

# Multi-server Demanding M/M/s

## 대기행렬의 Service Rate 변화곡선에 관한 연구

(A Study on the Service Rates of M/M/s Multi-server Demanding Queues)

金 聖 植\*  
張 眞 翼\*

### Abstract

An exact solution technique for M/M/s multi-server demanding queue is introduced, and a general form of the curves representing service rates of the system is presented. As the number of customers in the system increases, the service rate is shown to increase initially, then decrease, and finally converge to a certain value. It is also shown that this phenomenon persists regardless of the numbers of servers and customer types.

### 1. 서 론

대기행렬 시스템에 있어서 한 명의 고객이 다수의 서어비스 제공자(server)를 원하는 경우는 우리의 일상 생활에서 종종 볼 수 있다. 예로서 화재발생시 그 규모에 따라 진화에 소요되는 소방차의 수가 달라질 수 있고 컴퓨터에 들어오는 job들의 memory 사용 요구도 서로 다르다.

그러나 다른 많은 경우와 마찬가지로 이 경우에 대한 적당한 해결방법이 Queueing Theory에서 정립되어 있지 않은 상태다. Poisson형 도착, 지수분포형 service의 경우에만 이러한 문제를 연구한 실적([2], [4])이 있을 뿐 도착과정과 service가 상술한 분포가 아닌 경우에 대한 연구는 전무한 상태이다. 이러한 이유는 연구하고자 하는 사람들의 노력의 부족이라기보다는 G/G/1 모델과 M/G/s 모델의 경우처럼 시스템의 특성자체가 현재까지 연구된 Stochastic Processes의 기법들을 적용하기에는 너무나 복잡하기 때문이다. 어떤 새로운 혁신적인 방법의 출현이 없이는 이들 모델에 대한 Queueing Theory의 발전은 점점 복잡한 형태의 수식들의 도입에도 불구하고 매우 느릴

것으로 생각된다. 그러나 M/M/s의 경우는 이미 연구가 되어[4] 그 결과로 steady-state에서의 state probability와 measures of performance가 구해졌다.

여기서 사용된 approach에서는 state수가 행렬식의 차원울 결정한다. 따라서 server의 수가 증가할수록 system 내에 정의하여야 할 state의 수가 증가하며 어느 한도가 지나면, 주어진 컴퓨터의 용량으로는 이러한 행렬식을 취급할 수 없는 범위로 연관 행렬식의 크기가 커지게 된다. 이는 여기서 사용된 approach에서 뿐만 아니라 어느 시스템에서이든지 모델 형성 과정에서 모델중에 어떤 특별한 변수를 써서 나타내 주어야 할 state의 수가 증가하면 자연적으로 파생되는 문제이다. Priority Queue에서 priority의 수가 증가할 때에도 이러한 문제점이 발생한다. 이러한 문제들로 인하여 실제 계산가능한 서어비스 제공자의 수는 한정되어지며, 따라서 어떤 개선방법의 필요성이 대두된다.

본 논문에서는 [4]에 의하여 얻어진 결과들을 그래프로 옮겨 그럴 때 곡선의 공통적인 성질이 발견되었던 타 그러한 성질들을 갖게되는 이유를 증명하였다. 이는 앞으로 연구되어야 할 개선방법(approximation method)에 상당한 정보를 제공하는 것으로 믿어지는 바이다.

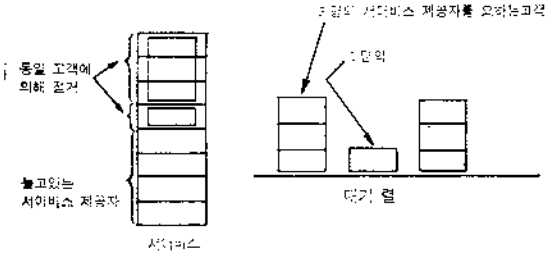
본 논문에서는 먼저 시스템의 특성을 설명하고 정확한 해를 구하는 과정을 제시한 후 이로부터 얻어진 해

\* 高麗大學校

로써 시스템의 service rate를 나타내는 그래프를 그렸다. 후에 이 그래프를 관찰하여 주어진 시스템의 성질을 추출한 후 이러한 성질을 갖는 이유를 증명한다.

2. 시스템의 성질 및 이의 해법

먼저 앞의 서론에서 언급된 시스템의 성질을 자세히 보기로 하자.



(그림 1) 대기행렬의 형태

- ① 도착 과정; 도착은 한명씩 관수(parameter)  $\lambda$ 를 갖는 Poisson 과정에 따라 일어나며 도착되는 고객이 요구하는 서어비스 제공자의 수는 type에 따라 각각 다르다. 한 고객이 type  $k$ 에 속할 확률이 주어졌다고 하고 이를  $P_k$ 로 표기한다. 따라서 type  $k$ 의 관수(input rate)는  $\lambda_k = \lambda P_k$ 가 되고,  $\sum \lambda_k = \lambda$ .
- ② 서어비스; 언젠가 쉬고있는 서어비스 제공자의 수가 대기열의 제일 선두에 있는 고객의 요구를 만족시키면 그 고객은 서어비스로 들어간다. 서어비스 시간은 평균  $1/\mu$ 를 갖는 지수분포를 이루고 있다.

기호;

- $s$ ; 서어비스 제공자의 수
- $K$ ; 도착하는 고객들의 type의 수
- $d_k$ ; type  $k$ 의 고객들이 요구하는 서어비스 제공자의 수. ( $d_1 < d_2 < \dots < d_K$ )
- $n$ ; 시스템 내의 고객의 수
- $t_1, t_2, \dots, t_n$ ; 시스템 내에  $n$ 명이 있을 때 이들이 시스템에 도착한 시간. ( $t_i \leq t_{i+1}, i=1, 2, \dots, n$ )
- $C_i$ ; 시간  $t_i$ 에 도착한 고객
- $c_i$ ; 고객  $C_i$ 의 type
- $g$ ; 서어비스 중에 있는 고객의 수.
- i.e.  $g = \max\{i; i \leq n | dc_1 + dc_2 + \dots + dc_i \leq s\}$

이외에 서어비스는 도착순으로 이루어지며 대기열 내에 있는 고객의 수는 제한이 없다. 먼저 이러한 시스템을 결정하는 정확한 분석방법을 소개함으로써 본 시스템의 성질을 고찰하고자 한다.

여기서 state들은 시스템내의 고객의 수와 이들의 type을 나타내는 벡터( $n, c_1, c_2, \dots, c_n$ ),  $n=0, 1, 2, \dots$ 에 의하여 완전히 나타내질 수 있다. 그러나  $n$ 이 커지면 벡터의 크기가 커지므로, 이로서는 어떤 수학적 model을 만들기가 거의 불가능하므로 다음과 같이 벡터 내에서 고려되어질 고객 type의 수를 정한다.

우선 서어비스 제공자의 입장에서 보면 고객의 type이 고객이 시스템 내에 들어오는 순간에 정해지든 아니면 대기열의 선두에 나서는 순간 정해지든 상관이 없다. 따라서 선두에 있는 것을 제외하고는 대기열 중 어떤 고객이 type  $k$ 에 속할 확률은 단순히  $P_k$ 이다. 그러므로  $n$ 이  $g+2$ 보다 클때

$$P(\text{시스템의 state}=(n, c_1, c_2, \dots, c_n)) = P[(n, c_1, \dots, c_{g+1})] P_{c_{g+2}} P_{c_{g+3}} \dots P_{c_n}$$

이고,

$g+1$ 이 취할 수 있는 최대값은

$$\max(g+1) = [s/d_1] + 1$$

([A]; A보다 작거나 같은 가장 큰 정수)

로 주어진다. 또한 모든 설비가  $[s/d_1]$ 의 고객을 서어비스하고 있을 때는 대기열 선두에 있는 고객은 type이 무엇이든지 설비내로 들어갈 수가 없으므로 이때 확률

$$P[C_{r+1} = \text{type } k] = p_k$$

로 계산된다.

여기서  $U = [s/d_1]$ 이라 하고  $f = \min(n, U)$ 라고 하자. 그러면  $n > f$ 일때는

$$P[\text{state}=(n, c_1, c_2, \dots, c_n)] = P[\text{state}(n, c_1, c_2, \dots, c_f)] \times P_{c_{f+1}} \dots P_{c_n}$$

으로 주어지므로 모든 steady-state 확률들은 state( $n, c_1, c_2, \dots, c_f$ ),  $n=0, 1, 2, \dots$ 의 확률이 구해지면 계산이 가능하다. 이러한 이유로 state들은 벡터( $n, c_1, c_2, \dots, c_f$ ),  $n=0, 1, \dots$ 로 표시될 수 있고 이들의 steady-state 확률들은  $q_{c_1, c_2, \dots, c_f}^{(n)}$ 으로 표시된다.

시스템 내에  $n$ 명의 고객이 있을 때 앞의  $f$ 명이 type을 가질 수 있는 방법의 수는  $K^f$ 가 된다. 예로서 한명이 있을 때는  $K$ 가지, 두명이 있을 때는  $K^2$ 등이다. 따라서 주어진 고객의 수에 대해 고려해 주어야 할 state의 수는  $K^n$  ( $n < U$ ) 또는  $K^U$  ( $n \geq U$ )로 주어진다.

이제  $J_n$  ( $n=1, 2, \dots, K^f$ )을 시스템 내에  $n$ 명의 고객이 있을 때 처음  $f$ 명의 type들을 나타낸다고 하고

$$J_1^n = [1, 1, \dots, 1]$$

$$J_2^n = [1, 1, \dots, 2]$$

⋮



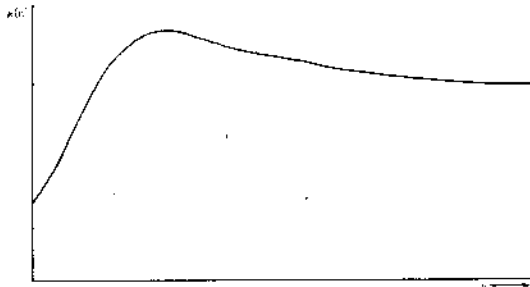
### 3. Service Rate의 변화 곡선

일반적으로 M/M/· 형 대기행렬의 steady-state 방정식은

$$\lambda(n)Q(n) = \mu(n+1)Q(n+1), \quad n=0, 1, 2, \dots$$

의 형태로 쓸 수가 있다. 여기서  $Q(n)$ 은 steady-state probability 들이고 우리의 경우  $Q(n)e$  로 구해진다.

따라서 이 시스템에서는  $\lambda = \lambda(n)$ ,  $n=0, 1, 2, \dots$  이므로  $\mu(n)$ 의 값을 알 수 있다면 최소한 시스템 내에  $n$  명이 있을 확률  $Q(n)$  ( $Q(n) = Q(n)e$ ) 을 구할 수 있다. 이제 실제 정확히 계산된  $Q(n)$  들로 부터 구해낸  $\mu(n) = \lambda Q(n-1)/Q(n)$ ,  $n=1, 2, \dots$  의 값을 그래프로 나타내면 <그림 2>와 같은 형태를 하고 있음을 볼 수 있다.



<그림 2> Service rate의 변화 곡선

- <그림 2>로부터 두가지 사실이 관측되는 바
- i)  $\mu(n)$ 들의 최고값은  $n$ 이  $U$ 에 근사하고  $n \leq U$ 인 곳에서 얻어지며
  - ii) 이 최고값에 도달한 후  $\mu(n)$ 은  $n$ 의 증가에 따라 감소하여 어떤 값(실제는  $\lambda \cdot sp(R)$ )으로 수렴한다.

먼저 수렴하는 사실 ( $\lim_{n \rightarrow \infty} Q(n+1)/Q(n) = \theta$ )을 확인하기 위하여 다음의 과정들을 보자. 행렬  $R$ 은 비음 행렬이고  $sp(R) < 1$ . 따라서 고유한 양(positive)의 특성치(eigen value),  $\theta = sp(R) < 1$ 이 존재하고 여기에 해당하는 양의 벡터(eigen-vector)  $\eta$ 가 페론-후르티니 우스의 정리에 의해 존재하므로

$$R\eta = \theta\eta \tag{5}$$

이제 벡터  $Z(n) = [z_1^n, z_2^n, \dots, z_r^n]$ 을  $(H(n)e)^{-1} \cdot H(n)$ 에서 얻어진 벡터라 하자.

그러면  $n > U$ 일때 (2)식에서

$$\pi(n+1)Z(n+1) = \pi(n)Z(n)R$$

로 쓸 수 있고 양면에  $\eta$ 을 곱하고 또 식 (5)에 의해  $\pi(n+1)Z(n+1)\eta = \theta\pi(n)Z(n)\eta$

이 되며

$$\therefore \theta = \pi(n+1)Z(n+1)\eta / \pi(n)Z(n)\eta \tag{6}$$

이 된다.

여기서 벡터  $Z(n)$ 의 한 원소  $z_i^n$ 은 분명히 시스템 내에  $n$ 명이 있을 때 고객들의 type 이  $i$ 번째 조합을 하고 있을 조건부 확률임을 알 수 있다. 다음의 이유들에 의하여  $Z(n)$ 은  $n \rightarrow \infty$  일때 일정한 값을 갖는다.

먼저  $g_{ij}(n, k)$ 를 정확히  $k(k > 1)$ 스텝에 시스템이 state  $(n, i)$ 를 떠나 처음으로 state  $(n, j)$ 에 도착할 확률이라 하자. ( $i, j=1, 2, \dots, K'$ ) 이제 행렬  $G(n, k)$ ,  $n=0, 1, 2, \dots$ 를  $g_{ij}(n, k)$ 들로 이루어진 행렬이라 하고

$$G(n) = \sum_{i=1}^{\infty} G(n, k)$$

라 하자. 그러면  $G(n)$ 은 substate  $i$ 에서  $j$ ( $i, j=1, 2, \dots, K'$ )로 갈 확률들을 나타내는 transition probability matrix이다.

이제 여기서 취급되는 Queueing Process는 positive recurrent이므로 system에  $n$ 명의 고객이 있다가 얼마간의 transition 후에 다시  $n$ 명의 고객이 있을 확률은 1이다. 다시말해 system이 한 state  $(n, i)$ 에 있다가 다시 어떤 state  $(n, j)$   $j=1, 2, \dots, K'$ 를 방문할 확률은 1이다. 따라서 행렬  $G(n)$ 의 열의 합은 1이 되므로  $G(n)$ 은 stochastic matrix이다. 이는  $g(n)G(n) = g(n)$ 과  $g(n)e = 1$ 을 만족시키는 벡터  $g(n)$ 이 존재한다는 것을 의미한다.  $g(n)$ 의 한 요소  $g_i(n)$ 은 limiting probability로 한 state  $(n, \cdot)$ 에서 어떤 state  $(n, \cdot)$ 로의 transition이 무수히 일어났을 때 고객이  $n$ 명이 있을 경우 system이 state  $i$ 에 있을 조건부 확률이다. 따라서  $g(n) = Z(n)$ 이 되며 벡터  $Z(n)$ ,  $n=U, U+1, \dots$ 은  $Z(n)G(n) = Z(n)$ 의 관계를 만족시킨다.

그러므로  $Z(n)$ 이  $n \rightarrow \infty$  일때 어떤 특정한 값으로 수렴하는 것은  $G(n)$ 이 어떤 특정 행렬로 수렴하는 것을 보이던 된다.  $x_{ij}(n, k)$ 와  $y_{ij}(n, k)$ 를 각각  $k$  스텝에 시스템이 state  $(n+1, i)$ 와 state  $(n-1, i)$ 에서 출발하여 state  $(n, j)$ 에 처음으로 도착할 확률이라 하자  $n > U$ 일때  $\underline{X}(n, k)$ 와  $\underline{Y}(n, k)$ 를 원소들이 각각  $x_{ij}(n, k)$ 와  $y_{ij}(n, k)$ 인 행렬들이라 하고

$$\underline{X}(n) = \sum_{i=1}^{\infty} \underline{X}(n, k)$$

$$\underline{Y}(n) = \sum_{i=1}^{\infty} \underline{Y}(n, k)$$

라 하자.

그러면  $n > U$ 일때 떠나는 state에 다시 돌아오는 것은 짝수의 스텝에서만 가능하다는 사실을 인지하여

$$G(n) = \sum_{i=1}^{\infty} G(n, k)$$

$$\begin{aligned}
 &= \sum_{k=1}^n G(n, k) \\
 &= M \sum_{k=1}^n \underline{Y}(n, k) + W \sum_{k=1}^n \underline{X}(n, k) \\
 &= M \underline{Y}(n) + W \underline{X}(n) \tag{7}
 \end{aligned}$$

이 된다.

정리 1; 모든  $n=U, U+1, \dots$ 에서  $\underline{X}(n)=\underline{X}$ .

여기서  $\underline{X}$ 는  $\underline{X}=M+W\underline{X}^2$ 을 만족시키는 행렬이다.

증명;  $\underline{X}^{(i)}(n, k)$ 를 자원  $x^{(i)}_{n, k}$ 들이 정확히  $k$ 스텝에 시스템이 state  $(n+i, i)$ 에서 시작하여 state  $(n, j)$ 에 처음으로 도달할 확률을 표시하는 행렬이라 하고

$$\underline{X}^{(i)}(n) = \sum_{k=1}^n \underline{X}^{(i)}(n, k)$$

라 하자. 그러면

$$\begin{aligned}
 \underline{X}(n) &= \sum_{k=1}^n \underline{X}(n, k) \\
 &= M + W \sum_{k=2}^n \underline{X}^{(2)}(n, k-1) \\
 &= M + W \sum_{k=1}^n \underline{X}^{(2)}(n, k) \\
 &= M + W \underline{X}^{(2)}(n) \tag{8}
 \end{aligned}$$

Convolution의 성질과 Neuts [6]의 증명에 의해

$$\underline{X}^{(2)}(n) = [\underline{X}(n)]^2,$$

또 식  $\underline{X}(n) = M + W[\underline{X}(n)]^2$ 의 유일해는  $\underline{X}(n) = \underline{X}$ ,  $n > U$ 일때 모든  $\underline{X}(n)$ 의 식 (8)을 만족시키므로 정리가 성립된다.

정리 2; 행렬  $\underline{Y}(n)$ 은  $n$ 이 커짐에 따라  $\underline{Y} = W + MYW$ 를 만족시키는 행렬  $\underline{Y}$ 로 수렴한다.

증명:  $n > U$ 일때

$$\begin{aligned}
 \underline{Y}(n) &= \sum_{k=1}^n \underline{Y}(n, k) \\
 &= \underline{Y}(n, 1) + \sum_{k=2}^n \underline{Y}(n, k) \\
 &= W + \sum_{k=2}^n \underline{Y}(n, k) \tag{9}
 \end{aligned}$$

또  $n-1$ 에서  $n$ 으로 3 이상의 스텝을 거친후 도달하기 위해서는 transition은  $n-2$ 를 먼저 방문함으로써 시작해야 한다. 따라서 식 (9)에서

$$\begin{aligned}
 \underline{Y}(n) &= W + M \sum_{k=1}^n \underline{Y}(n-1, k) W \\
 &= W + M \underline{Y}(n-1) W \tag{10}
 \end{aligned}$$

따라서 두 행렬  $\underline{Y}(n+1)$ 과  $\underline{Y}(n)$ 의 차의 절댓값은

$$\begin{aligned}
 |\underline{Y}(n+1) - \underline{Y}(n)| &= |M \underline{Y}(n) W - M \underline{Y}(n-1) W| \\
 &= |M(\underline{Y}(n) - \underline{Y}(n-1)) W| \\
 &= M |\underline{Y}(n) - \underline{Y}(n-1)| W \\
 \text{따라서 일반적으로} \\
 |\underline{Y}(U+l+2) - \underline{Y}(U+l-1)| &= M^l |\underline{Y}(U+2) \\
 &\quad - \underline{Y}(U+1)| W^l
 \end{aligned}$$

그런데  $D (= M+W)$ 는 irreducible stochastic 행렬이고  $sp(D) = 1$ 이므로  $sp(M)$ 과  $sp(W)$ 는 각각 양이고 1보다 작다. 따라서

$$\lim_{l \rightarrow \infty} M^l = \lim_{l \rightarrow \infty} W^l = 0$$

여기서 0는 영 행렬이다.

그러므로

$$\begin{aligned}
 \lim_{l \rightarrow \infty} |\underline{Y}(n) - \underline{Y}(n-1)| &= \lim_{l \rightarrow \infty} |\underline{Y}(U+l+2) \\
 &\quad - \underline{Y}(U+l+1)| = 0
 \end{aligned}$$

이러한 이유로  $\underline{Y}(n)$ 은 수렴하며 더구나  $\underline{Y} = \lim_{n \rightarrow \infty} \underline{Y}(n)$

이라고 하면 식 (10)은 항상 성립하므로

$$\underline{Y} = W + M \underline{Y} W$$

가 되어 정리가 성립한다.

이제 정리 1 및 정리 2를 이용하여 벡터  $Z(n)$ 들이 어떤 고정된 벡터  $Z$ 로 수렴함을 보자.

식 (7)에서  $G(n) = M \underline{Y}(n) + W \underline{X}(n)$ 이므로 정리를 1과 2에 의해

$$\begin{aligned}
 \lim_{n \rightarrow \infty} G(n) &= \lim_{n \rightarrow \infty} (M \underline{Y}(n) + W \underline{X}(n)) \\
 &= M \underline{Y} + W \underline{X}
 \end{aligned}$$

$$\begin{aligned}
 \text{따라서 } Z &= \lim_{n \rightarrow \infty} Z(n) = \lim_{n \rightarrow \infty} Z(n) G(n) \\
 &= ZG
 \end{aligned}$$

이요  $Ze = 1$ , 여기서  $G = M \underline{Y} + W \underline{X}$ 이다.

정리 3.

$$\lim_{n \rightarrow \infty} Q(n+1)/Q(n) = \theta$$

증명; 식 (4)로부터  $Q(n) = a \Pi(n) Y_{\nu}$ 의 형태를 쓰여질 수 있음을 보았다. 여기서

$$a = \sum_{n=U}^{\infty} \Pi(n) Y_{\nu} e$$

따라서  $n > U$ 이면

$$\begin{aligned}
 Q(n+1)/Q(n) &= Q(n+1)e/Q(n)e \\
 &= a \Pi(n+1) Y_{\nu} e / a \Pi(n) Y_{\nu} e \\
 &= \pi(n+1) Z(n+1) Y_{\nu} e / \pi(n) Z(n) Y_{\nu} e \tag{11}
 \end{aligned}$$

따라서

$$\begin{aligned}
 \lim_{n \rightarrow \infty} Q_{n+1}/Q_n &= \lim_{n \rightarrow \infty} \pi(n+1) Z(n+1) Y_{\nu} e / \\
 &\quad \pi(n) Z(n) Y_{\nu} e
 \end{aligned}$$

그런데

$$\lim_{n \rightarrow \infty} Z(n+1) Y_{\nu} e = \lim_{n \rightarrow \infty} Z(n) Y_{\nu} e = Z Y_{\nu} e \text{ 이므로}$$

$$\lim_{n \rightarrow \infty} Q_{n+1}/Q_n = \lim_{n \rightarrow \infty} \pi(n+1)/\pi(n)$$

식 (6)과  $\lim_{n \rightarrow \infty} Z(n) = Z$ 의 성질을 이용하면

$$\lim_{n \rightarrow \infty} \pi(n+1)/\pi(n) = \theta$$

이제  $\lambda/\mu(n+1) = Q(n+1)/Q(n)$ 이므로 평균 서비스율  $\mu(n)$ 은  $n \rightarrow \infty$  일 때 앞의 특성에서 지적되었던 대로  $\lambda/\theta$ 로 접근해 간다. 더구나  $\theta = sp(R) < 1$  (\*)인 것은  $\pi(n+1) = \pi(n) R\theta < \theta\pi(n)$ ,  $n \geq U$ 임을 의미한다.

따라서  $Q(n+1) \leq \theta Q(n)$ ,  $n > U$ 이고

$$Q(n+1)/Q(n) = \lambda/\mu(n) \leq \theta.$$

위 식은  $\mu(n) \geq \lambda/\theta$ 을 의미하므로  $\mu(n)$ 은  $n$ 의 증가에 따라 감소하며  $\lambda/\theta$ 로 수렴해간다. 이것으로 두번째 특성이 증명되었다.

서비스비율이 증가하여 어떤  $n$ ,  $n \leq U$ 에서 최고 값에 달하고 곧이어 감소하기 시작하는 현상을 설명하기 위하여 먼저 행렬들  $A_n$ ,  $B_n$ ,  $W$ ,  $M$  및  $R$ 의 성질을 알아볼 필요가 있다.

식 (3)을  $V_n$ 의 형태를 이용하여 다시 쓰면

$$H(n) = \begin{cases} H(n-1)A_{n-1}(I - V_{n+1}B_{n+1})^{-1} & , n < U \\ H(U-1)A_{U-1}(I - RM)^{-1} & , n = U \\ H(n-1)R & , n > U \end{cases}$$

로 쓰여진다. 이 모양을 이용하면 조건부 확률  $Z(n)$  들은

$$Z(n) = \begin{cases} \frac{\pi(n-1)}{\pi(n)} \cdot Z(n-1)A_{n-1}[I + V_{n+1}B_{n+1} + (V_n + B_{n+1})^2 + \dots] & n < U \\ \frac{\pi(U-1)}{\pi(U)} \cdot Z(U-1)A_{U-1}[I + RM + (RM)^2 + \dots] & n = U \\ \frac{\pi(n-1)}{\pi(n)} \cdot Z(n-1)[W + R^2M], & n > U \end{cases} \quad (12)$$

로 쓰여진다. 여기서 행렬  $V_n$ 과  $R$ 은 모든 원소들이 0보다 크고 행렬  $B_n$ 과  $M$ 은 비교적 sparse한 행렬이나 양(+)의 원소들은 골고루 분포되어 있다. 반면 행렬  $A_n$ 은 band diagonal이고  $W$ 는 diagonal이다. 따라서 벡터  $Z(n-1)$ 의 한 원소  $z_i^{n-1}$ 가 벡터  $Z(n)$ 에 미친 영향을 생각할 때 행렬들  $V_n$ ,  $B_n$ ,  $R$  및  $M$ 에 의해서 미쳐진 영향은  $Z(n)$ 의 모든 원소들에 비교적 골고루 미치게 되는 한편  $W$ 와  $A_n$ 에 의하여 미쳐진 영향은  $z_i^n$  부근 또는  $z_i^n$  자체에만 국한된다.

이제  $Z(n)$ 의 원소들  $z_i^n$ 은 시스템에  $n$ 명의 고객이 있을 때 처음  $f$ 명의 고객의 type이  $i$ 번째 조합을 이룰 조건부 확률이다.  $I$ 를 제외하고 식 (12)의 모든 행렬들은 열의 합이 1보다 작은 양(+)의 행렬들이다.

따라서  $(V_n B)^n$ 과  $(RM)^n$ 에 의한 효과는  $n$ 의 증가에 따라 급격히 감소되고 조건부 확률 벡터  $Z(n)$ 의 값은 주로 행렬  $A_n$ 과  $W$ 에 의해 결정이 된다.

이제 행렬  $A_n$ 과  $W$ 가  $Z(n)$ 에 미치는 효과를 보자. 행렬  $A_n$ 의 양(+)의 원소는  $\lambda_i/(l_i + \lambda)$ 의 형태로 주어지고  $W$ 의 원소는  $\lambda/(l_i + \lambda)$ 의 형을 하고 있다. 여기서  $l_i$ 는 처음  $f$ 명의 고객의 type이  $i$ 번째 조합을 이루고 있을 때 서비스를 받고 있는 고객의 수이다. 따라서 한 조합이 적은 수의 고객( $l_i$ )을 서비스에 가지고 있을수록 upward transition 확률이 커짐은 명확하다. 다시 말해  $Z(n-1)W$ 와  $Z(n-1)A_{n-1}$ 이 계산될 때 만일 한 조합  $i$ 가 다른 조합  $j$ 보다 적은 수의 고객을 서비스에 가지고 있다면  $z_i^{n-1}$ 에 곱해지는  $W$ 와  $A_{n-1}$ 의 원소가  $z_j^{n-1}$ 에 곱해지는 원소보다 크다는 것을 의미한다. 따라서 큰 수의 서비스 제공자를 요구하는 고객들로 이루어진 조합(서비스에 있는 고객의 수가 적은)의 상태에 시스템이 있을 조건부 확률은  $n$ 의 증가에 따라 증가하며 그렇지 않을 상태에 있을 확률은 반대로 감소한다. 이러한 추세는 어떤  $n$ 에서든지 보여진다. 그러나  $n$ 이 아주 적을 때는 서비스에 들어오는 고객의 모든 서비스 제공자 요구를 충족시키므로 처음에는 서비스율이 증가하게 된다. 이후  $n$ 이  $U$ 에 접근함에 따라 차차 서비스가 모든 고객의 요구를 만족시켜 줄 수 있는 기회가 줄어들며 위에서 말한 효과가 점점 확실하게 나타나게 되어 증가율은 떨어지며 마침내 감소를 시작하게 되는 것이다.

### 결론

여기서 고려된 주어진 시스템의 service rate의 변화형태에 대한 증명은 그 자체에 효용성이 있는 것은 아니다.

그러나 상기의 증명으로 이 변화형태가 M/M/s Multi-server demanding queue에서 공통적으로 일어난다는 사실이 확인되었다. 이 성질을 이용하여 적당한 변곡점을 찾아 service rate의 변화를 직선으로 생각하여 변화율의 근사치를 구하는 것이 가능하다고 생각되어 지는 바 이를  $\lambda Q(n) = \mu(n+1)Q(n+1)$ 의 식에 대입하면 시스템에  $n$ 명이 있을 확률  $Q(n)$ 의 계산(approximation)이 가능하다. 이러한 계산법은 컴퓨터 기억용량에 거의 제한을 받지 않으므로 시스템의 서비스 제공자 수와 고객 type의 수에 거의 무관하게 큰 시스템에도 적용 가능하리라 믿는 바이다.

### References

1. Gantmacher, F.R., *Application of the Theory of*

\* Karlin [3], Page 484

1. *Matrics*, Interscience Publishers, N.Y., 1959.
  2. Green, L., "Queues in Which Customer Require a Random Number of Servers," Presented at ORSA/TIMS Meeting in New York, May 1978.
  3. Karlin, S., *A First Course in Stochastic Processes*, Academic Press, N.Y., 1969.
  4. Kim, S.S., "M/M/s Queueing System Where Customers Demand Multiple Server Use," Ph.D. Dissertation, Department of Industrial Engineering, Southern Methodist University, 1979.
  5. Neuts, M.F., "The Markov Renewal Branching Processes," *Proc. Conf. on Math. Methods in the Theory of Queues*, Kalamazoo, MI., Springer-Verlag, N.Y., 1-21, 1974.
  6. Neuts, M.F., "The M/G/1 Queue with Several Types of Customers and Change-Over Times," *Advances in Applied Probability*, 9, 604-644, 1977.
  7. Ross, M.S., *Applied Probability Model with Optimization Applications*, Holden-Day, 1970.
-