

Multiclass loss systems with several server allocation methods

Seongryong Na^{a,1}

^aDepartment of Information and Statistics, Yonsei University

(Received March 4, 2016; Revised April 15, 2016; Accepted May 9, 2016)

Abstract

In this paper, we study multiclass loss systems with different server allocation methods. The Markovian states of the systems are defined and their effective representation is investigated. The limiting probabilities are derived based on the Markovian property to determine the performance measures of the systems. The effects of the assignment methods are compared using numerical solutions.

Keywords: multiclass loss systems, server allocation methods, Markov modeling, numerical iteration, performance analysis, loss probabilities

1. 서론

여러 종류의 고객을 서비스하는 멀티클래스 손실시스템은 실시간 멀티미디어 통신환경에 대한 유용한 확률모형이며 이들에 대한 성능분석의 기초가 된다. 얼랑의 선구적인 업적 이후에 손실시스템에 대한 이론과 응용은 대기행렬 분야에서 기본적인 연구 주제가 되어왔다. 다양한 고객을 가정하는 멀티클래스 손실시스템은 Kelly (1991)에서 자세히 다루어 졌고 이를 기초로 통신시스템 성능분석을 다룬 Koo 등 (2000, 2002)과 Lee 등 (2002)의 논문이 있다. 얼랑 B-공식에 대한 반복 계산식이 큰 의미를 가지는 것과 같이 멀티클래스 손실시스템의 현실 응용에서는 계산적인 문제가 중요하다. 멀티클래스 손실시스템의 극한확률에 대한 반복 계산식을 제안하는 연구 결과는 Ross (1995)와 Na (2006) 등이 있다.

최근에 현실적인 시스템 운용 상황을 반영하기 위하여 Na (2010)는 멀티클래스 손실시스템에 유니트 개념을 도입하고 기초적인 시스템에 대하여 확률적 특성을 연구하였다. 유니트는 물리적 혹은 기술적으로 서버 집단을 구별하고 이를 기준으로 서비스를 배정하는 개념인데 이에 대한 마코프 시스템의 복잡도가 급격하게 증가하고 극한확률의 해석적 계산이 매우 어려워지는 경향이 있다. 유니트 구조를 가지는 멀티클래스 손실시스템의 극한확률 계산은 반복적 수치해 방법이 현실적인 대안이 될 수 있다. 마코프 과정의 수치해에 대한 반복 계산법을 다룬 논문으로 O'Leary (1993), Stewart (2000), Zhao 등 (2012), Nesterov와 Nemirovski (2015) 등이 있다. 또한 Na (2014)에서 설명한 수치적 반복법을 이용하면 해석적인 극한확률 유도가 어려운 경우 균형방정식의 형태에 기초한 극한확률의 수치해 계산이 수월하게 이루어질 수 있다.

This work was supported by the Yonsei University Research Fund of 2011.

¹Department of Information and Statistics, Yonsei University, 1 Yonseidae Gil, Wonju 26493, Korea.

E-mail: nasr@yonsei.ac.kr

이 논문에서는 새로운 고객을 서비스하기 위하여 시스템 자원을 할당하는 서버 배정에 대하여 여러 방식을 살펴봄으로써 현실의 다양한 시스템 운용 상황을 반영한다. 시스템에 대한 마코프 모델링을 고려하고 이에 기초한 해석적 혹은 수치적 해를 구하는 방법을 연구한다. 고객의 서비스를 위한 서버 배정방식을 다르게 할 때 고객 종류에 따른 손실확률 등의 시스템 성능 척도에 어떤 변화가 있는가를 살펴 보는 것이 이 논문의 목적이다. 등확률배정, 순차배정, 최적배정 등의 유니트 구조의 손실시스템과 단일 유니트 구조에 해당하는 Kelly (1991) 시스템의 성능 척도를 산출하여 다양한 시스템의 효율성을 비교한다.

이 논문의 구성은 다음과 같다. 2절에서는 다양한 서버배정방식의 멀티클래스 손실시스템의 정의를 살펴보고 마코프 모형으로 표현하는 방법을 연구한다. 3절에서는 마코프 모델링의 극한확률을 수치적 절차를 통하여 구하는 방법을 알아본다. 기본적인 모형을 가정하고 제안된 서버배정방식의 효율을 비교한 결과를 4절에서 정리한다.

2. 서버배정방식 및 마코프 모델링

서비스를 위하여 각각 k_1, \dots, k_r 개의 서버를 필요로 하는 r 종류의 고객이 시스템에 도착하고 시스템은 전체 c 개의 서버로 구성된다. 한편 도착할 때 필요한 수의 서버를 배정받지 못하는 고객은 즉시 탈락하는 손실시스템을 가정한다. 마코프 모델링을 위하여 각 종류의 고객은 각각 도착률 $\lambda_1, \dots, \lambda_r$ 의 서로 독립인 포아송과정의 형태로 도착하고 각각 평균 $1/\mu_1, \dots, 1/\mu_r$ 의 지수분포의 서비스시간을 가짐을 가정한다. 본 연구에서는 Na (2010)에서 제시된 유니트 구조의 시스템을 가정하고 등확률배정, 순차탐색배정, 최적배정 등의 다양한 서버배정방식을 고려하는데 각 유니트는 M 개의 서버로 구성되고 전체 시스템은 U 개의 유니트로 구성되며 따라서 전체 서버의 수는 $c = MU$ 됨을 가정하고 각 단위 서비스는 동일한 유니트 안의 서버들이 담당한다. 유니트 구조의 시스템에서는 한 고객의 서비스를 위하여 같은 유니트 내의 서버를 이용해야 하는데 가령 근처의 여러 기지국에 소량의 여유 자원이 산재해도 어느 한 기지국에서 필요한 양의 서버 자원을 동시에 배정받아야 이동통신서비스가 가능함을 예로 들 수 있다. Figure 2.1(a)는 유니트 구조 시스템의 개념도인데 유니트는 고객 서비스를 위하여 논리적 혹은 물리적으로 전체 서버를 구분하여 구성한 서버의 집합이며 신규 고객에 가용 유니트 혹은 서버를 배정하는 방법에 따라 시스템 운용 방식이 달라지며 시스템 성능에도 변화가 생길 수 있다. 한편 Kelly (1991), Na (2006) 등에서 다루어진 완전랜덤배정은 신규 서비스를 위하여 전체 유휴 서버를 별도의 구분없이 사용하는 방식으로 전체 시스템이 한 개의 유니트에 해당하는 $M = c$ 혹은 $U = 1$ 의 경우이다.

시스템의 유니트 구조는 신규 서비스 배정에 제약이 있는 시스템에서 고려하게 된다. 신규 고객 서비스를 위하여 인접한 서버의 사용이 시스템 운용상 필요하다면 이러한 인접 서버를 유니트로 묶어서 관리하는 것이 필요할 수 있다. 실제 시스템에서는 하위 구조의 서브시스템 용량에 따른 자연스러운 유니트 구조가 가능할 수 있고 대용량 서비스에 필요한 서버의 크기가 유니트의 크기가 되는 경우가 많다. 시스템의 유니트 구조를 가정하면 신규 고객 서비스를 위한 유니트 선택이 선행되고 유니트 안에서 유휴 서버의 배정이 이루어진다. 먼저 등확률배정은 새로운 고객에게 가용 유니트를 등확률로 배정하는 방식이다. 순차탐색배정은 미리 정해진 유니트 순서에 따라 가용 유니트를 탐색하는 방식이다. 최적배정은 서버 자원의 효율적 사용을 통하여 고객 탈락을 줄이기 위한 방식인데 특히 대용량 서비스를 위한 자원을 예비하기 위하여 소용량 서비스 고객을 가능하면 점유율이 높은 유니트에 밀집하게 배정하는 원리이다. Figure 2.1(b)에는 앞으로 4절에서 다루게 되는 $M = 2, U = 2, k_1 = 1, k_2 = 2$ 의 시스템에 대하여 각 배정방식에 따라 종류 1의 고객에게 서로 다른 유니트가 배정되는 결과가 표현되어 있다. 이때 그림에서 검게 표시된 원은 사용중인 서버를 나타내며 EQ, SQ, OP는 각각 등확률배정, 순차배정, 최적배정을 의미한다.

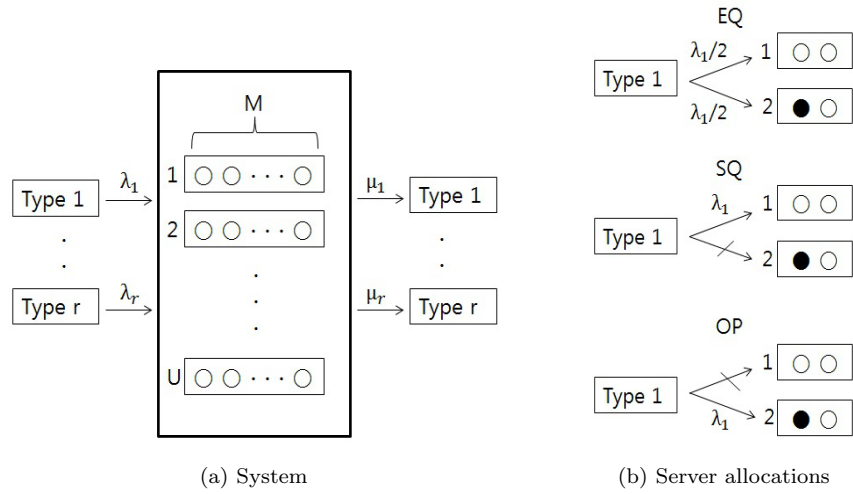


Figure 2.1. System with unit structure and server allocations according to allocation methods.

멀티클래스 손실시스템의 성능분석을 위하여 마코프 모델링을 고려한다. 유니트 구조에서의 고객 도착, 탈락 및 서비스 종료를 표현하기 위해서는 유니트별로 상태를 추적해야 한다. 시간 t 에 유니트 u 에서 서비스중인 종류 j 의 고객 수를 $N_j^u(t)$, $u = 1, \dots, U$, $j = 1, \dots, r$, $t > 0$ 기호로 표현한다. 마코프 확률변수 $N_j^u(t)$ 는

$$k_1 N_1^u(t) + \dots + k_r N_r^u(t) \leq M, \quad u = 1, \dots, U$$

범위 조건을 만족하는 값을 가질 수 있다. 손실시스템의 성능척도는 마코프 확률변수의 극한확률을 통하여 결정한다. 등확률배정의 유니트 구조 멀티클래스 손실시스템의 극한확률에 대한 정상방정식 혹은 균형방정식은 Na (2010)에 자세히 설명되어 있다. 다른 배정방식의 손실시스템에 대한 정상방정식 역시 유사한 방식으로 결정 가능하다.

Na (2010)에 설명된 바와 같이 유니트 구조 손실시스템의 극한확률의 해석적 접근은 쉽지 않으리라 예상된다. 따라서 극한확률의 수치해를 구하는 것이 대안이 되며 각 배정방식의 균형방정식을 유도하여 수치적 반복법을 적용하고자 한다. 이때 행렬대수의 이론이 필요없고 수렴속도가 빠른 Na (2014)에서 제시한 반복계산법이 이용에 편리하다. 이 방법은 O’Leary (1993), Stewart (2000) 등에서 언급한 연속시간 마코프과정에 기초한 거듭제곱법이 유사한 절대값을 가지는 고유치로 인하여 수렴속도가 느려질 수 있다는 단점을 해결하기 위하여 내재 이산시간 마코프과정을 이용한다. 이 방법은 이산시간 마코프과정의 확률분포가 극한확률로 지수적으로 수렴한다는 이론에 기초하며 (Ross, 1996) 이산시간 마코프과정에서 발생가능한 주기성 문제를 해결하는 방안을 동시에 제시한다.

3. 내재 마코프 체인을 이용한 수치적 반복법

내재된 이산시간 마코프과정의 정상방정식에 기초하여 연속시간 마코프과정의 극한확률을 계산하는 Na (2014)의 반복적 수치해 방법을 정리한다. 상태공간 $S = \{1, \dots, N\}$ 의 연속시간 마코프과정 $X(t)$, $t \geq 0$ 가 기약성을 만족하고 전이를 λ_{ij} , $i, j \in S$ 을 가진다고 가정하자. 이때 $X(t)$ 의 극한확률 π_j ,

$j \in S$ 에 대하여 다음의 정상방정식이 성립한다.

$$\pi_j \lambda_j = \sum_{i \in S} \pi_i \lambda_{ij}, \quad j \in S,$$

여기에서 $\lambda_j = \sum_{i \in S} \lambda_{ji}$ 식을 만족하는 탈출률에 해당한다. 한편 $X(t)$ 에 내재된 이산시간 마코프과정의 극한확률 혹은 정상확률 $\tilde{\pi}_j, j \in S$ 는

$$\tilde{\pi}_j = \sum_{i \in S} \tilde{\pi}_i P_{ij}, \quad j \in S$$

형태의 정상방정식을 만족하는데 $P_{ij} = \lambda_{ij}/\lambda_i, i, j \in S$ 는 전이확률이다. 연속시간 정상방정식을 이용하는 경우는 행렬대수 이론이 필요한 반면에 이산시간의 경우에는 정상방정식의 형태를 직접 이용한 단순한 반복을 통해 수치해를 손쉽게 얻을 수 있는 장점이 있다.

내재 마코프과정이 비주기적이면 극한확률 $\tilde{\pi}_j, j \in S$ 의 반복적 수치해는 다음과 같이 계산한다.

- 단계 1. 초기값: $\tilde{\pi}_j^{(0)}, j \in S$.
- 단계 2. 반복식: $n \geq 1$ 에 대하여 $\tilde{\pi}_j^{(n)} = \sum_{i \in S} \tilde{\pi}_i^{(n-1)} P_{ij}, j \in S$.
- 단계 3. 수치해: 정지조건 $\max_{j \in S} |\tilde{\pi}_j^{(n)} - \tilde{\pi}_j^{(n-1)}| < \epsilon$ 만족하는 n 에 대하여 $\tilde{\pi}_j^* = \tilde{\pi}_j^{(n)}, j \in S$.

등확률 혹은 상수분포의 초기값과 $\epsilon = 10^{-4}$ 혹은 10^{-5} 정지조건이 적절하리라 본다.

한편 내재 마코프과정이 주기 $d \geq 2$ 를 가지는 경우에는 주기 효과를 다음과 같이 평활한다.

- 단계 2*. 반복식 및 평활식: 단계 2의 $\tilde{\pi}_j^{(n)}$ 과 $n \geq d - 1$ 대하여

$$\tilde{\pi}_j^{(n)*} = \frac{\tilde{\pi}_j^{(n)} + \tilde{\pi}_j^{(n-1)} + \dots + \tilde{\pi}_j^{(n-d+1)}}{d}, \quad j \in S.$$

- 단계 3*. 수치해: 정지조건 $\max_{j \in S} |\tilde{\pi}_j^{(n)*} - \tilde{\pi}_j^{(n-1)*}| < \epsilon$ 의 n 에 대한 $\tilde{\pi}_j^* = \tilde{\pi}_j^{(n)*}, j \in S$.

다만 주기적 마코프과정의 반복적 수치해가 근사하는 정상방정식의 정상확률 $\tilde{\pi}_j, j \in S$ 는 극한확률의 성질을 가지지 않는다.

내재 마코프과정의 극한확률 혹은 정상확률의 반복적 수치해를 결정하면 연속시간 마코프과정의 극한확률 $\pi_j, j \in S$ 에 대한 반복적 수치해는 다음의 과정으로 결정한다.

- 단계 4. 수치해: $\pi_j^* = \tilde{\pi}_j^* \lambda_j^{-1} / (\sum_{i \in S} \tilde{\pi}_i^* \lambda_i^{-1}), j \in S$.

다음 절에서 반복적 수치해 π_j^* 에 기초한 수치실험을 연구한다.

4. 시스템 성능비교를 위한 수치실험 연구

동일한 시스템에 대하여 서버배정방식을 달리할 때 성능척도의 변화를 살펴보고자 한다. 단순한 형태의 멀티클래스 손실시스템을 고려하기 위하여 $r = 2$ 종류의 고객이 시스템에 도착하는 것을 가정한다. $k_1 = 1, k_2 = 2$ 와 $M = 2, U = 2$ 즉 $c = 4$ 의 유니트 구조 시스템을 가정한다. 트래픽 모수 $\lambda_1, \lambda_2, \mu_1, \mu_2$ 의 값은 추후 계산과정에서 명시한다. 본 연구에서는 유니트 구조에서의 등확률배정, 순차배정, 최적배정과 기존의 완전랜덤배정 등 4가지 서버배정방식을 비교연구한다.

등확률 유니트 배정의 마코프과정 표현과 정상방정식은 Na (2010)에 잘 연구되어 있다. 마코프 확률 변수는 각 유니트의 종류별 서비스 중인 고객 수로 정의되는데 평형상태의 유니트 u 의 두 종류의 고객

수를 각각 N_1^u, N_2^u , $u = 1, 2$ 라 하면 N_1^u, N_2^u 의 가능한 값은 $(0, 0), (1, 0), (2, 0), (0, 1)$ 이며 각각 0, 1, 2, 3으로 표현한다. 따라서 2개의 유니트로 구성된 우리의 시스템은 총 16개의 상태가 가능하며 전체 시스템의 상태를 두 유니트의 상태를 붙여서 ij , $i, j = 0, 1, 2, 3$ 방법으로 표현한다. 상태전이도는 Figure 4.1(a)와 같고 다음의 정상방정식이 성립한다.

$$\begin{aligned}
\pi(00)(\lambda_1 + \lambda_2) &= (\pi(10) + \pi(01))\mu_1 + (\pi(03) + \pi(30))\mu_2, \\
\pi(10)(\lambda_1 + \lambda_2 + \mu_1) &= \pi(00)\frac{\lambda_1}{2} + \pi(11)\mu_1 + \pi(20)2\mu_1 + \pi(13)\mu_2, \\
\pi(01)(\lambda_1 + \lambda_2 + \mu_1) &= \pi(00)\frac{\lambda_1}{2} + \pi(11)\mu_1 + \pi(02)2\mu_1 + \pi(31)\mu_2, \\
\pi(20)(\lambda_1 + \lambda_2 + 2\mu_1) &= \pi(10)\frac{\lambda_1}{2} + \pi(21)\mu_1 + \pi(23)\mu_2, \\
\pi(02)(\lambda_1 + \lambda_2 + 2\mu_1) &= \pi(01)\frac{\lambda_1}{2} + \pi(12)\mu_1 + \pi(32)\mu_2, \\
\pi(11)(\lambda_1 + 2\mu_1) &= (\pi(10) + \pi(01))\frac{\lambda_1}{2} + (\pi(21) + \pi(12))2\mu_1, \\
\pi(21)(\lambda_1 + 3\mu_1) &= \pi(11)\frac{\lambda_1}{2} + \pi(20)\lambda_1 + \pi(22)2\mu_1, \\
\pi(12)(\lambda_1 + 3\mu_1) &= \pi(11)\frac{\lambda_1}{2} + \pi(02)\lambda_1 + \pi(22)2\mu_1, \\
\pi(22)4\mu_1 &= (\pi(21) + \pi(12))\lambda_1, \\
\pi(23)(2\mu_1 + \mu_2) &= \pi(13)\lambda_1 + \pi(20)\lambda_2, \\
\pi(32)(2\mu_1 + \mu_2) &= \pi(31)\lambda_1 + \pi(02)\lambda_2, \\
\pi(13)(\lambda_1 + \mu_1 + \mu_2) &= \pi(03)\lambda_1 + \pi(10)\lambda_2 + \pi(23)2\mu_1, \\
\pi(31)(\lambda_1 + \mu_1 + \mu_2) &= \pi(30)\lambda_1 + \pi(01)\lambda_2 + \pi(32)2\mu_1, \\
\pi(03)(\lambda_1 + \lambda_2 + \mu_2) &= \pi(00)\frac{\lambda_2}{2} + \pi(13)\mu_1 + \pi(33)\mu_2, \\
\pi(30)(\lambda_1 + \lambda_2 + \mu_2) &= \pi(00)\frac{\lambda_2}{2} + \pi(31)\mu_1 + \pi(33)\mu_2, \\
\pi(33)2\mu_2 &= (\pi(03) + \pi(30))\lambda_2.
\end{aligned}$$

앞의 정상방정식을 직접 풀어서 극한확률 $\pi(ij)$ 를 해석적으로 구하는 것은 매우 어려운 작업이며 이에 대한 해결책으로 3절의 반복적 수치해 방법을 적용한다. 등확률 배정의 내재된 이산시간 마코프과정의 극한확률 혹은 정상확률 $\tilde{\pi}(ij)$ 의 정상방정식을 다음과 같이 결정한다.

$$\begin{aligned}
\tilde{\pi}(00) &= (\tilde{\pi}(10) + \tilde{\pi}(01))\frac{\mu_1}{\lambda_1 + \lambda_2 + \mu_1} + (\tilde{\pi}(03) + \tilde{\pi}(30))\frac{\mu_2}{\lambda_1 + \lambda_2 + \mu_2}, \\
\tilde{\pi}(10) &= \tilde{\pi}(00)\frac{\lambda_1/2}{\lambda_1 + \lambda_2} + \tilde{\pi}(11)\frac{\mu_1}{\lambda_1 + 2\mu_1} + \tilde{\pi}(20)\frac{2\mu_1}{\lambda_1 + \lambda_2 + 2\mu_1} + \tilde{\pi}(13)\frac{\mu_2}{\lambda_1 + \mu_1 + \mu_2}, \\
\tilde{\pi}(01) &= \tilde{\pi}(00)\frac{\lambda_1/2}{\lambda_1 + \lambda_2} + \tilde{\pi}(11)\frac{\mu_1}{\lambda_1 + 2\mu_1} + \tilde{\pi}(02)\frac{2\mu_1}{\lambda_1 + \lambda_2 + 2\mu_1} + \tilde{\pi}(31)\frac{\mu_2}{\lambda_1 + \mu_1 + \mu_2}, \\
\tilde{\pi}(20) &= \tilde{\pi}(10)\frac{\lambda_1/2}{\lambda_1 + \lambda_2 + \mu_1} + \tilde{\pi}(21)\frac{\mu_1}{\lambda_1 + 3\mu_1} + \tilde{\pi}(23)\frac{\mu_2}{2\mu_1 + \mu_2}, \\
\tilde{\pi}(02) &= \tilde{\pi}(01)\frac{\lambda_1/2}{\lambda_1 + \lambda_2 + \mu_1} + \tilde{\pi}(12)\frac{\mu_1}{\lambda_1 + 3\mu_1} + \tilde{\pi}(32)\frac{\mu_2}{2\mu_1 + \mu_2},
\end{aligned}$$

$$\begin{aligned}
\tilde{\pi}(11) &= (\tilde{\pi}(10) + \tilde{\pi}(01)) \frac{\lambda_1/2}{\lambda_1 + \lambda_2 + \mu_1} + (\tilde{\pi}(21) + \tilde{\pi}(12)) \frac{2\mu_1}{\lambda_1 + 3\mu_1}, \\
\tilde{\pi}(21) &= \tilde{\pi}(11) \frac{\lambda_1/2}{\lambda_1 + 2\mu_1} + \tilde{\pi}(20) \frac{\lambda_1}{\lambda_1 + \lambda_2 + 2\mu_1} + \tilde{\pi}(22) \frac{2\mu_1}{4\mu_1}, \\
\tilde{\pi}(12) &= \tilde{\pi}(11) \frac{\lambda_1/2}{\lambda_1 + 2\mu_1} + \tilde{\pi}(02) \frac{\lambda_1}{\lambda_1 + \lambda_2 + 2\mu_1} + \tilde{\pi}(22) \frac{2\mu_1}{4\mu_1}, \\
\tilde{\pi}(22) &= (\tilde{\pi}(21) + \tilde{\pi}(12)) \frac{\lambda_1}{\lambda_1 + 3\mu_1}, \\
\tilde{\pi}(23) &= \tilde{\pi}(13) \frac{\lambda_1}{\lambda_1 + \mu_1 + \mu_2} + \tilde{\pi}(20) \frac{\lambda_2}{\lambda_1 + \lambda_2 + 2\mu_1}, \\
\tilde{\pi}(32) &= \tilde{\pi}(31) \frac{\lambda_1}{\lambda_1 + \mu_1 + \mu_2} + \tilde{\pi}(02) \frac{\lambda_2}{\lambda_1 + \lambda_2 + 2\mu_1}, \\
\tilde{\pi}(13) &= \tilde{\pi}(03) \frac{\lambda_1}{\lambda_1 + \lambda_2 + \mu_2} + \tilde{\pi}(10) \frac{\lambda_2}{\lambda_1 + \lambda_2 + \mu_1} + \tilde{\pi}(23) \frac{2\mu_1}{2\mu_1 + \mu_2}, \\
\tilde{\pi}(31) &= \tilde{\pi}(30) \frac{\lambda_1}{\lambda_1 + \lambda_2 + \mu_2} + \tilde{\pi}(01) \frac{\lambda_2}{\lambda_1 + \lambda_2 + \mu_1} + \tilde{\pi}(32) \frac{2\mu_1}{2\mu_1 + \mu_2}, \\
\tilde{\pi}(03) &= \tilde{\pi}(00) \frac{\lambda_2/2}{\lambda_1 + \lambda_2} + \tilde{\pi}(13) \frac{\mu_1}{\lambda_1 + \mu_1 + \mu_2} + \tilde{\pi}(33) \frac{\mu_2}{2\mu_2}, \\
\tilde{\pi}(30) &= \tilde{\pi}(00) \frac{\lambda_2/2}{\lambda_1 + \lambda_2} + \tilde{\pi}(31) \frac{\mu_1}{\lambda_1 + \mu_1 + \mu_2} + \tilde{\pi}(33) \frac{\mu_2}{2\mu_2}, \\
\tilde{\pi}(33) &= (\tilde{\pi}(03) + \tilde{\pi}(30)) \frac{\lambda_2}{\lambda_1 + \lambda_2 + \mu_2}.
\end{aligned}$$

한편 상태전이도에서 내재 마코프과정이 $d = 2$ 의 주기를 가지게 됨을 알 수 있는데 인접한 상태로 한 스텝씩 전이가 가능한 생성-소멸-과정의 일반적인 특성이기 때문이다. 따라서 단계 2*에 설명된 바와 같이 위의 이산시간 정상방정식의 우변에 값을 대입하여 좌변의 새로운 값을 계산하고 직전 값과 평활하여 $\tilde{\pi}(ij)$ 에 대한 근사값을 반복적으로 계산하는데 단계 3*의 정지조건을 만족할 때까지 계속한다. 마코프과정의 이론에 의하여 빠른 수렴속도를 기대하며 단계 4의 방법으로 연속시간 마코프과정의 극한확률 $\pi(ij)$ 의 수치해를 간단히 결정할 수 있다. 순차배정과 최적배정의 상태전이도는 각각 Figure 4.1(b), (c)와 같다. 앞의 등확률배정과 같이 상태전이도를 기초로 순차배정과 최적배정의 연속시간 및 이산시간 정상방정식을 수립하고 수치적 반복법으로 극한확률 $\pi(ij)$ 의 수치해를 구한다.

종류 1의 고객은 서비스를 위하여 $k_1 = 1$ 개 서버를 필요로 하고 모든 서버가 사용중일때 탈락이 발생하므로 탈락확률은 $B_1 = \pi(22) + \pi(23) + \pi(32) + \pi(33)$ 식으로 결정된다. 종류 2의 고객은 $k_2 = 2$ 개 서버를 요구하고 빈 유닛트가 있을때 서비스가 가능하므로 탈락확률은 $B_2 = 1 - \pi(00) - \pi(01) - \pi(02) - \pi(03) - \pi(10) - \pi(20) - \pi(30)$ 식과 같다.

유닛 구별이 없는 완전랜덤방식의 경우에는 시스템 전체에서 서비스 받는 각 종류 고객의 수가 마코프 확률변수이다. 상태전이도는 Figure 4.1(d)와 같으며 각 상태는 시스템에서 서비스 중인 두 종류 고객의 수이다. 유닛 구조의 경우와 다르게 해석적인 형태의 극한확률이 가능한데 평형상태의 두 종류의 고객의 수를 각각 N_1, N_2 라 하면 극한분포는 다음을 만족한다.

$$P(N_1 = n_1, N_2 = n_2) \propto p(n_1; 4, \rho_1)p(n_2; 2, \rho_2),$$

여기에서 음이 아닌 정수 n_1, n_2 는 $k_1 n_1 + k_2 n_2 = n_1 + 2n_2 \leq 4$ 식을 만족해야 하며 $\rho_1 = \lambda_1/\mu_1$,

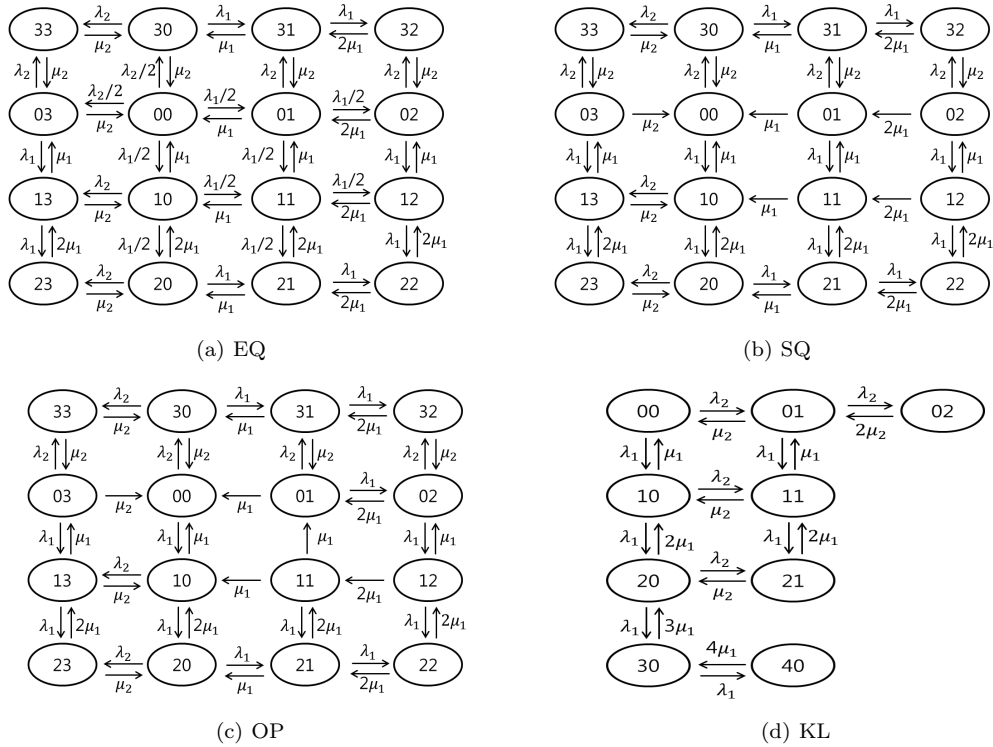


Figure 4.1. Transition diagrams.

$\rho_2 = \lambda_2/\mu_2$ 는 각 종류 고객의 트래픽 강도이다. 또

$$p(n; c, \rho) = \frac{\rho^n}{n!} \left(\sum_{j=0}^n \frac{\rho^j}{j!} \right)^{-1}, \quad n = 0, 1, \dots, c$$

식은 M/M/c/c 시스템의 극한확률이며 이에 대한 반복 계산식은 Na (2006)를 참고할 수 있다. 물론 이 절의 $c = 4$ 에 해당하는 시스템에서는 반복 계산식의 도움없이 직접적인 $p(n; c, \rho)$ 계산이 가능하지만 서버의 수가 커지면 완전랜덤방식의 정확한 극한분포 계산을 위하여 반복 계산식의 이용이 필수적이다. 종류 1 고객의 탈락확률은 $B_1 = P(N_1 + 2N_2 = 4)$, 즉 모든 서버가 사용될 확률과 같고, 종류 2 고객의 탈락확률은 $B_2 = P(N_1 + 2N_2 \geq 3)$, 즉 유휴서버가 2개 미만일 확률과 같다.

몇 가지 트래픽 환경을 가정하고 고객 종류별로 배정방식에 따른 탈락확률의 차이를 살펴본다. 먼저 트래픽 밀도의 단순 비교를 위하여 $\mu_1 = \mu_2 = 1$ 조건을 가정하고 따라서 $\rho_1 = \lambda_1$, $\rho_2 = \lambda_2$ 식이 성립한다. 한편 서비스 종류를 통합한 시스템 차원의 환산된 통합 트래픽을 $\tilde{\rho} = k_1\rho_1 + k_2\rho_2 = \rho_1 + 2\rho_2$ 식으로 정의하며 이를 기준으로 배정방식에 따른 탈락확률 차이를 살펴본다. 통합 트래픽의 값이 1얼랑 혹은 2얼랑이고 λ_1 과 λ_2 의 비율이 근사적으로 각각 1:2, 1:1, 2:1, 4:1일 때 B_1 , B_2 값을 수치적 반복법으로 구한 결과를 Table 4.1에 정리하였다. 여기에서 EQ, SQ, OP, KL은 각각 등확률배정, 순차배정, 최적배정, 완전랜덤배정을 의미한다. 통합 트래픽이 1얼랑의 경우에 B_1 값은 5% 이하이고 B_2 는 9.47%에서 14.5% 범위의 값을 가지는 것을 보게 된다. 한편 시스템 용량 4얼랑의 절반에 해당하는 2얼랑의 통합 트래픽을 가정하는 경우에는 B_1 , B_2 모두 큰 폭으로 증가한다.

Table 4.1. Loss probabilities of type 1 and type 2 customers

λ_1	λ_2	$\tilde{\rho}$	$B_1(\text{EQ})$	$B_1(\text{SQ})$	$B_1(\text{OP})$	$B_1(\text{KL})$	$B_2(\text{EQ})$	$B_2(\text{SQ})$	$B_2(\text{OP})$	$B_2(\text{KL})$
0.200	0.400	1	0.0484	0.0489	0.0491	0.0492	0.0992	0.0962	0.0952	0.0947
0.334	0.333	1	0.0371	0.0382	0.0386	0.0389	0.1128	0.1048	0.1025	0.1003
0.500	0.250	1	0.0282	0.0301	0.0305	0.0312	0.1288	0.1119	0.1080	0.1011
0.668	0.166	1	0.0225	0.0245	0.0249	0.0260	0.1450	0.1170	0.1113	0.0965
0.400	0.800	2	0.1249	0.1267	0.1278	0.1285	0.2481	0.2437	0.2408	0.2389
0.668	0.666	2	0.1021	0.1059	0.1079	0.1104	0.2800	0.2683	0.2623	0.2546
1.000	0.500	2	0.0893	0.0946	0.0970	0.1020	0.3196	0.2969	0.2867	0.2653
1.336	0.332	2	0.0854	0.0905	0.0927	0.0991	0.3618	0.3280	0.3134	0.2712

각 트래픽 조건에 따라 행을 살펴 보면 B_1 값은 EQ, SQ, OP, KL 방법의 순으로 증가하고 B_2 값은 감소한다. EQ의 경우에는 종류 1의 고객이 여러 유니트에 분산되어 서비스되는 경향이 있으므로 종류 2의 고객이 탈락할 가능성이 높아지고 따라서 B_2 값이 크고 반대로 B_1 값이 작다. 반면 SQ, OP 경우에는 종류 1의 고객이 집중되는 경향이 있는 관계로 종류 2의 고객의 서비스 성공이 증가하는 것으로 해석가능하고 유니트 구분이 없는 KL 경우에는 B_2 값이 더욱 작게되는 효과가 나타난다.

동일한 배정방식에 따라 열을 따라 탈락확률을 살펴 보면 λ_1 의 비율이 증가함에 따라 B_1 은 감소하고 B_2 는 증가하는 경향을 보임을 보게 된다. 종류 1의 고객 비율이 증가할수록 여러 유니트에 분포해서 서비스될 가능성이 커지고 종류 2의 고객의 서비스를 방해할 개연성이 커지는 것으로 해석가능하다. 그러나 $B_2(\text{KL})$ 의 값을 보면 이러한 경향이 절대적이지 않을 수 있는데 이는 통합 트래픽의 인위적인 정의에서 기인하는 것으로 보인다.

유니트 구조의 손실시스템에 대하여 여러 트래픽 환경을 가정하고 수치적 반복법을 적용하여 얻은 손실확률을 이용하여 서버배정방법의 차이를 비교하였다. 고객의 종류에 따른 탈락확률 등 시스템 성능척도에 서버배정방법이 작지 않은 차이를 야기할 수 있음을 제한적인 수치실험이 보여주고 있다. 앞에서 설명한 것과 같이 서버배정방법과 트래픽의 구성비율 변화에 따라 시스템 성능척도에 일정한 규칙의 변화가 나타나는데 이는 수치적 반복법의 적용을 통한 극한확률의 용이한 계산이 이루어짐으로 확인 가능하다.

5. 결론

본 연구는 현실세계를 더욱 효과적으로 반영하는 새로운 대기행렬시스템을 멀티클래스 손실시스템에 대하여 도입하며 이에 대한 확률통계적 특성을 규명한다. 현실성을 반영하기 위하여 유니트 구조를 가진 시스템을 정의하고 완전랜덤배정 이외에 등확률배정, 순차배정, 최적배정 등의 서버배정방법을 고려한 손실시스템의 마코프 모델링을 다루었고 서버배정방법에 따라 시스템 효율성에 차이가 발생하는가를 살펴보았다. 유니트 구조의 멀티클래스 손실시스템의 마코프 모델에 대한 해석적인 해는 가장 단순한 구조의 경우에도 현실적으로 불가능고 다만 수치적 해법이 가능하리라 판단되며 이 논문에서는 Na (2014)에서 소개된 수치적 반복법을 여러 서버배정방법에 대한 마코프 모델의 극한확률 계산에 적용하였다. 제한된 수치실험의 결과를 토대로 서버배정방법이 시스템 성능에 차이를 야기함을 확인하였고 서버의 수가 많이 소요되는 서비스는 완전랜덤배정, 최적배정, 순차배정, 등확률배정의 순서로 유리함을 볼 수 있었다.

마코프 모델의 반복적 수치해법의 실제 응용 사례로 다양한 서버배정방법의 유니트 구조 멀티클래스 손실시스템이 가능함을 알 수 있고 반복적 수치해의 계산을 통하여 다양한 큐잉시스템의 효율비교가 가능

해지리라 예상된다. 연구결과는 실시간 서비스를 제공하는 통신시스템의 성능분석 등에 활용되어 효율적인 시스템 설계의 근거가 될 수 있다. 본 논문에서 다루는 마코프 모델은 통신시스템 등의 모델링에 유용하리라 예상하고 반복적 수치해법은 시스템의 확률 특성의 규명을 위한 편리한 도구가 될 것이다. 실제 시스템 연구에 이 논문의 유니트 개념과 서버배정방식을 적용하는 것과 함께 새로운 구조 및 서버배정방식의 통신시스템 연구에 이 논문에서 사용된 마코프 기법과 배정방식의 개념을 확장 응용하고 또한 반복적 수치해법을 적용하는 것은 향후 흥미로운 연구 주제가 될 것이다.

References

- Kelly, F. P. (1991). Loss networks, *Annals of Applied Probability*, **1**, 319–378.
- Koo, I., Yang, J., Ahmad, A., and Kim, K. (2002). Erlang capacity analysis of hybrid FDMA/CDMA systems supporting multi-class services according to channel assignment methods, *International Journal of Communication Systems*, **15**, 867–880.
- Koo, I., Yang, J., and Kim, K. (2000). Analysis of Erlang capacity for DS-CDMA systems supporting multi-class services with the limited number of channel elements. In *Proceedings of IEEE Conference on Wireless Communications and Networking 2000, USA*, 355–359.
- Lee, J., Na, S., Choi, Y., and Lim, S. (2002). Performance analysis on channel distribution algorithm of selector in CDMA2000 1x system. In *Proceedings of Korea Information and Communications Society, Summer 2002*.
- Na, S. (2006). Exact performance evaluation of multi-class loss systems with applications to telecommunications networks, *Telecommunications Review*, **16**, 1084–1091.
- Na, S. (2010). Markov modeling of multiclass loss systems, *The Korean Journal of Applied Statistics*, **23**, 747–757.
- Na, S. (2014). Numerical iteration for stationary probabilities of Markov chains, *Communications for Statistical Applications and Methods*, **21**, 513–520.
- Nesterov, Y. and Nemirovski, A. (2015). Finding the stationary states of Markov chains by iterative methods, *Applied Mathematics and Computation*, **255**, 58–65.
- O’Leary, D. P. (1993). Iterative methods for finding the stationary vector for Markov chains, *Linear Algebra, Markov Chains, and Queueing Models* (C.D. Meyer and R.J. Plemmons (ed.)), The IMA volumes in mathematics and its applications, **48**, 125–136, Springer, New York.
- Ross, K. W. (1995). *Multiclass Loss Models for Broadband Telecommunication Networks*, Springer, London.
- Ross, S. M. (1996). *Stochastic Processes* (2nd Ed), Wiley, New York.
- Stewart, W. J. (2000). Numerical methods for computing stationary distributions of finite irreducible Markov chains. In W.K. Grassmann (Ed.), *Computational Probability*, International series in operations research and management science, **24**, 81–111, Springer, New York.
- Zhao, D., Li, H., and Su, D. (2012). A numerical algorithm on the computation of the stationary distribution of a discrete time homogeneous finite Markov chain, *Mathematical Problems in Engineering*, **2012**.

여러 서버배정방식의 멀티클래스 손실시스템 연구

나성룡^{a,1}

^a연세대학교 정보통계학과

(2016년 3월 4일 접수, 2016년 4월 15일 수정, 2016년 5월 9일 채택)

요약

이 논문에서는 서로 다른 서버배정방식을 가지는 멀티클래스 손실시스템을 연구한다. 마코프 성질을 유지하도록 상태를 정의하고 상태를 효율적으로 표현하는 방법을 살펴본다. 시스템 성능척도를 산출하기 위하여 마코프 성질에 기초한 극한확률 유도를 연구한다. 해석해 혹은 수치해를 구하여 배정방식이 시스템 성능에 미치는 영향을 비교한다.

주요용어: 멀티클래스 손실시스템, 서버배정방식, 마코프 모델링, 수치적 반복법, 성능분석, 손실확률

이 논문은 2011학년도 연세대학교 학술연구비의 지원에 의하여 이루어진 것임

¹(26493) 강원도 원주시 연세대길 1, 연세대학교 정보통계학과. E-mail: nasr@yonsei.ac.kr