 식 (5-5)를 식 (5-6)과 같이 정리할 수 있다.

$$P(1,w,3) = \frac{[f(w)-1](\mu_1+\mu_2)(w-1)C(w)}{(\lambda_2-\lambda_2 w)K(w)f(w)w^2}$$

여기서,

$$C(w)= \left\{\frac{(\lambda_1+\lambda_2)\lambda_1}{\mu_1} w \right.$$
$$\left. +\left[1-\frac{\lambda_1+\lambda_2}{\mu_1}(w-1)\right](\lambda_2 w-\mu_2)\right\}\cdot P_{0,0,1}$$
$$+\frac{\lambda_1\mu_2}{\mu_1}P_{0,0,0}$$
$$K(w)$$
$$=\lambda_1+\frac{1}{\mu_1}\left(\frac{\mu_2}{w}-\frac{\mu_1+\mu_2}{f(w)}\right)\cdot\left(\lambda_1+\lambda_2+\mu_2-\lambda_2 w-\frac{\mu_2}{w}\right) \tag{5-6}$$

식 (5-6)을 w에 관하여 미분하고 로피탈의 정리를 이용한 후 w 대신 1을 대입하면 식 (5-7)을 얻는다.

$$\frac{d}{dw}P(1,w,3)|_{w=1}$$
$$=[(-4f^{(2)}(1)K^{(1)}(1)+f^{(1)}(1)\{3K^{(2)}(1)$$
$$+4K^{(1)}(1)(f^{(1)}(1)+1)\})C(1)$$
$$+10f^{(1)}(1)K^{(1)}(1)C^{(1)}(1)]\cdot\frac{\mu_1+\mu_2}{12\lambda_2[K^{(1)}(1)]^2}$$

여기서,

$$K^{(1)}(1)=[-\mu_2+(\mu_1+\mu_2)f^{(1)}(1)]\lambda_1/\mu_1+\lambda_2-\mu_2$$
$$K^{(2)}(1)= [2\mu_2+(\mu_1+\mu_2)\{f^{(2)}(1)-2(f^{(1)}(1))^2\}]\frac{\lambda_1}{\mu_1}$$
$$+2[-\mu_1+f^{(1)}(1)(\mu_1+\mu_2)]\frac{\mu_2-\lambda_1}{\mu_1}+2\mu_2$$
$$f^{(1)}(1)=\frac{\lambda_2}{\mu_1+\mu_2-\lambda_1},\qquad f^{(2)}(1)=\frac{2\lambda_2^2(\mu_1+\mu_2)}{(\mu_1+\mu_2-\lambda_1)^3},$$
$$C^{(1)}(1)=[(\lambda_1+\lambda_2)\frac{\mu_2+\lambda_1-\lambda_2}{\mu_1}+\lambda_2]P_{0,0,1} \tag{5-7}$$

식 (4-6), (4-8), (5-4), (5-7)을 식 (5-8)에 대입하면 class 2의 평균대기고객수를 얻을 수 있다.

$$L_2= \Pr(\text{서버 2만 } busy)\cdot\frac{d}{dw}P(0,w,2)|_{w=1}$$
$$+\Pr(\text{서버 1,2 모두 } busy)\cdot\frac{d}{dw}P(1,w,3)|_{w=1} \tag{5-8}$$

다음으로, class 2의 평균대기시간 W_2은 식 (5-8)과 Little의 법칙을 이용하여 식 (5-9)와 같이 유도할 수 있다.

$$W_2=\frac{L_2}{\lambda_2} \tag{5-9}$$

6. 결과 요약 및 추후 연구 과제

본 연구에서는 고객의 집단을 두 가지 class로 나누어 각 class 별로 사용할 수 있는 서버의 개수를 다르게 두는 M/M/2 대기행렬모형의 변형된 형태를 분석하였다. 일반적인 M/M/2 대기행렬모형에서는 모든 고객이 균등하게 두 서버를 임의로 선택하여 서비스를 받을 수 있는 반면, 본 모형에서 한 고객 class는 두 서버 중 하나를 선택하여 서비스를 받을 수 있지만, 다른 한 고객 class는 하나의 서버에서만 서비스를 받도록 제한을 두었다. 그리고, 본 연구에서는 서비스 규칙으로 비축출형 우선순위를 적용하였다. 즉, 우선순위가 높은 class의 고객이 우선순위가 낮은 class의 고객에 대하여 우선권을 가지고 서비스를 받되, 우선순위가 낮은 고객이라도 일단 서비스를 받고 있으면 밀려나지 않고 계속 서비스를 받는 서비스 규칙을 적용하였다.

본 연구에서는 시스템의 전체 고객수를 고려하기보다는 대기열에 기다리는 고객수를 중심으로 안정상태에서의 각 성능치들을 분석하였다. 서버의 상태에 따른 대기고객수 GF를 유도하였고, 대기고객수와 서버 이용률과의 관계를 이용하여 서버 이용률을 유도하였으며, 대기고객수 GF로부터 각 class 별 평균대기고객수와 평균대기시간을 유도하였다.

추후 연구 과제로는 좀 더 현실적인 모형을 분석하기 위해 축출형 우선순위를 가지는 모형과 3개 이상의 고객 class를 가지는 모형으로의 확장이 필요하다.

참고문헌

[1] 이호우; *대기행렬이론*, 개정판, 시그마프레스, 1998.

[2] Cooper, R. B.; *Introduction to Queueing Theory*, 3rd ed., CEEPress Books, 1990.

[3] Choi, B. D., Kim, B., and Chung, J. M., "M/M/1 queue with impatient customers of higher priority", *Queueing Systems* 38, pp. 49-66, 2001

[4] Kella, O. and Yechiali, U.; "Waiting times in the non-preemptive priority M/M/c queue," *Commun. Statist.-Stochastic Model* 1(2), pp. 252-262, 1985.

[5] Wagner, D.; "Waiting times of a finite-capacity multi-server model with non- preemptive priorities," *European Journal of Operational Research* 102, pp. 227-241, 1997.