

Overflow Probabilities in Multi-class Feedback Queues¹⁾

Mi Jung Song²⁾ · Kyungsoon Bae³⁾ · Jiyeon Lee⁴⁾

Abstract

We consider M/M/1 feedback queues with multi-class customers. We assume that different classes of customers have different arrival rates, service rates and feedback probabilities. Using the h -transforms of McDonald(1999) we derive an importance sampling estimator for an overflow probability that the total number of customers in the system reaches a high level before emptying.

Keywords : Fast Simulations , h -transforms, Importance Sampling Estimators, Multi-Class Feedback Queues, Overflow Probabilities

1. 서론

다계층(multi-class) 대기행렬이란 서버에 여러 종류(계층)의 고객들이 섞여 도착하는 시스템으로서 계층에 따라 제공되는 서비스의 종류가 다르고 따라서 서비스 시간의 분포가 서로 다른 것이 일반적이다. 한 명의 서버에 의해 계층에 상관없이 FIFO(First In First Out)로 서비스가 진행되는 다계층 대기행렬은 준역행가능(quasi-reversible)하지 않기 때문에 그 정상확률분포를 곱의 형태로 얻을 수 없다(McDonald, 2004). 다만 최근에 서비스를 마친 고객이 바로 시스템을 이탈하는 다계층 대기행렬에 대해서 Boxma and Takine(2003)는 각 계층의 고객 수에 대한 결합정상확률분포의 적률생성함수를 구하였고, Dabrowski et al.(2007)은 전체 고객 수와 현재 서비스를 받고 있는 고객의 계층에 대한 결합정상확률분포의 점근적 결과(asymptotic results)를 얻었다. Lee and Bae(2007)는 전체 고객의 수가 적정 수준을 초과하여 처

1) This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-204-C00016).

2) Graduate Student, Dept. of Statistics, Yeungnam University, Kyungsan 712-749, Korea

3) Graduate Student, Dept. of Statistics, Yeungnam University, Kyungsan 712-749, Korea

4) Corresponding Author : Professor, Dept. of Statistics, Yeungnam University, Kyungsan 712-749, Korea

E-mail : lee jy@yu.ac.kr

음으로 과부하가 발생하는 확률을 추정하는 빠른 시뮬레이션 알고리즘을 제안하였다. 그러나 서비스를 마친 후 다시 재입장(feedback)하는 고객이 있는 다계층 대기행렬에 대해서는 Choi et al.(2000)에 의해 각 계층의 고객 수에 대한 정상확률분포의 적률생성함수를 구하는 반복식(iterative formula)만이 유도되었다.

안정된(stable) 시스템의 경우에는 전체 고객의 수가 일정한 수를 초과하여 발생하는 과부하는 거의 일어나지 않는 희귀 사건(rare event)이다. 시스템이 복잡해질수록 이러한 희귀 사건의 발생 확률을 직접 계산하는 것은 거의 불가능하기 때문에 시뮬레이션을 통해 추정하게 된다. 그런데 일반적인 시뮬레이션 방법으로 추정하면 시간과 비용이 많이 들 뿐 아니라 추정 오차도 더불어 큰 단점이 있다. 이런 경우에 중요 샘플링 방법(importance sampling method, Glynn and Iglehart, 1989; Lee and Kweon, 2001)을 적용하면, 기존의 시스템을 희귀 사건의 발생을 촉진시키는 새로운 시스템으로 변환한 후 시뮬레이션을 수행함으로써 좀 더 빨리 희귀 사건의 확률을 추정할 수 있게 된다.

확률공간 (Ω, Σ, P) 에서 희귀 사건 $A \in \Sigma$ 의 발생 확률 $\theta = P(A) \ll 1$ 를 추정하고자 한다. 일반적인 시뮬레이션 방법에서는 N 번의 독립적인 시행으로부터 얻어지는 θ 의 추정량

$$\hat{\theta} = \frac{1}{N} \sum_{i=1}^N 1_A(\omega_i) \quad (1)$$

을 이용한다. 여기서 1_A 는 사건 A 의 지시함수(indicator function)로서

$$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

으로 정의되며, ω_i 는 (Ω, Σ, P) 에서 얻어지는 i 번째 시행의 결과이다. 그리고 추정량 $\hat{\theta}$ 의 평균제곱오차(MSE, Mean Square Error)는

$$\begin{aligned} \text{MSE}(\hat{\theta}) &:= E[(\hat{\theta} - \theta)^2] \\ &= \frac{\theta(1-\theta)}{N} \end{aligned}$$

로 얻어진다.

기존의 확률측도 P 가 동일한 공간 (Ω, Σ) 에서 정의되는 새로운 확률측도 \tilde{P} 에 대해 절대 연속(absolutely continuous)이면

$$\begin{aligned} \theta &= \int 1_A(\omega) dP(\omega) \\ &= \int 1_A(\omega) L(\omega) d\tilde{P}(\omega) \end{aligned}$$

이 성립한다. 여기서 $L := \frac{dP}{d\tilde{P}}$ 를 우도비(likelihood ratio) 또는 Radon-Nikodym

derivative라고 한다(Parekh and Walrand, 1989). 따라서 변화된 새 확률측도 \tilde{P} 에 의해 확률 θ 의 중요 샘플링 추정량

$$\tilde{\theta} = \frac{1}{N} \sum_{i=1}^N 1_A(\tilde{\omega}_i) L(\tilde{\omega}_i)$$

을 얻을 수 있다(Lee and Kweon, 2001). 단, $\tilde{\omega}_i$ 는 $(\Omega, \Sigma, \tilde{P})$ 에서 얻어지는 i 번째 시행의 결과이고 $L(\tilde{\omega}_i)$ 는 그 때의 우도비 값이다. 그리고 추정량 $\tilde{\theta}$ 의 평균제곱오차(MSE)는

$$\begin{aligned} \text{MSE}(\tilde{\theta}) &:= \tilde{E}[(\tilde{\theta} - \theta)^2] \\ &= \frac{1}{N} \left(\int 1_A(\omega) L^2(\omega) d\tilde{P}(\omega) - \theta^2 \right) \end{aligned}$$

로 얻어진다. 이 때, \tilde{E} 는 확률측도 \tilde{P} 에서 계산되는 기대값이다.

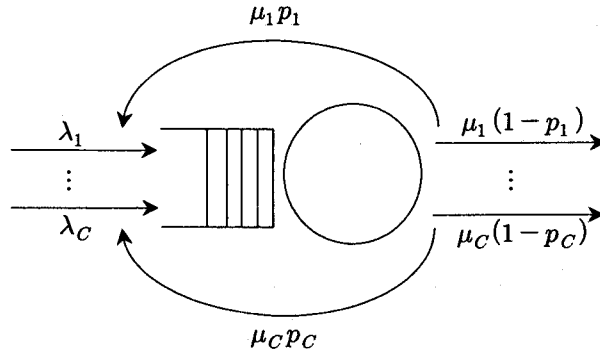
중요 샘플링 추정량을 적용함으로써, 안정된 시스템에서의 과부하 사건 대신에 확률측도를 변화시킨 불안정한 시스템에서의 과부하 사건을 시뮬레이션하여 확률을 빠르게 추정하는 방법을 빠른 시뮬레이션(fast simulation, Heidelberger, 1995)이라고 한다. 빠른 시뮬레이션 방법을 적용하기 위해서는 적용 후의 추정량의 분산이 감소하고, 변형된 시스템에서의 과부하 사건의 확률 값이 더 크게 되도록 적절하게 확률측도를 변화시키는 것이 중요하다. Parekh and Walrand(1989)는 과부하가 발생하는 최적 경로(optimal path)를 고려하여 최적의 변화된 확률측도를 찾는 직관적인(heuristic) 방법을 제안하였고, Lee and Kweon(2001)은 이 방법을 동시 입력이 있는 병렬 네트워크에 적용하여 과부하 확률을 추정하는 빠른 시뮬레이션을 얻었다. 한편 McDonald(1999)는 h -변환(h -transform)을 이용하여 새로운 확률측도를 찾는 방법을 제시하였다. 이 방법으로 Lee(2000)는 직렬 네트워크에서의 과부하 확률을 추정하고, Lee and Bae(2007)는 서비스 종료 후 바로 시스템을 이탈하는 다계층 대기행렬에서의 과부하 확률을 추정하는 빠른 시뮬레이션을 유도하였다.

본 논문에서는 McDonald(1999)의 h -변환 방법을 재입장이 있는 다계층 대기행렬에 적용하여 변화된 새로운 확률측도를 찾고 이를 이용하여 전체 고객의 수가 처음으로 일정한 수를 초과하는 과부하가 발생하는 확률을 빠르게 추정하는 중요 샘플링 추정량을 구하고자 한다. 2장에서는 재입장이 있는 다계층 대기행렬 모형을 소개하고, 빠른 시뮬레이션을 적용하기 위한 변화된 확률측도와 시스템에서의 과부하 확률을 추정한다. 3장에서는 일반 시뮬레이션과 빠른 시뮬레이션 결과를 예제를 통해 비교하고 마지막으로 4장에서는 그 결론을 정리한다.

2. 재입장이 있는 다계층 대기행렬

서버가 한 명 있는 대기행렬 시스템에 C 개의 서로 다른 계층의 고객들이 서로 독립적으로 도착한다. 이 때, 계층 c , $c = 1, 2, \dots, C$ 에 속하는 고객들은 도착률이 λ_c 인 포아송 과정(Poisson process)을 따라 도착하고 계층에 상관없이 FIFO로 서비스를 받는다.

다. 서비스 시간의 분포는 계층에 따라 서로 다르며, 계층 c 의 고객은 서비스 차례가 되면 평균이 $1/\mu_c$ 인 지수분포(exponential distribution)의 시간만큼 서비스를 받는다. 서비스가 끝난 후, 계층 c 의 고객들은 $1-p_c$ 의 확률로 시스템을 완전히 떠나거나 혹은 p_c 의 확률로 다시 서비스를 받기 위해 대기열의 마지막에 재입장한다. <그림 1>는 재입장이 있는 다계층 대기행렬과 그 모수들을 나타낸 것이다.



<그림 1> 재입장이 있는 다계층 대기행렬

이 때, 계층 c 인 고객들의 총 도착률(total arrival rate) Λ_c 는

$$\Lambda_c = \lambda_c + \Lambda_c p_c$$

을 만족하여

$$\Lambda_c = \frac{\lambda_c}{1-p_c}$$

로 얻어지고, 따라서 이 시스템의 안정성(stability)을 보장하기 위해서 총 로드 ρ 에 대해

$$\rho := \sum_{c=1}^C \frac{\lambda_c}{\mu_c(1-p_c)} < 1$$

을 가정한다. 그리고

$$\sum_{c=1}^C (\lambda_c + \mu_c) = 1 \tag{2}$$

를 가정하여 균일화 기법(uniformization method, Walrand, 1988)에 의해 본 시스템과 동일한 정상확률을 갖는 이산 시간형 마코프 체인(Markov chain) S 를 고려한다.

마코프 체인 $S = \{X(t), t = 0, 1, 2, \dots\}$ 의 상태 $X(t)$ 는 고객이 한 명도 없는 상태를 0 으로 나타내고, 적어도 1명의 고객이 있는 경우는 고객들의 계층을 이용하여 벡터 $X(t) = (X_0(t), X_1(t), \dots, X_{n-1}(t))$ 로 나타낸다. 이것은 시점 t 에서 총 고객의 수가

$|X(t)|=n$ 이고 $X_0(t)$ 는 대기열의 맨 앞에서 서비스를 받고 있는 고객의 계층, $X_1(t)$ 는 그 뒤에 기다리고 있는 고객의 계층, \dots , $X_{n-1}(t)$ 는 시스템에 마지막으로 도착하여 대기열의 끝에서 기다리고 있는 고객의 계층을 순서대로 표시한 것이다. 그리고 마코프 체인 S 의 전이확률(transition probability) $K(x, y)$ 는 $x=0$ 일 때,

$$K(0, y) = \begin{cases} \lambda_c & y = (c), c = 1, 2, \dots, C \\ \sum_{c=1}^C \mu_c & y = 0 \\ 0, & \text{그외 } y \end{cases}$$

이고, $|x|=n > 0$ 이고 $x=(x_0, x_1, \dots, x_{n-1})$ 일 때는

$$K(x, y) = \begin{cases} \lambda_c, & y = (x_0, x_1, \dots, x_{n-1}, c), c = 1, 2, \dots, C \\ \mu_{x_0}(1 - p_{x_0}), & y = (x_1, \dots, x_{n-1}) \\ \mu_{x_0} p_{x_0}, & y = (x_1, \dots, x_{n-1}, x_0) \\ \sum_{c \neq x_0} \mu_c & y = (x_0, x_1, \dots, x_{n-1}) \\ 0, & \text{그외 } y \end{cases}$$

가 된다.

Dabrowski et al.(2007)은 마코프 체인의 전이확률에 대한 조화함수(harmonic function)를 이용하여 원래의 안정된 마코프 체인을 불안정한 마코프 체인으로 확률추도를 변화시키고, 변화된 확률추도를 이용하여 과부하 발생 확률의 점근적 결과를 유도하였다. 이 때, 조화함수란 마코프 체인의 전이확률 $K(x, y)$ 에 대해

$$h(x) = \sum_y K(x, y)h(y), \quad \text{모든 } x \text{에 대해} \tag{3}$$

를 만족하는 함수를 말한다(McDonald, 1999). 조화함수 h 를 이용하면 다음과 같은 새로운 전이확률 $\tilde{K}(x, y)$ 를 정의할 수 있다.

$$\tilde{K}(x, y) = K(x, y) \frac{h(y)}{h(x)}. \tag{4}$$

마코프 체인 S 의 상태 값 $x=(x_0, x_1, \dots, x_{n-1})$ 에 대하여 $N_c(x)$ 를 $|x|=n$ 명의 고객들 중 계층이 c 인 고객의 수라고 하자. 그러면 다음과 같은 형태의 함수를 생각할 수 있다(Dabrowski et al., 2007).

$$h(x) = \prod_c e^{\gamma_c N_c(x)}. \tag{5}$$

이 함수가 $\{x: |x| > 0\}$ 의 영역에서 (3)을 만족하는 조화함수가 되기 위해서는 모든 계

층 c 에 대해

$$\sum_{d=1}^c \lambda_d e^{\gamma_d} + \mu_c e^{-\gamma_c} (1-p_c) + \mu_c p_c + \sum_{d \neq c} \mu_d = 1 \quad (6)$$

를 만족해야 한다. 여기서 특별히 $c=1$ 일 때, 식 (6)은

$$\sum_{d=1}^c \lambda_d e^{\gamma_d} + \mu_1 e^{-\gamma_1} (1-p_1) + \mu_1 p_1 + \sum_{d \neq 1} \mu_d = 1 \quad (7)$$

가 되고, 그 외 계층 c 에 대해 얻어지는 식 (6)을 위 식 (7)에서 빼면

$$\mu_c e^{-\gamma_c} (1-p_c) + \mu_c p_c - \mu_1 e^{-\gamma_1} (1-p_1) - \mu_1 p_1 = \mu_c - \mu_1$$

을 얻을 수 있다. 위 식을 $e^{-\gamma_c}$ 에 대해 정리하면 모든 계층 c 에 대해,

$$e^{-\gamma_c} = \frac{\mu_1 (1-p_1) (e^{-\gamma_1} - 1) + \mu_c (1-p_c)}{\mu_c (1-p_c)} \quad (8)$$

를 얻을 수 있다. 식 (8)을 식 (7)에 대입하고 조건식 (2)를 적용하면 γ_1 에 대한 식

$$\sum_c \frac{\lambda_c}{\mu_1 (1-p_1) (e^{-\gamma_1} - 1) + \mu_c (1-p_c)} = 1 \quad (9)$$

을 유도할 수 있다. 식 (9)의 좌변을 γ_1 의 함수로 보면 $\gamma_1=0$ 에서 1보다 작은 값 ρ 를 갖는 절대 증가함수이므로 식 (9)를 만족하는 양의 해 γ_1 이 항상 존재한다. 따라서 1보다 큰 e^{γ_1} 의 값을 항상 구할 수 있다. 또한, 식 (8)로부터 나머지 계층 $c \neq 1$ 에 대한 e^{γ_c} 의 값을 구할 수 있으며 모두 1보다 크다는 사실도 확인할 수 있다. 특별히 $C=2$ 일 때, 즉 고객의 계층이 두 개인 경우는

$$e^{\gamma_c} = 2\mu_c (1-p_c) / A_c, \quad c=1,2 \quad (10)$$

으로 계산된다. 단, $c=1,2$ 에 대해

$$A_c = \lambda_1 + \lambda_2 + (-1)^{c+1} \mu_1 (1-p_1) + (-1)^c \mu_2 (1-p_2) + \sqrt{[\lambda_1 + \lambda_2 + \mu_1 (1-p_1) - \mu_2 (1-p_2)]^2 - 4\lambda_1 \{\mu_1 (1-p_1) - \mu_2 (1-p_2)\}}.$$

그러므로 조건식 (6)을 만족하는 γ_c 로 구성된 함수 $h(\mathbf{x}) = \prod_c e^{\gamma_c N_c(\mathbf{x})}$ 는 $\{\mathbf{x}: |\mathbf{x}| > 0\}$ 의 영역에서 조화함수가 되므로 식 (4)에 대입하면 $|\mathbf{x}| = n > 0$ 이고 $\mathbf{x} = (x_0, x_1, \dots, x_{n-1})$ 일 때, 다음과 같은 새로운 전이확률을 얻을 수 있다.

$$\begin{aligned} \tilde{K}(\mathbf{x}, \mathbf{y}) &= K(\mathbf{x}, \mathbf{y}) \prod_c e^{\gamma_c [N_c(\mathbf{y}) - N_c(\mathbf{x})]} \\ &= \begin{cases} \lambda_c e^{\gamma_c}, & \mathbf{y} = (x_0, x_1, \dots, x_{n-1}, c), \quad c = 1, 2, \dots, C \\ \mu_{x_0} (1 - p_{x_0}) e^{-\gamma_{x_0}}, & \mathbf{y} = (x_1, \dots, x_{n-1}) \\ \mu_{x_0} p_{x_0}, & \mathbf{y} = (x_1, \dots, x_{n-1}, x_0) \\ \sum_{c \neq x_0} \mu_c, & \mathbf{y} = (x_0, x_1, \dots, x_{n-1}) \\ 0, & \text{그외 } \mathbf{y}. \end{cases} \end{aligned} \tag{11}$$

만약 $\mathbf{x} = \mathbf{0}$ 에서의 전이확률 $\tilde{K}(\mathbf{0}, \mathbf{y})$ 를

$$\tilde{K}(\mathbf{0}, \mathbf{y}) = \begin{cases} \lambda_c e^{\gamma_c}, & \mathbf{y} = (c), \quad c = 1, 2, \dots, C \\ 1 - \sum_{c=1}^C \lambda_c e^{\gamma_c}, & \mathbf{y} = \mathbf{0} \\ 0, & \text{그외 } \mathbf{y} \end{cases} \tag{12}$$

로 정의하면 식 (11)과 (12)의 전이확률을 갖는 새로운 마코프 체인 \tilde{S} 는 기존의 마코프 체인 S 로 설명되는 대기행렬(<그림 1>)과 모수는 다르지만 동일한 구조를 갖는 대기행렬을 나타낸다. 즉, 계층 c 의 고객이 도착률 $\tilde{\lambda}_c := \lambda_c e^{\gamma_c}$ 인 포아송 과정을 따르며 독립적으로 도착하고, 계층 c 의 고객의 서비스 시간은 평균이 $1/\tilde{\mu}_c := 1/\mu_c [p_c + (1-p_c)e^{-\gamma_c}]$ 인 지수분포를 따르고 계층 c 의 고객이 서비스를 마치고 시스템으로 다시 재입장할 확률이 $\tilde{p}_c := p_c / [p_c + (1-p_c)e^{-\gamma_c}]$ 인 다계층 대기행렬을 나타낸다. 한편 마코프 체인 \tilde{S} 에서 계층 c 인 고객의 총 도착률 $\tilde{\Lambda}_c$ 은

$$\tilde{\Lambda}_c = \frac{1}{\tilde{\mu}_c (1 - \tilde{p}_c)}$$

이고, 따라서 총 로드 $\tilde{\rho}$ 는

$$\tilde{\rho} := \sum_c \frac{\tilde{\lambda}_c}{\tilde{\mu}_c (1 - \tilde{p}_c)}$$

가 된다. 한편 식 (9)는

$$\sum_c \frac{\lambda_c}{\tilde{\mu}_c (1 - \tilde{p}_c)} = 1$$

으로 바꾸어 나타낼 수 있고, 모든 계층 c 에 대해 $e^{\gamma_c} > 1$ 이므로 총 로드 $\tilde{\rho}$ 에 대해

$$\tilde{\rho} = \sum_c \frac{\tilde{\lambda}_c}{\tilde{\mu}_c(1-\tilde{p}_c)} > \sum_c \frac{\lambda_c}{\mu_c(1-p_c)} = 1$$

임을 알 수 있다. 즉, 새 마코프 체인 \tilde{S} 의 대기행렬은 불안정한(unstable)한 시스템이 된다.

사건 H_ℓ 을 총 고객 수가 0을 벗어난다고 했을 때 다시 0이 되기 전에 처음으로 충분히 큰 수 ℓ 에 먼저 도착하는 사건이라고 하자. ℓ 을 시스템이 수용할 수 있는 최대 고객 수라고 하면 사건 H_ℓ 은 시스템에 과부하(overflow)가 발생하는 사건으로 과부하 확률 $\theta = P(H_\ell)$ 을 추정하고자 한다. 이 확률은 바쁜 기간(busy period) 동안의 최대 고객 수(maximum queue length)가 ℓ 이상일 확률과 동일하다.

상태 0을 벗어날 때, 제일 먼저 도착하는 고객의 계층이 c 일 확률은 λ_c/λ 로 얻어진다. 단, $\lambda = \sum_{c=1}^C \lambda_c$ 이다. T 를 고객 수가 0을 벗어난다고 했을 때 다시 0이 되기 전에 처음으로 ℓ 이 될 때까지의 총 전이횟수라고 하자. 그러면 사건 H_ℓ 에 해당되는 가능한 경로(trajjectory) $\omega := (\mathbf{x}(1) := (x_1), \mathbf{x}(2), \dots, \mathbf{x}(t))$ 에 대한 확률은

$$\begin{aligned} P(\omega) &:= P\{X(1) = (x_1), X(2) = (x_2), \dots, X(T) = (x(t)) \mid \text{상태 0을 벗어난다}\} \\ &= \frac{\lambda_{x_1}}{\lambda} \prod_{s=1}^{t-1} K(\mathbf{x}(s), \mathbf{x}(s+1)) \end{aligned}$$

이 된다. 위의 확률을 식 (11)과 (12)의 새로운 전이확률 \tilde{K} 를 이용하여 나타내면,

$$\begin{aligned} P(\omega) &= \frac{\lambda_{x_1}}{\lambda} \prod_{s=1}^{t-1} \tilde{K}(\mathbf{x}(s), \mathbf{x}(s+1)) \frac{h(\mathbf{x}(s))}{h(\mathbf{x}(s+1))} \\ &= \frac{\lambda_{x_1}}{\lambda} \frac{h(\mathbf{x}(1))}{h(\mathbf{x}(t))} \prod_{s=1}^{t-1} \tilde{K}(\mathbf{x}(s), \mathbf{x}(s+1)) \\ &= \frac{\tilde{\lambda}}{\lambda} \frac{1}{h(\mathbf{x}(t))} \frac{\tilde{\lambda}_{x_1}}{\tilde{\lambda}} \prod_{s=1}^{t-1} \tilde{K}(\mathbf{x}(s), \mathbf{x}(s+1)) \\ &= \frac{\tilde{\lambda}}{\lambda} \frac{1}{h(\mathbf{x}(t))} \tilde{P}(\omega) \end{aligned}$$

가 된다. 여기서 $\tilde{\lambda} = \sum_{c=1}^C \tilde{\lambda}_c$ 이고 $\tilde{P}(\omega)$ 는 마코프 체인 \tilde{S} 에서의 경로 ω 의 확률이다. 따라서 우도비 L 은

$$L = \frac{\tilde{\lambda}}{\lambda} \frac{1}{h(\mathbf{X}(T))} = \frac{\tilde{\lambda}}{\lambda} \prod_c e^{\gamma_c N_c(\mathbf{X}(T))}$$

가 된다. 그러므로 변화된 마코프 체인 \tilde{S} 를 이용한 확률 $\theta = P(H_\ell)$ 의 중요 샘플링 추정량은

$$\tilde{\theta} = \frac{1}{N} \frac{\tilde{\lambda}}{\lambda} \sum_{i=1}^N \left(1_{H_\ell}(\tilde{\omega}_i) \prod_c e^{\gamma_c N_c(\mathbf{x}(T))(\tilde{\omega}_i)} \right) \quad (13)$$

로 얻을 수 있다. 이 때, $\tilde{\omega}_i$ 는 마코프 체인 \tilde{S} 에서 얻어지는 시행의 결과이다.

3. 시뮬레이션

고객의 계층이 2개이고 모수가 $\lambda_1 = 0.1$, $\lambda_2 = 0.08$, $\mu_1 = 0.37$, $\mu_2 = 0.45$, $p_1 = 0.2$, $p_2 = 0.15$ 인 다계층 대기행렬의 총 로드는 $\rho = 0.54699$ 로서 안정된 시스템이다.

주어진 모수에 대해 식 (10)으로부터 $e^{\gamma_1} = 1.9623$, $e^{\gamma_2} = 1.6116$ 이 계산되어 변화된 모수는 $\tilde{\lambda}_1 = 0.1962$, $\tilde{\lambda}_2 = 0.1289$, $\tilde{\mu}_1(1 - \tilde{p}_1) = 0.1508$, $\tilde{\mu}_1\tilde{p}_1 = 0.0740$, $\tilde{\mu}_2(1 - \tilde{p}_2) = 0.2373$, $\tilde{\mu}_2\tilde{p}_2 = 0.0675$ 가 되고 따라서 총 로드 $\tilde{\rho} = 1.84457$ 인 불안정한 시스템을 얻는다.

일반 추정량 (1)과 빠른 시뮬레이션을 위한 중요 샘플링 추정량 (13)을 이용하여 각 ℓ 에 대한 과부하 확률 $P(H_\ell)$ 을 추정한 값을 <표 1>에 나타내었다. 이 때, 일반 시뮬레이션은 15억 개의 바쁜 기간을 생성해서 추정한 값이고, 빠른 시뮬레이션은 7000개의 바쁜 기간만 생성해서 추정한 값으로 두 방법에서 추정된 확률 값이 서로 비슷함을 알 수 있다.

<그림 2>는 바쁜 기간의 수를 70만 개로 동일하게 하였을 때, 두 추정량의 평균제곱오차를 추정하여 로그변환한 것이다. 일반 시뮬레이션의 평균제곱오차가 빠른 시뮬레이션의 평균제곱오차보다 훨씬 큰 것을 보여주고 있다. <표 2>는 95% 신뢰도에서 오차한계 10^{-5} 를 만족하는데 필요한 바쁜 기간의 수를 비교한 것으로, 빠른 시뮬레이션의 경우가 훨씬 작으며, ℓ 이 커질수록 필요한 바쁜 기간의 수가 현격히 줄어들어도 확인할 수 있다.

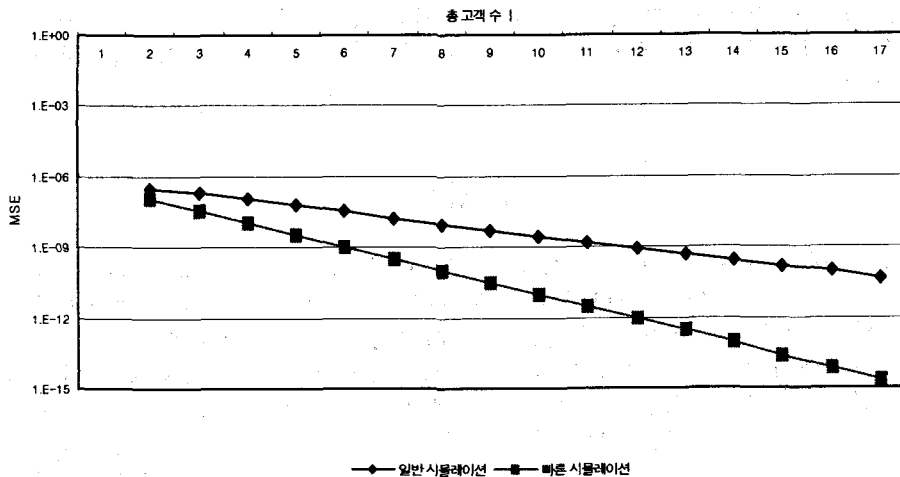
4. 결론

본 논문은 여러 개의 계층의 고객들이 시스템에 도착하여 계층에 따라 서로 다른 서비스를 받은 후 시스템을 떠나거나 혹은 다시 재입장하는 다계층 대기행렬 시스템에서 과부하가 발생하는 확률을 구하는 시뮬레이션 방법을 제시하였다.

안정된 시스템에서는 과부하가 발생하는 것은 희귀 사건이기 때문에 일반 시뮬레이션으로는 엄청난 시간과 비용이 든다. 본 논문에서는 McDonald(1999)의 h -변환을 이용한 중요 샘플링 방법을 적용하여 빠르게 추정할 수 있는 추정량을 제시하였고 예제를 통해 적은 수의 시뮬레이션으로도 효율적으로 과부하 확률을 추정할 수 있음을 확인하였다.

<표 1> 과부하 확률 추정값

ℓ	일반 시뮬레이션 ($N=1.5 \times 10^9$)		빠른 시뮬레이션 ($N=7 \times 10^3$)	
	추정값	95% 신뢰구간	추정값	95% 신뢰구간
1	1.000000000	(1.000000000, 1.000000000)	0.9986963514	(0.9964073216, 1.0009853813)
2	0.3523191160	(0.3522949414, 0.3523432906)	0.3542658109	(0.3479780329, 0.3605535889)
3	0.1616206387	(0.1616020101, 0.1616392672)	0.1618181961	(0.1581992416, 0.1654371507)
4	0.0814890600	(0.0814752147, 0.0815029053)	0.0816278191	(0.0796180658, 0.0836375725)
5	0.0428922973	(0.0428820436, 0.0429025510)	0.0429370723	(0.0418276327, 0.0440465119)
6	0.0230707240	(0.0230631265, 0.0230783215)	0.0231349471	(0.0225190370, 0.0237508572)
7	0.0125506300	(0.0125449962, 0.0125562638)	0.0125864382	(0.0122453034, 0.0129275730)
8	0.0068705860	(0.0068664057, 0.0068747663)	0.0068788826	(0.0066892117, 0.0070685535)
9	0.0037722187	(0.0037691163, 0.0037753210)	0.0037924474	(0.0036868672, 0.0038980276)
10	0.0020747307	(0.0020724280, 0.0020770334)	0.0020908486	(0.0020321251, 0.0021495721)
11	0.0011416567	(0.0011399477, 0.0011433656)	0.0011524264	(0.0011198019, 0.0011850508)
12	0.0006292020	(0.0006279330, 0.0006304710)	0.0006365830	(0.0006184362, 0.0006547297)
13	0.0003464820	(0.0003455402, 0.0003474238)	0.0003502063	(0.0003401661, 0.0003602464)
14	0.0001907433	(0.0001900445, 0.0001914422)	0.0001935464	(0.0001879377, 0.0001991552)
15	0.0001052133	(0.0001046943, 0.0001057324)	0.0001065902	(0.0001034859, 0.0001096944)
16	0.0000580540	(0.0000576684, 0.0000584396)	0.0000587194	(0.0000570019, 0.0000604370)
17	0.0000320353	(0.0000317489, 0.0000323218)	0.0000324594	(0.0000315000, 0.0000334189)
18	0.0000177833	(0.0000175699, 0.0000179967)	0.0000178577	(0.0000173256, 0.0000183897)
19	0.0000097760	(0.0000096178, 0.0000099342)	0.0000099072	(0.0000096086, 0.0000102058)
20	0.0000053347	(0.0000052178, 0.0000054516)	0.0000054555	(0.0000052899, 0.0000056211)
21	0.0000029600	(0.0000028729, 0.0000030471)	0.0000030149	(0.0000029226, 0.0000031073)
22	0.0000015933	(0.0000015295, 0.0000016572)	0.0000016557	(0.0000016047, 0.0000017068)
23	0.0000008880	(0.0000008403, 0.0000009357)	0.0000009072	(0.0000008791, 0.0000009352)
24	0.0000004793	(0.0000004443, 0.0000005144)	0.0000005012	(0.0000004856, 0.0000005168)
25	0.0000002587	(0.0000002329, 0.0000002844)	0.0000002754	(0.0000002668, 0.0000002840)



<그림 2> 두 추정량의 평균제곱오차 ($N=7 \times 10^5$)

<표 2> 95% 신뢰도에서 오차한계 10^{-5} 을 만족하기 위한 바쁜 기간의 수

ℓ	일반 시뮬레이션	빠른 시뮬레이션
2	8765360102	2805656090
3	5204994375	921456297
4	2870268001	283478892
5	1568362035	86787743
6	855757493	26632088
7	468257146	8190070
8	255012959	2521836
9	142430376	777401
10	77716857	239786
11	42484911	73994
12	22652067	22855
13	11850422	7058
14	6694193	2178
15	4005822	671
16	2579187	207
17	1317075	64

참고문헌

1. Boxma, O. J. and Takine, T. (2003). The M/G/1 FIFO queue with several customer classes, *Queueing Systems*, 45, 185-189.
2. Choi, B. D., Kim, B. and Cho, S. H. (2000). On the M/G/1 Bernoulli feedback queue with multi-class customers, *Computers and Operations Research*, 27, 269-286.
3. Dabrowski, A., Lee, J. and McDonald, D. (2007). Large deviations of multitype queues, preprint.
4. Glynn, P. W. and Iglehart, D. L. (1989). Importance sampling for stochastic simulations, *Management Science*, 35, 1367-1392.
5. Heidelberger, P. (1995). Fast simulation of rare events in queueing and reliability models, *ACM Transactions on Modeling and Computer Simulation*, 5, 43-85.
6. Lee, J. (2000). Fast simulation for excessive backlogs in tandem network, *The Korean Communications in Statistics*, 7, 499-512.
7. Lee, J. and Kweon, M. H. (2001). Estimation of overflow probabilities in parallel networks with coupled inputs, *The Korean Communications in Statistics*, 8, 257-269.
8. Lee, J. and Bae, K. (2007). Fast simulation of overflow probabilities in multiclass queues, *The Korean Communications in Statistics*, 14, 287-299.
9. McDonald, D. (1999). Asymptotics of first passage times for random

- walk in an orthant, *Annals of Applied Probability*, 9, 110-145.
10. McDonald, D. (2004). *Elements of Applied Probability for Engineering, Mathematics and Systems Science*, World Scientific, River Edge, NJ.
 11. Parekh, S. and Walrand, J. (1989). A quick simulation method for excessive backlog in networks of queues, *IEEE Transactions on Automatic Control*, 34, 54-66.
 12. Walrand, J. (1988). *An Introduction to Queueing Networks*, GL Jordan, Ed., Prentice Hall, Englewood Cliffs, NJ.

[2007년 9월 접수, 2007년 11월 채택]