Journal of the Korea Industrial Information Systems Research Vol. 22 No.1, Feb. 2017 :23-32
http://dx.doi.org./10.9723/jksiis.2017.22.1.023 ISSN:1229-3741

일량분석에 의한 이산시간 MAP/G/1 대기행렬시스템의 통합적 분석

(A Unified Approach for the Analysis of Discrete-time MAP/G/1 Queue: by Workload Analysis)

이 세 원[1)*]
(Se Won Lee)

요 약 본 논문에서는 이산시간 마코비안 도착과정(discrete-time Markovian arrival process, 이하 이산시간 MAP)을 갖는 대기행렬시스템의 주요성능척도들을 분석하기 위한 통합적 접근방법을 제시한다. 기존의 이산시간 MAP/G/1 대기행렬시스템의 연구들을 보면 동일한 시스템에 대하여 시스템 내 고객수와 대기시간 등을 분석할 때 서로 다른 방법으로 접근하였기에 이 둘을 동시에 살펴보고자 할 때는 추가적인 시간과 노력이 뒤따랐다. 따라서 하나의 시스템을 여러 방면에서 포괄적으로 분석할 수 있는 통합적인 접근방법은 시스템을 설계하고 관리하는 입장에서 볼 때 중요한 분석의 틀이 된다. 본 논문에서는 이산시간 MAP/G/1 시스템의 안정상태 일량 분포를 유도하고 이를 이용하여 임의고객의 대기시간, 체재시간 분포를 유도한다. 체재시간 분포로부터 이탈시점 고객수 분포를 구하고 이탈시점 고객수와 임의시점 고객수와의 관계로부터 고객수 분포를 유도한다.

핵심주제어 : 이산시간 시스템, 이산시간 MAP/G/1 대기행렬시스템, 통합적 접근방법

Abstract In this paper, We suggest a unified approach for the analysis of discrete-time MAP/G/1 queueing system. Many researches on the D-MAP/G/1 queue have been used different approach to analyze system queue length and waiting time for the same system. Therefore, a unified framework for analyzing a system is necessary from a viewpoint of system design and management. We first derived steady-state workload distribution, and then waiting time and sojourn time are derived by the result of workload analysis. Finally, system queue length distribution is derived with generating function from the sojourn time distribution.

Key Words : Discrete-time System, Discrete-time Markovian Arrival Process (D-MAP)/G/1 queue, Unified Approach

1. 서 론

사건 발생 간격들 사이에 상관성을 갖는 마코비안 도착과정(Markovian arrival process, 이하 MAP)은 Neuts[1]에 의해 소개된 이후 통신네트워크 트래픽의 분석에 폭넓게 사용되어 왔다 [2-4]. MAP은 매우 유연한 도착과정(확률과정)

* Corresponding Author : swlee@pknu.ac.kr
† 이 논문은 부경대학교 자율창의학술연구비(2014년)에 의하여 연구되었음(C-D-2014-1075)
Manuscript received Jan, 2, 2017 / revised Feb, 13, 2017 / accepted Feb, 25, 2017
1) 부경대학교 경영학부

으로서, 대기행렬모형에서 주로 사용되어 온 포아송 도착과정(Poisson arrival process)과 여러 재생과정들을 표현할 수 있는 일반성을 가지며, 자연적으로 많이 발생하는 bursty traffic 모델링의 가능성, M/G/1 관련시스템 분석에서 사용한 방법론의 확장성으로 인하여 널리 사용되고 있다[5]. 더구나 기존의 Poisson arrival process로는 bursty traffic의 도착간격들 사이의 상관성을 모델링하는 데에 한계가 있음이 Paxson and Floyd[6]에 의해 알려지면서 그 필요성은 더욱 높아지고 있다.

도착과정이 MAP을 따르는 연속시간, 이산시간 대기행렬시스템의 분석에 관한 연구들을 살펴보면 초기 연구들은 도착과정 그 자체로서 큰 의의를 지니는 연구들이 이루어졌고[1,7], 이 후 시스템 모델링을 통한 성능척도 분석, 복잡한 행렬 계산과 성능척도를 빠르게 구하기 위한 알고리즘 연구, 실제 데이터를 바탕으로 한 파라미터 추정에 관한 연구들이 주를 이룬다[8-10]. 이 중 시스템 모델링을 통한 성능척도 분석에서 사용하는 대표적인 방법으로는 행렬분석법(matrix analytic method), 부가변수법(supplementary variable technique), 분해성질(factorization)을 이용한 분석방법을 들 수 있다[11].

1960년대 Cinlar[12]와 Neuts[13]의 대기행렬시스템 연구와 Wallace[14]의 컴퓨터 모델 성능분석에서 처음으로 사용된 행렬분석법은 대상이 되는 시스템에 대한 비선형 행렬방정식(non-linear matrix equation)의 해를 구하는데 초점을 두고 있다. 이는 MAP 도착과정과 같이 고객의 도착과정에 잠재마코프체인(underlying Markov chain, 이하 UMC)과정과 같은 다른 과정이 포함되어 있는 이중 확률과정(doubly stochastic process)을 갖는 시스템의 분석에 용이한 방법으로서 지금도 MAP 관련 대기행렬시스템의 분석에 주로 사용하고 있다. 행렬분석법을 이용한 연구들은 연속시간, 이산시간 MAP/G/1 관련 대기행렬시스템의 근간이 되었지만, 성능척도를 계산하는 알고리즘에 초점이 맞추어져 있어 분석의 결과로부터 시스템의 운영특성을 파악하는데 어려움이 있고, 이탈시점 고객수를 분석하는 것을 시작으로 하여 임의시점의 고객수, 대기시간 분포를 따로 유도해야 하는 번거로움이 있다는 한계를 갖는다. 행렬분석법을 이용한 대기행렬시스템의 분석은 [13,15-18]에서 찾아 볼 수 있다.

Cox[19]에 의해 처음으로 도입되어 Hokstad[20]에 의하여 대기행렬시스템에 최초로 적용된 부가변수법은 마코비안이 아닌 확률과정에 어떤 변수를 추가하여 마코비안 확률과정으로 변환하여 분석하는 방법인데, 이를 이용한 연구들은 잔여(또는 경과) 서비스시간을 이용하여 시스템방정식을 세운 후 시스템을 분석한다. 임의시점에서의 고객수, 대기시간을 구할 수 있으며 주로 대기시간 분포를 구하는데 사용되었다. 이 분석방법은 중간에 유도되는 식들에 의해 다양한 성능척도들을 부산물로 얻을 수 있다는 장점이 있으나 분석을 위해서는 시스템방정식을 세워야 하고 이 방정식을 풀기가 간단하지가 않다는 단점도 가지고 있다[11].

Lee et al.[21]에 의해 최초로 발견된 분해성질은 일반형 휴가를 갖는 MAP/G/1 관련 대기행렬에서 임의시점에서의 고객수(또는 이탈시점에서의 고객수)를 임의의 유휴시점에서의 고객수와 다른 어떤 고객수의 합의 형태로 간단하게 표현할 수 있다는 성질이다. Lee et al.[21]은 고유값과 고유벡터를 이용하여 잔여서비스시간을 추가변수로 한 부가변수법으로부터 분해성질을 발견하였다. 이 성질을 이용하면 임의의 유휴기간의 고객수 분포로부터 손쉽게 임의시점(또는 이탈시점) 고객수 분포, 대기시간을 직접 구할 수 있다는 장점이 있어 여러 연구를 통해 사용, 증명되었다[22-25]. 하지만 바쁜기간 시작점에서 대기중인 고객들의 서비스시간이 서로 독립이어야 한다는 가정이 충족될 때만 사용할 수 있다는 한계가 있어 일량에 의해 제어되는 서버제어정책(예를 들면, D-정책)을 갖는 대기행렬시스템의 분석에서는 사용할 수 없다.

어떤 시스템을 분석할 때, 동일한 시스템임에도 불구하고 고객수와 대기시간의 분석에 있어 서로 다른 접근방법을 사용함으로써 추가적인 시간과 노력을 동반해야 하는 번거로움으로 인하여 한 시스템을 여러 방면에서 포괄적으로 분석하기 위해 통합적인 접근방법은 필요하다. 유휴기간 동안의 일량을 분석하여 시스템의 고객수, 대기

시간, 체재시간, 일량을 분석한 통합적 접근방법은 최초 Lee et al.[26]에 의해 연속시간 시스템인 일량제어정책을 갖는 M/G/1 시스템을 대상으로 소개되었다. 이후 Lee et al.[27]은 이산시간 시스템을 고려하여 일량제어정책을 갖는 Geo/G/1 시스템의 분석에도 유용하게 사용될 수 있음을 보였고, Lee et al.[28]은 일량제어정책을 갖는 이산시간 MAP/G/1 시스템의 분석에서 사용하였다. 본 논문의 접근방식은 분해성질을 이용한 방법에서 바쁜기간 시작점에서의 대기고객들의 서비스시간이 독립이어야 한다는 가정이 깨지더라도 유효하게 적용할 수 있는 유연한 분석방법이라 할 수 있다.

본 논문에서 다루는 내용 중에서 고객수와 대기시간의 분포에 관한 최종 결과들은 이미 각기 서로 다른 접근방법들을 통해 소개된 것이다. 그러나 이산시간 MAP/G/1 대기행렬시스템을 대상으로 일량의 분석과정을 자세히 다루고, 그로부터 시스템 주요성능척도 전반에 대해 분석한 논문은 Lee et al.[28] 외에는 찾아보기 어렵다. 따라서, 기초모델이라 할 수 있는 이산시간 MAP을 도착과정으로 갖는 대기행렬시스템의 자세한 분석절차를 일량분석을 바탕으로 하여 소개한다. 단, 평균일량, 평균대기시간, 평균체재시간, 시스템 내 평균고객수와 같은 성능척도들을 계산하는 방법들은 다소 복잡하지만 이미 잘 알려져 있는 계산 알고리즘(Lee[11])에 의해서 구할 수 있으므로 본 연구에서는 자세히 다루지 않는다.

2. 시스템 분석

본 논문에서 제시하는 접근방법은 안정상태 일량(workload, unfinished work) 분석을 시작으로 한다. 그 이유는 일단 일량 분석이 완료되면 이를 통해 대기시간, 체재시간의 분포를 비교적 간단히 유도할 수 있기 때문이다. 그리고, 이탈시점 고객수는 임의 고객의 체재시간과 밀접한 관계를 맺고 있으므로 체재시간의 분석결과로부터 이탈시점 고객수 분석을, 이탈시점 고객수 분석을 통해 최종적으로 임의시점 고객수 분석을 할 수 있다. 본격적인 분석에 앞서 논문 전반에서 사용할 확률과 기호들을 다음과 같이 정의하자.

$\boldsymbol{D}_0$, $\boldsymbol{D}_1$: 이산시간 MAP에서 underlying Markov chain(UMC)의 파라미터 행렬

m : UMV 위상의 개수

e : 모든 원소가 1인 열벡터(column vector)

$\boldsymbol{\pi}=(\pi_1,\pi_2,\ldots,\pi_m)$: UMC 위상의 정상확률벡터 (stationary probability vector)

$\boldsymbol{M}_{i,j}$: 행렬 $\boldsymbol{M}$의 (i,j)원소

$\lambda=\boldsymbol{\pi}\boldsymbol{D}_1\mathrm{e}$: 단위시간당 도착 고객수

S : 서비스시간(이산확률변수)

$s(x)=\Pr(S=x)$: S의 확률질량함수(pmf, probability mass function), $s(0)=0$.

$S^*(\omega)=\sum_{x=1}^{\infty}\omega^x s(x)$: S의 확률생성함수(PGF, probability generating function)

$E(\cdot)$: 확률변수의 평균, 기대치(expectation)

$\rho=\lambda E(S)$: 교통밀도(traffic intensity)

2.1 일량 분석

본 절에서는 D-MAP/G/1 대기행렬시스템의 안정상태 일량분포의 벡터생성함수(GF, generating function)를 유도한다. 일량 분석을 위해 다음과 같이 기호와 확률을 정의하자.

$U(k)$: k번째 슬롯 시작점에서의 일량

$J(k)$: k번째 슬롯 시작점에서의 UMC 위상

$u_i^{idle}(k)=\Pr[U(k)=0,\ J(k)=i]$

$u_{x,i}^{busy}(k)=\Pr[U(k)=x,\ J(k)=i,$ server is busy$]$, $(x=1,2,\ldots)$

$\boldsymbol{u}^{idle}(k)=\left(u_1^{idle}(k),u_2^{idle}(k),\ldots,u_m^{idle}(k)\right)$

$\boldsymbol{u}^{busy}(k)=\left(u_{1,x}^{busy}(k),u_{2,x}^{busy}(k),\ldots,u_{m,x}^{busy}(k)\right)$

$\boldsymbol{u}^{idle}=\lim_{k\to\infty}\boldsymbol{u}^{idle}(k)$

$\boldsymbol{u}_x^{busy}=\lim_{k\to\infty}\boldsymbol{u}_x^{busy}(k)$

본 논문에서 다루는 이산시간 시스템은 LAS-DA(late arrival system with delayed access)이므로 슬롯 중간에 도착한 고객은 슬롯 끝점에서 도착하는 것으로 가정하고 서비스시간은 1 이상

임을 고려하여 한 슬롯 동안의 일량 과정을 추적하면 위의 정의를 이용하여 다음과 같이 상태전이방정식을 얻는다. 이산시간 시스템 중 하나인 LAS-DA에 대한 자세한 설명은 Hunter[29]를 참고하기 바란다.

$$\begin{aligned} & u_i^{idle}(k+1) \\ &= u_i^{idle}(k)(\boldsymbol{D}_0)_{i,i} + \sum_{\substack{j=1 \\ j\neq i}}^{m} u_j^{idle}(k)(\boldsymbol{D}_0)_{j,i} \\ &+ u_{1,i}^{busy}(k)(\boldsymbol{D}_0)_{i,i} + \sum_{\substack{j=1 \\ j\neq i}}^{m} u_{1,j}^{busy}(k)(\boldsymbol{D}_0)_{j,i} \end{aligned} \tag{1}$$

$$\begin{aligned} & u_{0,i}^{busy}(k+1) \\ &= u_{1,i}^{busy}(k)(\boldsymbol{D}_0)_{i,i} + \sum_{\substack{i=1 \\ j\neq i}}^{m} u_{1,j}^{busy}(k)(\boldsymbol{D}_0)_{j,i} \end{aligned} \tag{2.1}$$

$$\begin{aligned} & u_{x-1,i}^{busy}(k+1) \\ &= u_i^{idle}(k)(\boldsymbol{D}_1)_{i,i} s(x-1) \\ &+ \sum_{\substack{j=1 \\ j\neq i}}^{m} u_j^{idle}(k)(\boldsymbol{D}_1)_{j,i} s(x-1) \\ &+ u_{x,i}^{busy}(k)(\boldsymbol{D}_0)_{i,i} + \sum_{\substack{i=1 \\ j\neq i}}^{m} u_{x,j}^{busy}(k)(\boldsymbol{D}_0)_{j,i} \\ &+ \sum_{y=1}^{x-1} u_{y,i}^{busy}(k)(\boldsymbol{D}_1)_{i,i} s(x-y) \\ &+ \sum_{y=1}^{x-1} \sum_{\substack{i=1 \\ j\neq i}}^{m} u_{y,j}^{busy}(k)(\boldsymbol{D}_1)_{j,i} s(x-y), \\ & \quad (x=2,3,\ldots) \end{aligned} \tag{2.2}$$

식 (1), (2.1), (2.2)의 벡터방정식은 다음과 같다.

$$\boldsymbol{u}^{idle}(k+1) = \boldsymbol{u}^{idle}(k)\boldsymbol{D}_0 + \boldsymbol{u}_1^{busy}(k)\boldsymbol{D}_0 \tag{3}$$

$$\boldsymbol{u}_0^{busy}(k+1) = \boldsymbol{u}_1^{busy}(k)\boldsymbol{D}_0 \tag{4.1}$$

$$\begin{aligned} \boldsymbol{u}_{x-1}^{busy}(k+1) &= \boldsymbol{u}^{idle}(k)\boldsymbol{D}_1 s(x-1) \\ &+ \boldsymbol{u}_x^{busy}(k)\boldsymbol{D}_0 \\ &+ \sum_{y=1}^{x-1} \boldsymbol{u}_y^{busy}(k)\boldsymbol{D}_1 s(x-y), \\ & \quad (x=2,3,\ldots) \end{aligned} \tag{4.2}$$

식 (3), (4.1), (4.2)에 $k\to\infty$ 으로 극한을 취하여 안정상태 방정식을 구하면

$$\boldsymbol{u}^{idle} = \boldsymbol{u}^{idle}\boldsymbol{D}_0 + \boldsymbol{u}_1^{busy}\boldsymbol{D}_0 \tag{5}$$

$$\boldsymbol{u}_0^{busy} = \boldsymbol{u}_1^{busy}\boldsymbol{D}_0 \tag{6.1}$$

$$\begin{aligned} \boldsymbol{u}_{x-1}^{busy} &= \boldsymbol{u}^{idle}\boldsymbol{D}_1 s(x-1) + \boldsymbol{u}_x^{busy}\boldsymbol{D}_0\ . \\ &+ \sum_{y=1}^{x-1} \boldsymbol{u}_y^{busy}\boldsymbol{D}_1 s(x-y), \\ & \quad (x=2,3,\ldots) \end{aligned} \tag{6.2}$$

이제 임의 슬롯시작점에서의 일량과 UMC 위상에 대한 정보를 포함하는 벡터 생성함수(GF)를 다음과 같이 정의하자.

$$\begin{aligned} \boldsymbol{u}^*(\omega) &= \boldsymbol{u}_{idle}^*(\omega) + \boldsymbol{u}_{busy}^*(\omega) \\ &= \omega^0 \boldsymbol{u}^{idle} + \sum_{x=1}^{\infty} \omega^x \boldsymbol{u}_x^{busy} \end{aligned} \tag{7}$$

식 (5)와 식 (6.1)로부터 유휴기간 임의슬롯 시작점에서의 UMC 위상은 다음과 같다.

$$\boldsymbol{u}^{idle} = \boldsymbol{u}_1^{busy}\boldsymbol{D}_0(\boldsymbol{I}-\boldsymbol{D}_0)^{-1} = \boldsymbol{u}_0^{busy}(\boldsymbol{I}-\boldsymbol{D}_0)^{-1} \tag{8}$$

식 (6.1), (6.2)의 양변에 ω^x를 곱하고 x에 대하여 모두 더하여 정리하면,

$$\begin{aligned} & \omega\boldsymbol{u}_0^{busy} + \sum_{x=2}^{\infty} \omega^x \boldsymbol{u}_{x-1}^{busy} \\ &= \sum_{x=2}^{\infty} \omega^x \boldsymbol{u}^{idle}\boldsymbol{D}_1 s(x-1) \\ &+ \sum_{x=1}^{\infty} \omega^x \boldsymbol{u}_x^{busy}\boldsymbol{D}_0 + \sum_{x=2}^{\infty} \omega^x \sum_{y=1}^{x-1} \boldsymbol{u}_y^{busy}\boldsymbol{D}_1 s(x-y) \end{aligned}$$

$$\begin{aligned} & \omega\boldsymbol{u}_0^{busy} + \omega\boldsymbol{u}_{busy}^*(\omega) \\ &= \boldsymbol{u}^{idle}\omega\boldsymbol{D}_1 S^*(\omega) + \boldsymbol{u}_{busy}^*(\omega)\boldsymbol{D}_0 \\ &+ \boldsymbol{u}_{busy}^*(\omega)\boldsymbol{D}_1 S^*(\omega). \end{aligned} \tag{9}$$

식 (9)에 식 (8)을 대입하여 정리하면 바쁜기간 임의슬롯 시작점에서의 일량과 UMC 위상에 관한 벡터생성함수는 다음과 같다.

$$\boldsymbol{u}^*_{busy}(\omega)=\omega\,\boldsymbol{u}_0^{busy}\big((\boldsymbol{I}-\boldsymbol{D}_0)^{-1}\boldsymbol{D}_1S^*(\omega)-\boldsymbol{I}\big) \times\left[\omega\boldsymbol{I}-\boldsymbol{D}_0-S^*(\omega)\boldsymbol{D}_1\right]^{-1} \quad (10)$$

식 (7)에 식 (8)과 식 (10)을 대입하여 정리하면

$$\begin{aligned}&\boldsymbol{u}^*(\omega)\\&=\boldsymbol{u}_0^{busy}\Big\{(\boldsymbol{I}-\boldsymbol{D}_0)^{-1}\\&\qquad+\omega\big((\boldsymbol{I}-\boldsymbol{D}_0)^{-1}\boldsymbol{D}_1S^*(\omega)-\boldsymbol{I}\big)\\&\qquad\quad\times\left[\omega\boldsymbol{I}-\boldsymbol{D}_0-\boldsymbol{D}_1S^*(\omega)\right]^{-1}\Big\}\\&=\boldsymbol{u}_0^{busy}(\boldsymbol{I}-\boldsymbol{D}_0)^{-1}(\omega-1)\\&\quad\times\left[\boldsymbol{D}_0+\boldsymbol{D}_1S^*(\omega)\right]\left[\omega\boldsymbol{I}-\boldsymbol{D}_0-S^*(\omega)\boldsymbol{D}_1\right]^{-1}\end{aligned} \quad (11)$$

식 (10)을 완성하기 위해서는 미지수벡터인 $\boldsymbol{u}_0^{busy}$를 구해야 한다. 서버가 바쁜 동안 시스템 내 일량이 0일 확률은 0이므로, 독자들은 벡터 $\boldsymbol{u}_0^{busy}$의 i번째 원소가 '서버가 바쁘고, 시스템 내 일량이 0이고, UMC 위상이 i일 확률'이 아님을 상기하기 바란다. 이 미지수벡터는 단위시간당 바쁜기간이 끝나는 평균횟수와 바쁜기간 종료점에서의 UMC 위상에 대한 정보를 동시에 담고 있는 벡터이다.

정리 1. $\boldsymbol{u}_0^{busy}=\dfrac{(1-\rho)\boldsymbol{\kappa}}{E(I)}$ **이다.**

(증명) $\boldsymbol{u}_0^{busy}$는 그 의미를 따져 보면, 단위시간당 바쁜기간이 끝나는 평균횟수와 그 때의 UMC 위상을 나타내는 벡터이다. 임의의 사이클의 평균길이를 $E(C)$라고 하면 단위시간당 바쁜기간이 끝나는 평균횟수는 $\dfrac{1}{E(C)}$이고, 임의의 바쁜기간 종료점에서의 UMC 위상을 나타내는 벡터를 $\boldsymbol{\kappa}$라고 하면 $\boldsymbol{u}_0^{busy}=\dfrac{\boldsymbol{\kappa}}{E(C)}$임을 알 수 있다. 또한 이산시간 MAP/G/1 대기행렬시스템은 일량보존시스템(work conserving system)이므로 사이클의 평균길이와 유휴기간의 평균길이 $E(I)$ 사이에 다음의 관계가 성립한다: $(1-\rho)E(C)=E(I)$. 따라서, $\boldsymbol{u}_0^{busy}=\dfrac{\boldsymbol{\kappa}}{E(C)}=\dfrac{(1-\rho)\boldsymbol{\kappa}}{E(I)}$이다. 여기서 유휴기간의 평균길이 $E(I)=\boldsymbol{\kappa}(\boldsymbol{I}-\boldsymbol{D}_0)^{-1}\mathrm{e}$ 이다.

<정리 1>로부터 임의슬롯 시작점에서의 일량분포에 대한 결합벡터생성함수는 식 (12)와 같이 정리된다. <정리 1>에서 언급한 $\boldsymbol{\kappa}$는 고객수 분석을 할 때 구하기로 한다.

$$\begin{aligned}&\boldsymbol{u}^*(\omega)\\&=\frac{(1-\rho)\boldsymbol{\kappa}}{E(I)}(\boldsymbol{I}-\boldsymbol{D}_0)^{-1}(\omega-1)\left[\boldsymbol{D}_0+\boldsymbol{D}_1S^*(\omega)\right]\\&\quad\times\left[\omega\boldsymbol{I}-\boldsymbol{D}_0-S^*(\omega)\boldsymbol{D}_1\right]^{-1}\end{aligned} \quad (12)$$

이산시간 MAP은 여러 가지 도착과정을 특수경우로 갖는 유연한 도착과정이다. 따라서 $\boldsymbol{D}_0=1-\lambda$, $\boldsymbol{D}_1=\lambda$, $m=1$인 경우, 이산시간 MAP/G/1 시스템의 일량 분포는 단위시간(슬롯)당 도착확률이 λ인 Geo/G/1 대기행렬시스템의 그것과 동일하다. 식 (12)로부터 확인해 보자. Geo/G/1 시스템에서는 UMC 위상이 하나 밖에 없으므로 $\boldsymbol{\kappa}=1$이다. 따라서,

$$\begin{aligned}u^*(\omega)&=\frac{(1-\rho)}{1/\lambda}(1-(1-\lambda))^{-1}(\omega-1)\\&\quad\times\left[1-\lambda+\lambda S^*(\omega)\right]\\&\quad\times\left[\omega-(1-\lambda)-\lambda S^*(\omega)\right]^{-1}\\&=(1-\rho)\frac{(\omega-1)(1-\lambda+\lambda S^*(\omega))}{\omega-1+\lambda-\lambda S^*(\omega)}\end{aligned} \quad (13)$$

식 (13)은 Takagi[30] p.27의

$$U^*_{Geo/G/1}(\omega)=\frac{(1-\rho)(1-\omega)(1-\lambda+\lambda S^*(\omega))}{1-\lambda+\lambda S^*(\omega)-\omega}$$ 와

동일함을 확인할 수 있다.

2.2 대기시간 분석

본 절에서는 2.1절의 일량분석 결과로부터 가상대기시간을 유도하여 정리하고 이를 바탕으로 실제대기시간분포를 유도한다. 안정상태에서 임의고객의 대기시간분포의 벡터생성함수는 다음의 관찰을 통해 유도할 수 있다. 먼저 가상대기시간을 살펴보자.

관찰 1. 이산시간 MAP/G/1 시스템에서 가상고객의 대기시간 W_q^V는 임의시점에서의 일량과 한 슬롯만큼 차이가 난다. 정확히 말하면 한 슬롯의 길이만큼 작다. 참고로 연속시간 MAP/G/1 대기행렬시스템에서는 임의시점에서의 일량과 가상고객의 대기시간이 같다.

관찰 2. 실제고객이 도착하면서 보는 고객수 확률은 임의시점 고객수 확률에 D_1/λ를 곱함으로써 얻을 수 있다. 따라서 실제대기시간의 벡터생성함수 $W_q^*(\omega)$ 또한 가상대기시간의 벡터생성함수에 D_1/λ를 곱하여 얻을 수 있다.

가상 도착고객의 대기시간 벡터생성함수 $W_q^{V*}(\omega)$는 (관찰 1)과 <정리 1>, 식 (10)을 이용하여 식(14)와 같이 유도할 수 있다.

$$\begin{aligned} W_q^{V*}(\omega) &= W_{q,idle}^{V*}(\omega) + W_{q,busy}^{V*}(\omega) \\ &= u_{idle}^*(\omega) + \omega^{-1} u_{busy}^*(\omega) \\ &= \frac{(1-\rho)\kappa}{E(I)}\Big\{(I-D_0)^{-1} \\ &\quad + \big((I-D_0)^{-1} D_1 S^*(\omega) - I\big) \\ &\quad\quad \times \left[\omega I - D_0 - D_1 S^*(\omega)\right]^{-1}\Big\} \\ &= \frac{(1-\rho)\kappa}{E(I)}(I-D_0)^{-1}(\omega-1) \\ &\quad \times \left[\omega I - D_0 - D_1 S^*(\omega)\right]^{-1} \end{aligned} \tag{14}$$

식 (14)에 $D_0 = 1-\lambda$, $D_1 = \lambda$, $m=1$, $\kappa=1$을 대입하면, Geo/G/1 시스템에서 가상대기시간의 변환을 얻는다. 이는 Takagi[30]의 결과와 동일하다.

$$W_{q,Geo/G/1}^V(\omega) = \frac{(1-\rho)(1-\omega)}{1-\lambda+\lambda S^*(\omega)-\omega}$$

이제 (관찰 2)와 식 (12)로부터 실제 도착고객의 대기시간 벡터생성함수는 다음과 같다.

$$\begin{aligned} W_q^*(\omega) &= W_q^{V*}(\omega)\frac{D_1}{\lambda} \\ &= \frac{(1-\rho)\kappa}{\lambda E(I)}(I-D_0)^{-1}(\omega-1) \\ &\quad \times \left[\omega I - D_0 - D_1 S^*(\omega)\right]^{-1} D_1 \end{aligned} \tag{15}$$

서버 제어정책을 갖는 시스템에서는 유휴기간 동안 도착하는 고객(특별고객(special customer)이라 하자)의 가상대기시간이 도착하면서 보는 시스템 내 일량과 다르다. 따라서 이런 경우에는 식 (15)를 그대로 사용할 수 없다. 그러나 2.1절의 일량 분석에서 유휴기간 동안의 일량과 바쁜기간 동안의 일량을 따로 분석하였으므로 고객의 유형을 유휴기간 동안에 도착하는 고객(특별고객)과 바쁜기간 동안에 도착하는 고객(보통고객)으로 나누어 식 (16)와 같이 표현하면 큰 문제가 되지 않는다. 이는 체재시간과 고객수 분석에도 동일하게 적용된다.

$$W_q^*(\omega) = W_{q,sc}^*(\omega) + W_{q,oc}^*(\omega) \tag{16}$$

2.3 체재시간 분석

안정상태에서 임의고객의 체재시간은 그 고객의 대기시간과 서비스시간의 합이다. 그리고 이산시간 MAP/G/1 대기행렬시스템에서도 이 둘은 확률적으로 서로 독립이다. 따라서 벡터생성함수 사이에는 다음의 관계가 성립한다.

$$\begin{aligned} W^*(\omega) &= W_q^*(\omega) S^*(\omega) \\ &= \frac{(1-\rho)\kappa}{\lambda E(I)}(I-D_0)^{-1}(\omega-1) \\ &\quad \times \left[\omega I - D_0 - D_1 S^*(\omega)\right]^{-1} D_1 S^*(\omega) \end{aligned} \tag{17}$$

2.4 고객수 분석

본 절에서는 안정상태 임의슬롯 시작점에서 시스템 내 고객수 분포의 벡터생성함수를 유도한다. Kim et al.[31]은 이산시간 BMAP/G/1 대기행렬시스템에서 이탈시점 고객수 $X(z)$와 임의시점 고객수 $Y(z)$ 사이에 성립하는 관계식을 증명하였다. 식 (18)은 Kim et al.[31]의 이산시간 MAP 형태이다.

$$Y(z)(D_0 + D_1 z - I) = \lambda(z-1)X(z) \tag{18}$$

따라서, 먼저 이탈시점고객수 벡터생성함수를

구하고 식 (18)을 이용하여 임의슬롯 시작점에서의 고객수 벡터생성함수를 유도한다. 임의의 이탈고객이 시스템에 남기는 고객수는 자신의 체재시간 동안 도착한 고객이다. 따라서 이산시간 MAP/G/1 대기행렬시스템에서 다음이 성립한다.

$$\begin{aligned} \boldsymbol{X}(z) &= \boldsymbol{W}^*(\omega)\big|_{\omega = \boldsymbol{D}_0 + \boldsymbol{D}_1 z} \\ &= \sum_{x=1}^{\infty} (\boldsymbol{D}_0 + \boldsymbol{D}_1 z)^x \boldsymbol{W}(x) \end{aligned} \tag{19}$$

식 (19)를 적용할 때에 스칼라 ω에 행렬 값이 대체되므로, 고객의 서비스가 도착 이후에 이루어진다는 사실을 상기하여 행렬 $\boldsymbol{D}_1$과 스칼라 함수인 $S^*(\omega)$의 곱셈순서에 주의를 기울여야 한다 [28]. 이제 식 (19)와 식 (17)을 이용하여 정리하면,

$$\begin{aligned} \boldsymbol{X}(z) &= \frac{(1-\rho)\boldsymbol{\kappa}}{\lambda E(I)} (\boldsymbol{I} - \boldsymbol{D}_0)^{-1} (\boldsymbol{D}_0 + \boldsymbol{D}_1 z - \boldsymbol{I}) \\ &\quad \times [\boldsymbol{D}_1 z - \boldsymbol{D}_1 \boldsymbol{A}(z)]^{-1} \boldsymbol{D}_1 \boldsymbol{A}(z) \\ &= \frac{(1-\rho)\boldsymbol{\kappa}}{\lambda E(I)} (\boldsymbol{I} - \boldsymbol{D}_0)^{-1} (\boldsymbol{D}_0 + \boldsymbol{D}_1 z - \boldsymbol{I}) \\ &\quad \times [z\boldsymbol{I} - \boldsymbol{A}(z)]^{-1} \boldsymbol{A}(z) \\ &= \frac{(1-\rho)\boldsymbol{\kappa}}{\lambda E(I)} (\boldsymbol{I} - \boldsymbol{D}_0)^{-1} (\boldsymbol{D}_0 + \boldsymbol{D}_1 z - \boldsymbol{I}) \\ &\quad \times \boldsymbol{A}(z) [z\boldsymbol{I} - \boldsymbol{A}(z)]^{-1} \end{aligned} \tag{20}$$

위 식에서

$$\boldsymbol{A}(z) = S^*(\omega)\big|_{\omega = \boldsymbol{D}_0 + \boldsymbol{D}_1 z} = \sum_{x=1}^{\infty} (\boldsymbol{D}_0 + \boldsymbol{D}_1 z)^x s(x)$$

이고 $\boldsymbol{A}(z)$와 $[z\boldsymbol{I} - \boldsymbol{A}(z)]^{-1}$는 서로 가환이다.

식 (20)의 마지막 항은 기존의 고객수를 분석할 때 주로 사용되었던 이탈시점 전이확률로부터 구한 고객수 분석결과와 동일하다. 뿐만 아니라 기존의 고객수 분석 연구들에서 등장하는 $\boldsymbol{\kappa}^*$(사이클 동안 서비스 받는 평균고객수와 사이클 시작점에서의 UMC 위상에 대한 정보가 담긴 열벡터)를 구하는 시간과 노력을 덜어준다.

마지막으로 식 (18)과 식 (20)을 이용하여 정리하면 임의 슬롯 시작점에서의 고객수 분포는 다음과 같다.

$$\begin{aligned} \boldsymbol{Y}(z) &= \frac{(1-\rho)\boldsymbol{\kappa}}{E(I)} (\boldsymbol{I} - \boldsymbol{D}_0)^{-1} (z-1) \\ &\quad \times \boldsymbol{A}(z) [z\boldsymbol{I} - \boldsymbol{A}(z)]^{-1} \end{aligned} \tag{21}$$

식 (20)을 이용하여 Geo/G/1 시스템의 임의슬롯 시작점 고객수 분포를 구해 보자. Geo/G/1 대기행렬시스템에서 이탈시점 고객수는 Burke의 정리에 의해 도착시점 고객수와 확률적으로 같고, BASTA(Bernoulli arrival sees time average) 성질에 의해 도착시점 고객수 확률과 임의슬롯 시작점에서의 고객수가 같다. 따라서, 식 (21)을 이용해서 구한 결과와도 같음을 알 수 있다.

$$\begin{aligned} & X_{Geo/G/1}(z) \\ &= (1-\rho)\lambda^{-1}(1-\lambda+\lambda z-1)A(z)[z-A(z)]^{-1} \\ &= \frac{(1-\rho)(z-1)A(z)}{z-A(z)} \\ &= Y_{Geo/G/1}(z) \end{aligned}$$

2.5 임의의 사이클 시작점에서의 UMC 위상 벡터 $\boldsymbol{\kappa}$

앞의 모든 분석결과는 임의의 사이클 시작점(또는 끝점)에서의 UMC 위상벡터인 $\boldsymbol{\kappa}$로 표현하였다. $\boldsymbol{\kappa}$를 구하는 방법은 기존의 이산시간 MAP 관련 대기행렬시스템에서 사용한 방법을 그대로 사용하지만 간단히 소개하기로 한다.

바쁜기간 시작점에서의 고객수와 유휴기간 시작점(즉, 사이클의 시작점)과 바쁜기간 시작점의 UMC 위상을 포함한 행렬을 $\boldsymbol{N}_B(z)$, 사이클 동안 도착한 고객수와 그 사이클의 시작점, 끝점의 위상을 포함한 행렬을 $\boldsymbol{K}(z)$라고 정의하자. 이산시간 MAP/G/1 대기행렬시스템에서는 서버가 유휴할 때 고객이 도착하면 바로 바쁜기간을 시작하고, 첫 번째 고객의 서비스시간 동안 도착하는 고객들은 각각 하나의 기본기간(fundamental period)을 형성한다. 따라서 $\boldsymbol{K}(z)$는 다음과 같다.

$$\begin{aligned} \boldsymbol{K}(z) &= \boldsymbol{N}_B(z)\big|_{z = \boldsymbol{G}(z)} \\ &= (\boldsymbol{I} - \boldsymbol{D}_0)^{-1} \boldsymbol{D}_1 z\big|_{z = \boldsymbol{G}(z)} \end{aligned} \tag{22}$$

여기서,

$G(z)=zA(G(z))=z\sum_{x=1}^{\infty}(D_0+D_1G(z))^x s(x)$ 이다.

식 (22)의 z에 1을 대입하면 사이클 시작점과 종료점의 UMC 위상만을 고려한 행렬 K를 구할 수 있고, 우리가 구하고자 하는 벡터 κ는 행렬 K의 정상확률(stationary probability) 벡터이므로 잘 알려진 다음의 관계식으로부터 구한다.

$$\kappa=\kappa K,\ \kappa e=1\ . \tag{23}$$

2.6 주요 성능척도

시스템의 주요 성능척도로서 평균값을 고려할 수 있다. 편의상 어떤 행렬 또는 벡터 생성함수 $H(z)$에 대하여 다음을 정의하자.

$$H=H(\omega)|_{\omega=1},\ H^{(n)}=\frac{d^n}{d\omega^n}H(\omega)\bigg|_{\omega=1}.$$

식 (12)로부터 평균일량은 다음과 같다.

$$E(U)=u^{*(1)}e$$

평균대기시간, 평균체재시간, 시스템 내 평균고객수는 식 (15), (17), (21)로부터 아래와 같이 구할 수 있다. 분석결과를 얻었다 하더라도 평균값을 계산하는 과정은 다소 복잡하며, 서론에서 언급한 바와 같이 이는 Lee[11]에서 자세한 계산절차를 소개하고 있으므로 본 논문에서는 다루지 않기로 한다.

$$W_q=W_q^{*(1)}e,\ W=W^{*(1)}e,\ L=Y^{*(1)}e.$$

3. 결 론

본 연구에서는 Bruneel and Kim[33]에 의해 통신·교통시스템에 적합하다고 잘 알려진 이산시간 시스템을 대상으로 하여 마코비안 도착과정(Markovian arrival process, MAP)에 의해 시스템에 한 명씩 도착하는 고객들을 단일 서버가 선입선출로 서비스하는 대기행렬시스템을 분석하였으며, 포괄적인 분석을 위해 일량 분석에 초점을 맞추어 접근하였다.

그 이유는 일단 일량 분석이 완벽하게 이루어지면, 일량 분석으로부터 가상대기시간의 분석을, 가상대기시간으로부터 실제대기시간의 분석을, 실제대기시간으로부터 체재시간의 분석을, 체재시간의 분석으로부터 이탈시점 고객수 분석을, 이탈시점 고객수 분석으로부터 임의시점 고객수 분석을 순차적으로 어렵지 않게 유도할 수 있기 때문이다. 이는 이탈시점 고객수가 자신의 체재시간 동안 도착한 고객수로 표현되는 모든 이산시간 MAP/G/1 관련 시스템에 그대로 적용할 수 있다. 단, Lee et al.[34]에서 고려한 FOR(fewest operation remaining)이나 후입선출(LSFS, last come last served)과 같이 우선순위(priority) 서비스규칙에 의해 서비스 순서가 뒤바뀌는 경우는 제외한다.

References

[1] Neuts, M. F., “A Versatile Markovian Point Process”, Journal of Applied Probability, Vol. 16, No. 4, pp. 764-779, 1979.

[2] Casale, G., Zhang, E. Z. and Smirni, E., “Trace Data Characterization and Fitting for Markov Modeling”, Performance Evaluation, Vol. 54, No. 2, pp. 61-79, 2010.

[3] Buchholz, P., “An EM-algorithm for MAP Fitting from Real Traffic Data.”, International Conference on Modelling Techniques and Tools for Computer Performance Evaluation. Springer Berlin Heidelberg, 2003.

[4] Klemm, A., Lindemann, C. and Lohmann, M., “Modeling IP Traffic Using the Batch Markovian Arrival Process”, Performance Evaluation, Vol. 54, No. 2, pp. 149-173, 2003.

[5] Artalejo, J.R., Gómez-Corral, A. and Qi-Ming H., “Markovian Arrivals in Stochastic Modelling: a Survey and Some New Results”,

Statistics and Operations Research Transactions, Vol. 34, No. 2, pp. 101-156, 2010.

[6] Paxson, V. and Floyd, S., "Wide Area Traffic: The Failure of Poisson Modeling", IEEE/ACM Transactions on Networking, Vol. 3, No. 3, pp. 226-244, 1995.

[7] Ramaswami, V., "The N/G/1 Queue and Its Detailed Analysis", Advances in Applied Probability, Vol. 12, No. 1, pp. 222-261, 1980.

[8] Okamura, H., Dohi, T. and Trivedi, K. S., "Markovian Arrival Process Parameter Estimation with Group Data". IEEE/ACM Transactions on Networking, Vol. 17, No. 4, pp. 1326-1339, 2009.

[9] Breuer, L. and Kume, A., "An EM Algorithm for Markovian Arrival Processes Observed at Discrete Times", In International GI/ITG Conference on Measurement, Modelling, and Evaluation of Computing Systems and Dependability and Fault Tolerance, Springer Berlin Heidelberg, pp. 242-258, 2010.

[10] Carrizosa, E. and Ramírez-Cobo, P., "Maximum Likelihood Estimation in the Two-state Markovian Arrival Process", arXiv preprint arXiv:1401.3105, 2014.

[11] Lee, H. W., Queueing Theory, Sigma Press, 2006 (in Korean)

[12] Cinlar, E., "Queues with semi-Markovian Arrivals", Journal of Applied Probability, Vol.4, No. 2, pp. 365-379, 1967.

[13] Neuts, M.F., "Matrix-analytic Methods in Queueing Theory", appears in: Advances in Queueing : Theory, methods and Open Problems, J.H. Dshalalow Editor, CRC Press, pp. 265-292, 1995.

[14] Wallace, V.L. The Solution of Quasi Birth and Death Processes Arising from Multiple Access Computer Systems, SEL Technical Report No. 35, Michigan Univ. Ann Arbor Systems Engineering Lab, 1969.

[15] Ramaswami, V., "A Stable Recursion for the Steady State Vector in Markov Chains of M/G/1 Type", Stochastic Models, Vol. 4, No. 1, 1988.

[16] Lucantoni, D. M., "New Results for the Single Server Queue with a Batch Markovian Arrival Process", Stochastic Models, Vol. 7, No. 1, pp. 1-46, 1991.

[17] Lucantoni, D. M., "The BMAP/G/1 Queue: A Tutorial", Lecture Notes in Computer Science, Vol. 729, pp. 330-358, 1993.

[18] Blondia, C. and Casals, O., "Statistical Multiplexing of VBR Sources: A Matrix-analytic Approach", Performance Evaluation, Vol. 16, pp. 5-20, 1992.

[19] Cox, D.R., "The Analysis of non-Markovian Stochastic Processes by the Inclusion of Supplementary Variables", Mathematical Proceedings of the Cambridge Philosophical Society, Vol. 51, No. 3, pp. 433-441, 1955.

[20] Hokstad, B.T., "A Supplementary Variable Technique Applied to the M/G/1 Queue", Scandinavian Journal of Statistics, Vol. 2, No. 2, pp. 95-98, 1975.

[21] Lee, H.W., Ahn, B.Y. and Park, N.I., "Decompositions of the Queue Length Distributions in the MAP/G/1 Queue under Multiple and Single Vacations with *N*-policy", Stochastic Models, Vol. 17, No. 2, pp. 157-190, 2001

[22] Lee, H.W. and Ahn, B.Y., "Operational Behavior of the MAP/G/1 Queue under *N*-policy with Single Vacation and Setup", Journal of Applied Mathematics & Stochastic Analysis, Vol. 15, No. 2, pp. 167-196, 2002.

[23] Lee, H.W., Shin, Y.J. and Lee, S.W., "Two Approaches for Discrete-time MAP/G/1 Queues", Journal of Industrial Engineering and Management Systems, Vol. 2, No. 2, pp. 188-192, 2003.

[24] Lee, H.W., Shin, Y.J., Lee, E.Y. and Chae, K.C., "An Eigenvalue Approach to Discrete-time MAP/G/1 Queues", International Journal of Industrial Engineering, Vol. 10, No. 4,

pp. 644-650, 2003.
[25] Lee, H.W. and Park, N.I., "Using Factorization for Waiting Times in BMAP/G/1 Queues with N-policy and Vacations", Stochastic Analysis and Applications, Vol. 22, No. 3, pp. 755-773, 2004.
[26] Lee, H.W., Lee, S.W., Seo, W.J., Cheon, S.H. and J. Jeon, "A Unified Framework for the Analysis of M/G/1 Queue Controlled by Workload", Lecture Notes in Computer Science, Vol. 3982, pp. 718-727, 2006.
[27] Lee, S.W., Lee, H.W. and Baek, J. W., "Analysis of Discrete-time Geo/G/1 Queue under the *D*-policy", 6th International Conference on Queueing Theory and Network Analysis, 2011, Seoul, Korea, QTNA2011.
[28] Lee, S.W., Lee, H.W. and Baek, J. W., "Analysis of Discrete-time MAP/G/1 Queue under Workload Control", Performance Evaluation, Vol. 69, No. 2, pp. 71-85, 2012.
[29] Hunter, J.J., Mathematical Techniques of Applied Probability, Vol. II, Discrete-time Models: Techniques and Applications, New York: Academic Press, 1983.
[30] Takagi, H., Queueing Analysis: Vol III, Discrete-Time Systems, North-Holland, 1993.
[31] Kim, N.K., Chang, S.H. and Chae, K.C., "On the Relationships among Queue Lengths at Arrival, Departure, and Random Epochs on the Discrete-time Queue with D-BMAP Arrivals", Operations Research, Vol. 30, pp. 25-32, 2002.
[32] Latouche, G. and Ramaswami, V., Introduction to Matrix Analytic Methods in Stochastic Modeling, Society for Industrial and Applied Mathematics, 1999.
[33] Bruneel, H. and Kim, B. G., Discrete-Time Models for Communication Systems Including ATM, Kluwer Academic Publishers, 1993.
[34] Lee, K.-G., Choi, S. and Ryu, S.-W., "Development of Simulation System for Evaluating Performance of the Flat Display Manufacturing Line with Repetitive Process", The Journal of Information Systems, Vol. 17, No. 4, pp. 301-319, 2008.

이 세 원 (Se Won Lee)

• 종신회원
• 성균관대학교 산업공학과 학사
• 성균관대학교 산업공학과 석사
• 성균관대학교 산업공학과 박사
• 부경대학교 경영대학 경영학부 조교수
• 관심분야: 대기행렬이론, 확률모형, 시스템 최적화