Predictive Connection Admission Control for Broadband ATM Satellite Systems

Yeong Min Jang* ***Regular Member***

ABSTRACT

In this paper, we propose a predictive (transient) connection admission control (CAC) scheme for satellite systems that supports on-board packet switching of multimedia traffic with predefined quality of service (QoS) requirements. The CAC scheme incorporates the unique characteristics of satellite systems, e.g. large propagation delays, no onboard buffer, and low computational requirements. The CAC scheme requires the estimation of the On-Off traffic characteristics (λ, μ) of the traffic sources. These estimated values are used to predict the transient cell loss ratio at each downlink. In case the QoS requirements are not met the proposed CAC scheme rejects the new connection. The numerical results obtained suggest that the proposed scheme is an excellent candidate for real time burst and call level connection prediction and control in broadband on-board satellite networks.

I. Introduction

Geostationary Earth Orbit (GEO) Satellites with wide area coverage are expected to provide interconnections of Global Area Business Networks supporting various applications such as point-to-point VSAT-type services, LAN interconnections, multimedia, and teleconferencing services. To accommodate these diverse services with inherent traffic fluctuations while efficiently utilizing the spectrum, packet switching capabilities should be implemented at the satellite [1]. NASA Lewis Research Center is currently investigating a geostationary communication satellite MCSPS (Multi Channel Signal Processing Satellite) which can support packet switching capability and provide direct-to-the-user services [2].

Recall that the bandwidth in a packet switching network is allocated on demand. Therefore, congestion problems may occur when the demand for resources on-board the satellite exceeds its capacity. Congestion can rapidly neutralize the delay and severely limit the advantages obtained by dynamic resource sharing achieved by packet switching. To achieve both efficient utilization of the space-segment resources and acceptable user QoS, high performance network management protocols that incorporate congestion/flow control must be provided. The congestion problem is not unique to satellite communications and in fact has been extensively studied for terrestrial broadband ATM networks [3,4]. However, in the design of CAC schemes for satellite systems, attention should be paid to the following unique characteristics of satellite communication:

- Satellite networks have long propagation delays between the earth-stations and the satellite (typically 125 ms), mandating predictive congestion control schemes.
- The on-board processing should be minimized (processing power and power consumption limitations), i.e., we should avoid complex, and computational - intensive procedures to be executed on-board.
- On-board storage is very expensive, i.e. congestion control schemes implemented in an on-board processing (OBP) satellite system should

---

* 한국전자통신연구원 무선방송기술연구소 (yjang@etri.re.kr)

논문번호 : 99490-1211, 접수일자 : 1999년 12월 11일

require minimal buffering capabilities.

- All the traffic in the satellite system is routed through the satellite (which is acting like a hub).

The long propagation delay characteristic will cause feedback-based control to be slow and ineffective. Ramamurthy and Sengupta [5] concluded that predictive control policies perform significantly better than the statistic rate and adaptive control policies in all cases.

The authors in [6] proposed the predictive approach for congestion control in satellite networks. They use statistical and neural approaches under Poisson traffic cases. There are several predictive papers for ATM and wireless networks [7,8].

We propose a predictive CAC scheme which accounts for all the above characteristics of satellite communication. The proposed scheme achieves both the efficient utilization of the space-segment resources while providing the required QoS (in terms of the tolerable cell loss ratio) in a multimedia environment. We estimate the state of the underlying process and predict the future cell loss ratio based on estimation and measurement of the bursts in progress and the duration of each burst.

This transient and predictive dynamics is used to avoid congestion while efficiently utilizing the available network capacity. Optimal criteria are proposed on how to optimally select the number of connections, given the transient cell loss ratio as the QoS. These optimal values maximize the bandwidth utilization while maintaining the requested QoS.

In the next subsection we present the system model and in subsection A, we introduce an outline of the proposed predictive CAC algorithm.

A. System Model

The ground stations are interconnected through the satellite using K uplinks and K TDM downlinks (which determine the downlink time division into slots which equal the cell size).

The propagation delays for uplinks and downlinks are d_1 seconds and d_2 seconds, respectively. The uplinks use TDMA, MF-TDMA, FDMA, CDMA or MF-CDMA techniques. The satellite has a switching fabric (with no buffering capabilities) capable of routing cells that successfully arrive on the uplinks, to their destination downlink. Due to statistical fluctuations of the traffic and the limited bandwidth of the downlink (C bits/second), congestion may occur at the downlinks causing packets to be discarded. Therefore, it will be mandatory to implement a CAC scheme in an OBP satellite system. See Fig. 1 for a detailed description of the system configuration.

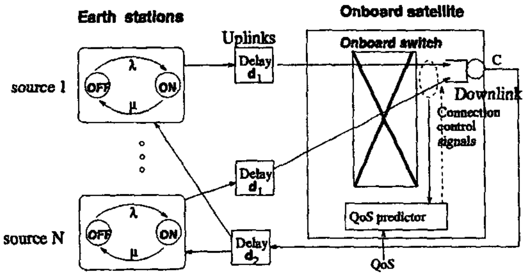

Fig. 1 Congestion prediction and control architecture for broadband satellite communication.

Although the traffic characteristics of the future OBP satellite networks are hard to predict with complete accuracy, there are a number of voice and video models reported in the literature [9,10]. Due to the statistical nature of the multimedia traffic, the modeling of incoming traffic characteristics plays an important role. As suggested in [9,10], both voice and video traffic sources can be characterized by a set of On-Off models. The QoS for each user is expressed in terms of the cell loss ratio that the connection can tolerate.

Each source is therefore, modeled as an On-Off traffic model. We assume that the "ON" and "OFF" periods of each source are both exponentially distributed with parameters μ and λ, respectively. In this traffic model, when the process is in the "ON" state, it generates cells at a constant rate of R_p bits/second. When the process is in the "OFF" state, it does not generate any cells. We assume that each one of the N connections sharing a downlink has the same traffic parameters (λ, μ, R_p).

B. Outline of the CAC Algorithm

An outline of the predictive CAC algorithm is as follows:

1. *Parameter Estimation*: For each traffic source using the downlink, we collect sample data of the On-Off traffic for a number of consecutive cell samples and compute these estimates, $\hat{\lambda}$ and $\hat{\mu}$, using the Maximum Likelihood (ML) method. Details are provided in Section II.

2. *Compute (predict) the transient cell loss ratio*: Based on the estimated parameters, $\hat{\lambda}$ and $\hat{\mu}$, the number of connections, N, and the number of active sources at time $t=0$, i, we compute the transient cell loss ratio denoted by $CLR(\hat{\lambda},\hat{\mu},N,i,t)$. In case $CLR(\hat{\lambda},\hat{\mu},N,i,t) < QoS_{CLR}$, the satellite controller sends congestion abatement messages. Otherwise, the algorithm proceeds to 3 below. Details are provided in Section III.

3. *Find the optimal number of connections, N^*, for given QoS_{CLR} requirements*: Whenever the predictive transient CLR is higher than the QoS_{CLR} requirement, the satellite controller will compute (using results obtained in 1. and 2. above) the number of connections depending on the network environment. The satellite controller will send choke messages to the earth stations including these parameters. In case the connections are established by sources, the satellite controller will reject the (N+1)-th connection request, given there are N connections. Details are provided in Section IV.

A flow chart of the CAC algorithm is depicted in Fig. 2.

In each ground station, the following actions are taken:

1. *Congestion onset message received*: reject the new connection request.

2. *Congestion abatement message received*: allow the new connection request.

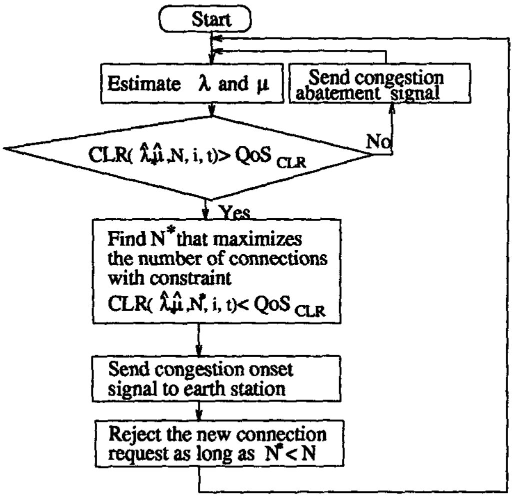

Fig. 2 Flow diagram of proposed CAC algorithm.

The congestion management at the ground station is beyond the scope of the paper.

Ⅱ. Parameter Estimation Technique

The parameters that we estimate must be such that they can be easily evaluated and monitored by traffic measurement on-board the satellite which has limited power, bandwidth and buffering capabilities. The traffic estimation process for each connection is based on the measurement of M consecutive cell samples. We next describe the Maximum Likelihood method for estimating λ and μ.

Recall that we have assumed that the sequences of active and idle periods (observed at the downlink) are exponentially distributed random variables (see Figure 3).

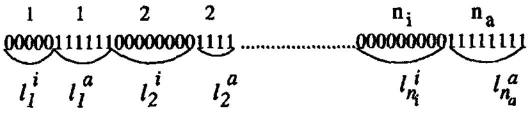

Fig. 3 Maximum Likelihood diagram.

Let $l_1^a,\cdots,l_k^a,\cdots,l_{n_a}^a$ and $l_1^i,\cdots,l_k^i,\cdots,l_{n_i}^i$ be the length of successive active and idle periods, respectively. The estimation is based on the measurement of a fixed number n_a of active periods and a fixed number n_i of idle periods of the On-Off sources. It is assumed that all idle and active periods are *i.i.d.* random variables. Since we will obtain M data samples, we have:

$$\sum_{k=1}^{n_a} l_k^a + \sum_{k=1}^{n_i} l_k^i = M. \tag{1}$$

Let $f(l_k^a;\mu)$ be the density function of the active period where μ is the only parameter to be estimated from a set of sample values $l_1^a, l_2^a, \cdots, l_{n_a}^a$. The likelihood function of this sample is

$$f(l_1^a;\mu)f(l_2^a;\mu).....f(l_{n_a}^a;\mu) = \mu^{n_a}\exp\left(-\mu\sum_{k=1}^{n_a} l_k^a\right). \tag{2}$$

The maximum likelihood estimation of the source parameter μ is found by maximizing the joint density, i.e., differentiating Eq. (2) with respect to μ:

$$\frac{\partial}{\partial\mu} f(l_1^a;\mu)f(l_2^a;\mu)\cdots f(l_{n_a}^a;\mu) = 0. \tag{3}$$

Then the ML estimate of μ:

$$\hat{\mu} = \frac{n_a}{\sum_{k=1}^{n_a} l_k^a}. \tag{4}$$

Similarly, we obtain the maximum likelihood estimator of λ:

$$\lambda = \frac{n_i}{\sum_{k=1}^{n_i} l_k^i}. \tag{5}$$

Ⅲ. Transient Cell Loss Ratio

It is very important that the decisions to accept or reject connections are made in real time. That is, a simple and fast CAC should be able to predict transient cell loss ratio rather than steady state cell loss ratio. We use a statistical bufferless fluid flow model [11] to predict the ratio that a cell loss occurs at time $t(=d_1+d_2)$, (i.e., the round-trip propagation delay) based on the traffic statistical behavior and the number of active sources at time 0. We assume N identical binary On-Off sources that share the capacity C of a downlink. Next we introduce the following definitions:

- $\rho(\lambda,\hat{\mu},N,i,t)$- the mean arrival rate at time t, given i sources are active at time 0 and there are N connections (traffic sources), assuming each traffic source is characterized by $(\lambda,\hat{\mu})$.
- $CLR(\lambda,\hat{\mu},N,i,t)$- the predicted transient cell loss ratio at time t, given i sources are active at time 0 and there are N connections (traffic sources), assuming each traffic source is characterized by $(\lambda,\hat{\mu})$.
- $OV(N,i,t)$- the mean excess traffic at time t, given i sources are active at time 0. and there are N connections (traffic sources), assuming each traffic source is characterized by $(\lambda,\hat{\mu})$.
- $p(\lambda,\hat{\mu},t)$- the transition probability that a source is active at future time t, given the source is idle at current time 0.
- $q(\lambda,\hat{\mu},t)$- the transition probability that a source is active at future time t, given the source is active at current time 0.

In the link overflow model, cell losses due to overflow occur if and only if the aggregated peak rate from x active sources, xR_p, exceeds link capacity C. The transient cell loss ratio is given by the ratio of mean excess traffic $OV(N,i,t)$ and mean traffic load $\rho(\lambda,\hat{\mu},N,i,t)$ at time t (see Figure 4).

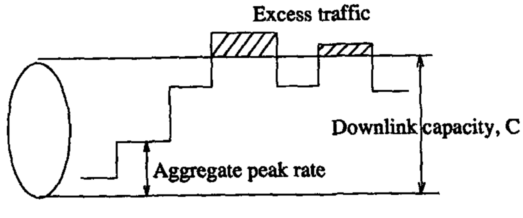

Fig. 4 Computation of the cell loss ratio

$CLR(\lambda,\hat{\mu},N,i,t)$ is given by:

$$CLR(\lambda,\hat{\mu},N,i,t) = \frac{OV(N,i,t)}{\rho(\lambda,\hat{\mu},N,i,t)}, \tag{6}$$

$$\rho(\hat{\lambda}, \hat{\mu}, N, i, t) = R_p[iq(\hat{\lambda}, \hat{\mu}, t) + (N-i)p(\hat{\lambda}, \hat{\mu}, t)]. \tag{7}$$

Let $Y(t)$ denote the number of active sources at time t.

$$OV(N, i, t) = \sum_{x \in \{x|(xR_p - C) \geq 0\}}^{N} P(Y(t) = x | Y(0) = i)(xR_p - C), \tag{8}$$

where $P(Y(t) = x | Y(0) = i)$ is given by:

$$\begin{aligned} &P(Y(t) = x | Y(0) = i) \\ &= \sum_{k=0}^{x} \binom{i}{x-k} [q(\hat{\lambda}, \hat{\mu}, t)]^{x-k} [1 - q(\hat{\lambda}, \hat{\mu}, t)]^{i-x+k} \\ &\quad \times \binom{N-i}{k} [p(\hat{\lambda}, \hat{\mu}, t)]^{k} [1 - p(\hat{\lambda}, \hat{\mu}, t)]^{N-i-k}, \end{aligned} \tag{9}$$

in which we assume that $\binom{i}{j} = 0$ for $i < j$.

To compute $p(\hat{\lambda}, \hat{\mu}, t)$ and $q(\hat{\lambda}, \hat{\mu}, t)$ we use the two-state Markov chain of each On-Off source. $p(\hat{\lambda}, \hat{\mu}, t)$ and $q(\hat{\lambda}, \hat{\mu}, t)$ can derived using the forward Chapman-Kolmogorov matrix differential equation:

$$p(\hat{\lambda}, \hat{\mu}, t) = \frac{\hat{\lambda}}{\hat{\lambda} + \hat{\mu}}(1 - \exp(-(\hat{\lambda} + \hat{\mu})t)), \tag{10}$$

$$q(\hat{\lambda}, \hat{\mu}, t) = \frac{\hat{\lambda}}{\hat{\lambda} + \hat{\mu}} + \frac{\hat{\mu}}{\hat{\lambda} + \hat{\mu}} \exp(-(\hat{\lambda} + \hat{\mu})t). \tag{11}$$

Ⅳ. Optimal Number of Connections

When the traffic source is indirectly connected to thesatellite system via another network, the satellite can not control the traffic parameters directly. It can only control the connection admission or the number of connections per downlink. When a subscriber requests to establish a connection with network, the satellite management function which executes the bandwidth allocation will first predict the cell loss ratio using our algorithm. Then the management function will deny or accept the connection. We determine the optimal number of connections, N^*, as follows:

$$N^* = \max\{n | CLR(\hat{\lambda}, \hat{\mu}, n, i, t) < QoS_{CLR}\}, \tag{12}$$

where $N^* \geq i$.

Ⅴ. Numerical Results

To investigate the quality of the ML estimators we have set up a simulation of binary sources. When the process is in the active state, it generates information at a constant rate of R_p bits/s. In Table Ⅰ and Ⅱ, we use $M = 1,000$ and $M = 10,000$ sample points, respectively. The percentage errors from the real values of λ and μ are also included. We observe that for low values of λ and μ (e.g. $\lambda =$ 0.0085 for $M = 10,000$) the ML estimation method exhibits large errors. However, as we increase λ and μ the percentage of error dramatically decreases

The number of samples M clearly makes a difference in the estimation quality. The more points we collect, the more accurate the estimator is.

For the numerical calculations of the transient cell loss ratio, we consider the following system: $N = 40$ PCM coded voice sources with $R_p = 64kbps$, $\lambda = 0.5$, $\mu = 0.833$ multiplexed onto a downlink of capacity $C = 1.544Mbps$. We first compare the results we obtained for the transient cell loss ratio with the steady state cell loss ratio (which is independent of the number of active sources at the beginning of the period).

In Figure 5 we depict the transient cell loss ratio as a function of the prediction time (in seconds) for various values of the initial conditions, $Y(0)$. We observe that after approximately 4.5 seconds the transient cell loss ratio will converge to the steady state cell loss ratio (obtained using the result presented in [6]). At 0.25 seconds, the round trip delay in a GEO satellite system, we observe significant difference in the results obtained as a function of the different initial conditions (the number of active sources at the beginning of the estimation interval). For example, at $t=0.25$sec we observe that the transient cell loss ratio given $Y(0)=0$ is approximately 10^{-12} while the steady state probability is approximately 10^{-4}. We, therefore, conclude that the computation of transient cell loss ratio is more accurate and can lead to significant different values of the cell loss ratio and consequently of the CAC decisions.

Table. I ML ESTIMATOR RESULTS FOR M=1,000.

Actual parameters		M=1,000	
μ	λ	$\hat{\mu}$ (%error)	$\hat{\lambda}$ (%error)
0.0132	0.0085	0.0136(3%)	0.0123(45%)
0.0132	0.085	0.0122(7.5%)	0.1010(1.9%)
0.0132	0.85	0.0142(7.5%)	0.7647(10%)
0.132	0.0085	0.1846(40%)	0.0139(63%)
0.132	0.085	0.1311(6%)	0.0724(15%)
0.132	0.85	0.1246(5.6%)	0.8720(.52%)

Table. II ML RESULTS FOR M=10,000.

Actual parameters		M=10,000	
μ	λ	$\hat{\mu}$ (%error)	$\hat{\lambda}$ (%error)
0.0132	0.0085	0.0155(17%)	0.0092(8%)
0.0132	0.085	0.0143(8%)	0.0828(2.5%)
0.0132	0.85	0.0143(8%)	0.8159(2%)
0.132	0.0085	0.1204(8%)	0.0097(14%)
0.132	0.085	0.1398(5.9%)	0.0852(0.2%)
0.132	0.85	0.1329(0.7%)	0.8523(0.2%)

In Figure 6 we plot the transient cell loss ratio as a function of the number of connection for different values of the ratio between the number of active sources at $t=0$ and the number of connections ($Y(0)/N$). We observe that for given values of (λ, μ) per connection, the transient cell loss ratio increases as a function of the number of connections. This is due to the fact that the overall traffic intensity per downlink increases as the number of connections increases. As the ratio $Y(0)/N$ increases, the transient cell loss ratio will increase. These figures show the importance of considering the *transient* cell loss ratio, quantity which we compare with the required QoS and consequently make our CAC decision.

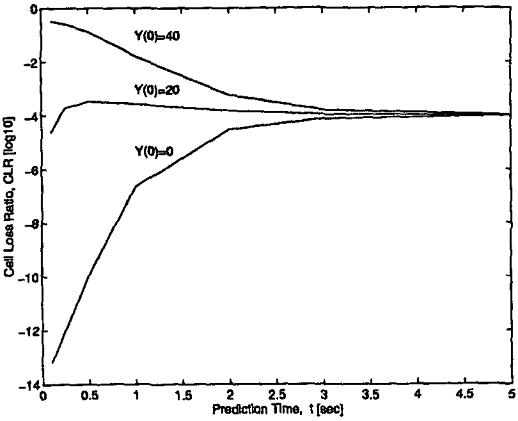

Fig. 5 Transient cell loss ratio versus prediction interval for multiple values of the number of active sources at $t=0$.

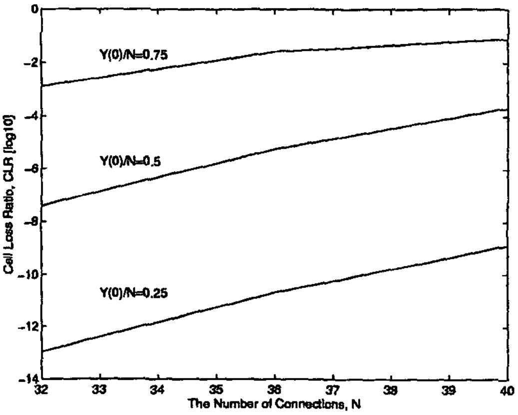

Fig. 6 Transient cell loss ratio as a function of the number of connections for different values of the ratio between the number of active sources at $t=0$ and the number of connections.

We first obtain the performance of the proposed CAC algorithm (as depicted in Figure 2) for obtaining the maximum number of connections, N^*. We assume the following network parameters: $t=0.25$sec, $R_p=64Kbps$, $C=1.544Mbps$, $\lambda=0.05$, $\mu=0.833$, and $QoS_{CLR}=0.01$.

In Figure 7 we observe that as the initial number of active sources increases, the maximum number of connections will decrease. This is due to the fact that as the number of initial active sources increases, the cell loss ratio increases. To satisfy the QoS we, therefore, have to decrease the total number of connections, N^*.

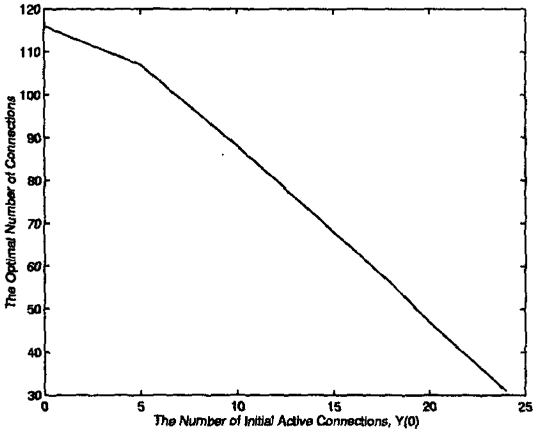

Fig. 7 The optimal number of connections versus initial number of active sources.

Ⅶ. Conclusions

In this paper we introduced a predictive CAC procedure for future on-board satellite systems that support multimedia traffic.

The proposed scheme considers the unique features of satellite communication, e.g., the congestion control scheme assumes no buffer on-board the satellite, accounts for the large propagation delays and has a relatively low computational complexity. Theoretical and simulation results suggest that the proposed ML estimator is well suited for voice traffic and video telephone traffic. The algorithm can be implemented in DSP chips to speed up its execution and reduce the power consumption on-board the satellite. Numerical results show the importance of using a transient analysis to compute the cell loss ratio rather than a steady state analysis. Therefore, the proposed method provides an accurate connection prediction and control scheme executed on-board the satellite.

Reference

[1] W. D. Ivancic, M. J. Shalkhauser, and J. A. Quintana, "A network architecture for a Geostationary communication satellite, "*IEEE Comm. Magazine*, July 1994.

[2] P. P. Chu, W. D. Ivancic and H. Kim, "On-board closed-loop congestion control for satellite based packet switching networks," *NASA Technical Memorandum 106446*, AIAA-94-1062. 1994.

[3] W. Matragi and K. Sohraby, "Combined reactive/preventive approach for congestion control in ATM networks," *IEEE ICC*, June 1993.

[4] T. Murase and et al., "A call admission control scheme for ATM networks using a simple quality estimate," *IEEE Journal on Selected Area in Communications*, Dec. 1991.

[5] G. Ramamurthy and B. Senpupta, "A predictive hop-by-hop congestion control policy for high speed networks," *IEEE INFOCOM*, April 1993.

[6] E.A. Bobinsky, "Predictive on-board flow control packet satellites," *AIAA-92-2004-CP*, pp 1776-1786, March 1992.

[7] P. Tsingotjidis, J. F. Hayes and H. Kim, "Estimation and prediction approach to congestion control in ATM networks," *IEEE GLOBECOM}*, Dec. 1994.

[8] Y. M. Jang, A. Ganz and J. F. Hayes, "A predictive congestion control for broadband wireless LAN," *IEEE International Conference on Wireless Communications*, July 1995.

[9] P.T. Brady, "A statistical analysis of on-off patterns in 16 conversations," *Bell System Technical Journals*, Jan. 1968.

[10] B. Maglaris and et al., "Performance models of statistical multiplexing in packet video communications," *IEEE Tran. on Communications*, July 1988.

[11] D. Anick, D. Mitra and M.M. Sondhi, "Stochastic theory of a data-handling system with multiple sources," *Bell System Technical Journal 61*, 1982.

장 영 민(Yeong Min Jang) 정회원

1985년 : 경북대학교 전자공학과 학사
1987년 : 경북대학교 전자공학과 석사
1999년 : University of Massachusetts 컴퓨터 과학과 박사
1987～현재 : 한국전자통신연구원 무선방송기술 연구소 선임연구원
<주관심 분야> 이동통신시스템, 위성통신 및 무선통신망