

Performance Analysis of LAN Interworking Unit for Capacity Dimensioning of Internet Access Links

Chul-geun Park*, Dong-hwan Han** Regular Members

ABSTRACT

We build and analyze some types of queueing model to discuss capacity dimensioning of access links of a LAN interworking unit connected to the Internet backbone network. We assume that the IWU has a FIFO buffer to transmit IP packets to the Internet through the backbone. In order to analyze the system, we use a Poisson process and an MMPP process as input traffic models of IP packets and we use a general service time distribution as a service time model. But we use both an exponential service time and a deterministic service time in numerical examples for simple and efficient performance comparisons. As performance measures, we obtain the packet loss probability and the mean packet delay. We present some numerical results to show the effect of arrival rate, buffer size and link capacity on packet loss and mean delay.

I. 서 론

A rapid growth of the Internet leads to an increasing number of Internet service providers as well as has a great impact on the Internet infrastructure: CATV Internet ATM (Asynchronous Transfer Mode) backbone to access the Internet. For an efficient design of such an infrastructure including capacity dimensioning of the access link, it becomes essential to study the characteristics of Internet traffic performance of the IWU (Interworking Unit) such as routers and servers in a LAN(Local Area Network).

Queueing Analysis is a popular tool for the performance evaluation of telecommunication networks and systems. To yield useful data, however, this requires accurate models of the networks and systems under study. Accordingly, the characteristics of Internet traffic have to be investigated more carefully. Fortunately, many studies have already been done on analyzing Internet traffic. B.A. Mah^[1] has shown that the

global Internet traffic is dominated WWW(World Wide Web) and the document sizes of WWW traffic have heavy talled distributions. M. Nabe et al. [2] have also shown that WWW traffic is generating a major part of Internet traffic. They have studied the characteristic of WWW traffic using the access log data and have found that the document size follows a lognormal distribution. They have also found that the request interarrival time follows an exponential distribution during the busiest hour. Based on these results, they have built an M/G/1/PS(Processor Sharing) queueing model to discuss a design methodology of the Internet access network with a time sharing proxy server.

Currently, Internet traffic is engineered and managed using traditional circuit methods, but the properties displayed by Internet traffic are not adequately captured by conventional model.[3] Erlangian stochastic fundamental differences in application traffic patterns and usages between Internet and other services such as voice service and cellular mobile radio service, the lack of more suitable methods

^{*} 선문대학교 전자정보통신공학부,

논문번호: 99405-1009, 접수일자: 1999년 10월 9일

[₩]본 연구는 한국통신 통신망연구소 1999년 HAN/B-ISDN 위탁과제 지원으로 수행되었습니다.

is resulting in capacity and performance problems. Several papers^[1-3] have indicated that the document sizes of WWW traffic have heavy tailed distribution, but all documents of Internet traffic have to be packetized in IP(Internet Protocol) layer. This generation process of IP packets may have different traffic characteristics compared with generation process of the documents.

A recent study^[4] on Internet traffic properties has revealed that the arrival process of IP packets is bursty in nature. Z. Niu at al.^[5] have used an MAP(Markovian Arrival Process), a versatile point process by which typical bursty arrival process like the MMPP(Markov Modulated Poisson Process) is treated as special case to model the nature of IP packet arrivals in a LAN.

In this paper, we have a concern on queueing model and performance analysis of the LAN IWU such as a router for capacity dimensioning of Internet access links connected to the Internet access backbone. We assume that the IWU has a FIFO(first in first out) buffer to transmit IP packets to the Internet through the backbone network. We also use an MMPP process to model packet arrivals of the LAN IWU.

The overall organization of this paper is as follows. In the section II, we describe the system model of the IWU with FIFO buffer and study the performance analysis of this system by using an MMPP/G/1/K queue. In the section III, we take some numerical examples for the performance measures of our queueing model and show some examples for capacity dimensioning of access links of the IWU connected to the Internet. We finally have conclusions in the section IV.

II. System model and queueing analysis

System model

The access link model and the queueing model of the LAN IWU under consideration are depicted in Fig. 1 and Fig. 2, respectively. The IWU is

connected to the Internet backbone network such as ATM network through the access link. The transmitting buffer size of the IWU is K-1.

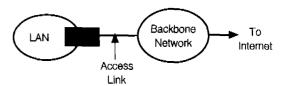


Fig. 1 IWU and access link model

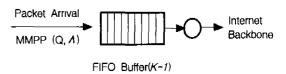


Fig. 2 Queueing model of IUW

We use a generally distributed service time with distribution H(x) and mean h, thus modeling the changing size of IP packets. The mean of the service rate is h^{-1} . The arrival process of IP packets is an MMPP. Packets arriving when the data buffer is full are blocked and lost and the packets queued in the buffer are served on a FIFO basis. Then we derive an MMPP/G/1/K queueing model for the performance estimation of the LAN IWU such as a router.

2. Queueing analysis

It is well known^[4,5] that a Poisson process is not appropriate to model burstiness of IP packets arrivals. A doubly stochastic Poisson process is a generalization of the Poisson process in which the rate $\lambda(t)$ is a non-negative valued stochastic process. The MMPP is an elementary and useful doubly stochastic Poisson process in which the arrival rate is given by $\lambda[f(t)]$, where f(t), $t \ge 0$, is an m-state irreducible MC (Markov Chain). The arrival rate can take only m values $\lambda_1, \dots, \lambda_m$. We use an m-state MMPP to model the input traffic into the IWU data buffer. The m-state MMPP is characterized by the infinitesimal generator Q of the underlying MC and by an arrival rate matrix $\Lambda = diag(\lambda_1, \dots, \lambda_m)$. Let N(t) be the number of

arrivals during the time interval [0,t) and let J(t) be the state of the underlying MC at time t. We denote the transition probabilities of the process $\{N(t), J(t), t \ge 0\}$ by

$$P_{ij}(n, t) = P\{N(t) = n, J(t) = j \\ |N(0) = 0, J(0) = i\}.$$
 (1)

The $m \times m$ matrix P(n, t) of the transition probabilities has the probability generating function as follows, $^{[6]}$

$$P^*(z,t) = e^{Qt + (z-1)At}, |z| \le 1.$$
 (2)

Let X(t) be the number of packets in the data buffer at time t. The MMPP/G/1/K queue as a queueing model of the LAN IWU can be described by means of a two dimensional state variable $\{\overline{X}(t) = (X(t), J(t)), t \ge 0\}$, where J(t) represents the state of the underlying Markov process that modulates the Poisson arrivals. Our aim is to determine the limiting distri- bution of the state $\overline{X}(t)$ whose state space consists of $\{0,1,\cdots,K\}\times$ $\{1, \dots, m\}$. T do this, we will use the classical embedded MC approach, considering the service completion epochs. Let t_n be the n-th embedded time point. Let $X_n = X(t_n +)$ and $J_n = J(t_n +)$ be the state of the queueing system and the state of the underlying MC just after t_n respectively. Then $\{(X_n, I_n), n \ge 0\}$ forms a finite MC with state space $\{0, 1, \dots, K-1\} \times \{1, 2, \dots, m\}$. The one-step transition matrix \mathcal{F} is given by

$$P = \begin{pmatrix} UA_0 & UA_1 & UA_2 & \cdots & UA_{K-2} & U\overline{A}_{K-1} \\ A_0 & A_1 & A_2 & \cdots & A_{K-2} & \overline{A}_{K-1} \\ 0 & A_0 & A_1 & \cdots & A_{K-3} & \overline{A}_{K-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_0 & \overline{A}_1 \end{pmatrix}, \quad (3)$$

where $A_k = \int_0^\infty P(k, t) dH(t)$, $k = 0, 1, \dots$, and \overline{A}_k

= $\sum_{i=k}^{\infty} A_i$, $1 \le k \le K-1$ and the matrix U which accounts for the evolution of f(t) during server's idle periods and whose (i, j) entry denotes the conditional probability of reaching phase j at the end of an idle period, starting from phase i, is given by

$$U = (\Lambda - Q)^{-1} \Lambda. \tag{4}$$

Let $\pi_{h,j}$ define the limiting probability distributions of $\overline{X}_n = (X_n, I_n)$ as follows

$$\pi_{k,j} = \lim_{n \to \infty} P(\overline{X}_n = (k,j)), \tag{5}$$

and let $\pi = (\pi_0, \pi_1, \dots, \pi_{K-1})$, with $\pi_k = (\pi_{k,1}, \dots, \pi_{k,m})$, $0 \le k \le K-1$. Then the steady-state probability vector π is obtained by the equations

$$\pi P = \pi, \text{ with } \pi e = 1, \tag{6}$$

where e is a column vector with all ones.

By suitably rearranging the equations of the linear system (6), it follows that

$$\pi_{0}UA_{0} + \pi_{1}A_{0} = \pi_{0},$$

$$\pi_{0}UA_{1} + \pi_{1}A_{1} + \pi_{2}A_{0} = \pi_{1},$$

$$\pi_{0}UA_{i} + \sum_{k=1}^{i+1} \pi_{k}A_{i+1-k} = \pi_{i}, i = 2, 3, \dots, K-2,$$

$$\pi_{0} \sum_{k=1}^{\infty} UA_{n} + \sum_{k=1}^{K-1} \pi_{k} \sum_{k=K-k}^{\infty} A_{n} = \pi_{K-1}.$$
(7)

From the first equation in (7), we have

$$\pi_1 A_0 = \pi_0 (I - U A_0) = \pi_0 C_1$$

From the second equation in (7), we have

$$\pi_2 A_0 = \pi_0 (C_1 A_0^{-1} - U A_1 - C_1 A_0^{-1} A_1) = \pi_0 C_2.$$

Similarly, we can take this procedure successively to obtain the following relation

$$C_{k} = C_{k-1}A_{0}^{-1} - UA_{k-1} - \sum_{i=1}^{k-1} C_{i}A_{0}^{-1}A_{k-i},$$

$$k = 1, 2, \dots, K-1.$$
(8)

Then we have $\pi_k A_0 = \pi_0 C_k$, $0 \le k \le K-1$, with $C_0 = A_0$. Consequently, once the vector π_0 is known, it is possible to determine also the vectors π_k , $1 \le k \le K-1$, by using the above recursive relation (8). To calculate π_0 summing both sides over equations in (7), we have the following

$$\sum_{k=0}^{K-1} \pi_k(I-A) = -\pi_0(I-U)A,$$

where $A = \sum_{k=0}^{\infty} A_k = \int_0^{\infty} e^{Qt} dH(t)$. Let θ be the invariant vector of the underlying MC Q. Adding

 $\theta = \sum_{k=0}^{K-1} \pi_k e \theta$ to the both sides and multiplying a matrix $(I - A + e \theta)^{-1}$, we have

$$\sum_{k=0}^{K-1} \pi_k = \theta - \pi_0 (I - U) A (I - A + e \theta)^{-1}.$$

From the relation $\pi_k A_0 = \pi_0 C_k$, $k = 0, 1, 2, \cdots$, K-1 with $C_0 = A_0$, we have

$$\pi_0 \left[\sum_{k=0}^{K-1} C_k A_0^{-1} + (I - U) A (I - A + \alpha \theta)^{-1} \right] = \theta. \quad (9)$$

Besides the above algorithm, various efficient algorithms that explore the special structure of the system of equation (6) have been developed to obtain the distribution π . [7-9]

Now we will find the limiting distribution of $\overline{X}(t) = (X(t), f(t))$ at an arbitrary time. Let $x_{n,j}$ be the limiting probability as

$$x_{n,j} = \lim_{t \to \infty} P\{\overline{X}(t) = (n,j) | \overline{X}(0) = (0,i)\}, \quad (10)$$

and let $x = (x_0, x_1, \dots, x_K)$ with $x_n = (x_{n,1}, \dots, x_{n,m})$. By supplementary variable method, we can find the steady-state probability vector x. Let the random variable η be the steady-state of server, namely $\eta = 1$ if the server is busy and $\eta = 0$ if it is idle. Then the probability that the system is busy is given by

$$\eta_1 = P\{\eta = 1\} = h/[h + \pi_0(\Lambda - Q)^{-1}e],$$

where h is the mean service time. Thus the relation between the state probability distribution at an embedded time point and an arbitrary time point is given by [7,8]

$$x_{0} = \eta_{1}\pi_{0}(\Lambda - Q)^{-1}h^{-1},$$

$$x_{n} = \eta_{1}\left[\pi_{n} + \sum_{k=0}^{n-1}\pi_{k}U^{n-1-k}(U-I)\right] \times (\Lambda - Q)^{-1}h^{-1}, \quad n = 1, 2, \dots, K-1,$$

$$x_{K} = \theta - \sum_{n=0}^{K-1}x_{n}.$$
(11)

The packet loss probability P_B for an arbitrary packet is given by

$$P_B = (x_K \Lambda e) / (\sum_{k=0}^K x_k \Lambda e). \tag{12}$$

We can use Little's law with the effective arrival rate $\lambda^* = \theta \Lambda e$ of the MMPP (Q, Λ) in order to find the mean packet delay W,

$$\lambda^* (1 - P_B) W = \sum_{k=1}^{K} k x_k e.$$
 (13)

When arrivals follow a Poisson process, we have quite simple expressions for the steady-state probability vectors and the performance measures. In this case, the corresponding equation to (8) is as follows^[10]

$$c_k = c_{k-1} * a_0^{-1} - a_{k-1} - \sum_{i=1}^{k-1} c_i a_0^{-1} a_{k-i}, k = 1, 2, \dots, K-1$$

where $a_k = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^k}{k!} dH(t)$, $k = 0, 1, \dots$, and so $\pi_k a_0 = \pi_0 c_k$, $k = 0, 1, \dots, K-1$ with $c_0 = a_0$. Then π_0 can be obtained by the following normalization condition

$$\sum_{k=0}^{K-1} \pi_k = \pi_0 \left[\sum_{k=0}^{K-1} c_k \right] = 1.$$

Since we can look a Poisson process on an MMPP with $\Lambda = \lambda$, Q = 0 and I = 1, by replacing (Q, Λ) with $(0, \lambda)$ in (11), (12) and (13), we have the steady-state probability x_n

$$x_{n} = \frac{\pi_{n}}{\pi_{0} + \lambda h}, \quad n = 0, 1, 2, \dots, K - 1,$$

$$x_{K} = 1 - \sum_{n=0}^{K-1} x_{n}.$$
(14)

Hence the packet loss probability P_B and the mean packet delay V in the M/G/1/K queueing system are given by

$$P_B = x_K, \tag{15}$$

$$\lambda(1-P_B)W = \sum_{k=1}^{K} kx_k. \tag{16}$$

III. Numerical example

In this section we present some numerical results to show the effect of the arrival rate, the buffer size and the mean service time(in fact, the link capacity) on the packet loss and the mean packet delay. We give some examples to have a

clue for capacity dimensioning of access link of backbone network to the Internet. To determine the applicability of our queueing model and the analytical results to practical IWU, we use both exponential and deterministic distribution with a mean of 250 bytes^[4] to model the size of IP packets. Thus if the capacity of the access link is 2Mbps, the mean service time becomes h=1[ms] and so the service rate becomes 1. We use a simple two-state MMPP (Q, A) and a Poisson process with effective arrival rate $\lambda^* = \theta \Lambda e$ of the MMPP as two kind of input processes, in order to compare an MMPP/G/1/K queueing model with an M/G/1/K queueing model. The two-state MMPP has the following representation

$$Q = \begin{pmatrix} -r_1 & r_1 \\ r_2 & -r_2 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

The corresponding effective arrival rate of the above MMPP is $\lambda^* = (\lambda_1 r_2 + \lambda_2 r_1) / (r_1 + r_2)$.

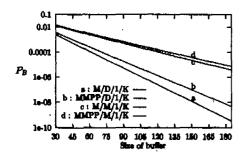


Fig. 3 Loss prob. vs. buffer size $(r_1=0.01, r_2=0.05, \lambda_1=0.3, \lambda_2=0.4, h=3.0)$

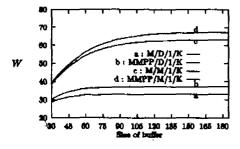


Fig. 4 Mean delay vs. buffer size $(r_1-0.01, r_2-0.05, \lambda_1-0.3, \lambda_2-0.4, h-3.0)$

In Figs. 3 and 4, we show the packet loss probability and the mean packet delay for the

IWU queueing model with a Poisson input and a deterministic service time(a), the model with an MMPP input and a deterministic service time(b), the model with a Poisson input and an exponential service time(c) and the model with an MMPP input and an exponential service time(d). as the size K-1 of buffer varies, when the other system parameters $\lambda_1, \lambda_2, r_1, r_2$ and h are fixed. From these figures, we see that increasing the size of buffer makes the packet loss probability improved and the mean packet delay deteriorated. We can also see that an MMPP input and an exponential service time have a negative effect on the IWU performance in comparison with a Poisson input and a deterministic service time.

In Fig. 3, we can recognize that to achieve the packet loss probability 10^{-6} , the models (a) \sim (d) require the buffer sizes 83, 98, 170 and 180 respectively. Thus an MMPP model with an exponential service time requires the largest buffer. In Figs. 3 and 4, we see that the packet loss probability is more sensitive to the change of the buffer size than the mean service time, when the buffer size is sufficiently large. This fact holds obviously, when the offered load is low.

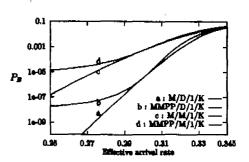


Fig. 5 Loss prob. vs. effective arrival rate $(r_1=0.01, r_2=0.05, \lambda_1=0.3, K=51, h=3.0)$

In Figs. 5 and 6, we show the packet loss probability and the mean packet delay for 4 cases of two Poisson input models(a,c) and two MMPP input models(b,d), as the effective arrival rate varies (in precise, λ_2 varies, because r_1 , r_2 and λ_1 are fixed). In these figures, changes of λ_2 entail changes of the burstiness, which is defined as the

ratio of peak arrival rate to mean arrival rate.

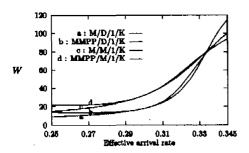


Fig. 6 Mean delay vs. effective arrival rate $(r_1-0.01, r_2-0.05, \lambda_1-0.3, K-51, h-3.0)$

When the effective arrival rates are 0.25 and 0.3, the corresponding burstiness values are 1.2 and 1.0 respectively. Therefore, we can conceive that the larger the burstiness becomes, the bigger the discrepancy of the loss performance becomes. From Fig. 5, we can check that the burstiness has a negative effect on the loss performance of our system and that the mean packet delay is very sensitive to the change of the effective arrival rate, when the traffic load is very heavy.

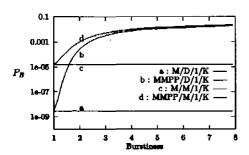


Fig. 7 Loss probability vs. burstiness $(r_2=0.05, \lambda_1=0.25, K=66, h=3.0)$

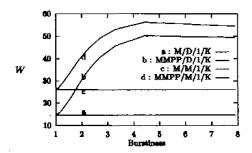


Fig. 8 Mean delay vs. burstiness $(r_2=0.05, \lambda_1=0.25, K=66, h=3.0)$

Figs. 7 and 8, we show the packet loss probability and the mean packet delay for 4 cases (a,b,c,d), as the burstiness varies, when the effect arrival rate $\lambda^* = 0.29$ is fixed (a,c) and the parameters r_2 and λ_1 are fixed. From these figures, we see that the burstiness factor of IP packet arrivals have to be considered seriously in capacity dimensioning of Internet access links. We can also see that the effect of the service time distribution on the packet loss probability is insignificant when the burstiness value is large.

In Figs. 9 and 10, we show the packet loss probability and the mean packet delay for 4 cases (a, b, c, d), as the mean service time varies(in fact, we can consider that the mean service rate varies, because these two parameters are reciprocal to each other).

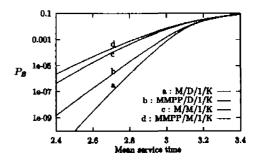


Fig. 9 Loss prob. vs. mean service time $(r_1=0.01, r_2=0.05, \lambda_1=0.3, \lambda_2=0.42, K=51)$

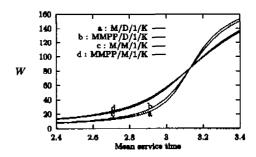


Fig. 10 Mean delay vs. mean service time $(r_1=0.01, r_2=0.05, \lambda_1=0.3, \lambda_2=0.42, K=51)$

Moreover the mean service time has a direct relation on the link capacity, especially when the mean packet length is fixed. In Fig. 7, to achieve the packet loss probability 10^{-5} , the models (a),

(b), (c) and (d) require the service rates 1/2.48=0.403, 1/2.57=0.39, 1/2.77=0.36 and 1/2.83=0.35 respectively. This fact implies that the MMPP input model with an exponential service time requires the largest link capacity among the four models.

In Figs. 11 and 12, we show the packet loss probability and the mean packet delay for 4 cases (a, b, c, d), as the buffer size varies, when the offered traffic is heavy(the traffic intensity $\rho \approx 1.0$). From these figures, we see that the mean packet delay is more sensitive to the change of buffer size than the packet loss probability, when the offered traffic is heavy. In these figures, a pair of (a) and (b), and the other pair of (c) and (d) have the same performance respectively, because the burstiness is near one.

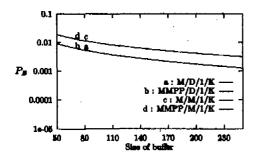


Fig. 11 Loss probability vs. buffer size $(r_1=0.01, r_2=0.05, \lambda_1=1.0, \lambda_2=0.99, h=1.0)$

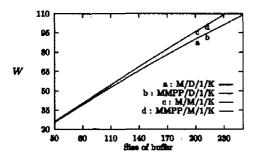


Fig. 12 Mean delay vs. buffer size $(r_1=0.01, r_2=0.05, \lambda_1=1.0, \lambda_2=0.99, h=1.0)$

IV. Conclusion

In this paper, we have built some types of queueing model for a LAN interworking unit and analyze this queueing models to discuss capacity dimensioning of access links. We use four types of queueing model such as MMPP/G/1/K, M/M/1/K, MMPP/D/1/K and M/D/1/K to compare the performances of a LAN IWU in different conditions. We have given some numerical results to show the effect of arrival rate, buffer size and mean service rate on the packet loss probability and the mean packet delay. We have seen that increasing size of buffer makes the packet loss probability improved but it makes the mean packet delay deteriorated.

We have recognized that the MMPP input and the exponential service time have a negative effect on the IWU performance in comparison with the Poisson input and the deterministic service time. We have also conceived that the larger the burstiness, the bigger the discrepancy of packet loss performances grows. Besides these facts, we have seen that the mean packet delay is more sensitive to the change of buffer size than the packet loss probability, when the offered load is very heavy and the packet loss probability is more sensitive to the change of buffer size, otherwise. From these facts, we conclude that the busrtiness factor of IP packet arrivals have to be considered seriously and both the loss probability and the mean delay have to be handled simultaneously in capacity dimensioning Internet access links.

References

- B. A. Mah, "An Empirical model of HTTP Network", ITC 15, pp. 593-601, 1997
- [2] M. Nabe, M. Murata, and H. Miyahara, "Analysis and modeling of World Wide Web traffic for capacity dimensioning of Internet Access lines", *Performance Evaluation*, Vol. 34, pp. 249-271, 1998
- [3] J. J. Gordon, K. Murti and A. Rayes, "Overview of Internet Traffic Issues on the PSTN", ITC 15, pp. 643-652, 1997
- [4] V. Paxson and S. Floyd, "Wide Area Traffic the Failure of Poisson Modeling", Proc.

SIGCOMM'94, pp. 257-268, 1994

- [5] Z. Niu, Y. Takahashi, and N. Endo, "Performance Evaluation of SVC-based IP-Over-ATM Networks", IEICE, Vol. E81-B, No. 5, pp. 948-957, 1998
- [6] W. Fischer and K. Meier-Hellstern, "The Markov modulated Poisson process (MMPP) cookbook", *Performance Evaluation*, 18, pp. 149-171, 1992
- [7] Blondia, C., "The N/G/1 finite capacity queue", Stochastic Models, Vol. 5, No. 2, pp 273-294, 1989
- [8] A. Baiocci and N. Blefari-Melazzi, "Steadystate Analysis of the MMPP/G/1/K Queue", IEEE Trans. on Comm., Vol. 41, No. 4, pp. 531-534, 1993
- [9] Raif O. Onvural, Asynchronous Transfer Mode Networks: performance issues, second edition, Artech House, 1997
- [10] H. Takagi, Queueing Analysis, vol. 2; Finite Systems, North-Holland, 1993

박 철 근(Chul-geun Park)

정회원



1983년 2월 : 부산대학교 수학과 *풑*업(이학사)

1986년 2월 : 한국 과학기술원 응용수학과 졸업 (이학석사)

1995년 8월 : 한국 과학기술원 수학과 졸업 (이학박사)

1986년 4월~1997년2월 : 한국 통신 통신망연구소 (선임연구원)

1997년 3월~현재: 선문대학교 전자정보통신공학부 정보통신공학전공(조교수)

<주관심 분야> 트래픽 공학, 통신망해석, 큐잉이론

한 동 환(Dong-hwan Han)

정희원



1987년 2월 : 서강대학교 수학과 **졸업(이학**사)

1989년 2월 : 한국 과학기술원 용용수학과 졸업 (이학석사)

1993년 8월 : 한국 과학기술원 수학과 졸업 (이학박사)

1994년 3월~현재: 선문대학교 자연과학부 수학전공 (부교수)

<주관심 분이> 확률과정론, 큐잉이론, 통신망해석