論文 89-26-10-1

Performance on the Beam-Switched Demand Assigned Multiple Access for the Packet Satellite Communication

(패킷 위성통신의 빔스위칭 요구할당 다중 접속 방식에 대한 성능 연구)

金 德 年* 金 在 明*
(Doug Nyun Kim and Jae Moung Kim)

要 約

본 논문에서는 패킷 위성통신에 있어서 요구할당 다중 접속 방식의 새로운 방법을 제안한다. 스팟빔 스위칭 능력과 신호처리 능력을 가진 위성체는 지상 제어국과 협력하여 위성 시스팀의 전송 최대 효율을 얻을 수 있도록 할 수 있다. 이를 실현하기 위하여 각 지상 터미날은 '요구추종' 전송 모드 기법을 사용한다.

Lee & Mark가 제시한 다중 접속방식은 전송지연 특성을 높이기 위하여 요구할당 방식과 임의할당 방식을 병행하여 사용하였음에도 본고에서 제안하는 요구할당 접송방식은 Lee & Mark 방식보다 동일한 전송효율에 도달하는데 최대 약 200msec의 전송지연 절감효과가 있다.

제안한 접송방식의 성능을 검증하기 위하여 전송지연 대 전송효율 곡선과 전송지연 대 통신량 곡선을 구하였다. 수치 결과에 의하면 제안된 접송방식은 증계기수가 각 지역수의 반보다 작을 때에도 저전송지연 효과가 있음을 확인하였다.

Abstract

This paper aims at investigating the Demand Assigned Multiple Access (DAMA) system for the packet-switched Satellite Communication. An onboard processor of the multisport beam satellite incorporates the ground controller to maximize the packet transmissions for each slot. 'Request Following' transmission mode is introduced as a transmission strategy of ground station under the control of its zone controller.

The combined scheme of reservation channel access and contention channel access was proposed by Lee & Mark[3] for improving the Delay-Throughput performance. Our scheme provides less communication delay of approximately max. 200msec for achieving the corresponding throughput than the Lee & Mark's work does.

Delay versus Throughput curves as well as Delay versus Traffic parameter curves are obtained. Numerical results obtained through the analysis and by the computer simulation show that the proposed scheme provides the low average packet delay even under the condition that the number of transponders (M) is below the half of the number of zones(N).

I. Introduction

Multiple access communication using the

---

*正會員, 韓國電子通信硏究所
(Electronics and Telecommunications Research Institute)

接受日字 : 1989年 4月 21日

multispot beam packet switching satellite has such an advantage that the frequency can be reused over the traffic area of interest. Many schemes have been proposed for efficient utilization of transponder or channel bandwidth. One approach is to design an intelligent satellite with an on-board processor that controls the schedule of assignment for transmission. Acampora[1] proposed 'frame by frame' assignment for request from the ground stations. Thereafter, many researchers[6,7] investigated the on-board processor scheduling techniques that would handle the variable frame length. In such scheduling algorithm minimizing the schedule length has been the key issue. The algorithms so proposed have still the additional delay of frame time before the assignment for chanel allocation is made besides the round-trip transmission delay due to request. In this sense the schedule assignment on a slot by slot basis[2] would receive more attention for the multibeam switching satellite system. Sometimes the access scheme can be designed to adapt the traffic fluctuation so it combines the reservation channel access with the contention channel access strategy, resulting in the maximum channel utilization [3,4]. However no attempt has been made to reduce the long round-trip delay in demand assigned type multiple access system before the actual information packet is transmitted. Such an inefficient time delay can be removed by employing ground controller for each zone which controls its zone queue permitting only one packet to transmit in a given time slot.

We propose a 'Request Following' transmission mode which, incorporating the ground controller, can play a decisive role in minimizing the packet delay. In section 2 we describe the basic structure of the proposed model and scenario of the DAMA system. Mathematical formulation is elaborated in section 3, the results of numerical analysis are summarized in section 4 and concluding statement is in section 5.

II. DAMA Protocol with Request Following Mode

Packet-switched multiple access communication system with an on-board processing satellite, where M transponders serve N ($\geqslant$ M) separate zones, will be discussed in this section. Each transponder operates in two different frequency bands -uplink frequency and downlink frequency- at any given time interval and the up(down) link frequency is uniquely assigned for each zone. Same frequency bands can be used through the communication area of interest as illustrated in Figure 1 so that the multifold frequency reuse can be accomplished. By so doing the satellite system can take advantage of the orthogonaity employed for the mixed time, frequency, space division multiple access.

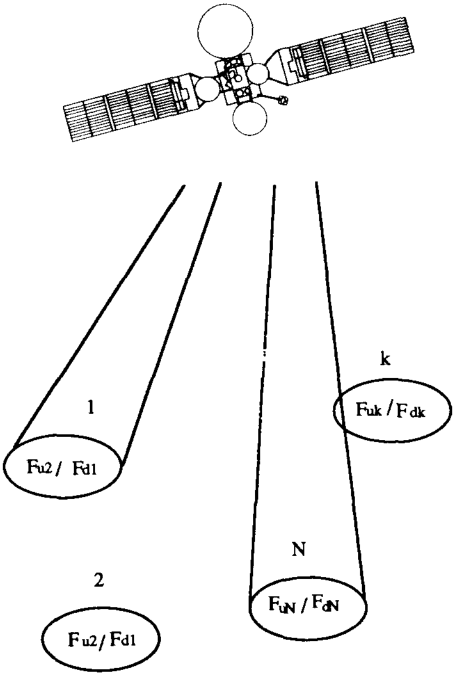

Fig. 1. Multibeam satellite.

Only a small portion of channel bandwidth is allocated for transmission of requests and acknowledgements of channel assignment. Request packets are separated with each other either by TDM or by FDM. It is assumed that bandwidth for requests is negligible when compared with packet channel bandwidth. Acknowledgements from the satellite are sent back to the ground stations in a broadcasting mode. Assume that the bandwidth for these acknowledgements is also negligible. Information packet channel is divided into slots. Each packet is of the same size as slot. Information packet transmitted is synchronized to the beginning of the slot.

Each request contains the source and the destination identifications. Each zone has ground controller which controls the timings for transmission of each request and information packet. The ground controller acquires all the information, such as the synchronization, the satellite drift, etc. from the satellite. Each station maintains a queue of requests and has an infinite buffer. Once a station has an information packet, whether the packet is newly generated or retransmitted, it inquires of the ground controller via terrestrial network about timings which will indicate the exact transmission time of the information packet. The ground controller maintains the status of waiting queue in its zone. The station transmits the information packet on the designated time slot right after the request is sent to the satellite. That is, request is first sent and the information packet is transmitted in synchronization with the satellite time slot.

The station keeps a copy of the transmitting packet at the time of transmission. The interval between the request packet and the information packet will be at most $\tau + 1$ slot and larger than the τ slot, where the τ is the maximum processing time of the processor on-board (Figure 2). The on-board processor measures the request for the next slot usage which are bound to arrive in a given time interval.. We assume here its interval is given by 1 slot. The processor has the strategy of assignment so that the utilization of the transponders could be maximized. No priority queues are allowed. Based on the information from the request queue and the assignment strategy, the on-board processor determines the access right for the next slot during the processing time.

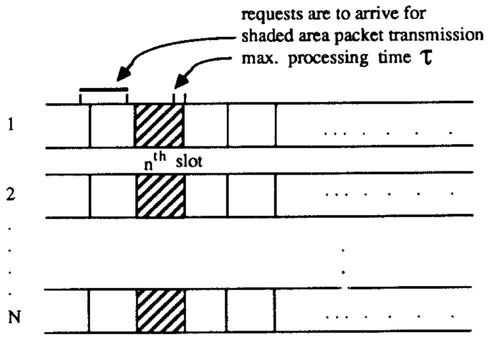

Fig. 2. Satellite time slot structure.

The result of assignment is sent back to the earth stations in a broadcasting mode for the acknowledgement. Assigned packets will be immediately served by the appropriate transponders whose switching matrix is set up by the processor. Upon receiving the positive acknowledgement, the ground station discards a copy of old packet, otherwise the ground station reinquires of the ground controller for retransmission.

From the scheme described above, following statements are observed.

- There is no contention between the stations, but the priorities on the transmission assignment is controlled by the ground controller. The waiting time elapsed before transmission must be counted.
- There is no buffer storage on the satellite. On-board processor schedules the switching sequence on a slot basis for next slot usage. Unassigned packets are cleared and need to request for the future assignment.
- 'Request following' transmission straegy is employed for removing the unnecessary round-trip delay.

Since no two information packets in a zone can be transmitted at the same time slot, the number of total requests received at the satellite in a given request interval is no greater than the number of zones. Processing time is always limited to the time processing for the maximum number of requests plus switching time so that all switchings are completed before the selected information packets are entering into that time slot (Figure 2). Maximum assignment algorithm proposed is similar to the procedure studied by Lee and Mark [3], but is simpler as follows. If more than one packet have requests for the same destination, only one is selected at random. If total number of destinations is greater than M, total number of transponders, only M packets are randomly selected without any priority.

It is interesting to note the similarities between Lee and Mark's proposal and ours. Both use the request channel before actual packet transmission. The on-board processor measures the incoming requests and assigns the schedule for maximizing the transponder utilizaticn. Nevertheless, our scheme has substantial differences in accessing the satellite channel, which leads to enhance the Delay-Throughput performance. In[3], if a terminal transmits in contention channel access mode it has to retransmit on a reservation mode later for assuring successful transmission. That wastes channel capacity. On the contrary, under our scheme the satellite only counts the demanding requests for slot assignment and the scheduling result is broadcast on a separate channel. This eliminates the unnecessary retransmission in case of success. It takes at least a round-trip delay in packet reservation under the scheme of [3], while the Request Following mode guarantees no such reservation delay if the packet is successful. Our scheme is originally designed such that no buffer storage for requests is required in the on-board processor. Thus we can extend to the on-board processor with buffer storage as Lee and Mark proposes to adaptively assign the unsuccessful packets.

Ⅲ. Delay-Throughput Analysis

In this section average packet delay relating to the satellite channel throughput is considered. Several assumptions are combined to the proposed model to justify the algorithm mathematically.

- Total source is a Poisson process with parameter G packets/slot which includes the newly generated packets as well as the packets retransmitted.
- The channel traffic during any time slot is an independent and stationary process.
- The whole communication regions are divided into zones such that each zone has an equally traffic of G/N.

Infinitely many small users generate bursty traffic of packet data. Thus the arrival traffic in an infinitesimal amount of time can be characterized collectively as Poisson statistics. The second assumption is necessary for the steady state analysis. In the long run, in and outflow of data traffic will obey the same statistics and are independent of the particular time slot. The third one is based on the ground that the whole communication area can be elaborately divided such that every zone has a uniform traffic. It can be generalized to include nonuniform zone traffic. We will not treat the extension case in this paper.

Transmission mode is shown in Figure 3. Once any station has a packet, the packet is queued in a distributed fashion under the supervision of the ground controller with First In First Out (FIFO) discipline.

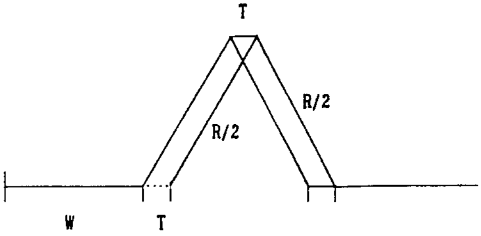

Fig. 3. Transmission model.

Suppose there is a packet (called test packet) which is to be transmitted from the ground station of a certain zone (called test zone). After waiting a certain amount of time in the ground queue, the test packet will be transmitted following after the request packet to the designated satellite cheannel (called test slot). The on-board processor will determine whether the packet is transmitted. If the test packet were allowed to transmit, the appropriate transponder would serve the packet, otherwise after one way propagation delay, the test packet will stand at the end of the line of the zone queue for retransmission. Thus the average packet transmission delay is given by

$$\begin{aligned} D &= W+R+T+r\,(W+R+T) \\ &= (W+R+T)\,(1+r) \qquad (1) \end{aligned}$$

where

W : Average queuing delay in the ground zone queue
R : Round-trip delay
T : Average interval between the request packet and the information packet
r : Average number of retransmissions

From Figure 2 it is apparent that $T = (\tau + 1)/2$ where τ is the maximum satellite processing time. In order to calculate W, average waiting time in a zone queue, Pollaczek-Khintchine formula for M/D/1 queue is applied with the parameter G/N. Average queue length L is given by

$$L = \frac{2\rho - \rho^2}{2(1-\rho)} \qquad (2)$$

where the traffic intensity $\rho=G/N$. Using Little's formula[8],

$$W=\frac{2N-G}{2(N-G)}, \qquad N>G. \tag{3}$$

Average number of retransmissions r can be estimated if the probability of success in transmitting is known. Let P_s be the probability that a test packet is successful when transmitted. Then,

$$r=\sum_{k=1}^{\infty} k(1-P_s)^k P_s$$

$$=(1-P_s)/P_s \tag{4}$$

Let us define the conditional probability q(i/k) as the probability that i packets are assigned out of k requests. q(i/j+1) can be expressed as

q(i/j+1) = q(i/j) Prob. (the additional packet not assigned)
|(N-i) channel pairs are avaliable)

+ q(i-1/j) Prob. (the additional packet assigned |
(N-i+1) channel pairs are available).

Thus

$$q(i|j+1)=q(i|j)\ i/N+q(i-1|j)\ (N-i+1)/N \quad \text{when } 1\le i\le M-1,$$

and

$$=q(i|j)+q(i-1|j)\ (N-i+1)/N \quad \text{when } i=M. \tag{5}$$

In calculating the q(i/j) recursively, we need to find the boundary conditons, q(1/k) and q(k/k).

$$q(1|k) = \text{Prob. }\{\text{only one packet assigned when k requests are present}\}$$
$$= \text{Prob. }\{\text{all request same destination}\}$$
$$= 1/N^{k-1} \tag{6}$$

and

$$q(k|k) = \text{Prob. }\{\text{all are assigned when k requests are present}\}$$
$$=\frac{\{\text{total number of combinations that avoid scheduling conflict}\}}{\{\text{total number of all possible requests}\}}$$
$$=\frac{N(N-1)\ (N-2)\ .\ .\ .\ (N-k+1)}{N^k}$$
$$=\frac{N\,!}{[N^k(N-k)\,!\,]} \tag{7}$$

From equation (5) and (6) every q(i/k) can be calculated recursively if N,M are given.

Now the probability of success P_s can be given by

$$P_s=\sum_{i=1}^{M}\sum_{k=1}^{N}\frac{i}{k}\,q(i|k)\ \mathrm{Fac}(N-1,k-1)\ P^{k-1}\,(1-P)^{N-k} \tag{8}$$

where

$$\mathrm{Fac}(N-1,k-1)=\frac{(N-1)\,!}{(N-k)\,!\ (K-1)\,!}$$

and P is the probability that there is at least a request in the ground zone queue. Fac(N-1,k-1) $P^{k-1}(1-P)^{N-k}$ is the probability that k-1 requests have arrived at the test slot from the remaining N-1 zones excluding the test zone. Therefore including the test packet from the test zone, k requests arrive at the test slot. In order to find P in equation (8), M/D/1 queue model is again used. It is known that the probability that there is at least a request in a zone queue is equal to the average number of requests arrived in a zone queue during 1 slot service time. Let X denote the queue length in a steady state and A denote the number of arrivals during one slot service time, then

$$P=\text{Prob. }(X>0)$$
$$=\int \mathrm{Exp}[A/T=t]\ b(t)\ dt. \tag{9}$$

where b(t) is the probability density function of service time. In M/D/1 queue $b(t)=\delta(t-1)$. Since arrival statistics is Poisson, Exp[A/T=t]=Gt/N. Thus

$$P=\frac{G}{N}\int tb(t)\ dt=\frac{G}{N} \tag{10}$$

and from equation (1) and (4)

$$D=(W+R+T)/P_s \tag{11}$$

The average packet delay D can be calculated if the parameters N, M, G are given. If throughput S is defined by the average number of assignments per slot for a transponder, S will be expressed as follows.

$$S = \sum_{m=1}^{M} m\ P_m/M \qquad (12)$$

where

$$P_m = \text{Prob.}\{m \text{ packets are successful}\}$$

$$= \sum_{k=m}^{N} q(m\,|\,k)\ \text{Fac}(N, k)\ P^k(1-P)^{N-k} \qquad (13)$$

Note that in calculating the throughput the channel bandwidth incurred by the request packet and channel assignment information is assumed to be negligible.

IV. Numerical Results

Based on the argument established in the previous section computer simulation is done as the traffic parameter G varies. From Figure 4 as an overview on the traffic flow, the total traffic G is the sum of the newly generated traffic G_n and the retransmitted traffic G_0 (traffic of the packets which have not been newly generated). The newly generated traffic must equal the average number of packets assigned in the statistical equilibrium, which is equal to M times S. Thus, for example, when M = 10, G_n is limited to MS and resultantly G_n should be less than 10. The values R and T are set by 27 and 1 in slot respectively throughout the calculation . Variations of these values affect the average packet delay very little since the probability of success in transmitting P_s in equation (8) is independent of these values.

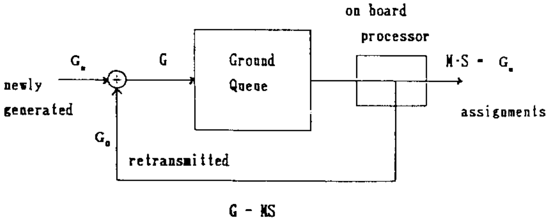

Fig. 4. Overview on the packet flow.

Table 1-3 show a comparison of Lee and Mark's and our results. When N=40 and M=20, our scheme is consistently superior to the Lee and Mark's proposal for all traffic conditions. This superiority remain unchanged even when we reduce the N and M to 10 and 5 respectively. Our scheme provides less communication delay of approximately max. 200 msec for achieving the corresponding throughput than the Lee and Mark's work does.

Figure 5 and 6 show the individual D and S characteristic as the traffic parameter G varies up to N. Having more transponders help stations experience less communication delay (Figure 5), at the same time it obviously lessens the utilization of satellite channel leading to the poor usage of transponders (Figure 6). This is due to the scheduling conflict among traffic. As the traffic is heavier it is more likely that the requests demand switching services for same destination.

Table 1. Comparison of Lee & Mark's and Our results with N=40, M=20, R=27.

S	D(msec)	
	Lee & Mark	Our results
0.2	518.8	306.9
0.4	529.4	324.7
0.6	539.6	345.6
0.8	550.5	376.7

Table 2. Comparison of Lee & Mark's and Our results with N=M=20, R=27.

S	D(msec)	
	Lee & Mark	Our results
0.2	463.4	323.8
0.4	495.7	373.7
0.6	535.3	501.0

Table 3. Comparison of Lee & Mark's and Our results with N=10, M=5, R=27.

S	D(msec)	
	Lee & Mark	Our results
0.2	459.8	304.1
0.4	491.8	323.0
0.6	521.7	344.9
0.8	542.4	387.9

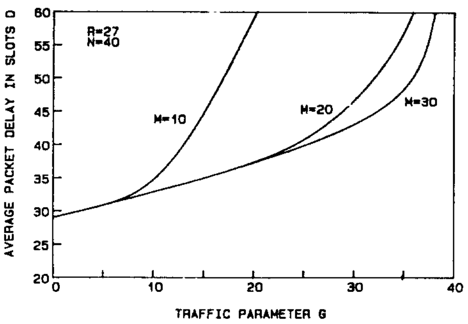

Fig. 5. D-G characteristic (N=40).

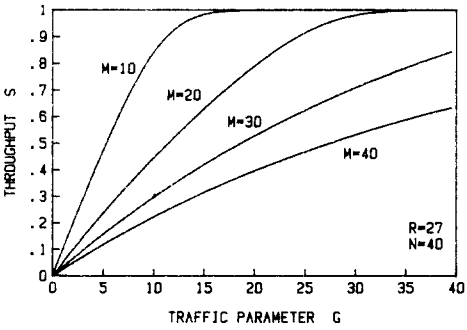

Fig. 6. S-G characteristic (N=40).

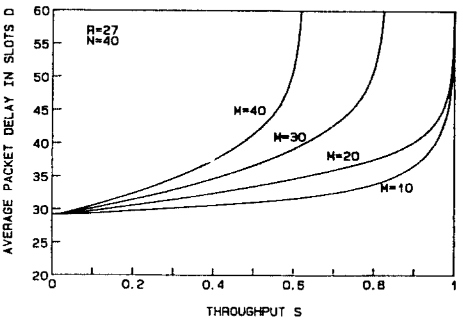

Fig. 7. D-S characteristic (N=40).

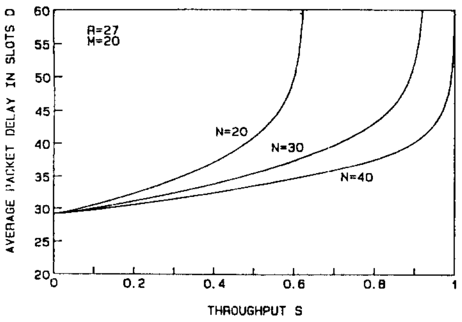

Fig. 8. D-s characteristic (M=20)

Let us denote S_{max} as the achievable maximum throughput of a system. Then from Figure 6 it is found that S_{max}(N=30,M=30) = .84 and max (N=40,M=40) = .63. It is interesting to notice that the maximum carried traffic MS_{max} is about 25 packeks/slot for both cases. The usage of more than 25 transponders gives no better improvement of satellite channel utilization.

To give us a little insight to find a low-delay and high-throughput system, D-S characteristics are shown in Figure 7 and 8 when N is fixed at 40 and when M is fixed at 20 respectively. The number of transponders need not be necessarily as large as the number of zones N. It is observed that the smaller M shows the better average Delay-Throughput performance to a certain limit. This is a great advantage because the reduction of the number of transponders makes the system more cost effective. The round-trip delay is assumed to be 27 slots so that the .27 sec round-trip delay means that 1 slot amounts to .01 sec.

In figure 7 the system supports only maximum throughput of .63 when M = 40. In other words only 25 out of 40 transponders are put into use in average, while all transponders can be used efficiently when M = 10. Figure 8 indicates that the increase of N is substantial to reduce the average packet delay, so that the available transponders should be utilized to the full extent.

A comparison between Figure 7 and Figure 8 also indicates that the D-S characteristic when N=M=40 is almost of the same performance as that of N=M=20. When N=40 and M = 10, D-S characteristic is shown to be stable if throughput is below than .97 although D-G characteristic is not so good as when N= 40 and M = 20 as illustrated in Figure 5. For a given system traffic the number of transponders can be selected as close to the maximum input traffic while achieving the maximum throughuput. This is the salient feature of the proposed multispot on-board the satellite access system because the satellite channel bandwidth can be fully utilized by the traffic on demand without paying much delay. The system with N=40 and M=20 supports the traffic of 36 packets/slot still maintaining the packet delay less than .6 sec.

Finally we illustrate D-S curves with serveral pairs of (N,M) in Figure 9. Although their M/N ratios are close one another, the figure shows how the D-S characteristic can be improved by gradually decreasing the ratio M/N. In this comparison the system view point that it can handle all traffic conditions without paying much delay. Since the ratio M/N characterizes D-S curves, once we design the number of transponders in accordance with the traffic demand we can determine the number of partitions in communication service area to achieve the low-delay and high-throughput capability.

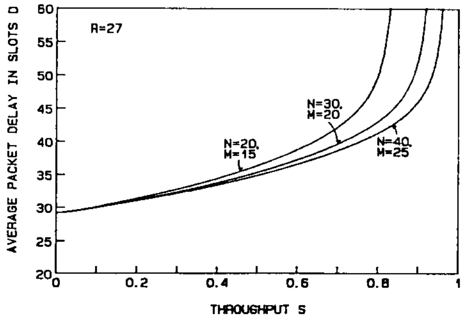

Fig. 9. D-S curves for the proposed scheme.

V. Conclusion

The beam switched satellite multiple access communication system was discussed. The proposed scheme achieves a stable system with very high-throughput and low-delay characteristic. Our scheme provides better Delay-Throughput performance than the hybrid channel access proposed by Lee and Mark. This achievement is not possible without the ground zone controller. The optimum number of transponders and the optimum number of zones can be found for the given traffic and the delay tolerance. Note that we do not discuss the efficient zone partitioning that employs the frequency reuse concept, which can enhance the system performance even better.

References

[1] A.S. Acampora and B.R. Davis, "Efficient Utilization of Satellite Transponders via Time-division Multibeam Scanning, "*Bell System Tech. Journal,* vol. 57, October 1978.

[2] Jayant Gadre and T.E. Stern, "A comparison of multiple access protocols for packet switching in satellite-switched multibeam systems," ICC 80, 58.4.1.

[3] H.W. Lee and J.W. Mark, "Packet-switched data communication over a multibeam satellite with on board processing," NTC 81 A-10.4.1.

[4] Sanjay K. Bose, "Packet communication system for a multibeam switched satellite repeater," ICC 82, 1E4.1.

[5] T. Suda, H. Miyahara and T. Hasegawa, "Optimal bandwidth assignment on up and downlinks of satellite with buffer capacity," *IEEE Trans. on Comm.,* vol. 28, October 1980.

[6] D.P. Gerakoulis,T.N. Saadawi and D.L. Schilling, "Minimizing the schedule length in an SS/TDMA system with minimum swtichings," ICC 1985, 5.7.1, pp. 157-161.

[7] G. Bongiovanni, D.T. Tong and C.K. Wong, "A general multibeam satellite switching algorithm," *IEEE Trans. Comm.,* vol. com-29, July 1981.

[8] J.D.C. Little, "A proof of the queuing formula L = W," oper. res., vol. 9, pp. 383-387, 1961.

Acknowledgement

The authors would like to acknowledge Mrs. Mi-Young Jung Ahn for her assistance in programming.

著 者 紹 介

金 德 年(正會員) 1952年 9月 15日生. 1975年 2月 서울대학교 전기공학과 졸업. 1981年 12月 미국 뉴욕 주립대(Stong Brook) 전자과 석사학위 취득. 1988年 6月 미국 Auburn 대 전자공학 박사. 1988年 9月~현재 한국 전자통신연구소 위성통신 기술개발실 선임연구원. 주관심분야는 스펙트럼 추정, Array 신호처리, 악세스 프로토콜, 위성통신 시스템 등임.

金 在 明(正會員) 1951年 12月 17日生. 1974年 2月 한양대학교 전자공학과 졸업. 1981年 2月 미국 남가주대 통신공학과 석사. 1987年 2月 연세대학교 전자공학과 박사. 1977~1979年 한국통신기술연구소 근무. 1982年~현재 한국전자통신연구소 전송기술 개발부 연구위원. 주관심분야는 디지틀 전송 시스템, 위성통신, 이동통신등임.