

# 통화서비스의 품질기준에 관한 연구

## A Study on the Quality of Real Call Time Service

趙 漢 璧\*  
金 在 連\*\*

### ABSTRACT

In this paper, the capacity of system which provides the real time response call service to the customer is studied. The quality of call service depends on the response time of the service in the system. Therefore, the focus of this paper is to investigate the capacity of system under the restriction of response time. In this paper, the system is modeled by queueing network. The analytical method is applied to solve this queueing network. The solution of the model has product form solution. To get the reasonable capacities, nonlinear programming problem is formulated and is solved by GINO.

### 1. 서 론

현대를 정보화사회라고 많은 분야에서 언급하고 있다. 컴퓨터와 통신을 결합시켜 정보화사회에서 필요한 기술을 개발 및 발전시키고 있으며 이와 함께 정보서비스의 요구도 점증하고 있다. 컴퓨터와 통신의 결합은 새로운 망기능을 지닌 망요소들이 도입되므로써 가능해지고 있으며, 이로 인해 보다 강력한 정보처리와 기능성 및

유연성을 지닌 망으로 발전되고 있다. 이 고도화된 망을 통하여 고객에게 서비스를 보다 빠르고 편하게 제공할 수 있으며 그 다양성도 날로 증가하고 있다.

이런 망을 분석하는 기법중의 하나는 큐잉이론이며, 큐잉망은 멀티프로그램, 시분할 컴퓨터 시스템의 중요한 모델이다. 지난 몇년간 이 적용분야에 대한 연구에서 컴퓨터 시스템의 중요한 양상을 포함하는 다양한 모델이 개발되었

\*한양대학교 산업공학과

\*\*한양대학교 산업공학과 교수

다.<sup>(5)</sup> 또는 최근 컴퓨터 시스템이 발달함에 따라 시스템의 특성을 연구하기 위하여 큐잉망을 이용한 분석이 확대되고 있다. 일반적으로 큐잉시스템을 묘사하는데 필요한 특성으로 고객의 도착형태, 서버의 서비스 형태, 대기 정책, 시스템 용량 등을 들 수 있다.

Baskett et al은 다등급 Jackson망을 확장하여, 프로세스분할(process sharing), 다수(ample)서비스 및 LCFSPR 서비스정책에 대하여 Product-form 해를 얻는 BCMP모형을 개발하였다. BCMP모형에서 평형상태 확률(Equilibrium state probabilities)은 다음과 같은 일반적인 형태를 갖는다.<sup>(1)</sup>

$$P(S) = C d(S) f_1(x_1) f_2(x_2) \cdots f_n(x_n)$$

- S : 시스템 상태
- $x_i$  :  $i$ 번째 서비스센터에 있는 고객의 형상
- $d(S)$  : 모델상태의 함수
- $f_i$  :  $i$ 번째 서비스센터의 정책에 따르는 함수
- C : 표준화 상수

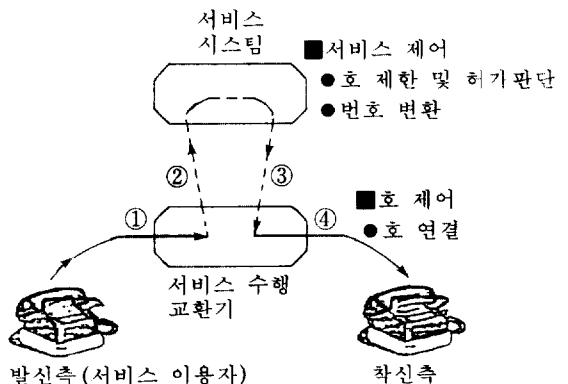
본 논문에서는 착신자 요금부담 서비스를 제공함에 있어 고객의 서비스품질을 보증한다는 측면에서 통화 처리시간을 일정시간이내로 만족시키면서 서비스를 제공하는 시스템(이하 서비스 시스템)의 용량을 결정하는 문제를 다루었다. 본 시스템은 BCMP큐잉망으로 모델링하였고 분석적 방법을 이용하여 해결하였다. 또한 용량을 산출할 수 있도록 속도와 용량과의 관계를 분석한 결과를 비선형계획법을 이용하여 제시하였다.

## 2. 시스템 개요

최근 종합정보통신망(ISDN)을 수행하기 위하여 공통신호망이 광범위하게 연구되고 있다.<sup>(2)</sup> 서비스시스템은 공통신호망을 이용하

망 서비스중에서 부가서비스 제공을 위한 고도 서비스망의 구성 요소로서 서비스 가입자의 제어정보를 이용하여 서비스를 제어하는 시스템이다. 서비스시스템은 적절한 전화처리 정보를 결정하여 다음번에 이 전화를 발송한 원래 신호중계기로 공통신호망을 통해 반응을 돌려보내는 기능을 수행한다.

우리나라에서도 국민의 소득수준 및 문화수준의 향상, 경제활동의 활성화, 통신의 편리성 인식 등으로 현재 선진국에서 제공하고 있는 착신자 요금부담 서비스(Freephone서비스, 800서비스), Credit call 서비스 등의 다양하고 편리한 망서비스에 대한 가입자의 수요가 발생할 것이다. 응용 예로서는 예약접수, 상품주문, 인쇄, 고객불평접수 등을 들 수 있겠다. 다이얼된 서비스 번호는 실제 트랜잭션 루팅을 위한 번호가 아니므로 트랜잭션 변경기능을 가지고 서비스 가입자의 여러 정보를 저장한 집중화된 데이터 베이스로 다이얼된 서비스 번호가 전달되어 접속 가능한 착신측 전화번호로 변환된다. 이때 서비스시스템으로 연결되는 전화망의 서비스수행 교환기는 서비스시스템에 이 서비스 번호에 대한 문의를 하고 서비스시스템으로부터 트랜잭션 접속 가능한 착신측 전화번호를 받아 최종적으로 트랜잭션의 루팅이 이루어진다. 서비스 트랜잭션 연결순서는 [그림 1]과 같다.<sup>(9)</sup>



[그림 1] 통화 서비스 흐름

본 논문의 관심사항은 서비스 수행 교환기 사이 즉 ②와 ③의 반응시간이며 서비스시스템 각 부분의 H/W 요구사항을 정의하여 적절한 용량을 결정하는 것이다.

서비스시스템의 대부분 요구사항은 다른 망노드 시스템의 요구사항과 유사하다. 서비스시스템은 신뢰도, 정확도, 사용자에게 의해 결정된 처리율(throughput)에 대한 엄격한 요구사항을 충족시켜야 한다.

800서비스 사용자를 위해 Bellcore는 다음 특성을 추천한다.<sup>(10)</sup>

- (1) 고장시간 3분/년 넘지 않을 것
- (2) 200 질문/초 이상
- (3) 평균 반응시간 0.5초 & 1초 이하에 99%

서비스시스템에서의 트랜잭션(transaction)의 특징은 판독(read)트랜잭션이 대부분이며 본 논문에서는 판독 트랜잭션만이 존재하는 것으로 고려하였다. 트랜잭션의 실행순서는 FCF-S이며 트랜잭션의 우선권이 없기 때문에 선점유(preemption)가 없다. 따라서 트랜잭션간에 충돌이 거의 발생하지 않는 것으로 고려하였다.

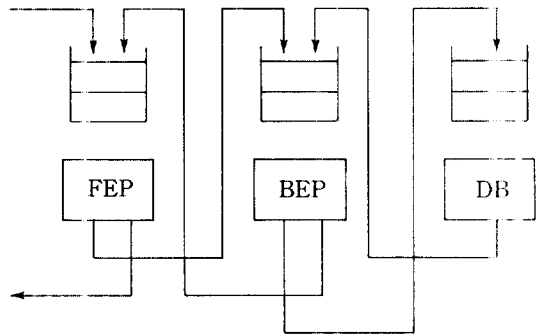
서비스시스템은 성능, 운용 및 상호연결관리의 요구사항을 만족시키기 위한 노드소프트웨어와 각 서비스에 대한 응용 소프트웨어를 포함한다. 높은 가용성과 결점 방지는 주 서비스시스템과 부 서비스 시스템을 동시에 사용하므로써 달성된다. 신호망에서는 시스템에 고장이 발생하는 경우 부 서비스 시스템에 전송토록하여 신뢰도를 높히도록 설계되고 있다.

### 3. 모델설정

이 절의 주요목적은 특정 모델에서 처리가 요구되는 질의들이 추천된 반응시간을 만족시키면서 수행될 수 있는지를 보이는 것이다. 서비스시스템의 여러 처리용량 윤곽을 검토하여 최적

의 서비스시스템 용량을 찾고자 한다.

본 논문에서 분석하고자 하는 시스템은 실시간처리를 필요로 하는 시스템으로 [그림 2]와 같이 모형화 할 수 있다.



[그림 2] 서비스시스템의 구조

처리해야할 입력 메시지는 질의(query)로서 대부분을 판독트랜잭션으로 가정하였다. 도착한 트랜잭션에 대한 경로를 따라가 보자. 도착한 트랜잭션은 3노드 중의 한곳에 있을 것이고 이산파라미터의 마르코프체인으로 모델링할 수 있다. 이 모형은 BCMP모형으로서 다음과 같은 특성을 갖는다고 가정하였다.

- 시스템은 안정적 확률과정으로 모형화 된다.
- 작업(job)은 확률적으로 독립이다.
- 작업은 마르코브체인에 따라 움직인다.
- 시스템은 확률적으로 평형상태에 있다.
- 시스템에 도착하는 트랜잭션 시간간격은 지수분포를 따른다.
- 시스템은 에르고딕(ergodic)이다-즉 long-term time average는 확률적 평형에 대해 계산된 값으로 수렴한다.
- 각 고객은, 서비스센터에서 서비스를 기다리거나 받는 동안, 한 등급에 속한다. 그러나 서비스 요청의 종료시 고정된 확률에 따라 등급과 서비스센터는 변할 수 있다.

### 4. 해 법

본 논문에서 적용할 해법은 다음과 같은 순서로 구성된다.

- 1) 파라미터 결정
- 2) 전이확률행렬 정의 및 방정식 해결(eir 구함)
- 3)  $\rho_i$  계산
- 4) 시스템내 작업수 계산
- 5) 반응시간 계산을 위한 비선형계획모형 설정
- 6) 용량계산

#### 가. 파라미터 결정

$\mu_3$ 등이고 모두 지수분포를 따른다. 서비스 정책은 선입선출(First In First Out)정책이다.

#### 나. 전이확률 정의 및 방정식 수립

에르고딕 체인에 대한 전이확률을 정의하고 방정식을 만든다. 전이확률은 다음과 같다.

$$P_{ir, js} = \begin{matrix} & 01 & 11 & 21 & 31 & 22 & 12 \\ \begin{matrix} 01 \\ 11 \\ 21 \\ 31 \\ 22 \\ 12 \end{matrix} & \begin{bmatrix} & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ 1 & & & & & & \end{bmatrix} & \end{matrix}$$

본 논문에서 필요한 파라미터는  $\lambda, \mu_1, \mu_2,$

이에 대한 독립균형방정식은 다음과 같다.

균 형 방 정 식	상 태
$1/2 * \mu_1 P(1, 0, 0) = \lambda P(0, 0, 0)$	$P(0, 0, 0)$
$1/2 * \mu_1 P(i+1, 0, 0) + 1/2 * \mu_2 P(i-1, 1, 0) + \lambda P(i-1, 0, 0)$ $= (\lambda + \mu_1) P(i, 0, 0)$	$P(i, 0, 0)$
$\mu_3 P(0, j-1, 1) + 1/2 * \mu_1 P(1, j, 0) + 1/2 * \mu_1 P(1, j-1, 0)$ $= (\lambda + \mu_2) P(0, j, 0)$	$P(0, j, 0)$
$1/2 * \mu_2 P(0, 1, k-1) + 1/2 * \mu_1 P(1, 0, k)$ $= (\lambda + \mu_3) P(0, 0, k)$	$P(0, 0, k)$
$\lambda P(l-1, m, 0) + 1/2 * \mu_2 P(l-1, m+1, 0)$ $+ 1/2 * \mu_1 P(l+1, m-1, 0) + 1/2 * \mu_1 P(l+1, m, 0)$ $+ \mu_3 P(l, m-1, 1)$ $= (\lambda + \mu_2 + \mu_1) P(l, m, 0)$	$P(l, m, 0)$
$\mu_3 P(0, n-1, p+1) + 1/2 * \mu_2 P(0, n+1, p-1)$ $+ 1/2 * \mu_2 P(1, n-1, p) + 1/2 * \mu_1 P(1, n, p)$ $= (\lambda + \mu_2 + \mu_3) P(0, n, p)$	$P(0, n, p)$
$\lambda P(q-1, 0, r) + 1/2 * \mu_1 P(q-1, 1, r)$ $+ 1/2 * \mu_1 P(q, 1, r-1) + 1/2 * \mu_2 P(q+1, 0, r)$ $= (\lambda + \mu_1 + \mu_3) P(q, 0, r)$	$P(q, 0, r)$
$\lambda P(s-1, t, u) + 1/2 * \mu_1 P(s+1, t, u) + 1/2 * \mu_1 P(s+1, t-1, u)$ $+ 1/2 * \mu_2 P(s-1, t+1, u) + 1/2 * \mu_2 P(s, t+1, u-1)$ $+ \mu_3 P(s, t-1, u+1) = (\lambda + \mu_1 + \mu_2 + \mu_3) P(s, t, u)$	$P(s, t, u)$

따라서 본 시스템의 Product-Form Solution 을 얻기 위한 방정식을 표시하면 다음과 같다.

$$\sum_{(i,r) \in E_k} eir Pir, js + qjs = ejs \quad (j, s) \in E_k \text{ (} E_k \text{는 시스템에 속한 전이)}$$

$$(qjs \text{는 외부입력확율})$$

$$\begin{bmatrix} & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ 1 & & & & & & \end{bmatrix} [e01, e11, e21, e31, e22, e12]^T = ejs$$

이에 대한 수식을 풀면  $e11=e21=e31=e22=e12=1$  이 된다. 실질적으로 윗식에서  $eir$ 은 시스템에서 업무가 등급  $r$ 로 서비스센터  $i$ 를 방문하는 평균회수라고 해석된다. 윗식의 해는 시스템이 개방인 경우 유일하게 주어진다. (7) BCMP모형의 결과를 적용하는 모든 경우에 독립균형방정식은  $\{eir\}$ 에 대한 방정식을 정의하는 것으로 줄어든다.

$$= \left( \sum_{n_1=0}^{\infty} \lambda^{n_1} h_1(n_1) \right) \left( \sum_{n_2=0}^{\infty} \lambda^{n_2} h_2(n_2) \right)$$

$$\left( \sum_{n_3=0}^{\infty} \lambda^{n_3} h_3(n_3) \right)$$

$$= \left( \frac{1}{1 - \frac{2 * \lambda}{\mu_1}} \right) \left( \frac{1}{1 - \frac{2 * \lambda}{\mu_2}} \right) \left( \frac{1}{1 - \frac{\lambda}{\mu_3}} \right)$$

$$P_i(n_i) = C \lambda^{n_i} h_i(n_i) \prod_{j=1}^3 \left( \sum_{n_j=0}^{\infty} \lambda^{n_j} h_j(n_j) \right)$$

$$= \lambda^{n_i} h_i(n_i) / \sum_{m=0}^{\infty} \lambda^m h_i(n_i)$$

다.  $\rho_i$ 계산

각 부분의 이용율은 다음과 같이 계산된다. 시스템의 통합상태를 각 창구의 총 작업수로 단순화하면 즉  $S=(n1, n2, n3)$ 라 하면  $P(S)=Cd(S) h_1(n_1)h_2(n_2)h_3(n_3)$ 가 된다.

가 된다.

$$h_1(n_1) = (2/\mu_1)^{n_1}$$

$$h_2(n_2) = (2/\mu_2)^{n_2}$$

$$h_3(n_3) = (1/\mu_3)^{n_3}$$

$$\rho_1 = 2 * \lambda * \left( \frac{1}{\mu_1} \right) * \left( \frac{1}{K1} \right)$$

$K1$ 은 시스템 용량

$$\rho_2 = 2 * \lambda * \left( \frac{1}{\mu_2} \right) * \left( \frac{1}{K2} \right)$$

$K2$ 은 시스템 용량

$$\rho_3 = \lambda * \left( \frac{1}{\mu_3} \right) * \left( \frac{1}{K3} \right)$$

$K3$ 는 DB의 속도배수

다음 표준화상수  $C$ 를 구한다.

이므로

$$C^{-1} = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \sum_{n_3=0}^{\infty} \left( \prod_{i=1}^3 \lambda^{n_i} h_i(n_i) \right)$$

$$P_i(n_i) = (1 - \rho_i) \rho_i^{n_i}$$

가 된다.

라. 시스템내 작업수 계산

시스템내 평균 작업수는 평균대기길이로서 다음과 같이 계산된다.

$$E[N_j] = \frac{\rho_j}{1 - \rho_j}$$

이것으로부터 시스템내 평균작업수와 평균반응시간이 계산된다.

$$E[N] = \sum_{j=1}^m \frac{\rho_j}{1 - \rho_j}$$

마. 반응시간 계산을 위한 비선형계획모형 설정

시스템내 각 부분의 평균 반응시간은 다음과 같이 구할 수 있다.

$$E[R_j] = \frac{1}{\lambda} \frac{\rho_j}{1 - \rho_j}$$

$$E[R] = \frac{1}{\lambda} \sum_{j=1}^m \frac{\rho_j}{1 - \rho_j}$$

본 논문에서는 시스템의 용량을 결정하는 것이 목적이므로 다음과 같은 모형이 성립될 것이다.

$$\text{Min } A * K1 + B * K2 + C * K3$$

$$E[R] \leq \text{요구반응시간}$$

$$\rho_j \leq \text{이용율 한도 } j=1, 2, 3$$

바. 용량계산

위의 비선형계획문제를 만족시키는 해가 본 모형의 해가 될 것이다.

이상으로 주어진 반응시간을 충족시키면서 트랜잭션을 처리할 수 있는 용량 결정에 관한 해법을 살펴보았다.

## 5. 수치 예제

현재 서비스시스템의 성능 특성이 완전히 결정되지 않은 상태이므로 정확한 데이터를 입력하는 것은 어려운 일이며, 따라서 개략적인 데이터를 입력하여 해를 구하였다.

입력자료는 다음과 같다.

$$\lambda = 200 \text{질의/초}$$

$$\frac{1}{\mu_1} = \frac{8000}{(K1 * 1000000)} \text{초}$$

$$\frac{1}{\mu_2} = \frac{15000}{(K2 * 1000000)} \text{초}$$

$$\frac{1}{\mu_3} = \frac{10000}{(K3 * 1000000)} \text{초}$$

입력자료를 바탕으로 각 부분의 이용율을 구하였다.

$$\rho_1 = 400 * \frac{8000}{(K1 * 1000000)} = \frac{32}{10 * K1}$$

$$\rho_2 = 400 * \frac{15000}{(K2 * 1000000)} = \frac{6}{K2}$$

$$\rho_3 = 200 * \frac{10000}{(K3 * 1000000)} = \frac{2}{K3}$$

요구 반응시간을 만족시키기 위하여 다음과 같은 비선형 계획모형을 설정하였다.

$$\text{Min } 10 * K1 + 10 * K2 + 10 * K3$$

$$\begin{aligned}
 E[R] &= \frac{1}{\lambda} \sum \frac{\rho_j}{1-\rho_j} \leq 0.5 \\
 &= \frac{1}{200} \left( \frac{32}{10 * K1 - 32} + \frac{6}{K2 - 6} \right. \\
 &\quad \left. + \frac{2}{K3 - 2} \right) \leq 0.5 \\
 \rho_1 &= \frac{32}{10 * K1} \leq 0.7 \\
 \rho_2 &= \frac{6}{K2} \leq 0.7 \\
 \rho_3 &= \frac{2}{K3} \leq 0.7
 \end{aligned}$$

본 비선형계획문제는 GINO를 이용하여 해결하였다. 그 결과 K1은 4.57 K2는 8.57 K3는 2.86을 얻었다. 해를 분석한 결과 각 부분의 용량은 해당 부분의 이용율에 영향을 받고있음을 알 수 있었다. 일 예로 데이터베이스의 속도배수가 3과 4일 때에 다른 프로세서의 용량은 거의 같은 값을 나타내었다. 목적함수식은 시스템의 설치비용으로 해석될 수 있으나 본 논문에서는 정확한 자료를 구하지 못하여 동일한 값으로 입력하였다.

## 6. 결 론

본 논문에서는 고객에게 제공하는 통화서비스의 품질기준을 만족시키기 위한 서비스시스템의 용량문제 및 반응시간을 연구하였다. 그 결과 FEP는 5 MIPS, BEP는 9 MIPS 그리고 데이터베이스는 기준속도의 3배로 구축하면 반응시간을 평균 0.5초 이내로 충족시키고 각 프로세서 및 데이터베이스의 이용율은 70% 이내로 만족시킬 수 있는 것으로 예측되었다. 계산과정에서 사용한 입력자료에는 트랜잭션 전송시간이 포함되지 않았기 때문에 가상의 반응시간을 설정하였다.

본 논문에서 반응시간 및 용량 계산 해법을 도출하였으며 용량결정까지의 과정을 수치 예로 보였다. 본 해법을 통하여 트랜잭션의 흐름이 종착점에서 서비스를 마치고 원래 위치로 복귀하는 종류의 서비스를 수행하는 시스템의 성능을 고려할 수 있게 되었다. 향후 연구과제는 트랜잭션의 종류가 다수인 경우와 트랜잭션의 전송시간을 포함하여 보다 정확한 시스템의 성능 및 반응 시간을 예측하는 것이다.

參 考 文 獻

1. Baskett, F., Chandy, K.M., Muntz, R.R. and Palacios, F.G. (1975), "Open, Closed and Mixed Network of Queues with Different Classes of Customers", J. ACM, Vol. 22, No. 2, pp. 248-260.
2. Choi, J.K., Jang, Y.M. and Lee, Y.H. (1988), "Queueing Network Analysis of STP System" JCCW'88, pp. 267-276.
3. Denning P.J., Buzen, T.P. (1978), "The Operational Analysis of Queueing Network Models", Computing Surveys, Vol. 10, No. 3, pp. 225-262.
4. Gross, D., Harris, C. (1985), "Fundamentals of Queueing Theor", 2nd Edition, John Wiley & Sons.
5. Heidelberger, P., Lavenberg, S.S. (1984. 12), "Computer Performance Evaluation Methodology", IEEE Tr. Computer, Vol. C-33, No. 13, pp. 1195-1220.
6. Kleinrock, L. (1975), Queueing System, Vol. I, Theory, John Wiley & Sons.
7. Reiser M. and Kobayashi, H. (1975), "Queueing Networks with Multiple Closed Chains: Theory and Computational Algorithms", IBM J of R & D, Vol. 19, pp. 283-294.
8. Trivedi, K.S. (1982), "Probaility and Statistics with Reliability, Queueing, and Computer Science Application", Prentice-Hall.
9. 고병도(1988), "SCP Overview", 한국전자통신연구소 지능망연구실.
10. 김지선(1988), "지능망 서비스제어 시스템 개념서", 한국전자통신연구소 지능망연구실.