Journal of the Korean Institute
of Industrial Engineers
Vol.22, No.3, September, 1996

분기 및 접합형태를 갖는 대기행렬 모형의 근사적인 분석

An Approximate Analysis of the Fork-Join Queue

양원석*, 채경철**

W. S. Yang*, K. C. Chae**

Abstract

We present an approximate algorithmic approach for the m-server fork-join queueing system with infinite queue capacity. We analyze the system size and queue waiting time. We assume Poisson arrival process and independent exponential service times.

1. 서론

분기및 접합형태를 갖는 대기행렬 모형(fork-join queue)은 m개의 병렬 채널(parallel channel)로 구성되어 있고 각 채널은 1명의 서어버(server)와 대기장소로 구성되어 있다(그림 1 참조).

도착과정은 m명씩 구성된 집단이 포아송 과정에 따라 도착하며, 시스템에 도착한 집단은 각 고객으로 분해된 후 각 채널에서 1명씩 서비스를 받는다. 서비스를 마친 고객은 크기가 무한대인 가상적인 공간에서 아직 서비스를 못 끝낸(동일 집단에 속한) 다른 고객들을 기다린다고 가정한다. 집단내의 모든 고객들이 서비스를 받으면 이 집단은 시스템을 떠난다. 서비스 시간은 서로 독립인 지수

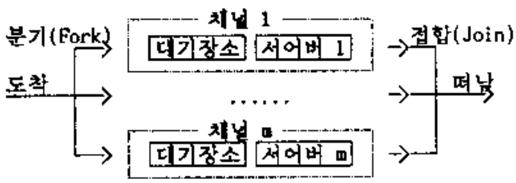

그림 1. 분기및 접합형태를 갖는 대기행렬 모형

* 한국과학기술원 산업경영학과 박사과정
** 한국과학기술원 산업경영학과 교수

분포를 따른다고 가정한다.

분기및 접합형태를 갖는 대기행렬 모형은 병렬처리(parallel processing), 유연생산(flexible manufacturing), 데이터베이스(database) 그리고 통신시스템등 여러 분야에 응용된다 [1,3,4,6,7,8]. 2-채널 시스템은 라플라스 변환을 통해 정확하게 분석되었다 [1,2,14]. 그러나 3-채널 이상의 시스템은 체류시간(sojourn time)만이 근사적으로 분석되었다 [3,4,8].

본 논문에서는 행렬기하(matrix-geometric) 방법에 기초하여 각 채널의 용량이 무한대인 분기 및 접합형태를 갖는 대기행렬 모형의 각 채널내 고객수의 안정상태 확률과 대기시간(queue waiting time)을 근사적으로 분석한다.

2. 수학적 모형: QBD 모형

본 논문에서 다루는 QBD(quasi-birth-and-death) 모형은 다음과 같다: m명의 고객으로 구성된 집단이 도착률이 λ인 포아송과정에 따라 도착한다. 채널 j 서어버의 서비스 시간은 서비스율이 μ_j인 지수분포를 따르고 각 서어버의 서비스 시간은 서로 독립이다. 서비스 정책은 선입선출방식(FIFO)이다. 채널 1의 용량은 무한대이고 나머지 채널들의 용량은 유한인 K이다. 채널 $2,\cdots,m$중 적어도 한 곳이 꽉 차있는 상태에 도착한 집단은 시스템에 입장하지 못하고 바로 시스템을 떠난다고 가정한다.

이러한 가정 하에서 각 채널당 고객수의 추계적 과정은 유사 출생사망과정(QBD process)이 되며 행렬기하(matrix-geometric) 방법 [9]을 통해 수치적(numerical)으로 분석할 수 있다. λ에 따라 K가 충분히 크면 시스템에 입장을 못하는 고객의 비율이 작아지기 때문에 QBD 모형을 각 채널의 용량이 무한대인 분기및 접합형태를 갖는 대기행렬 모형의 근사모형으로 이용할 수 있다. QBD 모형은 Rao와 Posner [10]의 모형을 일반화한 것이다. (Rao와 Posner는 m=2인 경우를 다루었음.)

안정상태에서 시스템 상태(state)를 벡터 $\boldsymbol{n}=(n_1,\cdots,n_m)$으로, 상태 $\boldsymbol{n}$에서 상태 $\boldsymbol{n}'$으로의 전이율(transition rate)을 $q_{\boldsymbol{n},\boldsymbol{n}'}$으로 표기한다. 여기에서 n_j 는 채널 j에 있는 고객수이다. $n_1 \geq 0,\ 0 \leq n_j \leq K,\ j=2,\cdots,m$. 편의상 상태공간(state space) $\{(i,n_2,\cdots,n_m) \mid 0 \leq n_j \leq K, j=2,\cdots,m\}$을 초상태(superstate) $\boldsymbol{i}$라 부르고 $\boldsymbol{i}$라 표기한다. 그리고 다음과 같은 지수(indicator) 함수를 정의한다.

$$\delta(A)=\begin{cases} 1 & \text{조건 } A\text{를 만족한다} \\ 0 & o/w \end{cases}$$

0이 아닌 전이율은 다음과 같은 세 가지 형태로 분류된다.

(i) S_i에서 S_{i+1}로의 전이

$$q_{(n_1,\cdots,n_m),(n_1+1,\cdots,n_m+1)}=\lambda\ \ \delta(n_j<K,\ j=2,\cdots,m)$$

(ii) S_i에서 S_{i-1}로의 전이

$$q_{(n_1,\cdots,n_m),(n_1-1,n_2,\cdots,n_m)}=\mu_1\ \delta(n_1>0)$$

(ii) S_i에서 S_i로의 전이

$$q_{(n_1,\cdots,n_j,\cdots,n_m),(n_1,\cdots,n_j-1,\cdots,n_m)}=\mu_j\ \ \delta(n_j>0)\ j=2,\cdots,m.$$

$$q_{\boldsymbol{n},\boldsymbol{n}}=-[\lambda\ \ \delta(n_j<K,\ j=2,\cdots,m)+\sum_{j=1}^{m}\mu_j\,\delta(n_j>0)]$$

먼저, 상태들을 초상태의 순서로 늘어놓는다. (즉, n_1에 대해 오름차순으로 늘어놓는다.) 그리고 초상태내의 상태들을 각각 $n_2,n_3,\cdots$에 대해 오름차순으로 늘어놓는다. 그러면 무한소 생성자(infinitesimal generator) Q는 다음과

같이 표현된다.

$$Q=\begin{bmatrix} A_-+A_0 & A_+ & 0 & 0 & \cdots \\ A_- & A_0 & A_+ & 0 & \cdots \\ 0 & A_- & A_0 & A_+ & \cdots \\ 0 & 0 & A_- & A_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \qquad (1)$$

여기에서 A_-, A_0, A_+는 크기가 $(K+1)^{m-1}$인 정방행렬이다. $i=1,2,\cdots$에 대해, A_+는 S_i에서 S_{i+1}, A_0는 S_i에서 S_i 그리고 A_-는 S_i에서 S_{i-1}로의 전이율 행렬이다. 아울러 A_-+A_0는 S_0에서 S_0로의 전이율 행렬이다. 예를 들어, $m=3$이고 $K=2$일 때 A_0, A_+ 그리고 A_-는 다음과 같다:(편의상 $\mu=\mu_1+\mu_2+\mu_3$라 함.)

$$A_0=\begin{bmatrix} -(\lambda+\mu_1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_3 & -(\lambda+\mu_1+\mu_3) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_3 & -(\mu_1+\mu_3) & 0 & 0 & 0 & 0 & 0 & 0 \\ \mu_2 & 0 & 0 & -(\lambda+\mu_1+\mu_2) & 0 & 0 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & \mu_3 & -(\lambda+\mu) & 0 & 0 & 0 & 0 \\ 0 & 0 & \mu_2 & 0 & \mu_3 & -\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_2 & 0 & 0 & -(\mu_1+\mu_2) & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu_2 & 0 & \mu_3 & -\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu_2 & 0 & \mu_3 & -\mu \end{bmatrix}$$

$$A_+=\begin{bmatrix} 0&0&0&0&\lambda&0&0&0&0 \\ 0&0&0&0&0&\lambda&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&\lambda&0 \\ 0&0&0&0&0&0&0&0&\lambda \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \end{bmatrix} \qquad A_-=\begin{bmatrix} \mu_1&0&0&0&0&0&0&0&0 \\ 0&\mu_1&0&0&0&0&0&0&0 \\ 0&0&\mu_1&0&0&0&0&0&0 \\ 0&0&0&\mu_1&0&0&0&0&0 \\ 0&0&0&0&\mu_1&0&0&0&0 \\ 0&0&0&0&0&\mu_1&0&0&0 \\ 0&0&0&0&0&0&\mu_1&0&0 \\ 0&0&0&0&0&0&0&\mu_1&0 \\ 0&0&0&0&0&0&0&0&\mu_1 \end{bmatrix}$$

Remark 1: 본 논문에서는 분석의 편의상 각 채널의 용량을 모두 같게 하였다. 그러나 각 채널의 용량이 $K_j(j=2,\cdots,m)$로 서로 다른 경우에도 무한소 생성자는 식 (1)과 같은 형태로 표현된다. 이 경우 행렬 A_-, A_0 그리고 A_+는 크기가 $\Pi_{j=2}^{m}(K_j+1)$인 정방행렬이다.

식 (1)의 Q와 같은 구조를 갖는 연속적 마코프사슬(continuous time Markov chain)을 유사 출생사망과정이라 부른다 [9]. 안정상태에서 시스템의 상태가 상태 $(n_1,\cdots,n_m)$에 있을 확률을 $x_{(n_1,\cdots,n_m)}$이라 표기한다. 그리고 초상태 S_i에 속하는 상태들의 안정상태 확률을 원소로 갖는 행벡터(row vector)를 $\mathbf{x}_i$로 표기한다. 즉, $\mathbf{x}_i=(x_{(i,0,\cdots,0)},\cdots,x_{(i,K,\cdots,K)})$이다.

QBD 모형은

$$\lambda(1-P_b) < \mu_1 \tag{2}$$

일 때 안정(stable)하다(증명은 부록 참조). 여기에서 P_b는 고객이 시스템에 입장하지 못할 차단확률(blocking probability)로 $P_b=1-\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{K-1}\cdots\sum_{n_m=0}^{K-1}x_{(n_1,\cdots,n_m)}$이다.

Remark 2: QBD 모형의 각 채널을 독립적으로 관찰하면, 채널 1은 용량이 무한대인 단일 서어버 대기행렬 모형(single server queue)이고 채널 2,⋯,m은 용량이 유한인 단일 서어버 대기행렬 모형이다. 채널 2,⋯,m은 용량이 유한이므로 자체적으로 안정하나 채널 1은 실질도착률(effective arrival rate) $\lambda(1-P_b)$가 서비스율 μ_1보다 작아야 안정하다. 즉, 식 (2)는 채널 1이 안정할 때 전체 시스템이 안정함을 의미한다.

Remark 3: 차단확률은 $P_b=1-\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{K-1}\cdots\sum_{n_m=0}^{K-1}x_{(n_1,\cdots,n_m)}$이므로 $x_{(n_1,\cdots,n_m)}$으로 부터 쉽게 구할 수 있다. 그리고 P_b가 충분히 작으면 QBD 모형은 각 채널의 용량이 무한대인 시스템의 근사적인 효과를 나타낸다.

QBD 모형의 안정상태 확률벡터 $\mathbf{x}_i$는

$$\mathbf{x}_i=\mathbf{x}_0R^i,\quad i=1,2,\cdots$$

를 만족시킨다. 행렬 R은 $R^2A_-+RA_0+A_+=0$의 비음최소해(nonnegative minimal solution)이고 다음과 같은 반복적인 절차를 통해 구할 수 있다 [9].

$$R_0=0$$
$$R_{k+1}=-A_+A_0^{-1}-R_k^2A_-A_0^{-1},\ k=0,1,\cdots$$

그리고 $\mathbf{x}_0$는 다음 두 식에 의해 결정된다.

$$\mathbf{x}_0[(A_-+A_0)+RA_-]=0$$
$$\mathbf{x}_0(I-R)^{-1}e=1$$

여기에서 e는 원소들이 모두 1인 열벡터(column vector)이다.

2.1 시스템내 집단수

N_j를 채널 j에 있는 고객수 그리고 N을 시스템내 집단수라고 하자. $N=\max(N_1,\cdots,N_m)$이므로 안정상태에서 시스템내 집단수의 확률은 다음과 같이 표현된다.

$$P[N=k]=\begin{cases} x_{(0,\cdots,0)} & k=0 \\ \sum_{n_1=0}^{k}\cdots\sum_{n_m=0}^{k}x_{(n_1,\cdots,n_m)}-\sum_{j=0}^{k-1}P[N=j] & 1\le k\le K-1 \\ \mathbf{x}_k e & k\ge K \end{cases} \tag{3}$$

그리고 시스템내 집단수 분포의 j번째 모멘트(moment) $E[N^{(j)}]$는 식 (3)으로 부터 직접 구할 수 있다, $j=1,\cdots$.

Remark 4: P_b가 충분히 작지 않은 경우에는 $E[N]$을 각 채널의 용량이 무한대인 분기 및 접합형태를 갖는 시스템의 평균 집단수의 하한치(low bound)로 이용할 수 있다.

3. 대기시간

본 논문에서는 분기및 접합형태를 갖는 대기행렬 모형의 대기시간(queue waiting time)을 집단이 시스템에 도착(arrival)한 시점부터 집단내의 모든 (고객중에서 마지막으로 서비스를 받기 시작하게 되는) 고객이 서비스를 받기 직전까지의 시간으로 정의한다. $W_Q(\cdot)$를 도착시점에서 대기시간의 분포라고 하면, PASTA [13]에 의해, $W_Q(\cdot)$는 가상 대기시간(virtual queue waiting time)의 분포이기도 하다.

도착하는 집단의 관점에서 대기시간은, 집단의 도착 이후 더 이상의 도착이 없다는 가정하에서, 도착 시점부터 시스템내 고객수가 0이 되기 직전까지의 시간, 즉, 시스템 상태가 도착 상태에서 상태 $(0,\cdots,0)$로 흡수되기 직전까지의 시간이다. 3.1 절에서는 조건부 평균 대기시간의 축차관계식을 이용하여 평균 대기시간을 구한다. 그리고 3.2절에서는 흡수 연속 마코프사슬(absorbing continuous time Markov chain)을 이용하여 대기시간의 LST (Laplace-Stieltjes transform)와 평균 대기시간을 구한다.

3.1 평균 대기시간

시스템 상태가 $(n_1,\cdots,n_m)$일 때 도착한 임의 집단의 조건부 평균 대기시간을 $w_Q(n_1,\cdots,n_m)$라 하자. 채널 $2,\cdots,m$중 적어도 한 곳이 꽉 차있는 상태에 도착한 집단은 시스템에 입장하지 못하므로 $w_Q(n_1,\cdots,n_m)$은 $n_1 \ge 0$, $0 \le n_j < K$에서 정의된다. 어느 채널에서 먼저 서비스가 끝났는가에 조건을 걸면 조건부 평균 대기시간은 다음과 같은 축차관계식으로 표현된다.

$$w_Q(n_1,\cdots,n_m)=\sum_{k=1}^{m}\frac{\delta(n_k>0)\,\mu_k}{\mu_M}w_Q(\nu_1,\cdot,\nu_m)+\frac{1}{\mu_M} \tag{4}$$

여기에서

$$\mu_M=\Sigma_{j=1}^{m}\delta(n_j>0)\,\mu_j,\ \nu_j=\begin{cases}(n_j-1)\,\delta(n_j>0) & :j=k\\ n_j & :j\neq k,\end{cases}$$

$j=1,\cdots,m$ 이고 $w_Q(0,\cdots,0)=0$ 이다. PASTA [13]에 의해 임의 집단이 도착한 시점에서의 각 채널당 고객수의 분포는 $x_{(n_1,\cdots,n_m)}$이므로 평균 대기시간은 다음과 같이 표현된다.

$$E[W_Q]=\Sigma_{n_1=0}^{\infty}\Sigma_{n_2=0}^{K-1}\cdots\Sigma_{n_m=0}^{K-1}x_{(n_1,\cdots,n_m)}w_Q(n_1,\cdots,n_m) \tag{5}$$

3.2 대기시간의 LST와 평균 대기시간

$(0,\cdots,0)$을 흡수 상태로 갖는 흡수 연속 마코프사슬 $\{Y(t),t\ge 0\}$를 생각해 보자. 여기서 $Y(t)$는 시점 t에서 시스템의 상태(각 채널당 고객수)이고 도착시점이 시간 0으로 정의된다. 이 마코프사슬의 무한소 생성자는 다음과 같다.

$$\widetilde{Q}=\begin{bmatrix}C & 0 & 0 & 0 & \cdots\\ A' & D & 0 & 0 & \cdots\\ 0 & A & D & 0 & \cdots\\ 0 & 0 & A & D & \cdots\\ \vdots & \vdots & \vdots & \vdots & \ddots\end{bmatrix} \tag{6}$$

이때, $\widetilde{Q}$는 식 (1)에서 λ를 0으로 한 것이다. 그리고 A'과 A는 식 (1)의 A_{-}와 같다. 행렬 C와 A'은 각각 $C=\begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{b}_1 & C_1 \end{bmatrix}$ 그리고 $A'=[\mathbf{b}_2 C_2]$로 분해된다. 여기에서 $\mathbf{b}_1$과 $\mathbf{b}_2$는 각각 크기가 $(K+1)^m$-1, $(K+1)^m$인 열벡터이다. 또한 C_1은 크기가 $[(K+1)^m-1]$인 정방행렬이고 C_2는 크기가 $(K+1)^m \times [(K+1)^m-1]$ 인 행렬이다.

시점 t에서 시스템 상태가 상태 $(n_1,\cdots,n_m)$에 있을 확률을 $y_{(n_1,\cdots,n_m)}(t)$로 표기하고 $\mathbf{y}_i(t)=(y_{(n_1,\cdots,n_m)}(t) \mid n_1=i)$, $\mathbf{y}(t)=[\mathbf{y}_0(t),\mathbf{y}_1(t),\cdots]$라 하자. 여기에서 n_j는 채널 j에 있는 고객수이고 $n_1 \geq 0$, $0 \leq n_j \leq K$, $j=2,\cdots,m$이다. 그리고 $\mathbf{y}_i(t)$의 원소 $y_{(i,n_2,\cdots,n_m)}(t)$들은 $n_2,n_3,\cdots$에 대해 오름차순으로 정렬되었다. PASTA [13]에 의해 초기조건은 $\mathbf{y}(0)=[\mathbf{x}_0,\mathbf{x}_1,\cdots]$으로 주어진다.

시스템에 도착한 임의 집단의 대기시간은 $\{Y(t)\}$가 시점 0에서 시작하여 상태 $(0,\cdots,0)$로 흡수되기 직전까지의 시간이므로 대기시간의 분포 $W_Q(t)$는 다음과 같이 표현된다.

$$W_Q(t)=y_{(0,\cdots,0)}(t) \tag{7}$$

Kolmogorov forward 미분 방정식 $\mathbf{y}'(t)=\mathbf{y}(t)Q$로 부터 다음을 얻는다.

$$z_0'(t)=\mathbf{z}(t)\mathbf{b}_1+\mathbf{y}_1(t)\mathbf{b}_2 \tag{8.1}$$

$$\mathbf{z}'(t)=\mathbf{z}(t)C_1+\mathbf{y}_1(t)C_2 \tag{8.2}$$

$$\mathbf{y}'_k(t)=\mathbf{y}_k(t)D+\mathbf{y}_{k+1}(t)A \quad k \geq 1,\ t \geq 0 \tag{8.3}$$

여기에서 $\mathbf{y}_0(t)=[z_0(t),\mathbf{z}(t)]$, $z_0(t)=y_{(0,\cdots,0)}(t)$이다. $W_Q(t)$와 $y_{(n_1,\cdots,n_m)}(t)$의 LST를 $W_Q^*(s)$와 $y^*_{(n_1,\cdots,n_m)}(s)$로 표기한다. 이때, $W_Q^*(s)=\int_0^\infty e^{-st}dW_Q(t)$이고 $y^*_{(\cdots)}(s)=\int_0^\infty e^{-st}y_{(\cdots)}(t)dt$이다. 그리고 $\mathbf{z}(t)$와 $\mathbf{y}_i(t)$의 LST 벡터를 $\mathbf{z}^*(s)$와 $\mathbf{y}_i^*(s)$로 표기한다. (8.2)의 양변에 LST를 취하면

$$\mathbf{z}^*(s)=\mathbf{z}(0)(sI-C_1)^{-1}+\mathbf{y}_1^*(s)C_2(sI-C_1)^{-1} \tag{9}$$

이다. (8.3)의 양변에 LST를 취하고 반복적으로 풀면 다음을 얻는다.

$$\mathbf{y}_1^*(s)=\sum_{i=1}^{\infty}\mathbf{y}_i(0)[(sI-D)^{-1}A]^iA^{-1} \tag{10}$$

그리고 식 (7)과 (8.1)에 LST를 취하면 $W_Q^*(s)$는 다음과 같이 표현된다.

$$W_Q^*(s)=\mathbf{z}^*(s)\mathbf{b}_1+\mathbf{y}_1^*(s)\mathbf{b}_2 \tag{11}$$

(9), (10) 그리고 초기조건 $\mathbf{y}(0)=[\mathbf{x}_0,\mathbf{x}_1,\cdots]$을 (11)에 대입하면 $W_Q^*(s)$를 얻는다. 아울러, 식 (12)를 미분하면 평균 대기시간을 얻는다.

$$W_Q^*(s)=\mathbf{z}(0)(sI-C_1)^{-1}\mathbf{b}_1 + \sum_{i=1}^{\infty}\mathbf{x}_i[(sI-D)^{-1}A]^iA^{-1}[C_2(sI-C_1)^{-1}\mathbf{b}_1+\mathbf{b}_2] \tag{12}$$

$$E[W_Q]=\mathbf{z}(0)C_1^{-2}\mathbf{b}_1+\sum_{i=1}^{\infty}\mathbf{x}_i\left\{\sum_{\gamma=0}^{i-1}H^{\gamma}(-D)^{-1}H^{i-\gamma}\right\}A^{-1}\mathbf{b}_2 + \sum_{i=1}^{\infty}\mathbf{x}_i\left\{H^iA^{-1}C_2C_1^{-2}+\left[\sum_{\gamma=0}^{i-1}H^{\gamma}D^{-1}H^{i-\gamma}\right]A^{-1}C_2C_1^{-1}\right\}\mathbf{b}_1 \tag{13}$$

여기에서 $\mathbf{z}(0)=\{x_{(0,\cdots,1)},\cdots,x_{(0,K,\cdots,K)}\}$이고 $H=-D^{-1}A$이다.

4. 수치예

표 1은 서비스율(service rate)과 채널의 용량을 달리하며 3-채널인 분기및 접합형태를

갖는 대기행렬 모형의 시스템내 평균 집단수, 평균 대기시간 그리고 차단확률을 구한 것이다. 여기에서 $\rho_j = \lambda / \mu_j, j=1,2,3$ 이다. 그리고

표 1. 3-채널 분기및 접합형태를 갖는 대기행렬 모형의 수치예

ρ_1	ρ_2	ρ_3	K	$E[N]$	$E[W_Q]$	차단확률(P_b)
0.1	0.1	0.1	10	0.2011	0.0254	0.0000000002
0.2	0.2	0.2	10	0.4465	0.1108	0.0000001628
0.2	0.2	0.2	15	0.4465	0.1108	0.0000000001
0.3	0.3	0.3	10	0.7553	0.2755	0.0000081559
0.3	0.3	0.3	15	0.7553	0.2755	0.0000000200
0.4	0.4	0.4	10	1.1586	0.5509	0.0001224767
0.4	0.4	0.4	15	1.1595	0.5517	0.0000012799
0.4	0.4	0.4	20	1.1595	0.5518	0.0000000132

$E[N]$은 (3)에서 그리고 $E[W_Q]$는 (5)에서 얻었다. $\rho_1 = \rho_2 = \rho_3 = 0.1$ 경우처럼 ρ_j가 작을 때는 K가 작아도 P_b가 충분히 작아진다. 또한 2절에서 언급한 바와 같이, $\rho_1 = \rho_2 = \rho_3 = 0.4$이고 $K=10$인 경우와 같이, P_b가 충분히 작지 않을 때는 $E[N]$을 하한치로 이용할 수 있다.

5. 결론

본 논문에서는 QBD 모형을 이용하여 각 채널의 용량이 무한대인 분기및 접합형태를 갖는 대기행렬 모형의 각 채널내 고객수의 안정상태 확률과 대기시간을 근사적으로 분석하였다. 첫 번째 채널의 용량을 무한대로 그리고 나머지 채널들의 용량을 유한인 K라고 가정한 QBD 모형은 유사 출생사망과정으로 분석할 수 있다. 각 채널당 고객수의 안정상태 확률은 행렬기하해를 갖고 알고리즘을 통해 그 해를 구할 수 있다. 도착율에 따라 K가 충분히 크면 차단확률이 0에 가까와져 각 채널의 용량이 무한대인 시스템의 근사적인 효과를 갖는다.

2절에서는 알고리즘적인 접근을 통해 각 채널내 고객수의 안정상태 확률을 구하였다. 3절에서는 조건부 평균 대기시간의 축차관계식을 이용하여 평균 대기시간을 구하였다. 아울러, 흡수 연속 마코프사슬 이용하여 대기시간의 LST를 유도하고 평균 대기시간을 급수(series)의 형태로 나타내었다.

감사의 글

본 논문의 내용을 명확하게 설명할 수 있도록 건설적인 조언을 주신 심사위원께 감사드립니다.

참 고 문 헌

[1] Flatto, L. and S. Hahn, "Two Parallel queues created by arrivals with two demands I", SIAM J. Appl. Math., Vol.44, No.5, pp.1041-1053, 1984.

[2] Flatto L., "Two Parallel queues created by arrivals with two demands II", SIAM J. Appl. Math., Vol.45, No.5, pp.861-878, 1985.

[3] Kim, C. and A. K. Agrawala, "Analysis of the fork-join queue", IEEE Transactions on Computers, Vol.38, No.2, pp.250-255, 1989.

[4] Kumar, A. and R. Shorey, "Performance Analysis and Scheduling of Stochastic Fork-Join Jobs in a Multicomputer System", IEEE Transactions on Parallel and Distri-

buted Systems, Vol.4, No.10, pp.1147-1164, 1993.

[5] Liu, Y. C. and H. G. Perrors, "A Decomposition Procedure for the Analysis of a Closed Fork/Join Queueing Systems", IEEE Transactions on Computers, Vol.40, No.3, pp.365-370, 1991

[6] Narahari, Y. and P. Sundarrajan, "Performability Analysis of Fork-Join Queueing Systems", Journal of the Operational Research Society, Vol.46, No.10, pp.1237-1249, 1995.

[7] Nelson, R. and A. N. Tantawi, "Approximate analysis of fork/join synchronization in parallel queues", IEEE Transactions on Computers, Vol.37, No.6, pp.739-743, 1988.

[8] Nelson, R., D. Towsley and A. N. Tantawi, "Performance Analysis of Parallel Processing Systems", IEEE Transactions on Software Engineering, Vol.14, No.4, pp. 532-539, 1988.

[9] Neuts, M. F., Matrix-Geometric Solutions in Stochastic Models : An Algorithmic Approach, The Johns Hopkins University Press, Baltimore, 1981.

[10] Rao B. M. and M. J. M. Posner, "Algorithmic and Approximation Analysis of the Split and Match Queue", Commun. Statistics-Stochastic Models, Vol.1, No.3, pp.433-456, 1985.

[11] Rao B. M., "On the departure process of the split and match queue", Computers Opns. Res., Vol.17, No.6, pp.349-357, 1990.

[12] Setia, S. K., M. S. Squillante and S. K. Tripathi, "Analysis of Processor Allocation in Multiprogrammed, Distributed-Memory Parallel Processing Systems", IEEE Transactions on Parallel and Distributed Systems, Vol.5, No. 4, pp.401-420, 1994.

[13] Wolff, R. W., "Poisson Arrival See Time Averages", Operations Research, Vol.30, pp.223-231, 1982.

[14] Zhang, Z., "Analytical results for waiting and system size distributions in two parallel queueing systems", SIAM J. Appl. Math., Vol.50, No.4, pp.1176-1193, 1990.

96년 3월 최초 접수, 96년 6월 최종 수정

부록 : 식 (2)의 안정조건 증명

$\max(n_2,\cdots,n_m)=K$이면 고객들이 시스템 내로 입장하지 못하므로

$$P_b=P[\max(n_2,\cdots,n_m)=K]=1-P[n_2<K,\cdots,n_m<K]$$

$$=1-\sum_{n_1=0}^{\infty}\sum_{n_2=0}^{K-1}\cdots\sum_{n_m=0}^{K-1}x_{(n_1,\cdots,n_m)} \qquad \text{(A1)}$$

이다.

QBD 모형에서 채널 1을 떼어낸, 즉, 용량이 K인 채널 $2,\cdots,m$으로 구성된 분기 및 접합형태를 갖는 대기행렬 모형의 시스템 상태를 $(n_2,\cdots,n_m)$, 안정상태 확률을 $h_{(n_2,\cdots,n_m)}$이라 하자. 먼저, $h_{(n_2,\cdots,n_m)}$의 평형식(balance equation)은 다음과 같다.

$$[\lambda\,\delta(n_j\langle K,j=2,\cdots,m)+\Sigma_{j=2}^{m}\mu_j\,\delta(n_j\rangle 0)]\,h_{(n_2,\cdots,n_m)}$$

$$=\Sigma_{j=2}^{m}\mu_j\,\delta(n_j\langle K)h_{(n_2,\cdots,n_{j+1},\cdots,n_m)}+\lambda\,\delta(n_j\rangle 0,j=2,\cdots,m)h_{(n_2-1,\cdots,n_m-1)} \tag{A2}$$

그리고 $x_{(n_1,\cdots,n_m)}$의 평형식은 다음과 같다.

$$[\lambda\,\delta(n_j\langle K,j=2,\cdots,m)+\mu_1\,\delta(n_1\rangle 0)+\Sigma_{j=2}^{m}\mu_j\,\delta(n_j\rangle 0)]\,x_{(n_1,n_2,\cdots,n_m)}$$

$$=\mu_1 x_{(n_1+1,n_2,\cdots,n_m)}+\Sigma_{j=2}^{m}\mu_j\,\delta(n_j\langle K)x_{(n_1,\cdots,n_j+1,\cdots,n_m)}+\lambda\,\delta(n_j\rangle 0,j=1,\cdots,m)x_{(n_1-1,\cdots,n_m-1)} \tag{A3}$$

(A3)을 n_1에 대해 합하면 (A3) 좌변의 두 번째 항의 합이

$$\Sigma_{n_1=0}^{\infty}\delta(n_1\rangle 0)x_{(n_1,\cdots,n_m)}=\Sigma_{n_1=1}^{\infty}x_{(n_1,\cdots,n_m)}=\Sigma_{n_1=0}^{\infty}x_{(n_1+1,n_2,\cdots,n_m)}$$

이기 때문에, (A3) 우변의 첫 번째 항의 합과 소거되고, 이렇게 정리된 식을 (A2)와 비교하면

$$h_{(n_2,\cdots,n_m)}=\Sigma_{i=0}^{\infty}x_{(i,n_2,\cdots,n_m)} \tag{A4}$$

임을 쉽게 보일 수 있다.

QBD 모형은 $\pi A_+e\langle \pi A_e$이면 양의재귀적(positive recurrent)이다 [9]. 이때 π는 $A=A_+A_0+A_+$의 안정상태에서의 확률벡터이고 e는 원소들이 모두 1인 열백터(column vector)이다. 행렬 A를 계산하면 A는 용량이 K인 채널 $2,\cdots,m$으로 구성된 분기 및 접합형태를 갖는 대기행렬 모형의 무한소 생성자가 됨을 쉽게 알 수 있다. 그러므로 $\pi=(h_{(n_2,\cdots,n_m)}\mid n_j=0,\cdots,K)$이다. 이때, $h_{(n_2,\cdots,n_m)}$들은 $\boldsymbol{x}_i$의 원소들과 같은 순서로 정렬되어 있다. 먼저, $\pi A_e=\mu_1$이다. 그리고

$$\pi A_+e=\lambda\,\Sigma_{n_2=0}^{K}\cdots\Sigma_{n_m=0}^{K}h_{(n_2,\cdots,n_m)}\Pi_{j=2}^{K}\delta(n_j\langle K) \tag{A5}$$

이다. (A1)과 (A4)를 이용하여 (A5)를 정리하면 $\pi A_+e=\lambda(1-P_b)$를 얻는다.