Differentiated Charging for Elastic Traffic

Hoon Lee*, Yoon Uh*, Jong-Hoon Eom**, Min-Tae Hwang* and Yong-Gi Lee**

Regular Members

ABSTRACT

In this paper, the authors propose methods for determining the differentiated price for elastic traffic in IP (Internet Protocol) network. First, we investigate the behavior in the consumption of bandwidth of elastic traffic in IP network. Next, we propose a method to relate the bandwidth usage with the pricing for the elastic traffic, which is based partially or fully on the usage rate of the network bandwidth. After that, we propose a charging function for elastic traffic, which is based on the de facto usage of the bandwidth. Finally, we will illustrate the implication of the work via simple numerical experiments.

I. INTRODUCTION

Until recently almost all the customers paid a fixed amount of charge irrespective of the amount of data generated and transferred over the commercial Internet, which is called a subscription charge. The IP-VPN and ADSL services correspond to this kind of charging scheme. They paid neither transfer charges nor content charges. They only paid the access charge to the Internet irrespective of the usage of the network resources. This charging scheme has reasons in a shared network with best effort service architecture, because there exists no classes or priorities in the service. So, there exists a high probability that the greedy users can use up the network resources, especially the bandwidth and the buffer space, so that the lazy users experience a high delay upon their visit to the network.

Recently we could find new applications which require timely delivery of data such as the Internet phone or applications which favor guarantee of appropriate amount of bandwidth during data transfer such as web browsing or Intra/Extranet via VPN (Virtual Private Network). To cope with these differences in the requirements for the network performance, differentiated and class or quality-based service policies have been proposed in the world of network service providers, system manufacturers as well as the standardization organizations such as IETF (Internet Engineering Task Forces). In line with these approaches, the concepts of charging in the Internet services are undergoing changes toward the usage-based charging [11,15,17, and references therein].

We could find a lot of literature in this field. To name a few for the usage-based pricing, Firdman [4] advocated the necessity of usage-based pricing based on revenue from usage. He gave alternatives such as the usage-sensitive pricing, priority-based pricing, value-added pricing, etc. McKnight [15] gave a qualitative and comprehensive overview on pricing the next generation Internet services after flat rate scheme. Blot et al. [2] reported a functional framework called *NetCounter* on charging the individual connection in IP commercial network. Karsten [6] proposed a scheme for a linear price calculation in integrated service architecture advocating that the internal price calculation should be linear, based on resource usage in the network. However,

* Department of Information & Communication Engineering, Changwon National University
** Access Network Laboratory, Korea Telecom
논문번호 : K01185-0827, 접수일자 : 2001년 8월 27일

he assumed a reservation-based service differentiation scheme. Pras [17] gave a general but realistic discussion for the current state of the art in Internet accounting. He summarized the objectives, protocols and methods for Internet accounting as well as the architecture for future implementation. Kelly gave many contributions on Internet accounting from the mathematical point of view via the concept of effective bandwidth, which gave a way for the evolution into usage based charging of IP networks. In [7], he gave a formal discussion on charging the bursty connections such as the elastic traffic. Lee [9] gave a formal discussion on bandwidth sharing and its impact on user utility and pricing for IP network. In [9], the authors argued that the network service provider has to levy charge based on the usage of the network bandwidth illustrating the quantitative numerical results for elastic traffic with best effort service architecture. In [10,11], the authors extended the concept of usage-based charging in [9] to more specific applications, the VPN services. There, the usage rate charging is also advocated by showing some numerical results.

This paper is an extension of those works in [9,10,11]. The authors argue that the usage-based charging has to be well tuned to the objectives of the provisioning of the bandwidth. That is, when the bandwidth is in the form of reservation the charging has the form of fixed charging, whereas if the bandwidth is completely shared the charging scheme has the complete usage-based charging. This paper discusses this aspect in detail, and the authors give a formal framework and quantitative discussion for charging the elastic traffic in best effort IP network.

Before entering into the discussion for the charging the customers, we have to differentiate the concept between pricing, charging and billing: Pricing is the process of determining a cost per unit bandwidth the connection uses, whereas charging is the process of translating the customer's bandwidth usage information into an amount of money the customer has to pay. Finally, billing is the procedure of issuing the bill to the customer [17]. This paper discusses methods to determine the first two ones: pricing and charging.

This paper is composed as follows: In Section II we describe the attributes of elastic traffic from the bandwidth usage. In Section III we propose pricing and charging schemes for elastic traffic with two types: ABR-like and UBR-like traffic. Section IV gives the results for numerical experiments, where the implication of the proposed methods is shown with graphs. In Section V we summarize the paper and give some comments on further research areas.

II. ATTRIBUTE OF ELASTIC TRAFFIC AND IMPLICATION TO PRICING

It is well known that the elastic traffic (ET) is named from the property of the elastic services such as files of data, text, picture, www pages and other documents in that it can cope with a non-guaranteed variable throughput. Thus, ET can tolerate packet delays and it would rather wait for reception of traffic in the correct order, without losses. So, the traffic in elastic services needs a large buffer and an elastic bandwidth allocation mechanism like TCP in IP network or ABR (Available Bit Rate) services in ATM network. Examples of elastic traffic include traditional data services such as remote terminal, file transfer, name services and electronic mail. Note that the attribute in elastic traffic is very similar to that of ABR or UBR (Unspecified Bit Rate) traffic in ATM network. As such, the Internet user and IP network can negotiate the transfer of packets via two different methods: guaranteed minimum bandwidth and plus-alpha (if available) that corresponds to ABR-like-ET (for simplicity, we call it ABR-ET) and no bandwidth guarantee and no QoS contraction for UBR-like-ET (UBR-ET).
As to ABR-ET, the specification for QoS (Quality of Service) is expressed in terms of minimum throughput, which is represented by the minimum bandwidth that is used by a connection. Through-

put is simply computed to be the file size divided by transfer time [14]. Minimum throughput of elastic traffic is synonymous with the minimum bandwidth the network has to provide to the traffic, which is contracted with the customer before traffic is transferred. The contracted minimum bandwidth (CMB), which is denoted by μ, is allocated (in the form of reservation) a priori, and an additional bandwidth (we had described it as "plus-alpha") is provided by network if there is any available bandwidth unused by other connections in the network.

For UBR-ET, no contraction with respect to the bandwidth usage is needed in order to transfer the data, and so no bandwidth is reserved to the connection with UBR-ET, and packet transfer occurs only if there exists available bandwidth not used by other high priority traffic.

At any instant, customer may generate traffic less than or greater than the CMB. If there is sufficient bandwidth in the link the network can carry out all the traffic in excess of the contracted value, otherwise some packets are forced to wait in the queue for later transmission. Fig.1. shows a rough graph of the behavior of elastic traffic as a function of time. The solid line denotes the maximum link capacity and the dotted line denotes the agreed bandwidth, so M is the maximum rate the users can use the bandwidth, whereas μ is the CMR (Contracted Minimum Rate) of a connection, which is equivalent to CMB. Because the variation of the traffic volume generated from a connection for Internet access is very harsh [3], there may happen cases where a connection can or can't use the agreed bandwidth. So, the traffic curve goes up and down the CMR very frequently. Of course, the traffic rate of a connection should not exceed the maximum link capacity M at any time.

From the network operators' perspective, it is favorable that the user generates smooth traffic shuffling around the CMR. If traffic rate varies more drastically varying from zero to M, the network operator has to take any action concerning the prevention of occurrence of instantaneous overflow resulting from simultaneous connections from a number of customers, because little statistical multiplexing gain can be obtained in the aggregation of heavy-tailed traffic shown from the Internet traffic. Keeping this problem in mind, let us describe a necessity to prevent transmission of large amount of traffic as an action in terms of differential charging, where customers with heavy usage are obliged to pay high charge.

The reverse of above statement would be as follows: When the traffic volume generated from the customer is smaller than the contracted minimum rate, the network may cut down the price corresponding to the difference between the minimum rate and the actual rate. This discussion corresponds to the charging for UBR-like-ET in Section III-2.

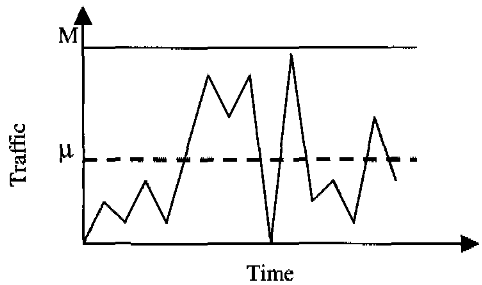


Fig. 1 Behavior of elastic traffic

Ⅲ. PRICING AND CHARGING THE ELASTIC TRAFFIC

The last discussion in Section II implies that the translation of bandwidth used by the elastic traffic into price can be divided into two cases: price for ABR-ET or UBR-ET. The former discriminates the value of the minimum bandwidth and additional bandwidth used by availability of the network at any instant, whereas the latter cares only the usage of bandwidth. A detailed discussion on these arguments is described in this section.

1. Price for ABR-ET

First, let us remind the attribute of ABR-ET. For ABR-ET, the minimum bandwidth can be looked upon as a reserved bandwidth, whereas the additional bandwidth, called as *plus-alpha*, is the excess bandwidth that can be used by a customer because there exists additional bandwidth the network operator can provide to the customer. Let us denote the former one by μ. The basic assumption behind the concept of fixed and residual pricing is that the customer has to pay a fixed amount of charge for the reserved bandwidth whether he/she uses it or not, whereas the implicit agreement in the provision of additional bandwidth is that the customer is ready to pay additional price for the additional bandwidth provided by the network. Let us call the price of residual bandwidth to be the residual price. From these discussions, we can find that the concept of fixed and residual pricing is well suited to the purpose of bandwidth provision of ABR-ET. We can easily find the similarity in levying the price for ABR-ET in an IP world and the concept of residual price in an ATM world. CANCAN [16] announced a recommendation for residual price for ABR traffic based on the Committed Information Rate (CIR) because the network operator reserves a minimum bandwidth relevant to CIR to an ABR connection.

The concept of fixed and residual pricing is illustrated in Fig.2. As we can find from Fig.2, the customer pays a fixed amount of charge irrespective of the usage of the network bandwidth so far as the measured traffic rate does not exceed the predefined minimum bandwidth μ. So, minute computation for pricing is not carried out by the network operator. However, the customer has to pay additional charge for the usage of the bandwidth in an amount he/she used in addition to μ when the bandwidth usage is greater than μ. There may exist various ways for levying prices differently to the usage of the bandwidth (see [9,10] and references therein). The typical curve for the residual price is a linear function connecting the two points of μ and M. The slope σ of residual price is determined by the price policy of the network operator. We assume that maximum price is Π for the usage of maximum allowed bandwidth of M. Then, σ is given by

$$\sigma = \frac{\Delta P}{\Delta \omega} = \frac{\Pi - C}{M - \mu} \qquad (1)$$

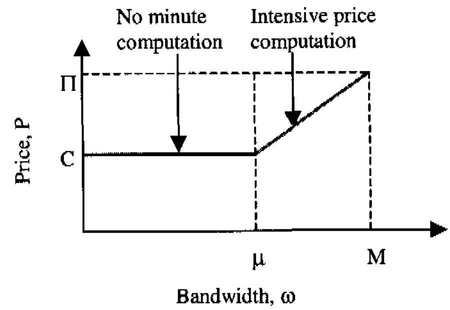


Fig. 2 Concept of fixed and residual pricing

2. Price for UBR-ET

In the current QoS-less IP world, the bandwidth of a link can never be reserved to a specific connection if there exists no priority scheme in the provisioning of bandwidth. The network bandwidth not used by any customer can be shared by the other customer if there is anyone who wants to send packets through the same link. Therefore, the fixed and residual pricing scheme discussed above should be modified, which is the aim of our second discussion.

We argue that the charge for the elastic traffic with UBR-ET type has to reflect only the usage of bandwidth faithfully. A probable option is to give a weight to the usage in order to take into account our discussion in Section II. To that purpose we exploit the concept of residual price for the CIR of ATM network. That is, the network operator levies a price of C Dollars to the customer for the usage of the bandwidth of μ. Then, we draw a line between the points (0,0) and (C, μ), and define the slope α between the two points by

$$\alpha = \frac{C}{\mu} \tag{2}$$

α is called the tariff parameter. We can easily find that the tariff parameter for the usage of bandwidth between (μ,M) is the same as that of Fig.2 with a condition given by

$$\sigma > \alpha . \tag{3}$$

The last inequality implies that the user has to pay higher price for the usage of bandwidth greater than μ, which is the basic philosophy of the differential pricing scheme that tries to suppress excess consumption of bandwidth. This is illustrated in Fig.3. In [5], concept of the marginal increment in expected cost for a marginal increment in load is introduced in the name of a shadow price. If we compare the concept of shadow price and equation (3), we can find that the two concepts are very similar.

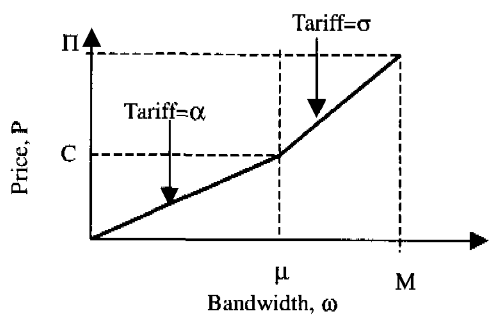


Fig. 3 Concept of differential pricing

3. Charging Function

Let us define some variables and parameters for the calculation of the charge imposed on the connection. Let u(t) be the traffic volume (unit: bits) which is generated by the customer at time t. Let T be the time interval of the measurement of the network usage from the beginning and to the end of the session. Then, the usage rate of the network bandwidth in a unit of bit per second is defined by the amount of bits transmitted in the network during a certain time period, and it is given by

$$\omega = \frac{1}{T}\int_0^T u(t)dt. \tag{4}$$

The duration of measurement T indicates the specification of the monitoring frequency and it is closely related to the speed of the network link, which is also related with delay characteristics expected from the network and the amount of bandwidth allowed by the network. The more sensitive the application is, the higher the monitoring frequency should be. T is also related with the accuracy of the measurement. In [1] a discussion on this value is given in a qualitative manner for three traffic classes: very frequent, frequent, and unspecified duration of measurement. The elastic traffic is categorized as an unspecified duration for the measurement of CDR value. Even for the elastic traffic with unspecified duration, the period of monitoring has a close relationship with the accuracy of the charging. However, we assume that the monitoring period is much shorter than the connection duration, from which we can accumulate enough data for the estimation of usage rate of the connection. Thus, in this paper, let us assume T=1 second for simplicity, via which ω is translated into bits per second. This, in fact, is the most familiar unit of speed of the network to us. Note that the above result considers only the one-way traffic. A similar result concerning the usage rate of the both-way traffic for the asymmetric link can be found in author's other work [11].

If we have the usage rate ω of network bandwidth, we can relate the usage rate into a charging function $F_X(\omega)$, where the lower index 'X' is 'A' or 'U' for the ABR-ET or UBR-ET, respectively, and they are given as follows:

For ABR-ET

$$F_{A1}(\omega) = C, \text{ if } \omega < \mu, \tag{5}$$

$$F_{A2}(\omega) = \sigma \times (\omega - \mu) + C, \text{ if } \omega \geq \mu, \tag{6}$$

where σ is given in (1).

For UBR-ET, they are given as follows:

$$F_{U1}(\omega) = \alpha \times \omega, \text{ if } \omega < \mu, \quad (7)$$

$$F_{U2}(\omega) = \sigma \times (\omega - \mu) + C, \text{ if } \omega \geq \mu, \quad (8)$$

where α is given in (2).

Ⅳ. NUMERICAL EXPERIMENTS

The price for a packet transfer for the current IP network in Korea is based on the fixed rate scheme and the only metric for the charging is the speed of the access link whether the user is accessed via ADSL or Leased-line [11]. As we may have a lot of options for the unit charge with respect to the use of the bandwidth for IP network depending on the network service operators, we refer to the current charge for the leased-line of KORNET, Korea Telecom's public IP network. From some numerical computation described in [11,12], we can find linearity from the trend of charge as a function of the speed of the access link. Based upon this finding, we can think two different methods for the determination of the tariff: the tariff by marginal price or absolute price.

The marginal price is defined as follows: Let τ be the marginal price paid for the consumption of marginal bandwidth, which is based on eq.(1) and is redefined, for convenience, by

$$\tau = \frac{P_a - P_b}{W_a - W_b} = \frac{\Delta P}{\Delta W}, \quad (9)$$

where P_x is the price and W_x is the capacity of bandwidth for a link with index x. ΔP and ΔW is the marginal increment of price and capacity of bandwidth for the link with index a and b, respectively. The formula (9) may be a good measure of price when the marginal price has the same value for each pair of two different link speeds.

On the other hand, the absolute price is defined by

$$\tau = \frac{P_x}{W_x}, \quad (10)$$

where the price is determined only by the price and bandwidth capacity of link with index x. From simple numerical experiments (we do not describe the detailed facts about the experiments for brevity, but we spare it to [12]), we could find that the marginal price is smaller than the absolute price under the current pricing system: the marginal price for the usage of one Megabits is 0.85Won (Korean currency) when τ is computed between 256 and 512 Kbps, whereas the absolute price for the use of 256Kbps is 1.524Won. Note, however, that this is just an example, and it may not be general situation. However, it is easy to find that the marginal price can be applied to the computation of σ in formula (1), whereas the absolute price corresponds to α in formula (2).

Bearing these relationships in mind, let us draw a graph for the charge with respect to various speed of the link. To that purpose let us assume three typical link speeds: 256Kbps, T1 (1.544Mbps) and E1 (2.048Mbps). The reference price for each link speed is assumed after the monthly price of KORNET [8]. Table 1 shows the maximum price Π for each link with maximum available speed M, which is defined in Section III.

Table 1. Price versus link speed

M (bps)	256K	T1	E1
Π(1,000Won)	985	2,490	3,270

From [13], we could find the linearity in the price of the link as a function of the link capacity, which implies that the marginal price can be used in the determination of the unit price. For the purpose of illustration, let us assume as follows:

$$\mu = 0.6 \times M, \quad (11)$$

$$C = 0.5 \times \Pi, \quad (12)$$

from which α and σ are determined from Table 1. Note that this assumption is derived from the authors' opinion, and it is closely dependent on the policy of charge that can be set by network operators.

Fig.4 shows the amount of charge F (unit: 1Won per Kilo bits) levied to the user of type ABR-like-ET for each link speed (maximum usable link speed) as a function of the normalized bandwidth (which corresponds to the usage rate, and is defined by $\overline{\omega}=\omega/M$) used by the connection. The threshold between the intensive price computation and no minute computation (see Fig.2) is μ_n=0.6, the normalized value of μ with respect to M.

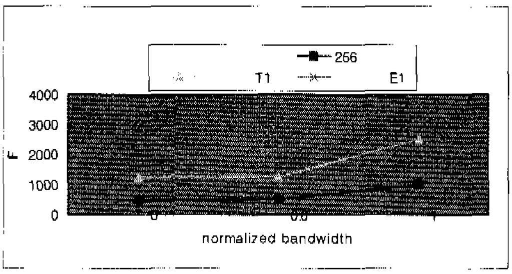


Fig. 4 Charge for ABR-ET

As we may have expected, fixed charge is levied to the connection unless the usage rate exceeds the threshold of 0.6. When the usage rate exceeds the threshold, the connection is charged by the tariff function (8). Connection with different access speed has to pay different amount of charge in accordance with the usage rate.

Fig.5 illustrates the charge (Unit: 1Won per kilo bits) to the user of type UBR-like-ET for each link speed (maximum usable link speed) as a function of the normalized bandwidth. We could find almost linear curve for each link with access

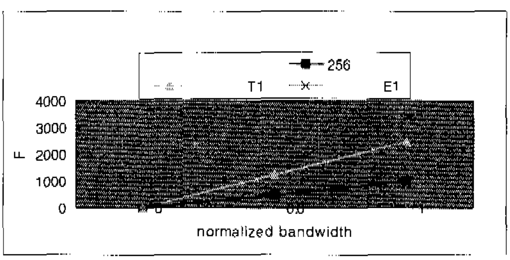


Fig. 5 Charge for UBR-ET

speed of 256Kbps, T1 and E1 even though each link has different value of slope. The last situation arises from the policy of tariff given by formula (11).

V. CONCLUSIONS

In this paper the authors proposed methods for determining the charges for the elastic traffic in IP network. We assumed two classes of services in levying price, ABR-ET and UBR-ET, and we proposed different functions for charging those traffic. Via numerical experiments we could show the differences in the charges for two schemes, which can provide intuition for the users to consider price and performance trade-off in choosing the services for transferring the elastic traffic over IP network.

This work is just a first step in the quantitative research for the charging the Internet based on the usage of the bandwidth. Thus, there remain lots of problems, as such areas for the further study would be wealthy: the determination of the optimal price for the use of links with different speeds, sophistication in the method for monitoring the traffic usage, time granularity for metering, etc. Our next research will be concentrated to these areas.

ACKNOWLEDGMENT

The views and conclusions contained in this paper are those of authors and should not be interpreted as representing the official policies, either expressed or implied, of Korea Telecom. This research is financially supported by Changwon National University in 2001.

REFERENCES

[1] O. Aboul-Magd and B. Jamoussi, "QoS and service interworking using CR-LDP", IEEE Communications Magazine May 2001.

[2] S.M. Blot, et al., "User-level billing and accounting in IP networks", Bell Labs

Technical Journal, Oct-Dec. 1999.

[3] J. Charzinski, "Problems of elastic traffic admission control in an HTTP scenario", Proceedings of IWQoS2001.

[4] E. Firdman, "Rx for the Internet: usage-based pricing", Data Communications on the web, http://www.data.com/business_case/rx_internet.html.

[5] R.J. Gibbens and F.P. Kelly, "Resource pricing and the evolution of congestion control", http://www.statslab.cam.ac.uk//~frank/evol.html.

[6] M. Karsten et al., "Provider-oriented linear price calculation for Integrated services", Proceedings of IWQoS'99, 1999.

[7] F.P. Kelly, "Charging and accounting for bursty connections", Internet Economics, ed. L.W. McKnight and J.P. Bailey (eds) MIT Press, 1996.

[8] The link price for VPN services for Korea Telecom IP network KORNET, http://www3.kornet.net/helper/entrance/index2.html, as of July 02, 2001.

[9] H. Lee, "Bandwidth sharing, utility and pricing for the best effort Internet services", Proceedings of the 3rd Int'l conference on advanced communication technology (ICACT 2001), February 8-10, 2001, Muju, Korea.

[10] H. Lee, "A usage based charging in IP-VPN", NCS2000, December 4-6, 2000.

[11] H. Lee and J.-H. Eom, "A usage-rate based charging for QoS-free traffic in IP-VPN", Proceedings of the ICC2001, June 11-15, 2001, Helsinki, Finland.

[12] H. Lee, "On the determination of tariff for KORNET leased-line", TM-NL-01-07-03, Network Lab., Changwon National University, July 4, 2001.

[13] H. Lee, "Analysis of link price for KORNET VPN", TM-NL-01-07-04, Network Lab., Changwon Nat'l Univ. July 4, 2001.

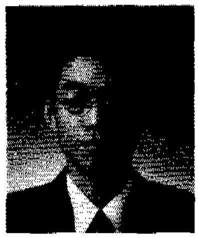
[14] K. Lindberger, " Balancing Quality of Service, pricing and utilization in multiservice networks with Stream and Elastic Traffic", ITC 16, Edinburgh, June 1999.

[15] Lee W. McKnight and J. Boroumand, "Pricing Internet services: after flat rate", Telecommunications Policy 24 (2000).

[16] D. Morris and V. Pronk Ed., "ATM charging Recommendations II, Final Recommendations, CANCAN Deliverables, August 1998.

[17] A. Pras, et al., "Internet accounting", IEEE Communications Magazine, May 2001.

Hoon Lee

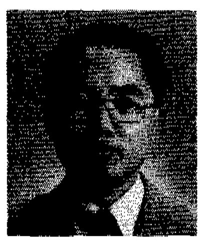


He was born in Daegu, Korea, on December 20, 1961. He received the B.E. degree in Electronics and M.E. degree in Communications from Kyungpook National University, Daegu, Korea, in 1984 and 1986, respectively. He received the Ph.D. degree in Electrical and Communication Engineering from Tohoku University, Sendai, Japan, in March 1996. From February 1986 to February 2001, he worked at Korea Telecom, where he has been engaged in the research on the teletraffic theory, network design, performance analysis and management of ATM and IP networks.

He is currently an Assistant Professor of Changwon National University, from March 2001. His current research interests include network design, performance analysis and provision of QoS for high speed telecommunication networks.

Dr. Lee is a member of IEEE, KICS and IEEK.

Yoon Uh



He was born in Jinyoung on December 1959.

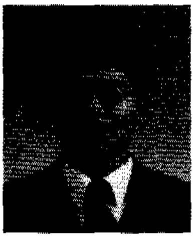
He received the B.E. and M.E. degrees from Hanyang University, 1982 and 1986, respectively. He received the Ph.D. degree from Tohoku University, Sendai, Japan, 1994.

From January 1986 to July 1987 he worked at LG Electric, LTD. From September 1987 to June 1998, he worked at ETRI as a senior research staff.

From September 1998, he is an Assistant Professor

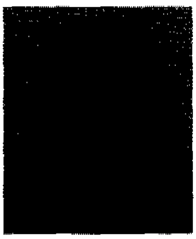
of Changwon National University.
His research interests include digital communication systems and coding theory.
Dr. Uh is a member of IEEE, KICS and IEEK

Min-Tae Hwang



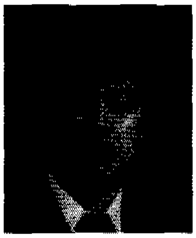
He received the B.E., M.E. and Ph.D. from Pusan National University in 1990, 1992 and 1996, respectively.
From 1996.2 to 1999.3 he was in ETRI. From 1999.3 to 2000.2 he was a Full Time Lecturer at Inje University. From 2000.3 he is an Assistant Professor at Changwon National University.
His research areas include Home networking, Next generation Internet, Multimedia MAC protocol Dr. Hwang is a member of KICS and KMMS.

Jong-Hoon Eom



He was born in Daegu, Korea, on July 10, 1963. He received the B.E. degree in Electronics and M.E. degree in Computer Engineering from Kyungpook National University, Daegu, Korea, in 1986 and 1992, respectively. He finished the course of Ph.D. degree in Computer Engineering from Kyungpook National University, Daegu, Korea, in 1996.
He is currently an Senior Member of Technical Staff from February 1993, in Korea Telecom, where he has been engaged in the research on the network engineering, network design and optical transport network. His current research interests include Metro Ethernet, Ethernet PON, Ethernet in the First Mile and QoS management of IP networks.
Mr. Eom is a member of KICS.

Yong-Gi Lee



He was born in Daegu, Korea, on August 31, 1958.
He received the B.E. degree in Electronics and M.E. degree in Communications from Kyungpook National University, Daegu, Korea, in 1981 and 1985, respectively. He received the Ph.D. degree in Electronic Engineering from Tohoku University, Sendai, Japan, in March 1996.
From 1985 he works at Korea Telecom, and now he is a director of Optical Networking Technology Division, Access Network Laboratory, Korea Telecom.
His research field includes WDM Technology and devices, Optical Internet Technology and Optical Test-bed Construction Technology Dr. Lee is a member of KICS and IEEK.