

일반적인 큐잉네트워크에서의 체류시간분포의 근사화[†]

윤복식*

An approximation method for sojourn time distributions in general queueing networks[†]

Bok Sik Yoon*

Abstract

Even though sojourn time distributions are essential information in analyzing queueing networks, there are few methods to compute them accurately in non-product form queueing networks. In this study, we model the location process of a typical customer as a GMPH semi-Markov chain and develop computationally useful formula for the transition function and the first-passage time distribution in the GMPH semi-Markov chain. We use the formula to develop an efficient method for approximating sojourn time distributions in the non-product form queueing networks under quite general situation. We demonstrate its validity through numerical examples.

1. 서론

큐잉네트워크(queueing network=대기망)에서의 체류시간(sojourn time)은 교환기에서의 지연시간, 생산시스템에서의 지연시간등에 직접적으로 관련되는 중요한 량이다. 따라서 이것

의 분석을 위한 많은 연구가 행해져 왔는데 (Boxma(1990) 참조) 큐잉네트워크가 곱형태의 경우에도 특수한 조건(예를들면 추월불가(overtake-free)등, Warland and Varaiya (1980), Schassberger and Daduna(1987)등 참조)을 만족할 경우에만 체류시간분포에 관한 이론적인 결과를 얻을 수 있고 일반적인 경우

[†] 본 연구는 부분적으로 학술진흥재단의 공모과제 연구비(1994)에 의해 이루어 졌음.

* 홍익대학교 기초과학과

에는 근사적인 방법에 의존하게 된다. 더우기 곱형태가 아닌 큐잉네트워크에서는 불가피하게 근사적인 접근방법을 이용하여 체류시간분포를 구해야 하는데 지금까지 이것을 구하기 위한 체계적인 연구 결과는 매우 적은 실정이고 (Whitt(1983), Shanthikumar and Buzacott (1984), Shanthikumar and Sumita(1988)), 이들 또한 1, 2차 모멘트를 이용한 평균적인 분석에 초점이 맞춰져 체류시간의 분포는 매우 부정확하게 얻어진다. 실제로 교환기의 성능분석의 예를 들면 95 percentile과 같은 값들이 요구되어 정확한 분포를 얻을 필요가 있는데 이것을 시뮬레이션으로 해결하는 데에는 융통성과 모형 및 시간제약의 문제 때문에 한계가 있어 체류시간분포를 계산해 줄 수 있는 효과적인 방법이 매우 필요하다.

본 논문의 주 목적은 곱형태의 제약에서 벗어난 일반적인 대기망에서의 체류시간분포를 간편하게 계산하는 방법을 확립하는 것이다. 필요 이상의 복잡성을 피하기 위해 서비스 받는 고객은 모두 동일한 확률적인 특성을 갖고 (즉 단일 클래스), 각 노드에서의 서비스 시간은 모두 독립적이라고 가정한다. 각 노드에서의 도착과정은 독립적인 재생과정이고 각 노드는 단일 서버라고 가정한다(이 가정은 반드시 필요하지는 않다). 대체적인 접근방법은 전형적인 하나의 고객의 위치과정을 관찰하여 이를 세마이마코프 과정으로 근사화하고 각 노드에서의 체류시간을 GPH(윤복식(1994)) 분포로 추정하여 얻어진 GMPH(generalized matrix phase-type) 세마이마코프(Semi-Markov) 체인

(2장 참조)의 최초진입시간(first passage time) 분포를 구하는 방식을 따른다. GPH 분포를 양의 구간에서 정의되는 임의의 분포함수를 근사화 할 수 있는 매우 유용한 분포로 계산에 편리하게 이용될 수 있는 성질이 Shanthikumar (1985), 윤복식(1994)에서 검토된 바 있다.

본 논문의 서론에 이어 2장에서는 GMPH 세마이마코프 과정이 정의되고 특정상태에 최초진입시간 분포를 계산하는 과정이 확립된다. 3장에서는 2장의 결과를 대기망의 체류시간 분포에 적용하기 위하여 대기망에서의 전형적인 특정고객의 위치이동 과정이 세마이마코프 체인으로 근사화되고 GMPH 마코프 체인 모형을 적용하기 위하여 각 노드에서의 체류시간 분포를 GPH 분포로 추정하는 과정이 주어진다. 4장에서는 노드가 6개인 대기망을 예로 들어 시뮬레이션과의 비교를 통해 본 연구의 방법의 정확성을 검증한다.

2. GMPH 세마이마코프 과정

2.1. 세마이마코프 과정과 마코프 재생함수

세마이마코프(semi-Markov) 과정 $X=\{X(t), t \geq 0\}$ 의 상태공간이 $S=\{1, 2, \dots, M\}$ 라고 하고, 과정 X 의 전이 시점(transition epoch)을 T_n ($n=0, 1, 2, \dots$)(단, T_0)이라고 하면, 전이 확률 P_{ij} 는

$$\begin{cases} P_{ij} = P\{X(T_{n+1}) = j \mid X(T_n) = i\}, i, j \in S, i \neq j, \\ P_{ii} = 0, i \in S, \end{cases} \quad (1)$$

로, 전이 시간의 분포 H_{ij} 는

$$\begin{cases} H_{ij}(t) = P\{T_{n+1}-T_n \leq t \mid X(T_{n+1}) = j, X(T_n)=i\}, t > 0, i, j \in S, i \neq j, \\ H_{ij}(t) = 1, t \geq 0, i \in S \end{cases} \quad (2)$$

로 정의될 수 있고 초기분포가 주어지면 이들 파라미터로써 세마이마코프 과정을 확정할 수 있다. 이들로부터 세마이마코프 과정의 세마이마코프 커널(semi-Markov Kernel)과 마코프

재생함수(Markov renewal function)을 도출할 수 있는데 우선 프로세스 X의 semi-Markov Kernel R_{ij} 는

$$R_{ij}(t) = P\{T_{n+1}-T_n \leq t, X(T_{n+1}) = j \mid X(T_n)=i\}, t \geq 0, i, j \in S \quad (3)$$

로 정의되며

$$R_{ij}(t) = P_{ij}H_{ij}(t) \quad (4)$$

의 관계를 가짐을 알 수 있다. R(t)의 도함수 r(t)

$$r(t) = \frac{dR(t)}{dt}, \quad t \geq 0 \quad (5)$$

를 마코프재생 밀도함수(Markov renewal density)라고 부른다.

이때, 임의의 시간 t에서 상태 i에 머물러 있을 확률 $C_i(t)$ 는 semi-Markov Kernel R_{ij} 를 이용하여,

$$C_i(t) \equiv P\{T_{n+1}-T_n \leq t \mid X(T_n)=i\} = \sum_{j \in S} R_{ij}(t), t \geq 0, i \in S \quad (6)$$

와 같이 계산되어 진다. 또, $R^{(n)}(t)$ 과 $r^{(n)}$ 을 R과 r의 n차 matrix convolution 함수, 즉

$$R_{ij}^{(n)}(t) = \sum_{k \in S} \int_0^t R_{ik}(x)R_{kj}^{(n-1)}(t-x)dx, t \geq 0, n=2, 3, \dots, i, j \in S, \quad (7)$$

$$R^{(1)}(t) = R(t), t \geq 0,$$

와

$$\begin{aligned} r_{ij}^{(n)}(t) &= \sum_{k \in S} \int_0^t r_{ik}(x)r_{kj}^{(n-1)}(t-x)dx, \quad t \geq 0, n=2, 3, \dots, i, j \in S, \\ r^{(n)}(t) &= r(t), \quad t \geq 0. \end{aligned} \quad (8)$$

라 하면

$$r^{(n)}(t) = \frac{dR^{(n)}(t)}{dt}, t \geq 0, n=1, 2, \dots \quad (9)$$

의 관계를 쉽게 볼 수 있다.

Markov renewal function M(t)는

$$M(t) = \sum_{n=1}^{\infty} R^{(n)}(t), t \geq 0 \quad (10)$$

로 정의되고, 이것의 밀도함수 m(t)는

$$m(t) \equiv \frac{dM(t)}{dt} = \sum_{n=1}^{\infty} r^{(n)}(t), \quad t \geq 0 \quad (11)$$

가 된다.

위에서 정의한 함수는 하나의 semi-Markov Chain에서 일반적으로 규정되는 함수이고, matrix convolution를 통해 계산되어 진다. 이때, 전이 시간의 분포함수 H_{ij} 를 GPH 분포(윤복식의(1994))로 변환시키면, 다음 절에서 보듯이 matrix convolution에 포함된 적분의 계산 없이 위에서 정의한 여러 함수를 계산해 낼 수 있다.

2.2. GMPH 세마이마코프 과정

$$\begin{aligned}
 &H_{ij}, ij=1, 2, \dots \text{가} \\
 &H_{ij}(t) = \sum_{n=0}^{\infty} \hat{H}_{ij}(n) e^{-\lambda t} \frac{(\lambda t)^n}{n!}, ij \in S, ij \in S
 \end{aligned} \tag{12}$$

의 $GPH(\lambda, \hat{H})$ 함수의 형태를 따른다고 가정하면, semi-Markov Kernel $R(t)$ 는

$$R(t) = \sum_{n=0}^{\infty} \hat{R}(n) e^{-\lambda t} \frac{(\lambda t)^n}{n!}, t \geq 0, \tag{13}$$

단, $\hat{R}_{ij}(n) = P_{ij} \hat{H}_{ij}(n), n=0, 1, 2, \dots$

와 같이 계산된다. 그리고, 식(13)을 미분하면 도함수 $r(t)$ 는

$$r(t) = \frac{dR(t)}{dt} = \sum_{n=0}^{\infty} \hat{r}(n+1) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda, t \geq 0 \tag{14}$$

여기서, $\hat{r}(0) = 0,$

$$\hat{r}(n) = \hat{R}(n) - \hat{R}(n-1), n=1, 2, \dots,$$

가 된다. 여기서, (13)의 R 처럼 matrix의 각 component 함수가 GPH 분포를 따를때, 이 matrix 함수를 Generalized Matrix Phase Type(GMPH)이라고 부르고, $GMPH(\lambda, \hat{R})$ 로

표현하기로 한다. matrix 함수 $R(t)$ 가 GPH이면, 식 (8)의 $r^{(k)}(t)$ 를 다음과 같이 GMPH 형태로 나타낼 수 있다.

[정리 2.1]

$R(t)$ 가 $GMPH(\lambda, \hat{R})$ 를 따르면,

$$r^{(k)}(t) = \sum_{n=0}^{\infty} \hat{r}^{(k)}(n+1) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda, k=2, 3, \dots \tag{15}$$

$$\text{여기서, } \hat{r}^{(k)}(n) = \sum_{v=0}^n \hat{r}^{(k-1)}(n-v) \hat{r}(v), k=2, 3, \dots \tag{16}$$

(증명)

먼저, $r^{(2)}(t)$ 를 계산하자. (14)를 (8)에 대입하면,

$$\begin{aligned}
 r_{ij}^{(2)}(t) &\equiv \sum_{k \in S} \int_0^t r_{ik}(x) r_{kj}(t-x) dx \\
 &= \sum_{k \in S} \int_0^t \left(\sum_{v=0}^{\infty} \hat{r}_{ik}(v+1) \frac{e^{-\lambda x} (\lambda x)^v}{v!} \lambda \right) \left(\sum_{u=0}^{\infty} \hat{r}_{kj}(u+1) \frac{e^{-\lambda(t-x)} (\lambda(t-x))^u}{u!} \lambda \right) dx \\
 &= \sum_{k \in S} \sum_{v=0}^{\infty} \sum_{u=0}^{\infty} \hat{r}_{ik}(v+1) \hat{r}_{kj}(u+1) e^{-\lambda t} \lambda^{v+u+2} \frac{1}{v!u!} \int_0^t x^v (t-x)^u dx.
 \end{aligned}$$

이 된다. 이 때,

$$\int_0^t x^v (t-x)^u dx = t^{v+u+1} \frac{v!u!}{(v+u+1)!} \tag{17}$$

의 결과를 이용하고, $n=v+u+1$ 와 $z=v+1$ 로 두면 결국,

$$\begin{aligned}
 r_{ij}^{(2)}(t) &= \sum_{k \in S} \sum_{v=0}^{\infty} \sum_{u=0}^{\infty} \hat{r}_{ik}(v+1) \hat{r}_{kj}(u+1) \frac{e^{-\lambda t} (\lambda t)^{v+u+1}}{(v+u+1)!} \lambda \\
 &= \sum_{k \in S} \sum_{n=1}^{\infty} \sum_{z=1}^n \hat{r}_{ik}(z) \hat{r}_{kj}(n+1-z) \frac{e^{-\lambda t} (\lambda t)^{n+1}}{n!} \lambda
 \end{aligned} \tag{18}$$

가 된다. 또한 $r(0)=0$ 이므로,

$$\sum_{n=1}^{\infty} \sum_{v=1}^n \hat{r}_{ik}(v) \hat{r}_{kj}(n+1-v) = \sum_{n=0}^{\infty} \sum_{v=0}^{n+1} \hat{r}_{ik}(v) \hat{r}_{kj}(n+1-v) \tag{19}$$

를 얻고 (19)를 (18)에 대입하여 행렬의 곱으로 표현하면

$$\begin{aligned}
 r^{(2)}(t) &= \sum_{n=0}^{\infty} \sum_{v=0}^{n+1} \hat{r}(v) \hat{r}(n+1-v) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda \\
 &= \sum_{n=0}^{\infty} \hat{r}^{(2)}(n+1) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda
 \end{aligned}$$

를 얻는다. 이와 같은 계산을 반복하여, 결국 정리의 결과를 얻는다. ■

[정리 2.1]를 이용하면, Markov Renewal 함수 $M(t)$ 와 Markov Renewal 밀도 함수 $m(t)$ 를 다음의 결과에 GMPH 함수로 구해낼 수 있다.

[정리 2.2]

$$(i) m(t) = \sum_{n=0}^{\infty} \hat{m}(n+1) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda, \quad t \geq 0, \tag{20}$$

$$(ii) \hat{M}(t) = \sum_{n=1}^{\infty} \hat{M}(n) \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad t \geq 0, \tag{21}$$

단, $\hat{m}(0) = I, \hat{m}(n) = \sum_{v=0}^{n-1} \hat{m}(v) \hat{r}(n-v),$
 $n = 1, 2, \dots$ (22)

$$\hat{M}(n) = \sum_{v=1}^n \hat{m}(v), \quad n = 1, 2, \dots \tag{23}$$

(증명)

(i) (15)를 (11)에 대입한 다음 변수 변환을 하면, $t \geq 0$ 에 대해

$$\begin{aligned}
 m(t) &= \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} \hat{r}^{(k)}(n+1) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda \\
 &= \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} \hat{r}^{(k)}(n+1) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \lambda
 \end{aligned} \tag{24}$$

가 된다. 그리고,

$$\hat{r}^{(0)}(0) = I, \hat{r}^{(0)}(n) = 0, \quad n = 1, 2, \dots \tag{25}$$

로 두고, $\hat{m}(v)$ 를 다음과 같이 정의하면,

$$\hat{m}(v) = \sum_{k=1}^{\infty} \hat{r}^{(k-1)}(v), \quad v = 0, 1, 2, \dots \tag{26}$$

$$\hat{m}(0) = I$$

가 되고,

$$\hat{r}^{(0)}(n) = 0, \quad n = 1, 2, \dots$$

이므로, 결국

$$\begin{aligned}
 \hat{m}(n) &= \sum_{k=1}^{\infty} \hat{r}^{(k)}(n) \\
 &= \sum_{v=0}^n \sum_{k=1}^{\infty} \hat{r}^{(k-1)}(v) \hat{r}(n-v) \\
 &= \sum_{v=0}^n \hat{m}(v) \hat{r}(n-v) \\
 &= \sum_{v=0}^{n-1} \hat{m}(v) \hat{r}(n-v), \quad \text{for } n = 1, 2, \dots
 \end{aligned}$$

을 얻게 된다.

따라서 (20)과 (22)가 구해진다.

$$(ii) M(t) = \int_0^t m(x) dx, \quad t \geq 0 \text{이므로,} \tag{27}$$

(20)을 (27)에 대입하면,

$$\begin{aligned}
 M(t) &= \sum_{v=0}^{\infty} \hat{m}(v+1) \int_0^t e^{-\lambda x} \frac{(\lambda x)^v}{v!} \lambda dx \\
 &= \sum_{v=0}^{\infty} \hat{m}(v+1) \sum_{n=v+1}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!},
 \end{aligned}$$

이 되고, 변수 변환을 하면 $t \geq 0$ 에 대해

$$M(t) = \sum_{n=1}^{\infty} \left(\sum_{v=1}^n \hat{m}(v) \right) e^{-\lambda t} \frac{(\lambda t)^n}{n!},$$

가 된다. 따라서, (21)과 (23)도 구해진다. ■

2.3. 전이함수(transition function) 및 체류시간 분포의 계산

먼저, 프로세스 X의 상태 전이확률(Transition Probability) $Q_{ij}(t)$ 를

$$Q_{ij}(t) = P\{X(t)=j | X(0)=i\}, t \geq 0, i, j \in S, \quad (28)$$

와 같이 정의하자.

그러면, 다음과 같은 matrix convolution의 형태로 표현된다[Cinlar(1975), p337].

$$Q_{ij}(t) = \int_0^t m_{ij}(x) \bar{C}_j(t-x) dx + \delta_{ij} \bar{C}_j(t), \quad t \geq 0, i, j \in S, \quad (29)$$

여기서, $\bar{C}_j(t) = 1 - C_j(t), t \geq 0, j \in S \quad (30)$

(29)에서 Markov Renewal 함수 R_{ij} 이 GM PH(λ, \hat{R})이면, $\bar{C}(t)$ 가

$$C_j(t) = 1 - \sum_{k \in S} R_{jk}(t) = 1 - \sum_{k \in S} \sum_{n=0}^{\infty} \hat{R}_{jk}(n) \frac{e^{-\lambda t} (\lambda t)^n}{n!}$$

로 계산된다. 여기서,

$$\hat{D}_j(n) = 1 - \sum_{k \in S} \hat{R}_{jk}(n), n=0, 1, 2, \dots \quad (31)$$

로 두면, 다시

$$\bar{C}_j(t) = \sum_{n=0}^{\infty} \hat{D}_j(n) \frac{e^{-\lambda t} (\lambda t)^n}{n!}, j \in S \quad (32)$$

로 표현된다. 여기서 계산된 $\bar{C}(t)$ 와 (20), (22)를 이용하면, 다음의 정리에 의해 $Q(t)$ 가 계산된다.

[정리 2. 3]

R_{ij} 이 GMPH(λ, d)이면,

$$Q(t) = \sum_{n=0}^{\infty} q(n) \frac{e^{-\lambda t} (\lambda t)^n}{n!}, t \geq 0, \quad (33)$$

여기서, $q(0) = I, q(n) = \sum_{k=0}^n \hat{m}(k) \hat{d}(n-k), n=1, 2, \dots \quad (34)$

$$\hat{d}(n) = I - \text{diag}(\hat{R}(n)1), n=0, 1, 2, \dots \quad (35)$$

(증명)

(20)과 (32)를 적분식 (29)에 대입하면,

$$\begin{aligned} \int_0^t \hat{m}_{ij}(x) C_j(t-x) dx &= \int_0^t \sum_{k=0}^{\infty} \hat{m}_{ij}(k+1) \frac{e^{-\lambda x} (\lambda x)^k}{k!} \lambda \sum_{v=0}^{\infty} \hat{D}_j(v) \frac{e^{-\lambda(t-x)} (\lambda(t-x))^v}{v!} dx \\ &= \sum_{k=0}^{\infty} \sum_{v=0}^{\infty} \hat{m}_{ij}(k+1) \hat{D}_j(v) e^{-\lambda t} \lambda^{k+v+1} \int_0^t \frac{x^k (t-x)^v}{k! v!} dx \\ &= \sum_{k=0}^{\infty} \sum_{v=0}^{\infty} \hat{m}_{ij}(k+1) \hat{D}_j(v) \frac{e^{-\lambda t} (\lambda t)^{k+v+1}}{(k+v+1)!} \\ &= \sum_{n=1}^{\infty} \sum_{k=1}^n \hat{m}_{ij}(k) \hat{D}_j(n-k) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \end{aligned} \quad (36)$$

가 된다. 그리고, (36)과 (32)를 (29)에 대입하면,

$$\begin{aligned} Q_{ij}(t) &= \sum_{n=1}^{\infty} \sum_{k=1}^n \hat{m}_{ij}(k) \hat{D}_j(n-k) \frac{e^{-\lambda t} (\lambda t)^n}{n!} + \delta_{ij} \sum_{n=0}^{\infty} \hat{D}_j(n) \frac{e^{-\lambda t} (\lambda t)^n}{n!} \\ &= \delta_{ij} e^{-\lambda t} + \sum_{n=1}^{\infty} \frac{e^{-\lambda t} (\lambda t)^n}{n!} (\delta_{ij} \hat{D}_j(n) + \sum_{k=1}^n \hat{m}_{ij}(k) \hat{D}_j(n-k)), \end{aligned}$$

가 된다.

$\hat{m}(0)=I$ 이고, $j \in S$ 에 대해 $\hat{D}_j(0)=1$ 이므로 (36)은 결국

$$Q_{ij}(t) = \sum_{n=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^n}{n!} \sum_{k=0}^n \hat{m}_{ij(k)} \hat{D}_j(n-k), \quad t \geq 0, \quad i, j \in S. \quad (37)$$

가 된다. 따라서, (34)와 (35)를 이용하여 GM PH 함수 (33)이 계산된다.

상태 집합 A에 속하는 한 상태 i 에서 상태 집합 B에 속하는 임의의 노드로의 첫번째 전이하는데 소요되는 시간을 T_{iB} 라 하면,

$$T_{iB} = \inf\{t : X(t) \in B | X(0) = i\}, \quad i \in A$$

로 표현된다. 이 T_{iB} 는 위의 [정리 2.3]을 이용하여 구해낼 수 있다.

상태 집합 A에 속하는 상태에서만 전이를

하는 프로세스 X_A 를 정의하고, 이 프로세스의 Kernel을 R_{AA} 라고 하자. 즉, R_{AA} 는 원래 주어진 함수 R 에서 A집합의 원소 상태만을 따로 떼어 정의한 함수이다. 그러면, $P\{T_{iB} > t\}$ 는

$$P\{T_{iB} > t\} = (Q_A(t)1)_i \quad (38)$$

단, $Q_A(t)$ 는 X_A 의 전이 확률 가 된다. [정리 2.3]을 이용하여, 결과를 정리하면, 최종적으로 다음의 결과가 나온다.

[정리 2.4]

R 이 GMPH(λ, d)라면,

$$P\{T_{iB} > t\} = \left(\sum_{n=0}^{\infty} q_A(n) I \frac{e^{-\lambda t} (\lambda t)^n}{n!} \right)_i, \quad (39)$$

가 된다. 단,

$$q_A(0) = I, \quad q_A(n) = \sum_{k=0}^n \hat{m}_A(k) \hat{d}_A(n-k), \quad n = 1, 2, \dots \quad (40)$$

$$\hat{d}_A(n) = I - \text{diag}(\hat{R}(n)I), \quad n = 0, 1, 2, \dots \quad (41)$$

$$\hat{m}_A(0) = I, \quad \hat{m}_A(n) = \sum_{k=0}^{n-1} \hat{m}_A(k) \hat{r}_A(n-k), \quad n = 1, 2, \dots \quad (42)$$

[정리 2.4]에는 이산함수의 덧셈 연산만이 포함되어 있기 때문에 수치적 계산이 쉽게 된다.

3. 큐잉네트워크에서의 체류시간 분포

일반적인 상황하에서의 큐잉 네트워크에서의 지연시간 분포를 구하기 위해 Shanthikumar and Buzacott(1984)에서와 같이 전형적인 고객의 경

로이동을 세마이마코프 프로세스로 Z 로 근사화 하여 $\hat{T} = \inf\{t : \hat{Z}(t) = \Delta\}$ 의 분포를 구하는 것을 목표로 한다. 결국 \hat{Z} 의 체류시간 분포를 구해야 하는데 이것은

(1) 세마이마코프 프로세스를 결정하는 각 개별 노드에서의 체류시간의 분포를 구한다.

(2) 개별노드의 체류시간 분포들로부터 세마이마코프 프로세스의 Δ 상태로의 최초진입시간 (first passage time)의 분포를 구한다.

와 같은 과정으로부터 구해낼 수 있다.

위의 과정을 일반적인 도착분포에 적용시키기 위해 본 장에서는 GPH 분포를 이용한 근사화 과정을 확립한다. 3.2절에서 (1)의 과정이 상세

히 언급되고 (2)의 과정은 2.3절에서 이미 구체적으로 다루어졌다.

3.1. 위치이동 과정의 세마이마코프 체인으로의 근사화

본 절에서는 네트워크에서의 전형적인 고객의 위치이동과정을 세마이마코프 체인(semi-Markov chain)으로 근사화시켜서 총 지연시간을 계산한다. 먼저, $Z(t)$ 를 시간 t 에서 특정한 고객이 있는 노드라고 하자. 즉 시간 t 에 고객이 i 노드에 있으면, $Z(t)$ 는 i 값을 가진다. 고객이 네트워크에 처음 들어온 시간을 0 으로 가정하고, 노드 번호를 $i=1, 2, \dots, M$ 이라하고, 고객이 네트워크를 떠난 상태를 노드 Δ 로 표시하자. 이 때, $P\{Z(0)=i\}=a_i$, ($i=1, 2, \dots, M$, a_i 는 고객이 i 노드에 도착할 확률)이라면, 고객의 총 네트워크 체류시간 T 는 다음과 같이 나타낼 수 있다.

$$T = \inf\{t : Z(t) = \Delta\}$$

T 의 분포를 구하는 것은 매우 어려우므로, 원래의 프로세스 $Z(t)$ 를 세마이마코프 프로세스 $\hat{Z}(t)$ 로 근사화시킨다. $\hat{Z}(t)$ 는 원래의 네트워크에서 전이가 마코프 성질을 따르고 각 노드를 방문할 때마다 그 노드에서의 체류시간이 상호독립적이고 동일한 분포를 따른다는 가정하면 얻어진다.

H_i 를 임의의 고객이 i 노드에서 머무르는 지연시간의 안정상태 분포라고 하자. P_Δ 를 $Z(t)$ 에 내재된(embedded) 이산시간 마코프 체인(embedded process) 전이 확률 행렬이라 하면

$$P_\Delta = \begin{bmatrix} P & e - Pe \\ 0 & 1 \end{bmatrix}$$

여기서 $e=(1, 1, \dots, 1)^t$, $0=(0, 0, \dots, 0)$ 는 M 벡터이다.

과 같이 표현될 수 있다. 그러면, \hat{Z} 는 상태 공간 $S=\{1, 2, \dots, M, \Delta\}$ 와 노드 i 에 있을 초기 확률 a_i , $i=1, 2, \dots, M$, 그리고 세마이 마코프 커널(kernel) $R_\Delta = D_H P_\Delta (D_H$ 는 대각 원소 $d_{ii}=H_i$, $d_{ij}=0$ 를 가지는 대각 행렬)를 가진다.

3.2. 개별노드에서의 입력트래픽 모델링

3.2.1. 노드별 도착 트래픽의 파라미터 계산

N 개의 노드로 구성되는 열린 큐잉네트워크에서 각 노드에서의 도착간격분포와 서비스 분포의 평균과 변동계수(squared coefficient of variation)가 주어져 있다고 하자. 본절에서는 각 노드에 도착하는 총 트래픽의 과정을 이 두 개의 파라미터로 추정하는 과정을 Whitt (1983)과 유사하게 확립한다. N 개의 노드에서, 노드 i 로의 외부 도착율이 λ_{0i} , 외부 도착 트래픽의 변동계수가 c_{0i} 이고, 각 노드의 서비스율이 μ_i , 서비스 시간의 변동계수가 c_{si} , 노드별 전이 확률이 q_{ij} 로 주어져 있을 때, 다음의 식으로부터 각 노드별 도착율 λ_j 가 구해진다.

$$\lambda_j = \lambda_{0j} + \sum_{i=1}^N \lambda_i q_{ij}, \quad j=1, 2, \dots, N \quad (43)$$

이를 이용하여 필요한 파라미터들은 다음과 같이 구할 수 있다.

$\lambda_{ij} = \lambda_i q_{ij}$: 노드 i 에서 노드 j 로의 도착율

$P_{ij} = \lambda_{ij} / \lambda_j$: j 노드로 도착하는 트래픽중에서 i 에서 온 트래픽의 비율

$\rho_i = \lambda_i / \mu_i$: i 노드의 트래픽 부하량

이 파라미터값을 이용하여 다음의 선형연립방정식을 풀어 각 노드별 도착 트래픽의 변동계수 c_{ij} 를 구해낸다.

$$c_{ij}^2 = a_j + \sum_{i=1}^N c_{ii}^2 b_{ij}, \quad j=1, 2, \dots, N \quad (44)$$

여기서,

$$a_j = I + w_j \{ (p_{0j} c_{0j}^2 - I) + \sum_{i=1}^N p_{ij} [(I - q_{ij}) + q_{ij} \rho_i^2 x_i] \} \quad (45)$$

$$b_{ij} = w_j p_{ij} q_{ij} [I - \rho_i^2] \quad (46)$$

$$x_i = 1 + (\max\{c_{in}^2, 0.2\} - 1) \quad (47)$$

$$w_j = [1 + 4(1 - \rho_j)^2 (v_j - 1)]^{-1} \quad (48)$$

$$v_j = \left[\sum_{i=1}^N p_{ij}^2 \right]^{-1} \quad (49)$$

식 (43)과 (44)의 연립방정식 들을 풀어서 구한 각 노드별 도착율과 변동 계수를 이용하여 각 노드별 도착 트래픽의 간격시간의 평균과 분산을 구해낼 수 있다.

3.2.2. 도착 분포의 GPH 분포로의 근사화

임의의 분포가 평균과 분산만 주어져 있다고 하자. 이 임의의 분포를 GPH 분포로 근사화시켜보자. 다음의 임의의 GPH 분포가 주어져 있다고 하자.

$$X \sim GPH(\lambda, g)$$

$$\text{단, } L \sim G(n)$$

그러면,

$$X = \sum_{n=1}^L E_n (\text{단, } E_n \sim \exp(\lambda)) \quad (50)$$

이다. 따라서,

$$E[X] = E[L]E[E_n] = \frac{E[L]}{\lambda} \quad (51)$$

$$\begin{aligned} V[X] &= E[L]V[E_n] + E[E_n]^2V[L] \\ &= E[L]/\lambda^2 + V[L]\lambda^2 \\ &= (E[L] + V[L])/\lambda^2 \end{aligned} \quad (52)$$

가 된다.

결국 임의의 분포의 평균이 m , 분산이 v 로 주어져 있으면 이 분포를 GPH 분포로 근사화시키면,

$$m = E[L]/\lambda \quad (53)$$

$$v = (E[L] + V[L])/\lambda^2 \quad (54)$$

가 된다. 식 (53) (54)에 의해 $E[L]$ 과 $V[L]$ 의 값이 결정될 수 있어서, $E[L]$ 과 $V[L]$ 를 이용하여 L 의 분포만 근사화시키면, 임의의 분포가 GPH 분포로 근사화 될 수 있다. 본 논문에서는 잠정적으로 L 의 분포로 Negative binomial 분포를 선택하였는데, Negative binomial 분포는 두 개의 파라미터로써 결정이 되고, 변동계수의 값이 1보다 클수도, 작을 수도 있기 때문에 두개의 파라미터로써 근사화할 수 있는 이산분포로서 무리가 없을 것이다.

Negative binomial 분포는 두 개의 파라미터 n 과 p 로써 결정이 되고, 확률밀도함수는

$$P_r(X=k) = \binom{k+n-1}{n-1} p^n (1-p)^k, \quad k=0, 1, \dots \quad (57)$$

이다. (57)과 같이 주어진 n 과 p 의 파라미터를 갖는 Negative binomial 분포의 평균과 분산은 각각

$$E(X) = \frac{n}{p} - n \quad (58)$$

$$V(X) = \frac{n}{p^2} (1-p) \quad (59)$$

가 된다. 역으로 $E(X)$ 와 $V(X)$ 를 알고 있을 때의 n 과 p 는 각각

$$n = \frac{E(X)^2}{V(X) - E(X)} \quad (60)$$

$$p = \frac{E(X)}{V(X)} \quad (61)$$

가 된다. 그런데, n 은 항상 정수값은 가져야 하므로 식(60)을 반올림하여 n 을 결정한다. 그리고, (58)에 n 을 대입하여 다시 p 를 구해낸다. 이렇게 하면, 식 (59)의 분산에 약간의 오차가 생기지만, 식 (58)의 평균은 만족하는 적당한 n 과 p 가 결정된다. 결국 n 과 p 가 결정이 되면서, 평균과 분산이 주어진 미지의 이산확률분포가 Negative binomial 분포로 근사화 될 수 있다.

3.3. 전체적인 계산과정

<계산과정>

0. 큐잉네트워크 결정을 위한 입력 자료

본 연구에서, 하나의 Queueing network를 정의하는 데 기본적으로 다음의 입력 자료가 필요하다.

- 네트워크 노드수
- 전이확률
- 외부 도착 트래픽
(노드별 외부 도착률 λ_{0i} , 노드별 외부 도착간격시간의 변동계수 c_{0i}^2)
- 서비스 분포
(서비스율 μ_i , 서비스 시간의 변동계수 c_{si}^2 , 서비스 시간의 분포함수)

1. 각 노드에서의 체류시간 분포의 근사화 : GPH/GPH/1 queue에서의 시스템 체류시간 분포

1.1. 노드별 도착 트래픽의 파라미터 계산

3.2.1절의 식(1)-(7)을 이용하여 산출한 각 노드별 도착율과 변동 계수 이용하여 각 노드

별 도착 트래픽의 간격시간의 평균과 분산을 구한다.

1.2. 도착간격분포의 GPH 분포로의 근사화

3.2.2절의 방법 또는 보다 직접적인 방법(윤복식의(1994) 참조)을 이용하여 각노드에서의 트래픽 도착간격 분포를 GPH 분포로 근사화 한다.

1.3. 서비스 분포의 GPH로의 모델링

일반 분포로 주어진 서비스 분포를 GPH 분포로 근사 변환을 시킨다.

1.4. 각 노드별 대기시간 계산

GPH/GPH/1 Queue로 decomposition된 각 노드의 시스템 대기시간의 분포를 계산한다 (윤복식의(1994) 참조).

2. GPH semi-Markov chain에서의 first passage time 계산

각 노드에서의 대기시간의 분포와 전이확률을 이용하여 주어진 Queueing Network를 semi-Markov Chain으로 모형화 한다. 여기서, 주어진 노드에서의 다른 노드로의 전이하는 데 소요되는 시간은 위에서 구한 대기시간의 분포 H_i 로 대체한다. GPH 분포를 따르는 전이시간 분포 H_i 와 전이확률 p_{ij} 를 이용하여 임의의 노드에서 출발한 고객이 Δ 노드에 진입할 때까지의 시간의 분포를 2장의 결과를 이용하여 계산해 낸다.

4. 적용예제 및 타당성 검증

4.1. 간단한 큐잉네트워크 예제

본절에서는 체류시간 근사화의 정확도를 검

〈표 3〉 노드별 전이확률(비대칭)

	1	2	3	4	5	6	out
1		0.15	0.07	0.05	0.18	0.25	0.30
2	0.05		0.30	0.18	0.14	0.13	0.20
3	0.09	0.12		0.28	0.08	0.18	0.25
4	0.10	0.05	0.19		0.25	0.23	0.18
5	0.30	0.13	0.20	0.11		0.06	0.20
6	0.17	0.28	0.05	0.14	0.11		0.25

■ 트래픽 밀도

트래픽 밀도가 근사화의 정도에 미치는 영향을 실험하기 위해서 〈표 4〉와 같이 각 노드의 트래픽 밀도(ρ)를 0.4, 0.6, 0.8로 변화시키면서 실험을 수행하였다. 이 세가지 파라미터의 변화에 의해 총 12가지의 근사화 예제가 만들어

졌다.

〈표 4〉 노드의 트래픽 밀도

	경우 1	경우 2	경우 3
노드별 트래픽 밀도(ρ)	0.4	0.6	0.8

■ 각 노드의 서비스 분포

시뮬레이션 결과와 비교를 위해 SIMAN에서 지원 가능한 분포함수에 맞추어 지수 분포와

Erlang-2 분포를 사용하였다. 구체적인 서비스 분포는 〈표 5〉, 〈표 6〉과 같다.

〈표 5〉 노드별 서비스 분포(전이확률이 대칭일 때)

노드별 외부 도착	트래픽 로드(ρ)	노드 1	노드 2	노드 3	노드 4	노드 5	노드 6
대칭	0.4	exp(2.5)	er2(5.0)	exp(2.5)	er2(5.0)	exp(2.5)	er2(5.0)
	0.6	exp(1.67)	er2(3.33)	exp(1.67)	er2(3.33)	exp(1.67)	er2(3.33)
	0.8	exp(1.25)	er2(2.5)	exp(1.25)	er2(2.5)	exp(1.25)	er2(2.5)
비대칭	0.4	exp(2.33)	er2(4.46)	exp(2.1)	er2(5.1)	exp(2.33)	er2(4.46)
	0.6	exp(1.55)	er2(2.96)	exp(1.4)	er2(3.4)	exp(1.55)	er2(2.96)
	0.8	exp(1.16)	er2(2.22)	exp(1.05)	er2(2.56)	exp(1.16)	er2(2.22)

〈표 6〉 노드별 서비스 분포(전이확률이 비대칭일 때)

노드별 외부 도착	트래픽 로드(ρ)	노드 1	노드 2	노드 3	노드 4	노드 5	노드 6
대칭	0.4	exp(2.05)	er2(4.26)	exp(2.23)	er2(4.36)	exp(2.15)	er2(4.66)
	0.6	exp(1.37)	er2(2.84)	exp(1.48)	er2(2.90)	exp(1.43)	er2(3.10)
	0.8	exp(1.03)	er2(2.12)	exp(1.11)	er2(2.18)	exp(1.08)	er2(2.32)
비대칭	0.4	exp(1.95)	er2(3.70)	exp(1.88)	er2(4.44)	exp(2.05)	er2(4.16)
	0.6	exp(1.33)	er2(2.46)	exp(1.25)	er2(2.96)	exp(1.37)	er2(2.76)
	0.8	exp(0.98)	er2(1.86)	exp(0.94)	er2(2.22)	exp(1.03)	er2(2.08)

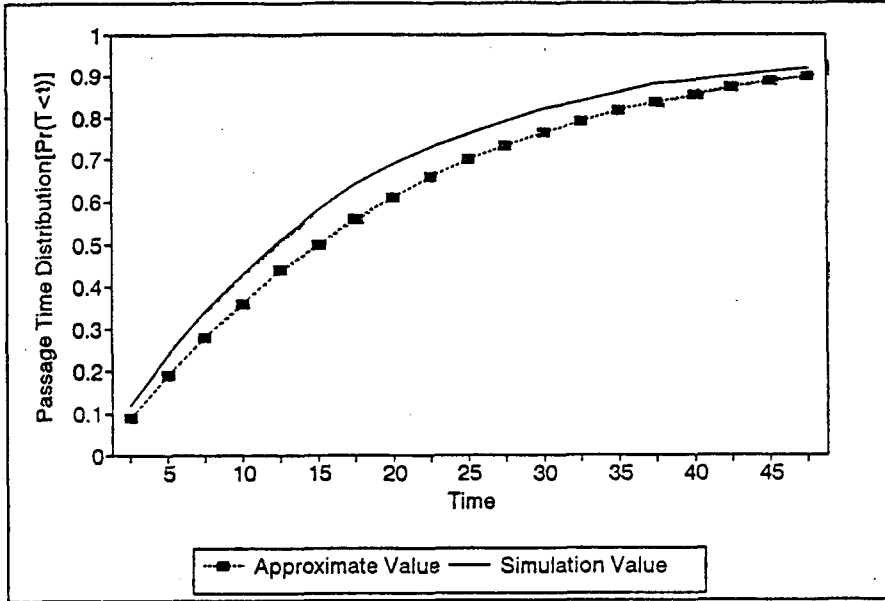
4.2. 시뮬레이션과의 비교

본절에서는 위의 여러가지 경우에 해석적 기법으로 구한 분포와 시뮬레이션을 통해 구한 분포와 비교가 이루어진다. 시뮬레이션 패키지는 SIMAN을 사용하였다. 아래의 그림에서 제목은 편의상

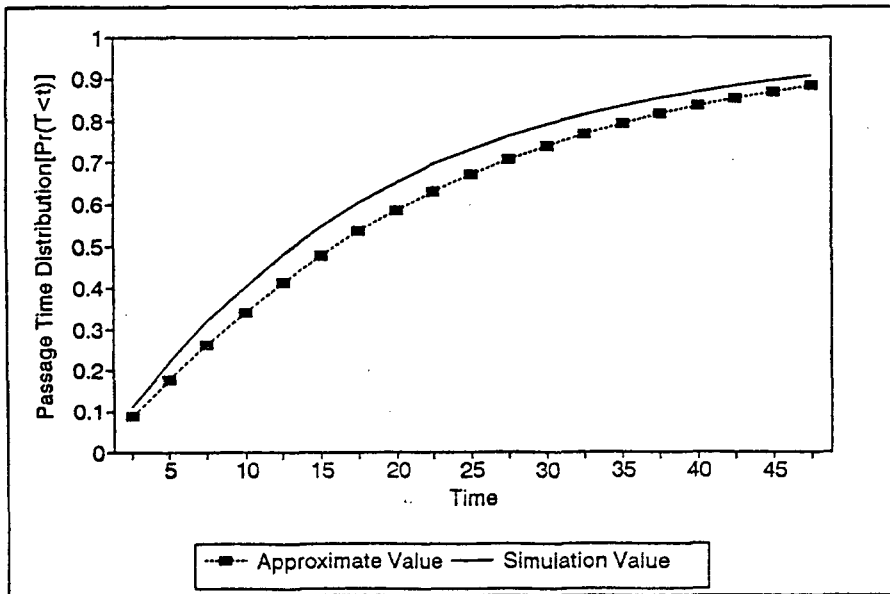
(전이확률 형태(대칭, 비대칭)-외부도착률 형태(대칭, 비대칭)-트래픽 로드(ρ))

의 형태로 주어진다. [그림 1]-[그림 4]에서 보듯이 체류시간 분포가 시뮬레이션 결과와 약

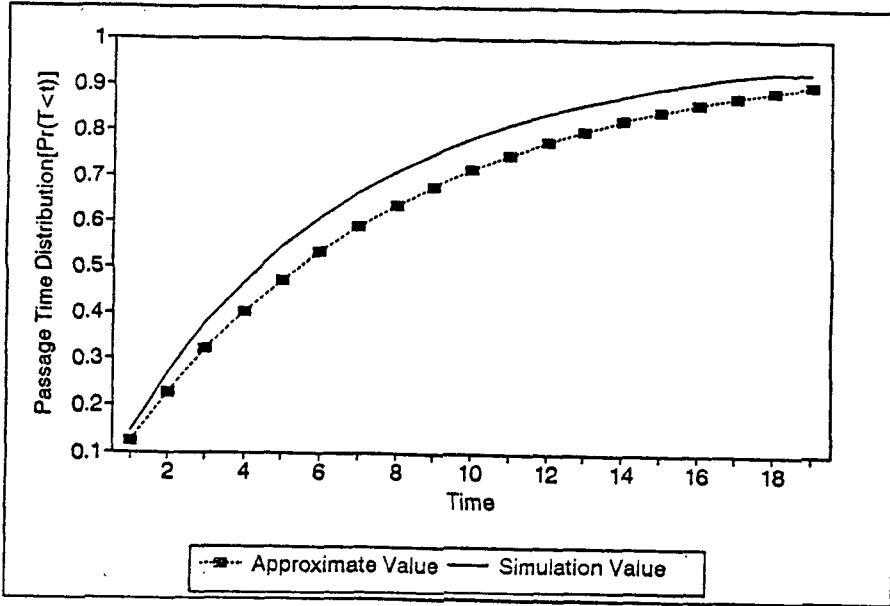
간의 차이를 보이는데 이것은 주로 PC를 사용하는데 따른 시스템의 기억용량(RAM)의 제한에서 비롯된다. 즉 기억용량의 한도를 맞추기 위해 전체 모형의 기본이 되는 GPH 근사화에서 λ 의 값을 12-15 정도로 작게 잡아 주기 때문에 입력 및 서비스 분포의 근사화가 비교적 부정확하게 이루어지는 것이 주 요인이다. 대용량 컴퓨터나 보다 기억용량이 확장성이 용이한 UNIX 시스템에서는 쉽게 정확도를 높일 수 있을 것이다. 다음절에서 정확성을 향상시키기 위한 방법들이 다양하게 언급된다.



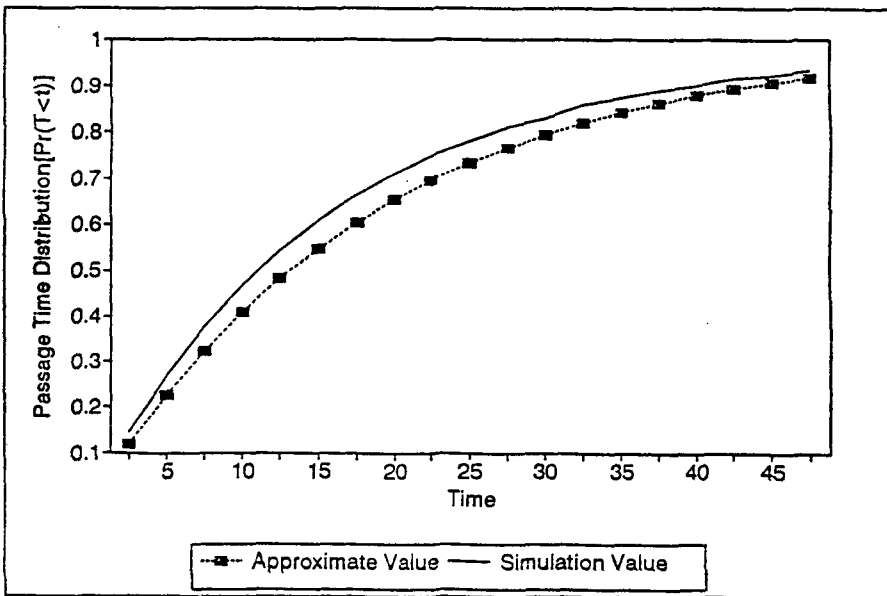
[그림 1] (대칭-대칭-0.8)의 경우



[그림 2] (대칭-비대칭-0.8)의 경우



[그림 3] (비대칭-대칭 -0.6)의 경우



[그림 4] (비대칭-대칭 -0.8)의 경우

4.3. 오차의 원인 및 해결 방법

본절에서는 3장에 주어진 전체적인 분포 근사화 과정에서 발생하는 오차의 주 원인을 분석하고 해결 방법에 대해 언급한다.

(1) 네트워크 각 노드에서의 이동 트래픽 모델링 과정상의 에러

-네트워크 내부에서 이동하고 있는 traffic의 특성을 평균과 분산으로 요약하고 이것들 사이의 선형적인 관계식을 구하는 과정에서의 단순화에 의한 오차: 이것은 노드간 트래픽들의 동적인 관계를 보다 정확히 표현하는 방법이 필요하나 매우 계산이 복잡해질 것이다.

-각 노드별 input traffic의 평균과 분산으로부터 GPH 분포로 근사화 과정에서 사용한 negative binomial 분포의 원래 분포와의 오차: 이것은 negative binomial 분포의 본래의 제약 때문에 적절한 방법이 될 수 없고 단지 계산상의 편의에서 선택된 것이므로 윤복식(1994)의 근사화 방법을 따라 gamma 분포와 같이 보다 일반적이고 자유로운 분포를 선택, 이산화하여 해결될 수 있을 것이다.

(2) 서비스 분포의 GPH 근사화 과정상의 오차

-일반 분포를 GPH 분포로 변환시킬 때, 오차가 생긴다.: 이것은 λ 의 값에 따라 충분히 작게 조정될 수 있다.

(3) GPH/GPH/1 대기시스템의 대기시간 계산 과정상의 에러

-대기시스템이 안정 상태에 도달할 때까지 iteration을 계속하는 과정에서, 각 분포의 합이 1이 안되기 때문에(메모리의 한계 때문에) 조금씩 오차가 누적된다. 그리고, 대기시간의

분포의 합 역시 1보다 작게 된다.: 이것은 가중치를 곱하여 분포를 1로 만들어 주는 장치를 마련하여 해결할 수 있다.

(4) Queueing Network의 위치과정을 세마이 마코프 체인으로 근사화하는데 따른 에러

-각 노드간의 체류시간의 상호관계와 같은 노드에서 순차적인 체류시간들 사이의 비독립성을 고려하지 않았다.: 이것은 모형화의 기본 과정에 기인한 것이므로 불가피한 오차인데 이것에 기인한 오차는 기존의 연구[6]에서 크지 않으리라고 간주할 수 있다.

5. 결론

본 연구에서는 큐잉네트워크의 체류시간분포를 구하는 분석적인 방법이 제시되었는데 전체적인 분석은 고객들이 큐잉 네트워크 각 노드 사이를 이동하는 과정(위치과정)을 세마이 마코프 과정으로 단순화시키고 이 세마이 마코프 체인에서 최초로 특정 상태로 전이되는 시간의 분포를 구하는 방식으로 행해졌다. 이때 개별 노드에서의 체류시간 분포를 미리 구해주어야 하는데, 이를 GPH/GPH/1 대기시스템으로 근사화하여 시스템에서의 지연시간을 구하는 방법이 제시되었고 정확성이 실험을 통해 검증되었다. 이 과정에서 GPH 분포로의 근사화에 대한 이론적, 실험적인 검토는 선행연구(윤복식의(1994))를 참조하기 바란다. 개별노드에서의 총 입력과정은 노드간의 동적인 상호관계를 평균과 변동계수에 관한 선형방정식을 통해 반영하여 근사적으로 구하는 방식을 따랐다. 이때 세마이 마코프 체인은 GMPH 세마이 마코프

체인이 되는데 이것에서 전이함수와 최초통과 시간(first passage time) 분포를 계산하는 과정이 또한 제시되었다. 전체적인 과정은 타당성 검증을 위해 컴퓨터 프로그램으로 작성되었고 간단한 적용예제를 만들어 다양한 상황에서 지연시간 분포를 계산하고 시뮬레이션 결과와 비교하였다. 비교 결과 본 연구의 방법이 지연시간 분포를 근사적으로 잘 표현하는 것을 알 수 있었으나 오차를 충분히 작게 하기 위해서는 컴퓨터의 기억용량이 충분히 클 필요가 있다는 것을 확인하였다. 또한 잠재적인 다양한 오차의 원인이 파악되었고 개선될 수 있는 방안이 제시되었다.

본 연구는 향후 종합적인 교환기 성능 분석 도구의 개발에 필요한 이론적, 방법론적인 토대를 마련하였는데 교환기의 성능 분석에서 지연시간 분포의 산출에 활용하는 것은 물론 통신망의 분석에서도 복잡한 트래픽 상황에서 시뮬레이션이 불가능하거나 번거로울 때 간편한 성능분석 도구로 활용될 수 있을 것이다. 또한 일반적 상황에서의 큐잉 네트워크의 근사적 분석 방법이 제시되어 이론적인 기여를 하였고, GPH 분포를 이용한 근사화는 대기시스템의 분석에 다양하게 사용될 수 있을 것이다.

본 연구에서 제시된 방법은 지연분포 이외의 척도를 포함한 전체적인 교환기 성능분석 방법의 개발에 직접 활용될 수 있으므로 이에 관한 계속적인 연구가 요구된다. 또한 본 연구를 기초로 교환기의 성능 분석에서 융통성, 유용성, 정확성, 간편성이 향상된 종합적인 방법이 제시될 수 있어 향후 광대역 환경하에서의 교환기 성능 분석에 까지 확장하는 연구가 필요할 것이다.

참고문헌

- [1] Boxma, O. J. and H. Daduna, Sojourn Times in Queueing Networks, in *Stochastic Analysis of Computer and Communication Systems*(ed. H. Takagi), Elsevier, 1990.
- [2] Ross S. M., *Stochastic Process*, Wiley, New York, 1983.
- [3] Schassberger, R., and H. Daduna, "Sojourn times in queueing networks with multiserver nodes," *J. Appl. Probab.*, vol. 24(1987), pp. 511-521.
- [4] Shanthikumar J. G., "Bilateral Phase-Type Distributions," *Naval Research Logistics Quarterly*, vol. 32(1985), pp. 119-136.
- [5] Shanthikumar J. G. and J. A. Buzacott, "The time spent in a dynamic job shop," *European J. Oper. Res.*, vol. 17(1984), pp. 215-226.
- [6] Shanthikumar, J. G. and U. Sumita, "Approximations for the time spent in a dynamic job shop with applications to due-date assignment," *Int. J. Prod. Res.*, vol. 26(1988), pp. 1329-1352.
- [7] Walrand, J. and P. Varaiya, "Sojourn times and the overtaking conditions in Jacksonian networks," *Adv. in Appl. Probab.*, vol. 12(1980), pp. 1000-1018.
- [8] Whitt, W., The Queueing Network Analyzer, *B. S. T. J.*, vol. 62(1983), pp. 2779-2814.
- [9] 윤복식, 박광우, 이창훈, "GPH 분포에 의한 확률적 근사화," 한국경영과학회지, 19권, 1호(1994), pp. 85-98.