

## 論 文

## 통신망에 있어서의 큐잉우선도에 관한 연구

正會員 金 榮 東\* 正會員 李 在 昊\*\* 正會員 宋 榮 宰\*\*\*

A Study on the Queueing Priority in  
Communication NetworkYoung Dong KIM \*, Jae Ho LEE \*\*, Young Jae SONG \*\*\* *Regular Members*

**요 약** 통신망에 있어서의 입력 메시지는 그 메시지가 가지는 우선도에 의해서 처리된다. 이 우선도 처리에 있어서 서로 다른 종류의 우선도를 가지는 메시지가 폭주하면 낮은 우선도의 메시지는 높은 우선도의 메시지가 끊이지 않는 한 우선도차를 넘는 지연을 갖게된다. 이 경우 낮은 우선도의 메시지에 대해서 연속지연을 보상하기 위한 평균지연보상방법을 제시하였다.

**ABSTRACT** The input messages in communication network are processed according to their priority. If the messages which are had different types of priority are excessively entered to communication network in such priority process, and if high priority messages are continuously entered, the messages which are had low priority are experienced delay beyond their priority. In this case, we proposed average delay compensation method to compensate continuous delay in low priority messages.

## 1. 서 론

데이터 통신망이나 컴퓨터 통신망에 있어서 그 통신망의 성능은 그 망을 통해서 전송되는 메시지의 시간지연에 의해서 평가된다. 이와같은 통신망의 전송에 관한 메시지의 지연은 그 메시지가 가지는 우선도에 의해서 그 시간지연이 좌우된다. 이 시간지연의 원인은 여러가지 근원으로 분류할 수 있으나, 우선도 차에 의한 지연은 우선도가 가지는 우선순위 이상으로 우선도 차를

가지게 한다. 이것은 서로 다른 메시지가 서로 다른 우선도로 입력될때 낮은 우선도의 메시지에 대한 지연보상에 관한 것으로 큐잉이론을 적용하여 해석한 것이다.

이와 같은 연구는 1964년에 Kleinrock<sup>(1)</sup>이 큐잉이론을 통신망에 적용했으며 1975년에 역시 Kleinrock<sup>(2)</sup>이 우선도가 없는 여러형태의 큐에 대한 지연을 해석했다. 우선도 할당에 관해서는 G. L. Fultz<sup>(3)</sup>가 1972년에 메시지 교환 컴퓨터 통신망에서의 우선도에 대해 연구했으며 1976년에 L. Kleinrock<sup>(4)</sup>이 고정 우선도에 대해서 연구했으며 역시 1976년에 R. Pichlitz와 C. McCoy<sup>(5)</sup>가 우선도가 통신망의 지연에 미치는 영향에 대해서 연구했으며 1977년에 A. Livine<sup>(6)</sup>이 컴퓨터통신망에 있어서의 우선도 큐에 대해 연구했으며 1982년에는 W. Wong, P. Savé, J. Field<sup>(7)</sup>가 패킷교환망에 있어서의 Scheduling에 우선도 큐잉이론

\* \*\*, 光云大學校電子通信工學科  
Dept. of Electronic Engineering, Kwangwoon  
University, Seoul, 132 Korea.

\*\*\* 慶熙大學校 電算科  
Dept. of Computer Science, Kyunghy University,  
Seoul, 132, Korea.  
論文番號 : 85-28(接受 1985. 7. 18)

을 적용했다.

이와 같은 연구는 단순히 각 입력 메시지에 그 메시지에 상응하는 우선도를 할당하고 그 우선도를 가지는 각 메시지 집단에 대한 평균지연과 평균큐길이의 해석에 치중한 것으로서 여러가지 메시지의 상태나 큐의 상태에 따른 적절한 우선도 할당에 대한 연구는 거의 시도되지 않았다. 따라서 본 논문에서는 통신망의 단일 노드의 큐의 상태에 따라서 적절하게 우선도를 재할당하는 큐잉처리를 강구하였다.

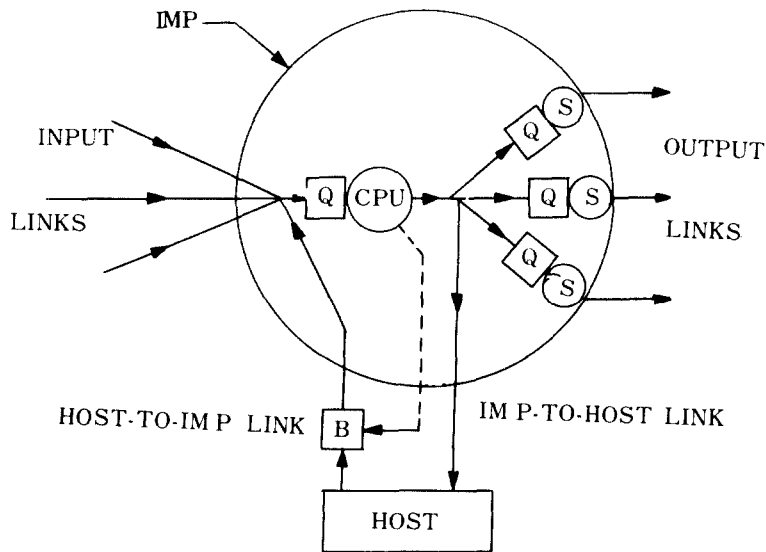
이와 같은 문제를 해결하는 방법으로는 큐에서 일정한 시간 이상 대기한 낮은 우선도 메시지의 우선도를 점차적으로 증가시켜 주는 방법, 우선도의 증가의 기울기를 서로 다르게 선택하는 방법, 큐의이용율( $\rho$ )에 따라서 메시지가 큐에 입력되기 전에 다시 적절하게 우선도 간의 메시지량을 재 조절하는 방법... 등이 있을 수 있다.

본 논문에서는 큐의 이용율에 따라서 우선도 간의 메시지량을 재 조절하는 방법을 택했다. 먼저 고정 우선도<sup>(8)</sup>를 분석해서 각 우선도의 지연을 지연곡선으로 구해 보았다. 다음으로 낮은 우선도의 지연을 평균적으로 보상하기 위한 방법의 이론적 해석을 제시했으며, 이 검증은 컴퓨터를 이용한 수치적인 모의 계산으로 대처하였다.

## 2. 노드 모델과 고정 우선도 해석

### 2.1 노드 모델

일반적으로 통신망 특히 컴퓨터 또는 데이터 통신망은 공간적으로 분리되어 있는 여러 개의 컴퓨터 시스템을 통신회선(또는 링크)을 이용해서 서로 연결해 놓은 것을 일컫는다. 이런 통신망에 있어서의 하나의 노드는 컴퓨터 시스템과 그 컴퓨터 시스템을 통신망에 연결시켜주는 통



- B = SWITCH WHICH CAN BLOCK TRAFFIC FLOW TO THE IMP
- CPU = CENTRAL PROCESSING UNIT ROUTINE
- Q = QUEUE
- S = SERVER (REPRESENTS THE FINITE RATE OF TRANSMISSION ON THE OUTPUT LINKS)

그림2-1-1 노드 구조  
Node structure.

신장치로서 구성되는데 그 구조를 보면 그림2-1-1과 같다<sup>(3)</sup>.

그림 2-1-1과 같은 한 노드에서 통신에 관련된 거의 모든 일들을 IMP(Inter face message processor)에서 처리된다. 이때 IMP가 처리하는 일 중의 하나가 통신망을 통해서 전송될 메시지에 우선도를 할당하는 일이다<sup>(3)</sup>. 이와 같이 IMP가 할당하는 우선도의 종류와 그 우선도에 해당하는 메시지의 일예를 표 2-1-1에 들어 놓았다. IMP에서의 모든 종류에 메시지는 큐와 채널을 통해서 처리된다. 그림2-1-1의 큐와 채널을보다 상세하게 나타내면 그림2-1-2와 같다.

표2-1-1 메시지 우선도 구조  
Message Priority structure

MESSAGE TYPE	CLASS	PRIORITY
ACKNOWLEDGMENT PACKETS	A	HIGHEST
DELAY TABLE UPDATING PACKETS	U	
SIGNLE-PACKET MESSAGES	P	
PACKETS FROM MULTI-PACKET MESSAGES	M	LOWEST

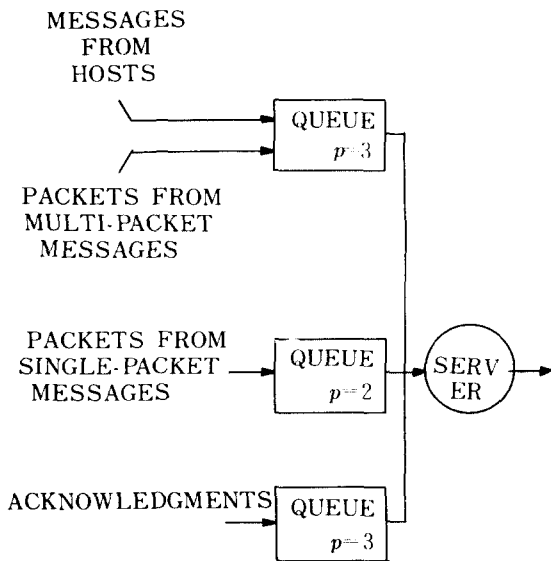


그림2-1-2 출력 채널 우선도 큐 모델  
Output channel priority queue model.

### 2.2 고정 우선도

고정 우선도의 지연을 해석하기 위해서 그림 2-1-2의 큐모형을 보다 용이한 형태의 모델로 변환하면 그림 2-2-1의 형태로 모델화 된다.

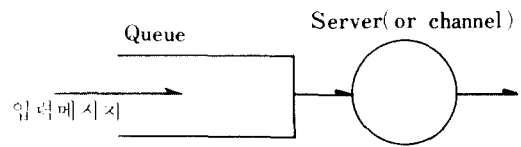
그림2-2-1(a)에서 메시지는 평균도착율  $\lambda$  인 poisson 분포로 입력되고 그 길이는 평균  $1/\mu$  의 지수분포를 가진다면 이 모델의 평균대기시간 ( $W_q$ ), 평균큐길이 ( $L_q$ ), 총지연시간 ( $T$ )는 다음과 같다<sup>(2)</sup>.

$$W_q = \frac{\rho / \mu}{1 - \rho} \quad (2-1)$$

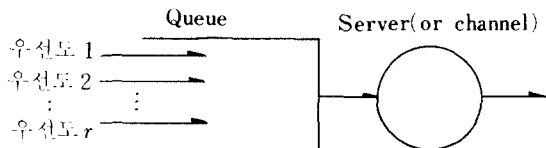
$$L_q = \frac{\rho}{1 - \rho} \quad (2-2)$$

$$T = W_q + \frac{1}{\mu} = \frac{1/\mu}{1 - \rho} \quad (2-3)$$

여기서  $\rho = \lambda / \mu$  로서 이용율(utilization factor)이다. 그리고 그림 2-2-1(a)에서 입력 메시지 집단이 여러개 존재하여 그 평균도착율이 각각  $\gamma_1, \gamma_2, \gamma_3, \dots$ 이며 그 분포가 poisson 분포라면  $\lambda = \gamma_1 + \gamma_2 + \gamma_3 + \dots$ 인 평균도착율을 가지는 단일 poisson 분포로 합성될 수 있다<sup>(9)</sup>.



(a) 단일 큐모형  
Single Queue Model



(b) r개의 우선도가 있는 큐 모델  
Queue Model with r Priority

그림2-2-1 큐모형  
Queue Model.

그림 2-2-1(b)와 같이  $r$  개의 메시지 집단이 있고 각각은  $P_k$ 의 우선도(non-preemptive priority;  $K = 1, 2, \dots, r$ ,  $K$ 가 작을 때 우선도는 높다)를 가지며 같은 우선도 내에서는 FCFS (First Come First Service) 방식으로 처리되며 각 우선도는 큐 외부에서 주어지며 일단 할당된 우선도는 채널에서 처리가 종료될 때까지 불변이라 생각하자. 또한 각 메시지는 그 집단별로 평균  $\lambda_k$ 의 Poisson 분포로 입력되고 그 길이는  $1/\mu_k$ 의 평균을 가지는 지수분포라 하자. 이때  $i$  번째 우선도 메시지의 대기시간(또는 시간지연)  $W_q^{(i)}$ 는 다음과 같다<sup>(8)</sup>

$$W_q^{(i)} = \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k} + \sum_{k=1}^{i-1} \rho_k W_q^{(k)}}{1 - a_i} \quad (2-4)$$

$$a_i = \sum_{k=1}^i \rho_k, \quad (\text{단 } a_i < 1)$$

이때 이용율  $\rho_k$ 는  $\rho_k = \lambda_k / \mu_k$  ( $K = 1, 2, 3, \dots, r$ )이다. 식(2-4)를 정리하면 다음과 같다 (10)

$$W_q^{(i)} = \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k}}{(1 - a_{i-1})(1 - a_i)} \quad (2-5)$$

또한 리틀(little)의 결과를 이용하면 각 우선도 별 평균 큐길이(또는 큐내에 대기하는 평균 메시지수)를 구할 수 있다.

$$L_q^{(i)} = \frac{\lambda_i \cdot \sum_{k=1}^r \frac{\rho_k}{\mu_k}}{(1 - a_{i-1})(1 - a_i)} \quad (2-6)$$

전체 입력 메시지에 대한 평균 큐길이는 다음과 같다.

$$L_q = \sum_{k=1}^r L_q^{(k)} \quad (2-7)$$

전체 우선도의 처리시간은 다음과 같다.

$$\frac{1}{\mu} = \sum_{k=1}^r \frac{\lambda_k}{\lambda} \cdot \frac{1}{\mu_k} \quad (2-8)$$

여기서,  $\lambda = \sum_{k=1}^r \lambda_k$ 이다. 만약  $\mu_i = \mu$ 라면 전체 입력 메시지에 대한 대

기시간  $W_q$ 는 다음과 같다.

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} \quad (2-9)$$

총 지연시간  $T$ 는 다음과 같다.

$$T = W_q + \frac{1}{\mu} = \frac{1/\mu}{1 - \rho} \quad (2-10)$$

식(2-3)과 식(2-10)은 동일하다. 즉 우선도가 없는 경우와 우선도가 있는 경우에 있어서의 총 지연은 같다. 그러나 높은 우선도를 가

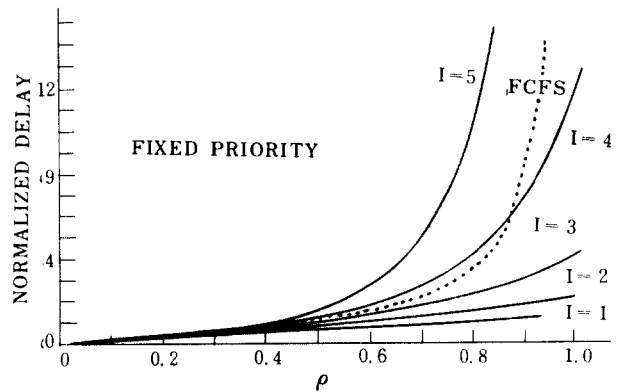


그림 2-2-2 FCFS와 우선도의 지연곡선  
Delay curve of FCFS & Priority.

LD	PRI.1	PRI.2	PRI.3	PRI.4	PRI.5	FCFS
0.10	0.5102	0.5315	0.5541	0.5782	0.6039	0.5556
0.15	0.7732	0.8225	0.8768	0.9366	1.0027	0.8824
0.20	1.0417	1.1322	1.2352	1.3528	1.4881	1.2500
0.25	1.3158	1.4620	1.6340	1.8382	2.0833	1.6667
0.30	1.5957	1.8133	2.0787	2.4069	2.8195	2.1429
0.35	1.8817	2.1380	2.5758	3.0767	3.7393	2.6923
0.40	2.1739	2.5800	3.1328	3.8700	4.9020	3.3333
0.45	2.4725	3.0153	3.7588	4.8159	6.3920	4.0909
0.50	2.7778	3.4722	4.4643	5.9524	8.3333	5.0000
0.55	3.0899	3.9614	5.2621	7.3294	10.9127	6.1111
0.60	3.4091	4.4856	6.1678	9.0144	14.4230	7.5000
0.65	3.7356	5.0481	7.1998	11.0997	19.3452	9.2857
0.70	4.0698	5.6525	8.3812	13.7147	26.5151	11.6666
0.75	4.4118	6.3025	9.7403	17.0454	37.4999	15.0000
0.80	4.7619	7.0028	11.3122	21.3675	55.5555	20.0000
0.85	5.1205	7.7583	13.1416	27.1046	88.5417	28.3333
0.90	5.4878	8.5747	15.2653	34.9379	160.7140	45.0000
0.95	5.8642	9.4584	17.8170	46.0272	395.8350	95.0002
1.00	6.2500	10.4167	20.8234	62.5001		

표 2-2-1 FCFS와 우선도의 지연  
Delay of FCFS & Priority.

지는 경우는 식(2-10)보다 총 지연이 작고 낮은 우선도를 가지는 경우는 식(2-10)보다 총

지연이 커진다. 이와 같은 경우에 있어서 각 우선도 별 대기시간의 관계가  $x$  (처리시간) = 5,  $x_i = x$  ( $i=1, 2, \dots, 5$ ),  $\rho = 0 \sim 1$ , 0,  $\lambda_i = \lambda/5$  인 경우에 대해 그림 2-2-2에 나타나 있다. 그림 2-2-2는  $W_q^{(i)}$ 를  $x_i$ 로 정규화시킨 즉  $W_q^{(i)}/x_i$ 에 대한 지연곡선이다. 또한 표 2-2-1은 시간지연을 계산한 값이다.

그림 2-2-2와 표 2-2-1에 의하면 우선도 1~3은 항상 FCFS (즉 우선도가 없는 경우)보다 그 지연이 작고 우선도 5는 FCFS보다 지연이 크며 우선도 4는  $\rho \leq 0.4$ 에서는 FCFS의 지연이  $\rho > 0.4$ 에서는 우선도 4의 지연이 작다는 것을 알 수 있다.

### 3. 낮은 우선도의 지연 보상

표 2-2-1을 살펴보면  $\rho = 0.3$ 이하에서는 각 우선도간의 지연이 2배를 넘지 못한다. 그러나  $\rho = 0.8$ 이상 일때는 우선도 1과 우선도 5간의 지연이 10배 이상된다. 이때 낮은 우선도의 지연을 보상하기 위해서는 서론에서 제시한 바와 같이 여러가지 방법이 있을 수 있으나 본 논문에서는 큐의 이용율( $\rho$ )에 따라서 평균적으로 지연을 보상해 주는 방법을 연구해 보았다. 이것은 우선도가 낮은 메시지의 양을 줄여 줌으로써 낮은 우선도 메시지의 지연을 줄이는 방법이다.

우선도가 높은 메시지는 거의 큐에 도착 즉시 처리되므로 메시지량이 다소 증가한다 해도 그 지연이 크게 증가하지는 않는다. 반면에 우선도가 낮은 메시지는 메시지량이 조금만 증가해도 그 지연은 크게 증가한다. 따라서 반대로 낮은 우선도의 메시지량을 큐의 이용율에 따라서 적절하게 줄여주고 그 줄어든 양에 높은 우선도를 부여하면 높은 우선도의 지연에는 거의 영향을 미치지 않으면서 낮은 우선도의 지연은 상당량 개선시킬 수 있다. 이때 낮은 우선도의 메시지를 그 전체에서 random하게 일정량만큼 그 우선도를 증가시켜서 지연을 줄이므로 평균 지연보상이라 한다.

이와 같은 경우에 있어서 각 우선도 메시지들

의 지연을 해석하기 위해서는 먼저 다음과 같은 가정을 하자.  $\alpha_i$  ( $i=1, 2, 3, \dots, r; 0 < \alpha_i < 1$ )의 비율로 낮은 우선도의 메시지를 높은 우선도의 메시지로 그 우선도를 증가시켜 주는 것이외에는 해석할 모델의 제반 조건이 2-2)절의 고정 우선도의 경우와 같다. 또 임의의 어떤 시간에서 큐에 우선도 1  $\rightarrow n_1$ , 우선도 2  $\rightarrow n_2, \dots$ 개 메시지가 큐에 대기하고 있다고 하자. 이때  $S_0$ 를 현재 처리중인 메시지가 처리 종료될 때 까지 필요로 하는 시간,  $S_k$ 를  $n_k$ 를 처리하기 위해 필요로 되는 총시간이라 하고 대기시간 동안에  $K < i$ 인 우선도를 가지는 메시지가  $n_k$ 개 도착했다 하면 이의 처리시간을  $S'_k$ 라 하자. 이때 우선도  $i$ 의 평균대기시간(평균지연)을 구해보면 아래와 같다.

1)  $i \neq 1, r$ 일때

$$W_q^{(i)} = E\{S_0\} + \sum_{k=1}^i E\{S_k\} + \sum_{k=1}^{i-1} E\{S'_k\} + E\{S_{i+1}\} - E\{S'_i\} \quad (3-1)$$

여기서  $E\{S_{i+1}\}$ ; 우선도  $i+1$ 에서  $i$ 로 우선도가 증가되는 메시지의 평균지연.

$E\{S'_i\}$ ; 우선도  $i$ 에서  $i-1$ 로 우선도가 증가되는 메시지의 평균지연.

그런데,  $E\{S_0\}$ 는 다음과 같다.

$$E\{S_0\} = P_r\{\text{System busy}\} \cdot E\{S_0 | \text{System busy}\}$$

여기서 system busy 확률은 다음과 같다.

$$\lambda \sum_{k=1}^r \frac{\lambda_k}{\lambda} \cdot \frac{1}{\mu_k} = \rho$$

또,  $E\{S_0 | \text{system busy}\}$

$$= \sum_{k=1}^r E\{S_0 | \text{system busy with } K \text{ type input}\} \cdot P_r\{\text{input has } K \text{ type priority}\}$$

$$= \sum_{k=1}^r \frac{1}{\mu_k} \cdot \frac{\rho_k}{\rho}$$

$$\therefore E\{S_0\} = \rho \cdot \sum_{k=1}^r \frac{1}{\mu_k} \cdot \frac{\rho_k}{\rho}$$

$$= \sum_{k=1}^r \frac{\rho_k}{\mu_k} \quad (3-2)$$

또,  $E[S_k]$ 를 구해보면,

$$\begin{aligned} E[S_k] &= E[n_k \cdot S_k^{(n)}] = E[n_k] \cdot E[S_k^{(n)}] \\ (\because n_k \text{와 } S_k^{(n)} \text{는 서로 독립, } S_k^{(n)} \text{은 평균처리 시간}) \\ &= \frac{E[n_k]}{\mu_k} = \frac{\rho_k W_q^{(k)}}{\mu_k} \\ &= \rho_k \cdot W_q^{(k)} \quad (3-3) \end{aligned}$$

$$\begin{aligned} E[S'_k] &= \frac{E[n'_k]}{\mu_k} \\ &= \frac{\lambda_k W_q^{(i)}}{\mu_k} \quad (\because i \text{ 메시지를 처리하는 동안 도착}) \\ &= \rho_k \cdot W_q^{(i)} \quad (3-4) \end{aligned}$$

또,  $E[S'_i]$ 를 구해보면, 다음과 같다.

$$E[S'_i] = \frac{E[n'_i]}{\mu_i} = \frac{\lambda'_i W_q^{(i)}}{\mu_i}$$

그런데,  $\lambda'_i$ 는  $\lambda_i$  도착중 우선도가 증가된 도착을 나타내고 그 증가율을  $\alpha_i$ 라 했으므로  $\lambda'_i$ 는 다음과 같다.

$$E[n'_i] = \lambda'_i \cdot W_q^{(i)} = \alpha_i \lambda_i W_q^{(i)}$$

따라서,

$$\begin{aligned} E[S'_i] &= \frac{E[n'_i]}{\mu_i} = \frac{\alpha_i \lambda_i W_q^{(i)}}{\mu_i} \\ &= \alpha_i \rho_i W_q^{(i)} \quad (3-5) \end{aligned}$$

그러므로 식(3-2)~(3-5)를 식(3-1)에 대입해서 정리하면,

$$\begin{aligned} W_q^{(i)} &= \sum_{k=1}^r \frac{\rho_k}{\mu_k} + \sum_{k=1}^i \rho_k W_q^{(k)} + \sum_{k=1}^{i-1} \rho_k W_q^{(i)} + \alpha_{i+1} \\ &\quad \rho_{i+1} W_q^{(i)} - \alpha_i \rho_i W_q^{(i)} \end{aligned}$$

$$\left(1 - \sum_{k=1}^{i-1} \rho_k - \alpha_{i+1} \rho_{i+1} + \alpha_i \rho_i\right) W_q^{(i)}$$

$$= \sum_{k=1}^r \frac{\rho_k}{\mu_k} + \sum_{k=1}^i \rho_k W_q^{(k)}$$

$$\therefore W_q^{(i)} = \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k} + \sum_{k=1}^i \rho_k W_q^{(k)}}{\left(1 - \sum_{k=1}^{i-1} \rho_k - \alpha_{i+1} \rho_{i+1} + \alpha_i \rho_i\right)} \quad (3-6)$$

이 성립된다.

식(3-6)을 정리하면 다음과 같다.

$$W_q^{(i)} = \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k} + \sum_{k=1}^{i-1} \rho_k W_q^{(k)}}{1 - \alpha_i - \alpha_{i+1} \rho_{i+1} + \alpha_i \rho_i} \quad (3-7)$$

식(3-7)에서  $\alpha_i = 0$  이라면 즉 우선도의 증가가 없다면 식(3-7)은 식(2-5)와 일치한다.

2)  $i = 1$  (가장 높은 우선도)

$$\begin{aligned} W_q^{(1)} &= E[S_0] + E[S_1] + E[S'_2] \\ &= \sum_{k=1}^r \frac{\rho_k}{\mu_k} + \rho_1 W_q^{(1)} + \alpha_2 \rho_2 W_q^{(1)} \\ \therefore W_q^{(1)} &= \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k}}{1 - \rho_1 - \alpha_2 \rho_2} \quad (3-8) \end{aligned}$$

3)  $i = r$  (가장 낮은 우선도)

$i = r$  일 경우에는 더 낮은 우선도 집단이 없으므로  $\alpha_{i+1} = 0$  이다. 따라서 지연  $W_q^{(r)}$ 는 다음과 같다.

$$W_q^{(r)} = \frac{\sum_{k=1}^r \frac{\rho_k}{\mu_k} + \sum_{k=1}^{r-1} \rho_k W_q^{(k)}}{1 - \alpha_{r-1} + \alpha_r \rho_r} \quad (3-9)$$

식(3-7)~(3-9)는 낮은 우선도의 메시지를  $\alpha_i$  비율로 우선도를 증가시켰을 경우의 지연의 이론적 해석을 보여주고 있다.

#### 4. 컴퓨터 해석

각 우선도간의 입력량이 동일하고 그 우선도가 고정일 경우에 표 2-2-1에서 알 수 있듯

이  $\rho$  가 크면(예,  $\rho \geq 0.8$ ) 우선도 5의 메시지는 우선도 1의 메시지에 비해서 그 지연이 10배 이상된다. 이때 우선도 5의 지연을 보상해 주기 위해서 각 우선도의 메시지량의 10%씩을( $\alpha_i=0.1$ ) random하게 한 단계씩 우선도를 증가시켜 주었다. 이렇게 하면 우선도 1은 입력량이 10%증가되고 우선도 5는 10% 감소되지만 2~4 우선도의

입력량에는 변화가 없다. 이와 같이 낮은 우선도의 지연을 보상한 경우에 있어서의 각 우선도의 정규화 지연곡선이 그림 4-1에 표시되어 있다.

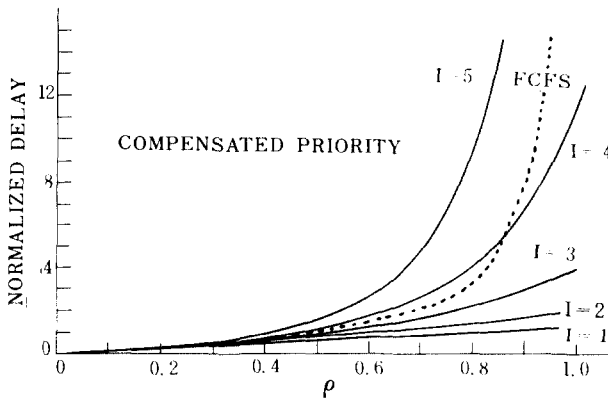


그림 4-1 보상된 지연곡선 compensated delay curve.

표 4-1 보상된 지연 Compensated delay.

LD	PRI.1	PRI.2	PRI.3	PRI.4	PRI.5
0.10	0.5112	0.5315	0.5541	0.5782	0.6026
0.15	0.7756	0.8226	0.8769	0.9367	0.9992
0.20	1.0460	1.1324	1.2354	1.3530	1.4809
0.25	1.3228	1.4624	1.6344	1.8387	2.0701
0.30	1.6060	1.8140	2.0795	2.4079	2.7967
0.35	1.8960	2.1892	2.5772	3.0783	3.7014
0.40	2.1930	2.5898	3.1350	3.8727	4.8409
0.45	2.4972	3.0180	3.7621	4.8202	6.2948
0.50	2.8090	3.4761	4.4693	5.9591	8.1791
0.55	3.1286	3.9668	5.2694	7.3395	10.6670
0.60	3.4562	4.4931	6.1780	9.0294	14.0262
0.65	3.7923	5.0581	7.2140	11.1216	18.6892
0.70	4.1371	5.6655	8.4006	13.7465	25.3916
0.75	4.4910	6.3195	9.7665	17.0914	35.4726
0.80	4.8544	7.0246	11.3474	21.4339	51.6001
0.85	5.2276	7.7859	13.1883	27.2009	79.8112
0.90	5.6110	8.6093	15.3471	35.0790	136.7490
0.95	6.0051	9.5016	17.8983	46.2372	288.1460
1.00	6.4103	10.4701	20.9402	62.8206	1256.4300

표 4-2 고정 우선도와 보상된 지연의 비교. Comparison between fixed priority and compensated delay.

LD	PRI.1	PRI.2	PRI.3	PRI.4	PRI.5
0.10 :	+0.2502070 :	+0.0061775 :	+0.0062016 :	+0.0062073 :	-0.2716650
0.15 :	+0.3559460 :	+0.0122820 :	+0.0122830 :	+0.0123116 :	-0.4009830
0.20 :	+0.4645700 :	+0.0206318 :	+0.0206499 :	+0.0206788 :	-0.5455870
0.25 :	+0.5759690 :	+0.0313153 :	+0.0313195 :	+0.0313544 :	-0.7086410
0.30 :	+0.6901140 :	+0.0443885 :	+0.0444096 :	+0.0444478 :	-0.8943550
0.35 :	+0.8070660 :	+0.0599420 :	+0.0599827 :	+0.0600303 :	-1.1084600
0.40 :	+0.9268940 :	+0.0780998 :	+0.0781411 :	+0.0781784 :	-1.3585200
0.45 :	+1.0497000 :	+0.0989164 :	+0.0989635 :	+0.0990531 :	-1.6550700
0.50 :	+1.1755700 :	+0.1225070 :	+0.1225780 :	+0.1226756 :	-2.0128400
0.55 :	+1.3048100 :	+0.1489830 :	+0.1490700 :	+0.1491820 :	-2.4532400
0.60 :	+1.4369600 :	+0.1784700 :	+0.1785520 :	+0.1786930 :	-3.0085500
0.65 :	+1.5727500 :	+0.2110440 :	+0.2111510 :	+0.2113120 :	-3.7298900
0.70 :	+1.7121100 :	+0.2468500 :	+0.2469790 :	+0.2471730 :	-4.7032100
0.75 :	+1.8551200 :	+0.2860040 :	+0.2861660 :	+0.2864170 :	-6.0852700
0.80 :	+2.0020000 :	+0.3286680 :	+0.3288390 :	+0.3291490 :	-8.1964000
0.85 :	+2.1528800 :	+0.3749660 :	+0.3751540 :	+0.3755470 :	-11.8150000
0.90 :	+2.3079300 :	+0.4250460 :	+0.4252620 :	+0.4257750 :	-19.5164000
0.95 :	+2.4673200 :	+0.4790570 :	+0.4793260 :	+0.4799890 :	-52.2632000
1.00 :	+2.5641000 :	+0.5128210 :	+0.5128210 :	+0.5128240 :	-100.0010000

그림 4-1의 각 변수의 값은 그림 2-2-1과 같으나  $\alpha_i (i = 1, 2, \dots, 5) = 0.1$ 이다. 표 4-1은 그림 4-1에 대한 비 정규화 지연을 나타내고 있다. 표 4-2는 표 4-1의 지연값을 표 2-2-1의 지연값과 비교해서 각 우선도의 지연의 변화율  $\rho = 0.1 \sim 0.95$ 의 범위에서 5개의  $\rho$ 에 대한 평균(예로써,  $\rho = 0.01 \sim 0.15, \dots$ ) 값으로 나타내고 있다.

표 4-1과 표 4-2를 고찰해 보면, 우선도 1~4의 지연은 증가했고 우선도 5의 지연은 감소했다. 우선도 1은 입력 메시지량의 증가로 인해서, 우선도 2~4는 입력 메시지량에는 변화가 없으나 우선도 1의 지연의 증가에 영향을 받아서 그 지연이 증가했다. 이때 우선도 2~4의 지연 증가량은 평균적으로 우선도 4 > 우선도 3 > 우선도 2가 된다. 그러나 그 증가량은 무시할 수 있을 정도로 작다.

그림 4-2는  $\rho = 0 \sim 1.0$ 에서 5개의 우선도에 대한 지연의 변화를 나타낸 곡선이다. 우선도 1은 최대 2.5%정도, 우선도 2~4는 최대 0.5%정도 지연이 증가했다. 반면 우선도는 5 최대

65% 지연이 감소했다. 따라서 우선도 5의 지연의 감소량이 우선도 1~4의 지연의 증가량(약 4%)을 훨씬 능가한다는 사실을 알 수 있다. 이와 같은 낮은 우선도의 지연의 감소량과 높은 우선도의 지연의 증가량 차의 크기는 그림 4-2에서 볼 때 이용율( $\rho$ )에 따라서 달라지며, 이용율이 작을 수록 그 차는 적어지고 이용율이 클수록 그 차는 커진다(예를 들면,  $\rho = 0.5; +1.45, -1.85, \rho = 0.6; +1.89, -2.75, \rho = 0.65; +2.11, -3.39, \rho = 0.7; +2.34, -4.24; +$ 증가량,  $-$ 감소량). 이렇게 볼 때 평균지연보상방법은  $\rho$ 가 클 때(예로써  $\rho \geq 0.7$ ) 그 효과가 커진다는 것을 알 수 있다.

지금까지는, 이용율( $\rho$ )이 0~1.0범위에 있어서의 지연의 증감에 대해서 고찰해 보았다. 다음으로 이용율은 일정하지만 각 우선도간의 입력 메시지량이 변할 경우의 지연의 증감을 알아보기 위해서 우선도 1:우선도 2의 입력 메시지량의 비율을 1~99%까지 변화시켜 보았다(이때 우선도는 2종류를 택했다). 이와 같은 경우의 결과로서 그림(4-3)은 특정  $\rho$ 에 대한 지연곡선을 표(4-3)은 우선도 1과 우선도 2의 입력량의 비율을 나타내고 있다.

표 4-3에서 높은 우선도와 낮은 우선도간의 입력 메시지량의 비율을 알 수 있다. 이때 적절한 한계는 낮은 우선도의 지연 개선율과 높은 우선도의 지연 증가율이 1:1인 경우를 택했으며 이때를 최대입력비율이라 했다.

이 최대입력비율로 부터  $\rho = 0.8; 4$ 배,  $\rho = 0.9; 6.69$ 배,  $\rho = 0.95; 13.18$ 배,  $\rho = 0.99; 48.5$ 배(높은 우선도의 낮은 우선도에 대한)이하로 입력되어야 한다는 것을 알 수 있다.

이상의 결과에서  $\alpha_i (i = 1, 2, \dots) = 0.1, \mu_i (x_i) = \mu (x)$ 인 경우에 있어서의 평균지연보상방법은  $\rho$ 가 클 때, 또 상하위 우선도의 입력 메시지량의 비율이 최대 48.5( $\rho = 0.99$ )이하일 경우에 효율적이라는 사실을 알 수 있다.

IMP에 평균지연보상방법을 구현할 때는 이용율  $\rho$  (또는 평균 큐길이)를 측정할 수 있거나 IMP에 입력될 메시지량을 예측할 수 있어야 한다. 또한 채널의 처리 속도를 예측할 수 있어야 한다.

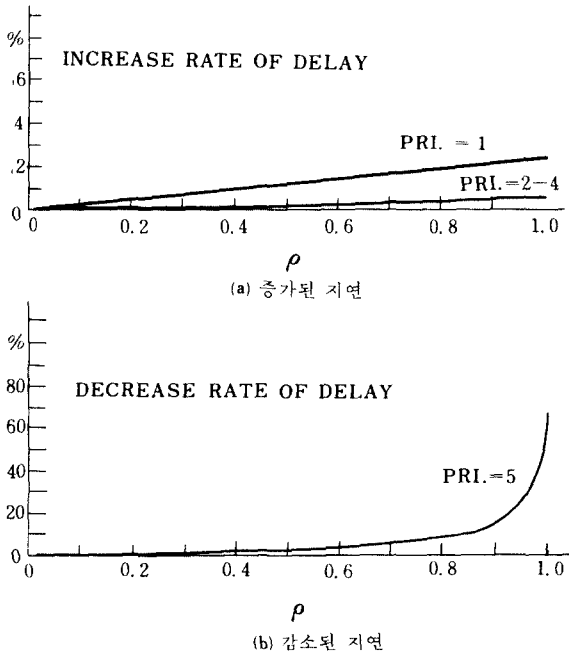


그림 4-2 각 우선도에서의 지연의 증가와 감소  
Increase and decrease in each priority.



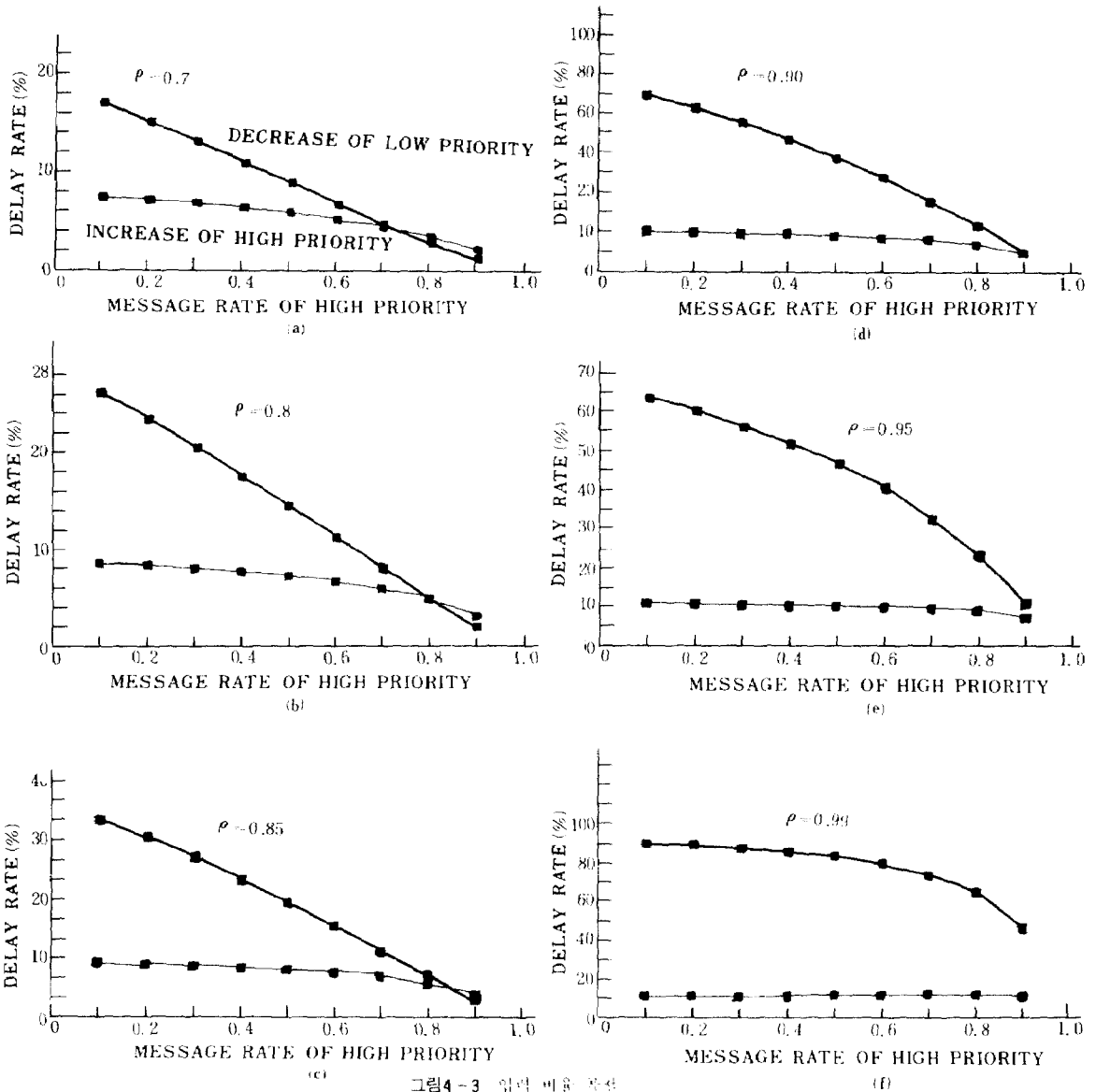


그림 4-3 입력 비율 우선  
Input rate curve.

표 4-3 입력비율  
Input rate.

$\rho$	$\rho$ (우선도 1 : 우선도 2)		우선도 1/우선도 2(배)		
	증가율	1 : 2	1 : 1	입력량    입력량	
0.80	0.4	: 0.4	0.64	: 0.16	1        4
0.90	0.648	: 0.252	0.783	: 0.117	2.57    6.69
0.95	0.817	: 0.133	0.883	: 0.067	6.14    13.18
0.99	0.96	: 0.03	0.97	: 0.02	32        48.5

이때  $\rho$  가 크면 낮은 우선도의 메시지가 후에 입력되기 전에 그 메시지량에서  $\alpha_i$  비율로 랜덤(random)하게 그 우선도를 증가시키면 된다. (이순시를 IMP 적용 알고리즘에 나타내었다)

이렇게 함으로써 낮은 우선도의 지연을 평균적으로 보장할 수 있다.

〈IMP 적용 알고리즘〉

- ① 컴퓨터 통신망의 IMP의 장기간 평균(또는 예측) 이용율을 측정한다.
- ② 각 우선도에 해당하는 메시지량의 장기간 평균(또는 예측값)을 측정해서 우선도간의 메시지량의 비율을 계산한다.
- ③ 과정①과 과정②를 이용해서 적정  $\alpha_i$ 를 계산한다.
- ④ 메시지가 도착하는 임의의 시간 구간  $T$ 에 있어서의 이용율( $\rho$ )을 측정한다.
- ⑤ 이용율이 높으면  $\alpha_i$ 만큼 도착 메시지중 random하게 우선도를 증가시켜서 우선도를 부여하고 그렇지 않으면 원래의 우선도를 부여한다.
- ⑥ 도착 메시지를 큐에 입력시킨다.
- ⑦ 일정시간( $T$ ) 경과후에 이용율을 재 측정하고 그 이용율로서 과정⑤와 과정⑥을 반복한다.

5. 결 론

본 논문에서는 고정 우선도(Fixed Priority, non-Preemptive Priority) 이론을 분석해서 각 우선도의 지연을 우선도가 5종류,  $\rho = 0 \sim 1.0$  범위에서 구해 보았다. 결과로서  $\rho$ 가 클 때 낮은 우선도의 지연이 높은 우선도의 지연에 비해서 상당히 크다는 것을 알 수 있었다. 이 경우에 낮은 우선도의 지연을 평균적으로 줄이기 위해서  $\alpha_i$ 라는 파라메타를 도입했다. 이 때  $\alpha_i$ 를 이용해서 낮은 우선도의 메시지의 지연을 감소시키는 평균지연보상방법을 이론적으로 해석했다. 아울러  $\alpha_i$ 가 공히 0.1인 경우,  $1/\mu_i$ 가 모두 같은

경우에 대해서 컴퓨터를 이용해서 해석해 보았다. 해석 결과에 의하면 이용율은 커야하며(예,  $\rho \geq 0.7$ ) 우선도간의 메시지량의 비율은 최대 48.5 배( $\rho = 0.99$ ) 이하일 때 효과적이라는 사실을 알 수 있었다. 이때 우선도 1~4의 지연의 증가는 최대 4%였으며 우선도 5는 최대 65% 지연이 감소했다.

參 考 文 獻

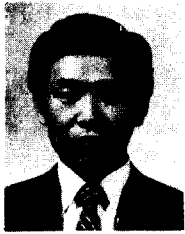
- (1) L. Kleinrock, "Communication Nets," New York: McGraw-Hill, 1964.
- (2) L. Kleinrock, "Queueing System," vol. 1: Theory, New York: Wiley Interscience, 1975.
- (3) G. L. Fultz, "Adaptive routing techniques for message switching computer communication networks", UCLA-ENG-7252, July, 1972.
- (4) L. Kleinrock, "Queueing system; vol. 2," Computer applications, New York: Wiley-Interscience, 1976.
- (5) R. Pikholtz, C. McCoy, "Effects of a priority discipline in routing for packet-switched networks", IEEE Trans. on Commun., vol. COM-24, no. 5, pp. 506~516, May, 1976.
- (6) A. Livine, "Dynamic routing in computer communication Networks," Ph. D. dissertation, Dep. of E., Polytechnic Institute of New York, 1977.
- (7) W. Wang, J. Savé, J. Field, "A study of fairness in packet-switched networks", IEEE Trans. on Commun. vol. COM-30, no. 2, pp. 346~353, Feb. 1982.
- (8) G. Harris, "Fundamentals of queueing Theory," New York: John Wiley & Sons, 1974.
- (9) H. Kobayashi, "Modeling and analysis: An introduction to system performance evaluation methodology, New York: Addison-Wesley, 1978.
- (10) Cobham, "Priority assignment in waiting line problems," Operation Research, 2, pp. 70~76, 1954.



金榮東 Young Dong KIM 正會員  
1960年12月22日生  
1984年2月: 光云大學電子通信工學科卒業  
1984年3月~現在: 光云大學大學院電子通信工學科 在學.



李在昊 (Jae Ho LEE) 正會員  
1934年5月26日生  
1968年2月: 光云大學通信工學科 卒業  
1978年2月: 檀國大學校大學院電子工學科(工學碩士)  
1970年~現在: 光云大學電子通信工學科 教授 本學會理事



宋榮宰(Young Jae SONG) 正會員

1947年4月20日生

1969年2月：仁荷大學校電子工學科卒業

1976年3月：日本Keio Univ. 大學院卒業(工學博士)

1982年：美國 Univ. of Maryland 客員教授

1976年～現在：慶熙大學校電算科副教授