

DT-정책 하에서 운영되는 이산시간 Geo/G/1 시스템의 대기시간과 체재시간 분석⁺

(Waiting Time and Sojourn Time Analysis of Discrete-time Geo/G/1 Queues under DT-policy)

이 세 원^{1)*}
(Se Won Lee)

요약 본 논문에서는 대기행렬이론의 대표적인 서버제어정책 중 D -정책과 T -정책이 혼합된 상황 하에서 운영되는 이산시간 대기행렬시스템을 다룬다. 베르누이 도착과정으로 시스템에 도착하는 고객들을 단일 서버가 선입선출로 서비스하며, 시스템 내에 서비스할 고객이 없으면 서버는 휴가를 떠났다가 돌아와서 대기 중인 고객들의 서비스시간의 총합(즉, 총 일량)이 D 를 초과할 때까지 반복하여 휴가를 갖는다. 본 연구에서 다루는 시스템의 운영은 2차전지를 사용하는 모바일 디바이스의 효율적인 자원 활용의 모델링에 사용할 수 있다. 또한 통신·교통 시스템 분석에 적합한 것으로 잘 알려진 이산시간 시스템을 대상으로 하여 유연한 혼합 제어정책(DT-정책) 하에서 대기행렬시스템의 안정상태 대기시간과 체재시간을 하나의 틀 안에서 유도하였다는 데 그 의의가 있다.

핵심주제어 : 이산시간 대기행렬시스템, 혼합제어정책, 일량제어정책, 복수휴가, 통합적 접근방법

Abstract In this paper, we studied a discrete-time queuing system that operates under a mixed situation of D -policy and T -policy, one of the representative server control policies in queuing theory. A single server serves customers arriving by Bernoulli arrival process on a first-in, first-out basis(FIFO). If there are no customers to serve in the system, the server goes on vacation and returns, until the total service time (i.e., total amount of workload) of waiting customers exceeds predetermined workload threshold D . The operation of the system covered in this study can be used to model the efficient resource utilization of mobile devices using secondary batteries. In addition, it is significant in that the steady state waiting time and system sojourn time of the queuing system under a flexible mixed control policy were derived within a unified framework.

Keywords : Discrete-Time Queueing System, Mixed Server Control Policy, Workload Control, Multiple Vacations, Unified approach

* Corresponding Author : swlee@pknu.ac.kr

+ 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2021R1F1A1062480).

Manuscript received April 04, 2024 / revised April 19, 2024

/ accepted April 21, 2024

1) 국립부경대학교 경영학부, 제1저자, 교신저자

를 초과할 때까지 반복해서 휴가를 떠난다. 휴가에서 돌아왔을 때 총일량이 D 를 초과하면 서버는 서비스를 제공하기 시작하여 바쁜 기간(busy period)을 보낸다(Fig. 1 참조).

D -정책, T -정책 모두 단일 정책 하에서 운영되는 대기행렬시스템에 관한 연구들을 많이 찾아볼 수 있으나 여기서는 두 개 또는 그 이상의 서버 제어정책을 혼합한 시스템에 대한 연구들을 소개하고자 한다. Lee et al. (2006)은 D -정책과 복수휴가 하에서 $M/G/1$ 대기행렬의 안정상태 고객수와 대기시간 분포를 구하고, 성능척도를 분석하였다. Lee and Seo (2008)는 $\text{Min}(N,D)$ -정책 하에서 $M/G/1$ 대기행렬시스템의 고객수, 일량, 대기시간 분포를 구하고 선형비용함수를 통해 단일 정책하에서의 최적의 N^* 와 D^* 를 $\text{Min}(N,D)$ -정책하에서의 값들과 비교하였다. 여기서 $\text{Min}(N,D)$ -정책이란 유희기간 동안 고객 수가 N 명이 되거나 이들의 서비스시간의 합이 D 를 초과하거나 하는 두 가지 상황 중에 하나라도 먼저 만족하면 바쁜기간을 시작하는 제어정책을 뜻한다. Lee et al. (2009)은 단수휴가와 D -정책 하에서 도착간격들이 상관성을 갖는 $\text{MAP}/G/1$ 대기행렬시스템을 분석하였다. Lee et al. (2010)은 $\text{Min}(N,D)$ -정책하에서의 $\text{MAP}/G/1$ 대기행렬을 분석하였다. 이들은 상관성을 갖는 도착과정을 랜덤 도착인 포아송 과정으로 가정하게 되면 성능척도를 과소평가할 수 있음을 보였다. Lee et al. (2011)은 삼중 제어정책인 (N,D,T) -정책 하에서 $M/G/1$ 시스템의 안정상태 고객수와 일량, 대기시간을 분석하고, 특수한 상황을 가정하여 고객수 분포를 그래프로 보였다. 이는 저자가 아는 범위 내에서 삼중 서버 제어정책을 다룬 첫 번째 연구이다. 앞에서 소개한 연구들은 주로 연속시간 대기행렬시스템을 다루고 있다.

이후 이산시간 대기행렬시스템의 분석에 대한 관심이 높아지면서 이와 함께 다중 서버 제어정책을 갖는 이산시간 대기행렬의 분석 또한 많이 연구되었다. Kim et al. (2013)은 고객들이 집단으로 도착하는 이산시간 $\text{Geo}^X/G/1$ 시스템의 안정상태 고객수를 확률적 분해성질을 이용하여 분석하였다. 이 연구에서는 N -정책과 휴가정책(단수휴가, 복수휴가)을 고려한 혼합 제어정책을 가정

하고 있다. Lan and Tang (2016)은 N -정책과 D -정책을 동시에 적용하는 상황에서 두 가지 도착율을 갖는 $\text{Geo}/G/1$ 시스템을 분석하였다. Lee (2018)는 일량제어와 복수휴가가 혼합된 정책 하에서 운영되는 이산시간 $\text{Geo}/G/1$ 시스템의 안정상태 일량을 분석하였다. Lee (2020)는 D -정책과 단수휴가를 갖는 이산시간 $\text{Geo}/G/1$ 시스템의 안정상태 고객수 분포를 유도하였다. 가장 최근의 연구로는 Luo et al. (2023)을 찾아볼 수 있는데, 이들은 수정된 $\text{Min}(N,D)$ 정책과 복수휴가정책 하에서 고객수를 분석하고 최적화 문제를 다루었다. 통상적으로 수정된 D -정책에서는 본 연구에서 다루는 D -정책 관련 대기행렬시스템 분석의 복잡성을 회피하기 위해 수학적 분석이 보다 용이한 형태로 바쁜기간 시작점에서 대기 중인 고객들의 서비스시간을 독립으로 가정하고 있다. 본 연구에서는 이러한 가정을 배제하고 원래의 D -정책 제어 방식에 따라 작동하는 시스템을 대상으로 한다.

본 연구는 Lee et al. (2006)이 대상으로 한 시스템의 이산시간 버전에 해당하는 연구로 볼 수 있으며, Lee (2018)의 후속 연구이다. 기존 연구와의 차별성은 대기시간과 체제시간을 하나의 틀 안에서 유도하였다는 것이다. 본 연구의 결과는 평균 성능척도의 계산과 함께 2차, 3차 모멘트의 계산을 통한 대기시간, 체제시간의 변동성과 왜도(skewness) 확인, 비용 최적화 모형, 전력 최적화 모형 등의 최적화 분석 연구의 기초자료로 사용할 수 있다.

이후의 논문 구성은 다음과 같다. 2장에서는 분석을 대상으로 하는 시스템에 대하여 기본 가정과 이론적 배경을 소개하고 대기시간, 체제시간을 분석하기 위해 꼭 필요한 선형 연구의 주요 결과들을 요약 정리한다. 3장에서는 대기시간을 분석한다. 분석에서 핵심이라 할 수 있는 조건부 결합확률을 정의하였는데 이는 본 연구가 선형연구와 차별성을 갖는 부분이라 할 수 있다. 이를 통해 유희기간 중에 도착하는 고객들의 상황을 4가지로 구분하여 안정상태 대기시간 분포의 변환을 유도하였다. 4장에서는 대기시간 분석에서 사용한 접근방식을 그대로 적용하여 비교적으로 수월하게 체제시간을 유도하는 방법을 보인다. 5장

에서는 유도한 대기시간의 변환으로부터 평균 대기시간을 구한 후 정리하여 그 결과를 보이고, 마지막으로 6장에서 결론과 함께 추후 연구과제에 대해 논의함으로 논문을 맺는다.

2. 시스템 및 선행연구의 주요결과

2.1 시스템

본 연구의 대상시스템의 특성은 다음과 같다.(Lee, 2006; Takagi, 1993)

- 1) 도착 과정: 단위 슬롯에 고객이 도착할 확률이 λ 인 베르누이 도착과정(Bernoulli process)을 따른다. 즉, 고객들의 도착간격은 iid 기하분포를 따른다.
- 2) 서비스 과정: 고객들의 서비스시간은 이산시간 일반분포를 따르고 단일 서버가 선입선출(FIFO, first in first out)로 서비스한다.
- 3) 제어 정책: 시스템에 더 이상 서비스할 고객이 없으면 서버는 이산확률변수 V 만큼의 휴가를 보낸 후 돌아와 시스템에 있는 대기 고객들의 서비스시간의 총합(총 일량)을 확인한다. 이 때 총 일량이 미리 정해 놓은 일량 임계값 D 를 초과하면 서비스를 제공하여 바쁜 기간을 시작하고, D 이하이면 다시 휴가를 떠난다.
- 4) 이산시간 시스템: 서비스시간이 1보다 크다고 가정하는 지연도착시스템(LAS-DA, late arrival system with delayed access)을 따른다.

분석을 위해 필요한 기호와 확률들을 정리하면 다음과 같다.

- λ : 한 슬롯 동안에 고객이 도착할 확률
- S : 임의의 고객의 서비스 시간, 확률변수
- S_n : S 의 n 차 중합(n -th convolution)
- V : 휴가의 길이, 확률변수
- $v_x = \Pr(V=x)$: 휴가의 길이가 x 일 확률

$$V^*(\omega) = \sum_{x=1}^{\infty} \omega^x v(x) : \text{휴가 길이의 확률생성함수}$$

(PGF, probability generating function)

S : 서비스시간의 길이(이산확률변수)

$s_k = \Pr(S=k), (s(0)=0)$: S 의 확률질량함수
(pmf, probability mass function),

$$S^*(\omega) = \sum_{k=1}^{\infty} \omega^k s_k : S \text{의 확률생성함수(PGF)}$$

$S^{(n)}(k) = \Pr(S_1 + S_2 + \dots + S_n \leq k)$: S 의

n 차 중합의 분포함수(DF), ($S^{(0)}(k) = 1$)

$s^{(n)}(k) = S^{(n)}(k) - S^{(n)}(k-1)$: S 의 n 차
중합의 확률질량함수(pmf)

$$M_k = \sum_{n=1}^{\infty} S^{(n)}(k) : \text{iid 서비스 시간으로 이루어진 재생과정의 재생함수(renewal function)}$$

$$m_k = \sum_{n=1}^{\infty} s^{(n)}(k) = M_k - M_{k-1} : \text{iid 서비스 시간으로 이루어진 재생과정의 재생 질량 함수}$$

I : 유희기간(idle period)의 길이

B : 바쁜 기간(busy period)의 길이

C : 사이클(cycle)의 길이

N_B : 바쁜기간 시작점에서의 고객수

U_B : 바쁜기간 시작점에서의 일량

$E(\cdot)$: 확률변수의 평균(expectation)

$\rho = \lambda E(S)$: 서버가 바쁠 확률

2.2 선행연구 주요 결과

본 절에서는 선행연구라 할 수 있는 Lee (2018)에서 대기시간과 체재시간의 분석에 꼭 필요한 일량 균휴과 과정을 소개한다. 일량 균휴과 과정을 사용하는 이유는 분석의 용이성 때문이다. 이는 각각의 휴가를 고려하여 분석하는 것보다 쉽고 효율적이다. 본 절의 논의는 Lee (2018)을 그대로 따른다.

Fig. 2에서 파란색 실선은 유희기간 동안의 균휴과 과정을 나타낸 것이다. 휴가 동안 1명 이상의 고객이 도착하여 휴가 시작점과 종료점에서

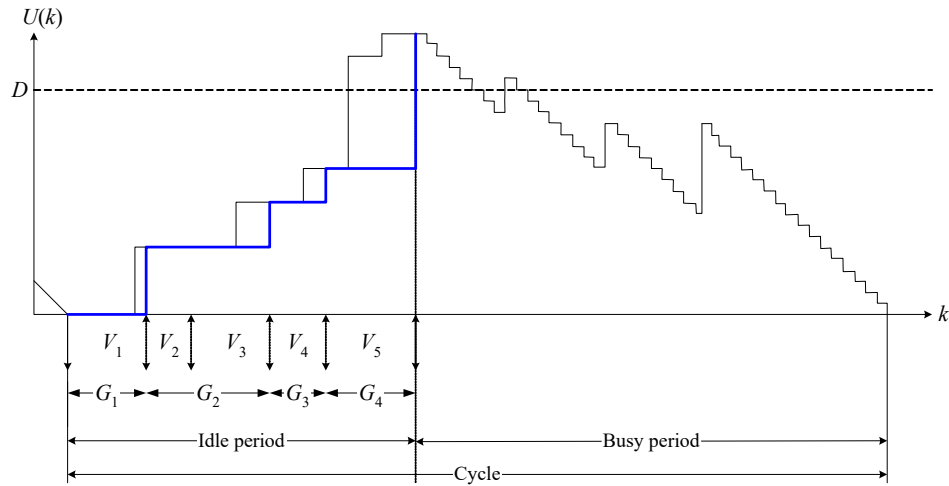


Fig. 2 Illustration of grand vacation process from the workload (Lee, 2018)

일량의 변화가 생기면 군휴가가 하나씩 생성되는데 Fig. 2에서는 총 4개의 군휴가가 생성되며 마지막 군휴가가 끝나는 시점이 서버가 바빠지기 시작하는 시점(즉, 바쁜기간 시작점)이 된다.

군휴가의 길이는 복수휴가가 있는 일반적인 Geo/G/1 대기행렬시스템의 유휴기간과 확률적으로 동일하며 군휴가의 길이를 확률변수 G 로 정의하면 임의의 군휴가의 평균길이는 Takagi (1993)로부터 다음과 같다.

$$E(G) = \frac{E(V)}{1 - V(1 - \lambda)} \quad (1)$$

여기서, $V(1 - \lambda)$ 는 하나의 휴가 동안 고객이 아무도 도착하지 않을 확률이다. 일량 군휴가 과정 분석을 위한 기호들을 다음과 같이 정의하자.

- G : 군휴가의 길이, 확률변수
- H : 하나의 군휴가 동안 증가한 일량
- A_G : 하나의 군휴가 동안 도착한 고객수
- N_G : 유휴기간 동안 발생한 군휴가의 개수

정의에 의하면 유휴기간의 길이 I 는 첫 번째부터 N_G 번째까지 군휴가의 길이의 합이다. 따라서 유휴기간의 길이는 식 (2)와 같이 표현할 수 있다.

$$I = G_1 + G_2 + \dots + G_{N_G}. \quad (2)$$

식 (2)의 N_G 는 정지시간(stopping time)이므로 Lee (2006)에 의하면 유휴기간의 평균 길이는 다음과 같다.

$$E(I) = E(N_G)E(G) \quad (3)$$

이제 $E(N_G)$ 를 구하기 위해 한 군휴가 동안의 일량 증가분 확률과정을 사용하자. 임의의 군휴가 동안 도착한 고객수를 A_G 라고 하자. 이 군휴가 동안의 일량 증가분 H 는 $S_1 + S_2 + \dots + S_{A_G}$ 이 되고, H 의 확률질량함수 $P(H = x) = h_x$ 는 식 (4)와 같다.

$$h_x = \frac{\sum_{y=1}^{\infty} \sum_{n=1}^y v_y \binom{y}{n} \lambda^n (1 - \lambda)^{y-n} s^{(n)}(x)}{1 - V(1 - \lambda)}, (x \geq 1) \quad (4)$$

iid 확률변수 H 들로 이루어지는 확률과정 H -process의 재생밀도함수와 재생함수를 식 (5)와 같이 정의하면 시스템의 유휴기간 동안 발생한 군휴가의 평균개수 $E(N_G) = 1 + M_D^H$ 가 된다.

$$m_k^H = \sum_{n=1}^{\infty} h_k^{(n)}, M_k^H = \sum_{n=1}^{\infty} H_k^{(n)} \quad (5)$$

식 (5)에서 $h_k^{(n)} = \sum_{y=1}^k h_y h_{k-y}^{(n)}$ 으로 식 (4)를 이용하여 재귀적으로 계산하여 구한다.

이 결과들은 유휴기간 동안 도착한 고객들의 대기시간과 체재시간 분포를 구할 때 사용된다. 구체적으로는 도착 시점에서의 상황과 연관된 여러 가지 변수들을 고려한 수식을 모형화하는 데 중요한 열쇠 역할을 한다.

3. 대기시간 분석

본 장에서는 시스템에 도착하는 고객들의 대기시간을 분석한다. 시스템에 도착하는 고객들은 서버가 유휴한 중(서비스를 제공하지 않는 기간)에 도착하는 경우와 바쁜 중에 도착하는 고객으로 분류할 수 있다. 서버가 유휴한 중에 도착하는 고객들을 특별고객(SP; special customer), 바쁜 중에 도착하는 고객들을 보통고객(OC; ordinary customer)이라고 하자. 시스템에 도착하는 고객이 특별고객일 확률을 p_{sc} , 보통고객일 확률을 p_{oc} 라고 하면 임의의 고객의 대기시간의 확률생성 함수(PGF) $W_q^*(\omega)$ 는 다음과 같이 표현할 수 있다. 여기서 고객들의 유형을 나누어 분석하는 것은 D -정책 하에서 작동하는 시스템에서는 바쁜 기간 시작점에서 대기 중인 고객들의 서비스시간이 서로 독립이 아니기 때문이다.(Lee, 2006) 이러한 특성으로 인해 D -정책 관련 대기행렬시스템의 분석과 계산은 다른 제어정책에 비해 복잡도가 높다.

$$W_q^*(\omega) = p_{sc} W_{q,sc}^*(\omega) + p_{oc} W_{q,oc}^*(\omega) \quad (6)$$

여기서, $W_{q,sc}^*(\omega)$ 와 $W_{q,oc}^*(\omega)$ 는 각각 임의의 특별고객과 보통고객의 대기시간 분포의 생성함수(GF)이다. 이제 식 (6)의 구성요소를 구해보자.

3.1 유휴기간 중 도착한 고객들의 대기시간

먼저 어떤 고객이 유휴기간 중에 도착할 확률, 즉 특별고객일 확률(p_{sc})은 BASTA(Bernoulli Arrival See Time Averages)에 의해 그 고객이 도착하면서 시스템이 유휴한 상태를 볼 확률 $1 - \rho$ 이다. 따라서 어떤 고객이 보통고객일 확률은 $\rho = \lambda E(S)$ 가 된다.

유휴기간 중에 도착한 고객의 대기시간은 그 고객(특별시험고객)이 도착하면서 보는 일량과 잔여 유휴기간의 합으로 표현할 수 있는데 이들은 서로 독립이 아니다. 왜냐하면 특별시험고객이 도착하면서 보는 일량이 일량 임계값 D 에 가까운 값이라면 이후 도착하는 고객들로 하여금 곧 바쁜 기간을 시작하게 될 확률이 높으므로 그 시점에서 잔여 유휴기간이 짧아질 것이기 때문이다.

또한 잔여 유휴기간은 특별시험고객이 도착한 휴가의 잔여 시간(remaining time)의 길이와 그 잔여 휴가길이 동안 도착하는 고객들의 총 일량에 영향을 받는다. 따라서 유휴기간 동안 도착하는 고객들의 대기시간 $W_{q,sc}^*(\omega)$ 를 구하기 위해서는 다음을 모두 동시에 고려하여야 한다.

- 도착하면서 보는 일량 a
- 시험 고객의 서비스시간 x
- 시험 고객이 도착한 잔여 휴가의 길이 r
- r 동안 도착한 고객들의 서비스시간의 합 y

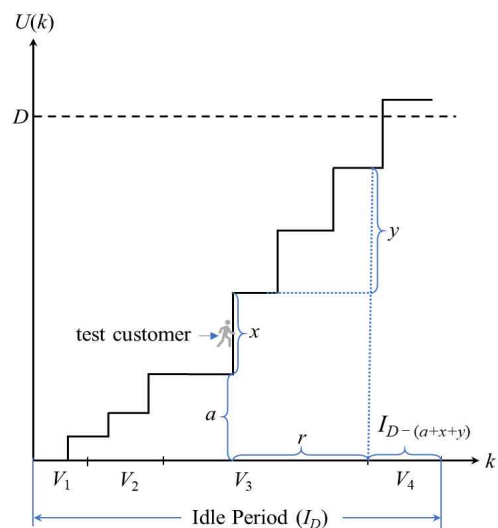


Fig. 3 Special test customer's arrival situation

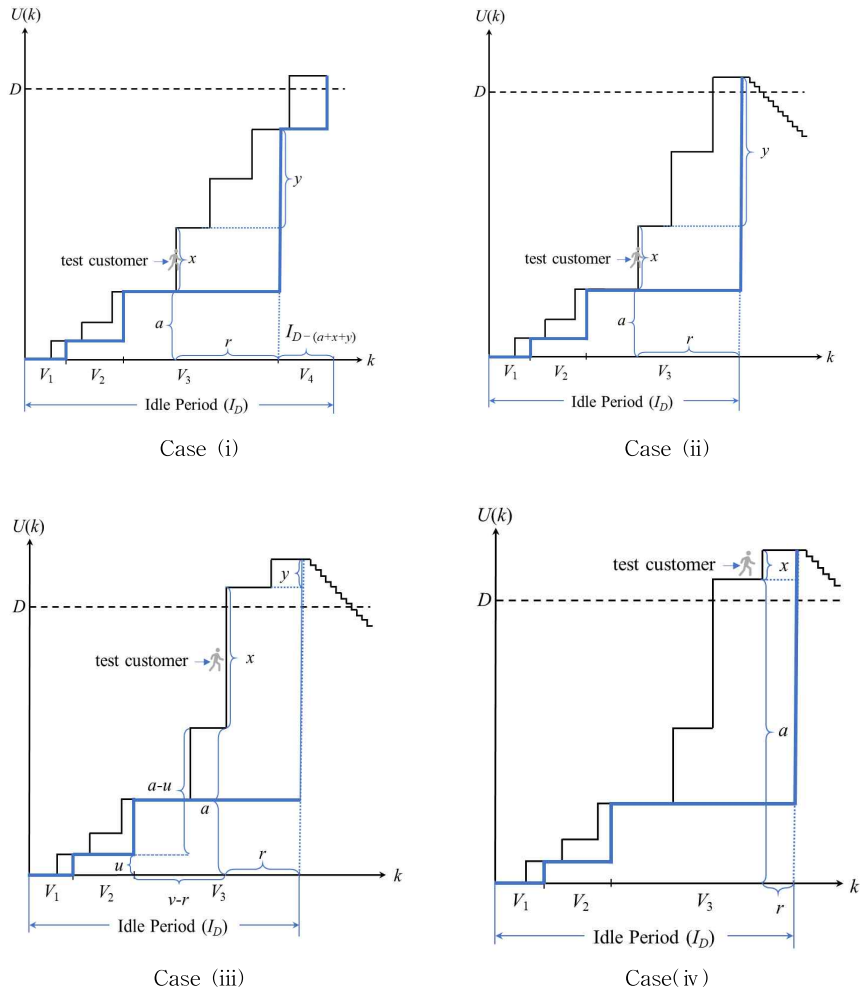


Fig. 4 Four cases of test customer with grand vacation process

Fig. 3은 이러한 상황을 이해하기 쉽도록 도식화한 것이다. 길이가 v 인 휴가에 특별시험고객이 도착하고 잔여 휴가의 길이가 r 이라는 조건 하에 도착 시점에서의 일량이 a , r 동안 도착한 고객들의 총일량이 y 일 조건부 결합확률을 $\phi(a, y | v, r)$ 라고 하자. 이 조건부 결합확률은 대기시간과 체제시간을 하나의 틀 안에서 구체화하는 열쇠(key)의 역할을 한다. 이는 Lee(2018)의 선행연구에서는 다룰 필요가 없었으나 대기시간, 체제시간, 더 나아가서는 고객수 분석에서 반드시 해결하여야 하는 핵심적이고 중요한 부분이다. 특별시험고객이 처할 수 있는 상황을 정리하

면 크게 다음의 4가지로 분류할 수 있다.

- (i) 시험 고객이 도착한 휴가 종료점에서의 일량이 D 를 넘지 않는 경우 ($a + x + y \leq D$)
- (ii) 시험 고객의 서비스시간을 포함한 일량은 D 를 넘지 않았지만 잔여 휴가 동안 도착한 고객들의 일량으로 휴가 종료점에서의 일량이 D 를 초과하는 경우 ($a + x \leq D, a + x + y \geq D + 1$)
- (iii) 시험 고객이 도착하면서 보는 일량은 D 이하였는데, 자신의 서비스시간으로 인해 일량이 D 를 초과하는 경우 ($a \leq D,$

- $a+x \geq D+1$)
 (iv) 시험 고객이 도착하면서 보는 일량이 이
 미 D 를 초과한 경우 ($a \geq D+1$)

앞서 정의한 $\phi(a, y|v, r)$ 를 사용하여 각 경우에 해당하는 대기시간을 변환 형태로 유도하자. 첫 번째 경우는 Fig. 4의 Case (i)에 해당하는데 시험 고객은 자기가 도착한 휴가의 잔여 시간 r 만큼 기다린 후 그 휴가 종료점에서의 일량이 임계값 D 를 초과하지 않았으므로, 일량 임계값을 $D-(a+x+y)$ 로 갖는 복수휴가 대기행렬시스템의 유희기간만큼 추가로 기다려야 한다. 서버가 바쁘기 시작하면 자기 앞에 도착한 고객들이 모두 서비스할 때까지 기다려야 하므로 a 만큼 더 기다리면 대기 공간을 벗어나 서비스를 받는다. 다른 경우에 대해서도 식을 쓰면 다음과 같다.

$$W_{q,sc_1}^*(\omega) = \omega^r \sum_{a=0}^D \sum_{x=1}^{D-a} \sum_{y=1}^{D-(a+x)} \omega^a s_x \phi(a, y|v, r) I_{D-(a+x+y)}^*(\omega) \quad (7.a)$$

$$W_{q,sc_2}^*(\omega) = \omega^r \sum_{a=0}^D \sum_{x=1}^{D-a} \sum_{y=D-(a+x)}^{\infty} \omega^a s_x \phi(a, y|v, r) \quad (7.b)$$

$$W_{q,sc_3}^*(\omega) = \omega^r \sum_{a=0}^D \sum_{x=D-a+1}^{\infty} \sum_{y=1}^{\infty} \omega^a s_x \phi(a, y|v, r) \quad (7.c)$$

$$W_{q,sc_4}^*(\omega) = \omega^r \sum_{a=D+1}^{\infty} \sum_{y=1}^{\infty} \omega^a \phi(a, y|v, r) \quad (7.d)$$

식 (7.a-d)를 모두 더하고 조건을 고려하여 다시 쓰면 다음과 같다.

$$W_{q,sc}^*(\omega) = \sum_{i=1}^4 \sum_{v=1}^{\infty} \sum_{r=0}^v W_{q,sc_i}^*(\omega) \Pr(v, r) \quad (8)$$

길이가 v 인 휴가에 특별시험고객이 도착하고 잔여 휴가의 길이가 r 일 확률 $\Pr(v, r)$ 은 2.1절에서 정의한 기호를 사용하여 정리하면 식 (9)와 같다.

$$\begin{aligned} \Pr(v, r) &= \frac{v \Pr(V=v)}{E(V)} \cdot \frac{1}{v} \\ &= \frac{\Pr(V=v)}{E(V)} = \frac{v_v}{E(V)}. \end{aligned} \quad (9)$$

이제 $\phi(a, y|v, r)$ 을 살펴보자. Fig. 4에서 두꺼운 파란 실선은 본 시스템의 분석에 핵심이라 할 수 있는 Lee(2018)의 일량 균휴가과정을 나타낸 것인데 균휴가 과정의 성질로부터 시험 고객이 도착한 휴가가 v 이고, 잔여 휴가 기간이 r 이면 이 고객이 도착하면서 보는 일량 a 는 도착한 휴가가 속해 있는 균휴가의 시작 시점에서의 일량 u 와 경과시간 $v-r$ 동안 도착한 고객들의 일량의 합 $a-u$ 가 되므로 이를 수식으로 표현하면 다음과 같다.

$$\begin{aligned} \phi(a, y|v, r) &= \sum_{u=0}^a \beta_u \sum_{k=0}^{v-r} \binom{v-r}{k} \lambda^k (1-\lambda)^{v-r-k} s^{(k-1)}(a-u) \\ &\quad \times \sum_{n=1}^r \binom{r}{n} \lambda^n (1-\lambda)^{r-n} s^{(n)}(y) \end{aligned} \quad (10)$$

여기서 β_u 는 어느 균휴가 시작 시점에서의 일량이 u 일 확률로서 Lee(2018)의 결과로부터 균휴가의 평균 개수는 $1+M_D^H$ 이고, 각 균휴가의 길이는 i.i.d.하므로 시험 고객이 첫 번째 균휴가에 도착하면서 일량 0을 보는 횟수는 그중에 한 번이다. 또한 균휴가의 시작 시점에서의 일량은 균휴가 동안의 일량과 같으므로 0보다 큰 일량 u 로 시작할 확률은 H-process가 일량 수준 u 에서 발생할 확률과 같다. 따라서 이를 정리하면 다음을 얻는다.

$$\beta_0 = \frac{1}{1+M_D^H} \quad (11.a)$$

$$\beta_u = \frac{m_u^H}{1+M_D^H}, \quad u=1, 2, \dots, D \quad (11.b)$$

따라서 시스템의 유희기간 동안 도착한 고객들의 대기시간 분포는 식 (9)를 이용하면 식 (12)와 같다.

$$W_{q,sc}^*(\omega) = \sum_{i=1}^4 \sum_{v=1}^{\infty} \sum_{r=0}^v W_{q,sc_i}^*(\omega) \frac{v_v}{E(V)} \quad (12)$$

3.2 바쁜기간 중 도착한 고객들의 대기시간

서버가 바쁜 중에 도착하는 시험 고객의 대기 시간은 BASTA에 의해 그 고객의 도착 시점에서 보는 일량과 같다. 따라서, 보통 고객들의 대기시간 변환은 식 (13)과 같다.

$$\begin{aligned} W_{q,oc}^*(\omega) &= U_{busy}(\omega) \quad (13) \\ &= U_{Geo/G/1}(\omega) \cdot U_B^-(\omega) \\ &= \frac{(1-\rho)(1-\omega)}{1-\lambda+\lambda S(\omega)-\omega} \cdot \frac{1-U_B(\omega)}{E(U_B)(1-\omega)} \end{aligned}$$

이체사이클 이론을 적용하면 식 (13)은 아래와 같이 표현할 수도 있다.

$$W_{q,oc}^*(\omega) = \frac{1}{\rho E(C)} \cdot \frac{1-U_B(\omega)}{1-\lambda+\lambda S(\omega)-\omega} \quad (14)$$

여기서 $U_B(\omega) = \sum_{x=D+1}^{\infty} \omega^x \left[h_x + \sum_{y=1}^D m_y^h h_{x-y} \right]$,
 $E(U_B) = \frac{\lambda E(S)E(V)}{1-V(1-\lambda)} \left[1 + \sum_{n=1}^D H_D^{(n)} \right]$ 이다. (Lee, 2018)

식 (6)에 식 (7.a-d), 식 (10.a-b), 식 (11.a-b)를 넣어 정리하면, 임의 고객의 대기시간 분포의 PGF는 다음과 같다.

$$\begin{aligned} W_q^*(\omega) \quad (15) \\ &= \frac{1-\rho}{E(V)} \sum_{i=1}^4 \sum_{v=1}^{\infty} v_v \sum_{r=0}^v W_{q,sc_i}^*(\omega) \\ &\quad + \frac{1-U_B(\omega)}{E(C)[1-\lambda+\lambda S(\omega)-\omega]} \end{aligned}$$

4. 체재시간 분석

본 장에서는 대기시간 분석의 틀 안에서 손쉽게

체재시간의 분포를 유도한다. 이는 통합적 접근방법의 장점 중 하나로서 어떤 시스템의 다각적인 분석을 큰 틀 안에서 해결할 수 있는 열쇠를 제공한다.

일반적으로 D-정책을 고려하여 운영되는 대기행렬시스템에서는 특별고객들의 서비스 시간이 iid하지 않으며, 서비스 시간과 체재시간이 독립이 아니기에 다른 대기행렬시스템에서 많이 사용하는 관계식이 성립하지 않는다.

$$W^*(\omega) \neq W_q^*(\omega)S^*(\omega) \quad (16)$$

따라서 체재시간의 분포 또한 대기시간과 같이 특별고객과 보통고객의 체재시간을 나누어 구한다. 체재시간(sojourn time)은 여러 분야에서 ‘흐름시간(flow time)’ 또는 ‘이행시간(transit time)’ 등으로 부르며 시스템의 효율성을 측정하고, 자원 할당과 서비스 품질 개선 등의 근거로 사용하기에 정확한 체재시간의 분석은 의미가 있다.

$$W^*(\omega) = p_{sc} W_{sc}^*(\omega) + p_{oc} W_{oc}^*(\omega) \quad (17)$$

이제 특별고객의 체재시간을 살펴보자. 서버가 유희한 중에 도착한 고객의 체재시간을 각각의 경우를 고려하여 정리하면 다음과 같다.

$$\begin{aligned} W_{sc_1}^*(\omega) &= \sum_{a=0}^D \sum_{x=1}^{D-a} \sum_{y=1}^{D-(a+x)} \omega^{a+x+r} \quad (18.a) \\ &\quad \times s_x \phi(a, y|v, r) I_{D-(a+x+y)}^*(\omega) \end{aligned}$$

$$\begin{aligned} W_{sc_2}^*(\omega) &= \sum_{a=0}^D \sum_{x=1}^{D-a} \sum_{y=D-(a+x)}^{\infty} \omega^{a+x+r} \quad (18.b) \\ &\quad \times s_x \phi(a, y|v, r) \end{aligned}$$

$$\begin{aligned} W_{sc_3}^*(\omega) &= \sum_{a=0}^D \sum_{x=D-a+1}^{\infty} \sum_{y=1}^{\infty} \omega^{a+x+r} \quad (18.c) \\ &\quad \times s_x \phi(a, y|v, r) \end{aligned}$$

$$W_{sc_4}^*(\omega) = \sum_{a=D+1}^{\infty} \sum_{y=1}^{\infty} \omega^a S^*(\omega) \omega^r \phi(a, y|v, r) \quad (18.d)$$

여기서 $\phi(a, y|v, r)$ 는 식 (10)과 같다. 따라서

$$W_{sc}^*(\omega) = \sum_{i=1}^4 \sum_{v=1}^{\infty} \sum_{r=0}^v W_{sc_i}^*(\omega) \frac{v_v}{E(V)}. \quad (19)$$

서버가 바쁜 중에 도착하는 시험 고객의 체제 시간은 보다 쉽게 유도할 수 있다. 바쁜 기간 동안 도착하는 고객들의 서비스시간은 iid하므로 임의의 보통고객의 체제시간은 자신의 대기시간과 서비스시간의 합으로 변환 상으로는 두 변환의 곱의 형태로 식 (20)과 같이 표현할 수 있다.

$$W_{oc}^*(\omega) = W_{q,oc}^*(\omega) \cdot S^*(\omega) \quad (20)$$

$$= \frac{1}{\rho E(C)} \cdot \frac{[1 - U_B(\omega)]S^*(\omega)}{1 - \lambda + \lambda S(\omega) - \omega}$$

식 (17)에 $p_{sc} = 1 - \rho$, $p_{oc} = \rho$, 식 (18.a-d), (19), (20), (7.a-d), (10.a-b), (11.a-b)를 대입하여 정리하면 임의의 고객의 체제시간을 얻는다.

5. 성능척도

성능척도로서 평균 대기시간을 구한다. 본 연구의 대상시스템은 이산시간 시스템이므로 평균 대기시간은 식 (15)의 대기시간 확률생성함수 (PGF)를 ω 에 대하여 1차 미분한 후 $\omega = 1$ 을 대입하여 얻는다. 다음 기호를 정의하자.

$$P^* = P^*(\omega)|_{\omega=1}, \quad P^{*(n)} = \left. \frac{d^n}{d\omega^n} P^*(\omega) \right|_{\omega=1} \quad (21)$$

먼저 특별시험고객의 대기시간을 구한 후 식 (21)을 사용하여 정리하면 다음과 같다.

$$W_{q,sc} = \left. \frac{d}{d\omega} W_{q,sc}^*(\omega) \right|_{\omega=1}$$

$$W_{q,sc_1}^*(1) = \sum_{a=0}^D (a+r) \sum_{x=1}^{D-aD-(a+x)} \sum_{y=1}^{D-aD-(a+x)} \times s_x \phi(a, y|v, r) I_{D-(a+x+y)} \quad (22.a)$$

$$+ \sum_{a=0}^D \sum_{x=1}^{D-aD-(a+x)} \sum_{y=1}^{D-aD-(a+x)} \times s_x \phi(a, y|v, r) E(I_{D-(a+x+y)})$$

$$W_{q,sc_2}^*(1) = \sum_{a=0}^D (a+r) \sum_{x=1}^{D-a} \sum_{y=D-(a+x)}^{\infty} s_x \phi(a, y|v, r) \quad (22.b)$$

$$W_{q,sc_3}^*(1) = \sum_{a=0}^D (a+r) \sum_{x=D-a+1}^{\infty} \sum_{y=1}^{\infty} s_x \phi(a, y|v, r) \quad (22.c)$$

$$W_{q,sc}^{*4(1)} = \sum_{a=D+1}^{\infty} (a+r) \sum_{y=1}^{\infty} \phi(a, y|v, r) \quad (22.d)$$

바쁜 기간 중에 도착하는 보통고객의 대기시간은 시험고객이 도착하면서 보는 일량과 확률적으로 같고 식 (20)을 ω 에 대하여 미분한 후 $\omega = 1$ 을 대입하면 로피탈의 정리를 거쳐 구한다. 정리하면 본 시스템에서의 평균 대기시간은 식 (23)과 같다. 체제시간의 평균 또한 동일한 절차를 거쳐 구할 수 있으므로 여기서는 다루지 않기로 한다.

$$W_q = \frac{1-\rho}{E(V)} \sum_{i=1}^4 \sum_{v=1}^{\infty} v_v \sum_{r=0}^v W_{q,sc_i}^*(1) \quad (23)$$

$$+ \rho \left[\frac{\lambda E[S(S-1)]}{2(1-\rho)} + \frac{E[U_B(U_B-1)]}{2E(U_B)} \right]$$

6. 결론 및 추후 연구과제

본 논문에서는 DT-정책을 갖는 이산시간 대기행렬시스템의 안정상태 대기시간과 체제시간을 분석하였다. 베르누이 도착과정으로 시스템에 도착하는 고객들을 단일 서버가 선입선출로 서비스하며, 시스템 내에 서비스할 고객이 없으면 서버는 휴가를 떠난다. 매 휴가 종료 시점에서 대기 중인 고객들의 서비스시간의 총합(즉, 총 일량)이 D 를 초과할 때까지 반복하여 휴가를 갖는다.

T -정책은 CPU가 내장되어 있는 스마트 모바일 디바이스가 최소한의 전력소모를 위해 사용하는 작동방식과 매우 유사하다. 모바일 디바이스에서 실시간으로 도착하는 신호와 작업을 처리함에 있어 core가 많은 자원들을 사용해야 하는데 일정 시간 후에 시스템을 모니터링하면서 대기 중인 작업이 있을 때 처리하는 방식은 효율적인 자원 활용에 도움이 된다. 거기에서 일량 임계값을 추가로 고려하여 유휴 중인 서버가 작업들을 쌓아 두었다가 총 일량이 D 를 초과하면 서비스

를 제공함으로써 시스템의 재가동 횟수를 더욱 적게 만들어 준 모델이 본 연구에서 다른 시스템이다.

본 연구에서는 시스템을 통합적으로 분석하는데 필요한 유희기간과 일량의 분석결과로부터 대기시간과 체재시간을 하나의 틀 안에서 해결할 수 있도록 유희기간 동안 도착하는 고객의 도착 시점에서의 상황을 추적하여 각 분포를 변환의 형태로 유도하고 성능척도로 평균 대기시간을 구하였다. 이론연구로서 선행연구에서 다루지 못한 대기시간과 체재시간을 분석할 수 있는 조건부 결합확률을 이용하여 분포를 구하였다는 것에 학술적 시사점을 찾을 수 있다.

추후 연구로는 성능척도로서 평균과 2차, 3차 모멘트의 계산을 통한 대기시간, 체재시간의 변동성과 왜도(skewness) 확인, 비용 최적화 모형, 전력 최적화 모형 등의 최적화 분석 연구를 고려하고 있으며 계산과 함께 시뮬레이션을 수행 중이다. 본 논문에서 다루는 시스템은 유희기간 동안 복잡한 메커니즘에 의해 작동하기 때문에 엄밀해(exact solution)를 구하는 데에는 상당한 계산량을 필요로 한다. 실제로 특별고객의 1번 유형의 대기시간 계산만 하더라도 다수의 \sum 연산을 수행해야 하기에 모멘트의 체계적인 계산과 이를 통한 최적화 문제의 해결은 학술적으로도 전적인 과제라 할 수 있다.

또한 베르누이 도착 과정이 아닌 일반 도착과정 하에서 시스템의 대기시간과 체재시간을 구할 수 있는 일반화된 시뮬레이션의 개발도 추후 연구과제로 고려해 볼 수 있다. 여기서, 시스템 가동률이 높을 때 대기시간을 근사적으로 구하기 위해 사용하는 Kingman's formula가 여전히 성립하는지를 확인해 보는 것도 의미가 있을 것이다. 시뮬레이션의 개발은 복잡한 이론 결과들을 검증하고 현장에서 손쉽게 적용하여 사용할 수 있도록 함에 있어 실무적으로 비중 있는 역할을 할 것으로 기대한다.

References

- Balachandran, K.R. (1973). Control Policies for a Single Server System, *Management Science*, 19, 1013-1018.
- Balachandran, K.R. and Tijms, H. (1975). On the D -policy for the M/G/1 Queue, *Management Science*, 21(9), 1073-1076.
- Heyman, D.P. (1977). The T -policy for the M/G/1 Queue, *Management Science*, 23(7), 775-778.
- Kim, S. J., Kim, N. K., Park, H.-M., Chae, K. C. and Lim D.-E. (2013). On the Discrete-time $Geo^X/G/1$ Queues under N -policy with Single and Multiple Vacations, *Journal of Applied Mathematics*, 2013, Article ID 587163.
- Lan, S. and Tang, Y. (2016). Analysis of a Discrete-time $Geo^{\lambda_1, \lambda_2}/G/1$ Queue with N -policy and D -policy, *Journal of Applied Mathematics and Computing*, 1-25.
- Lee, H. W. (2006). *Queueing Theory*, Sigma Press.
- Lee, H. W., Cheon, S. H. and Seo, W. J. (2006). Queue Length and Waiting Time of the M/G/1 Queue under the D -policy and Multiple Vacations, *Queueing Systems*, 54, 261-280.
- Lee, H. W. and Seo, W. J. (2008). The Performance of the M/G/1 Queue Under the Dyadic $\text{Min}(N, D)$ -policy and Its Cost Optimization, *Performance Evaluation*, 65, 742 -758.
- Lee, H. W., Cheon, S. H., Lee, S. W. and Baek, J. W. (2009). MAP/G/1 Queue under Workload Control and Postprocessing, *Stochastic Analysis and Applications*, 27(2), 331-352.
- Lee, H. W., Seo, W. J., Lee, S. W. and Jeon, J. (2010). Analysis of MAP/G/1 Queue under $\text{Min}(N, D)$ -policy, *Stochastic Models*, 26(1), 98-123.
- Lee, H. W., Seo, W. J. and Lee, S. W. (2011). Analysis of the M/G/1 Queueing System under the Triadic (N, D, T) -policies, *Quality Technology and Quantitative Management*,

Balachandran, K.R. (1973). Control Policies for

8(3), 333-357.

- Lee, S. W. (2018). Analysis of Discrete-time Geo/G/1 Queues under Workload control and Multiple Vacations, *Journal of Korea Society of Industrial Information Systems*, 23(2), 29-39.
- Lee, S. W. (2020). Queue Length Analysis of Discrete-time Queueing System under Workload Control and Single Vacation, *Journal of Korea Society of Industrial Information Systems*, 25(1), 89-99.
- Liu, R. and Deng, Z. (2014). The Steady-state System Size Distribution for a Modified D-policy Geo/G/1 Queueing System, *Mathematical Problems in Engineering*, 2014, Article ID 345129.
- Luo, L., Tang, Y.-H., Yu, M.-M. and Wu, W.-Q. (2023). Optimal Control Policy of M/G/1 Queueing System with Delayed Randomized Multiple Vacations under the Modified Min(N, D)-Policy Control. *Journal of the Operations Research Society of China*, 11, 857 - 874.
- Takagi, H. (1993). *Queueing Analysis: Vol III, Discrete-time Systems*, North-Holland.



이 세 원 (Se Won Lee)

- 중신회원
- 성균관대학교 산업공학과 학사
- 성균관대학교 산업공학과 석사
- 성균관대학교 산업공학과 박사
- 국립부경대학교 경영대학 경영학부 부교수

- 관심분야: 대기행렬이론, 확률모형, 시스템 최적화, 시뮬레이션