

PROCEEDINGS



**7th
Ka Band Utilization Conference
September 26-28, 2001
Santa Margherita Ligure - Genoa - Italy**

**Organized by:
IIC - Istituto Internazionale delle Comunicazioni
Genoa, Italy**

**Under the auspices of:
ASI - Agenzia Spaziale Italiana, Italy
CRC - Communications Research Centre, Canada
CRL - Communications Research Laboratory, Japan
NASA - National Aeronautics & Space Administration, USA**

A Control Architecture for Short- and Medium-Term Bandwidth Allocation in Satellite Channels with Fading

Raffaele Bolla*, Nedo Celandroni^o, Franco Davoli*, Erina Ferro^o, Mario Marchese[^]

*Department of Communications, Computer and Systems Science (DIST)
University of Genoa, Italy

[^]CNIT - Italian Consortium for Telecommunications
DIST - University of Genoa Research Unit
Via Opera Pia 13, 16145 Genoa, Italy

^oCNUCE-CNR, Area della Ricerca del C.N.R., Via Moruzzi 1, I-56010 Pisa, Italy

* lelus, franco@dist.unige.it; ^ mario.marchese@cnit.it

^o nedo.celandroni, erina.ferro@cnuce.cnr.it

Abstract: In previous papers we investigated two different architectures for bandwidth allocation and Call Admission Control (CAC) in a satellite environment, where real-time synchronous connections and best-effort data were transmitted. In this paper we present a new method derived from both architectures, aimed at keeping the call blocking probability of the real-time data below a given threshold, and at reducing the packet loss probability of best effort traffic. The environment now considered is much more dynamic with respect to the long-term statistics previously used. The present architecture is based on medium- and short-term control horizons and on the explicit real-time estimation of the fading conditions. In a real environment, based on the Italsat satellite national coverage payload characteristics, we tailored and compared both the new and the old architectures.

1. Introduction

In order to support different classes of traffic, characterized by diverse statistical nature and Quality of Service (QoS) requirements, over bandwidth limited channels, some form of control is necessary in the allocation of the available bandwidth. Moreover, to cope with possible variations in bandwidth demand and offered load, control actions should be devised to be dynamic (based on instantaneous and past information) or, at least, adaptive in nature. This is even more needed in satellite channels, where dynamically varying fading conditions can heavily affect the transmission quality, especially when working in Ka band, unless adaptive fade countermeasure techniques are adopted.

In [1] we described a centralized control algorithm for bandwidth allocation, in TDMA, based on a fixed boundary (B_0) between the capacity assigned to real-time traffic (stream traffic) and the portion devoted to non real-time traffic (datagram traffic). In any weather condition (clear sky or rain), new stream requests which would have overcome this boundary were refused by the call admission control (CAC). The datagram traffic could occupy the space unused by the stream traffic, up to the whole frame. The position of the CAC boundary B_0 was chosen a-priori, and once fixed it could not be moved. In faded conditions, data needed to be sent with higher redundancy, thus occupying a larger equivalent capacity. In this situation, already active connections were allowed to occupy a larger portion of the capacity up to a boundary B_1 , but from the CAC point of view the limit for the acceptance of new stream requests still remained fixed to B_0 . The algorithm was implemented and employed in a real environment. The experimental results are reported in [2]. In the following, we will refer to this architecture as CAC Fixed Boundary, or simply FB.

In [3] and [4], we proposed another assignment architecture, based on two-levels of bandwidth allocations and call admission control in a satellite environment with the same two basic traffic classes. The allocation mechanism aimed at keeping the call blocking probability

of the real-time traffic below a given threshold, and at reducing the packet loss probability of the datagram traffic. In doing this, a channel model suitable to reflect medium- or long-term fading statistics was adopted, disregarding the specific aspects involved in the estimation of its parameters. On this ground, the performance indexes to be used for resource allocation (call blocking and packet loss probabilities) were evaluated using stationary distributions. In this paper we investigate the application of a centralized control architecture, similar to the two-level architecture previously mentioned, applied within a more dynamic environment, based on medium- and short-term control horizons and on the explicit real-time estimation of the fading conditions. From here on this method will be indicated as FASTCOP (Fade Adaptive SaTellite Channel Optimization of Partitioning). To this aim, we adopt a resource allocation procedure in a TDMA framework based on two different nested time scales: i) a longer one (medium-term, in the order of minutes), where the control performed is of a stochastic nature and the system parameters (fading levels and system's traffic loads) are estimated by averaging over a time window; ii) a shorter one (in the order of the TDMA frame length), where the bandwidth is allocated deterministically to the stations for all connections in progress in the system, according to the "instantaneous" fading level, individually estimated over each point-to-point link between the transmitting and the receiving stations. In addition, we compare FASTCOP with the FB scheme, thus underlying in which cases one method performs better than the other. The comparison is made by using a real environment, where the characteristics of the Italsat satellite national coverage payload are considered.

2. FASTCOP system description

On the longer scale, each station sends the master the information about its traffic, i.e.: i) the mean real-time call traffic intensity ($\bar{\sigma}$), which may be computed by observing the number of Erlangs on a time interval of half an hour / an hour, or based on the statistical experience of previous days at the same time (this second case may be used at the startup of a station); ii) the mean bulk traffic intensity ($\bar{\gamma}$); iii) the factor \bar{f} , that is the mean redundancy factor (equal to 1 in clear sky conditions), necessary to cope with the signal fade. Both $\bar{\sigma}$ and $\bar{\gamma}$ values must be multiplied by \bar{f} , in order to evaluate the actual traffic load and, consequently, the bandwidth requirement. For bulk traffic, we consider a self-similar traffic model, which represents the superposition of an infinite number of on-off sources, with Pareto-distributed "on" time and exponentially-distributed "off" time, respectively [5, 6].

On receiving the $\bar{\sigma}$, $\bar{\gamma}$, and \bar{f} values, the master optimizes the allocation of capacity by minimizing a cost function $J(p_b, p_l)$, where p_b is the call blocking probability, derived from a stochastic knapsack model [7], and p_l is an upper bound on the bulk data loss probability, derived from [6], for self-similar input traffic, respectively. The optimization gives the maximum capacity to be devoted to the stream traffic (boundary B), and consequently the maximum number of calls acceptable by the system (I_{\max}). If stations are in different classes of fade, i.e. they have different \bar{f} values, they have generally different blocking probabilities, while the loss probability is, in principle, the same for all stations. In fact, the capacity allocated for best effort data on the shorter scale is made proportional to the actual traffic intensity, which takes into account the redundancy factor. More specifically, the average blocking probability is computed by using the stochastic knapsack stationary distribution, since we are in the presence of a multirate loss model. Actually, even though all stream connections have the same rate, their apparent rate may be different among groups of stations subject to different fading conditions. As regards datagram traffic, packets are supposed to be fragmented into cells (e.g., ATM) that fit each into a slot in the TDMA frame,

and the upper bound mentioned above is computed on the cell loss probability from a finite buffer. The optimization procedure is repeated with a period determined by the longer time scale. Ideally, such capacity reallocation interval should be long enough to justify the use of stationary distributions for the stream traffic, but not too long, in order to avoid basing the decision on stale values of f , which is in its own respect a time-varying parameter.

On the shorter scale, at each time frame each station sends the master the real-time estimation (derived by the instantaneous measurement of the attenuation) of the fading factor (f), together with an update of the situation relevant to the real-time synchronous connections (number of ongoing and of new requests). The master accepts new real-time calls if the boundary B has not yet been reached, and allocates each station a capacity sufficient to guarantee the new real-time traffic load, keeping into consideration the last f factor of the station. After the allocation to the real-time traffic is done, the remaining capacity is shared among the stations for the bulk traffic, keeping into account the individual f factors as well. Note that bulk traffic may temporarily suffer from a capacity assignment lower than the one resulting from the cost function minimization, due to the f factors' fluctuations. As a matter of fact, ongoing connections may exhibit a bursty behaviour, characterized by multiple rates, whereas their bit rate is considered constant, determined by the local value of \bar{f} , over their whole duration, in the model used for optimization. On a longer time scale, this effect is probably compensated, on average, as in some time intervals fluctuations of f may allow an assignment for the bulk traffic larger than the value resulting from the cost function minimization.

3. The real case study

We consider a fully meshed satellite network that uses bent-pipe geostationary satellite channels. This means that the satellite performs only the function of a repeater and it does not make any demodulation of data. The system operates in MF-TDMA (Multi Frequency-Time Division Multiple Access) mode. A master station maintains the system synchronization and it is responsible for the on-demand capacity allocation to the traffic stations. The master station performance is the same as the slave stations' one, thus the role of master can be assumed by any station in the system. This assures that the master operates in clear sky conditions for almost all the time, because when the current master's attenuation exceeds a given threshold, its role is assumed by another station that is in good conditions. The fade counter measure system adopted is based on the up-link power control and on bit rate changing. The multimedia data transmissions between the traffic stations occur in temporal slots assigned by the master, each one generally on different TDMA carriers (frequencies). The multi-frequency feature allows us to divide the system capacity into a number of channels, so that the traffic stations can be downsized with respect to a pure TDMA system. A traffic station cannot, however, transmit simultaneously on different TDMA carriers in the same temporal slot, because it is assumed to have only one modulator. The total transmission capacity of each station is thus limited to that of one carrier.

In Table 1 the most significant parameters of the system are reported. The values needed to compute the link budget are relevant to the transponder #1 of the Italsat national coverage payload, which operates in the 20/30 GHz band [8]. The information rate of 6.554 Mbit/s for each carrier is obtained with a 4/5 punctured convolutional encoder. The value of 7 dB of channel E_b/N_0 (bit energy to one-sided noise spectral density ratio), after 1 dB of modem implementation margin, is assumed as the threshold of the clear sky conditions. At the threshold conditions, after the Viterbi decoder [1], the bit error rate is 10^{-7} .

Stations' antenna diameter	1.8 m
Stations' power	13-dBW
Satellite G/T	5.9 dB/K
Satellite $E.I.R.P.$ (effective isotropic radiation power)	48 dB W
Number of carriers	3
Capacity of each carrier (QPSK modulation)	8.192 Mbit/s
Up-link power control range	5 dB
Minimum net E_b/N_0 in clear sky conditions	7 dB
Bit error rate (BER) guaranteed	10^{-7}
Possible data coding rates (convolutional)	4/5 (clear sky), 2/3, 1/2
Total information bit rate in clear sky conditions	19.66 Mbit/s
Information bit rate in clear sky conditions after system overhead	18 Mbit/s

Table 1. Most significant values of the MF-TDMA system considering the Italsat satellite.

In order to compute the resulting net values of E_b/N_0 at the earth station's receiver input, relation (1) is given. No automatic gain control feature operates on the transponder. For this reason the attenuation on the up-link affects both the up- and down-link C/N_0 values.

$$E_b/N_0 = C^{(res)} - 10 \log_{10} b_r - m_i \quad (1)$$

where:

$$C^{(res)} = C_r^{(up)} - A_{u_p} + C_r^{(dn)} - A_{u_p} - A_d - 10 \log_{10} \left(10^{(C_r^{(up)} - A_{u_p})/10} + 10^{(C_r^{(dn)} - A_{u_p} - A_d)/10} \right)$$

is the resulting C/N_0 (carrier power to one-sided noise spectral density ratio) at the earth station receiver,

$C_r^{(up)}$ is the reference (in clear sky) up-link $C/N_0 = 80.7$ [dBs⁻¹],

$C_r^{(dn)}$ is the reference (in clear sky) down-link $C/N_0 = 81.6$ [dBs⁻¹],

A_d is the dB down-link attenuation of the receiving station,

A_{u_p} is the dB up-link attenuation of the transmitting station, after up-link power control intervention:

$A_{u_p} = 0$, if the up-link attenuation $A_u \leq p_r$ (p_r is the up-link power control range = 5 dB); $A_{u_p} = A_u - p_r$, if $A_u > p_r$,

b_r is the data bit rate in bit/s

m_i is the modem implementation margin (assumed equal to 1 dB)

Table 2 contains the fade levels of the traffic stations, called *fade classes*, as function of the C/N_0 values. Each fade class imposes the adoption of the indicated transmission parameters (and then \bar{f} values) to limit the BER below the chosen threshold of 10^{-7} .

Fade classes	C/N_0 [dB]	\bar{f}	coding rate, b_r [Mbit/s]	Net E_b/N_0
1	>77.13	1	4/5, 8.192	7 dB
2	74.63-77.13	1.2	2/3, 8.192	4.5 dB
3	72.63-74.63	1.6	1/2, 8.192	2.5 dB
4	69.63-72.63	3.2	1/2, 4.096	2.5 dB
5	66.63-69.63	6.4	1/2, 2.048	2.5 dB
6	< 66.63	-	outage	-

Table 2. Redundancy factor \bar{f} versus fade classes.

The test used to make comparisons with the fixed boundary system shows 10 active stations, 8 of which are in clear sky, and 7 of them send data to stations in clear sky, while one sends data to one the two faded stations. One faded station sends data to the other one in fade, and thus these data experience fade on both up- and down-link. The other transmitting station in fade sends data to one that is in clear sky; thus, these data are affected by up-link attenuation only. The attenuation data are taken from a data set chosen from the results of the propagation experiment, in Ka band, carried out on the Olympus satellite by the CSTS (Centro Studi sulle Telecomunicazioni Spaziali) Institute, on behalf of the Italian Space Agency (ASI). The up-link (30 GHz) and down-link (20 GHz) samples considered were 1-second averages, expressed in dB, of the signal power attenuation with respect to clear sky conditions. The attenuation samples were recorded at the Spino d'Adda (North of Italy) station, in September 1992. In order to give some independence between the attenuation data of the two stations we shifted the samples relevant to one of them by a time interval of one hour.

3.1. The simulation results

We report in the following a simulative analysis of the proposed adaptive strategy, and its comparison with the fixed threshold policy. The basis of our real-life fading environment are the attenuation samples (averaged over 1s intervals), whose behaviour is depicted in Fig. 1.

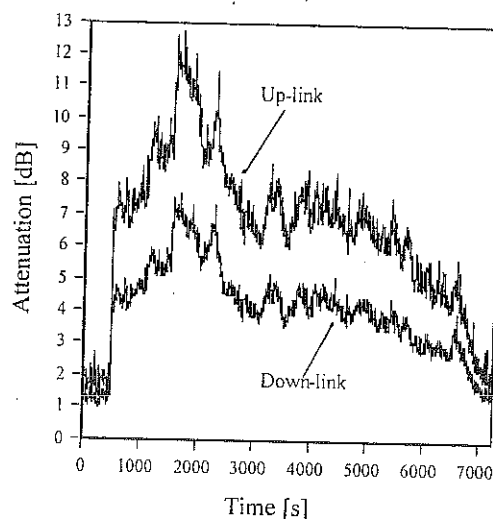


Fig. 1. Up- and down-link attenuation vs time.

The attenuation level determines, according to (1) and Table 2, the attribution of a station to a certain fade class. Actually, the value of attenuation (and, consequently, of \bar{f}) to be used over a reallocation period is determined by averaging the current estimation with a few previous values, weighted with a forgetting factor. In order to avoid too many oscillations in the "instantaneous" bandwidth assignment of a station, we have introduced a sort of hysteresis mechanism, whereby a station remains in the same fading class, unless the corresponding attenuation value exhibits a change above a given threshold (1 dB, in our case) for more than 3 seconds. On the other hand, as far as the outage is concerned, we adopted the definition of "unavailable time" given in ITU-T Recommendation G.821 [9].

The data reported in Table 3 have been used for the generation of the traffic.

A weight coefficient α is used in the cost, namely, the cost function J to be minimized with respect to the boundary is

$$J(p_b, p_l) = \alpha p_b + (1-\alpha) p_l \quad (2)$$

The minimum bandwidth unit (*mbu*) that can be allocated has been taken equal to 12.8 kbit/s. The stream source rate is assumed to be 64 kbit/s for all sources. In practice, the particular type of stream traffic considered may represent relatively short videoconference calls, as in video-telephony applications. The reallocation interval is also relatively short (100 s).

Station connection generation rate for stream traffic	$\lambda_s = 0.2$ requests / s
Average stream connection duration	$1/\mu = 60$ s
Shape parameter of the Pareto distribution	$\delta = 1.5$
Average datagram burst duration	$\bar{\tau} = 19.8$ s
Average idle period	$\bar{\phi} = 80.2$ s
Datagram traffic intensity	$\lambda_{asy} = M/(\bar{\tau} + \bar{\phi})$
Buffer dimension [cells]	8000

Table 3. Data used in the simulation.

A first comparison is shown in Figs. 2 and 3, with respect to blocking probability of stream traffic and datagram traffic loss probability, respectively. Both quantities are averaged over the whole simulation interval (10,000 s), and plotted against the datagram load per station (which is obtained by varying the value M of independent datagram sources per station). As regards blocking, the Fixed Boundary strategy is obviously insensitive to the load of datagram bulk traffic; lower (less than 0.5%) or higher (about 4%) values are obtained by two different values of the boundary, expressed in *mbu*. On the contrary, the dynamic allocation provided by FASTCOP tries to adapt to the datagram load, while coping at the same time with the varying fading conditions. This effect is also apparent on the cell loss probability in FASTCOP, which is always between the two extreme values provided by the widely different thresholds of the fixed allocation. The role played by the weighting coefficient α is also well underlined. The slight reduction experienced with $\alpha=0.5$ in Fig. 3 for increasing datagram load can be explained with the larger space allowed by the increase in blocking.

The dynamic behaviour in the presence of relevant variations in the fading conditions is highlighted in Figs. 4 and 5, where the same quantities as above, respectively, are plotted against time (1,000 s averages). The behaviour of the FASTCOP reflects indeed the same characteristics already outlined with respect to the global average.

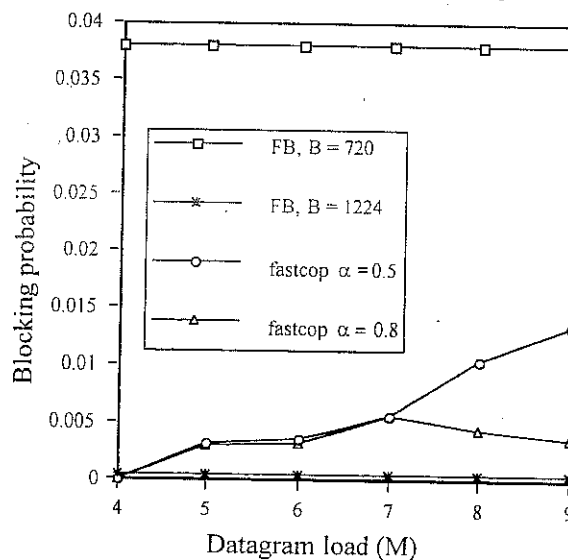


Fig. 2. Connection blocking probability vs datagram load, averaged over the whole simulation time.

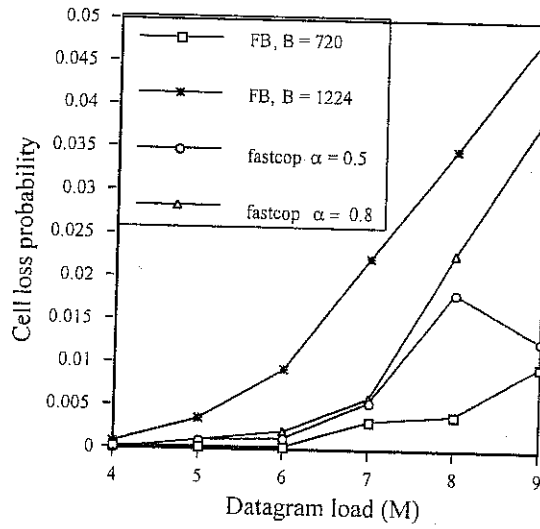


Fig. 3. Cell loss probability vs datagram load, averaged over the whole simulation time.

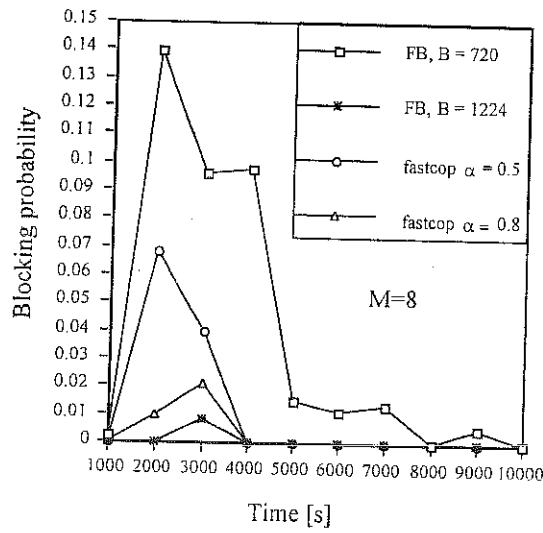


Fig. 4. Connection blocking probability vs time, averaged over 1000 s intervals (M=8).

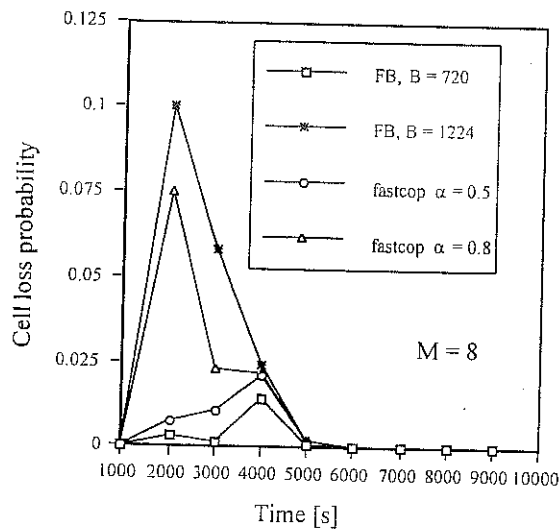


Fig. 5. Cell loss probability vs time averaged over 1000 s intervals (M=8).

4. Conclusions

We have presented and briefly analyzed a mechanism for the adaptive allocation of bandwidth in a satellite channel in the presence of fading and multimedia traffic. The proposed method is based on two control layers, operating at different time scales. With the frequency of the TDMA frame, a scheduler in the hub assigns the bandwidth that is necessary to support ongoing stream traffic connections, based on the instantaneous estimation of the fade attenuation, done by each earth station. With a much smaller frequency (typically, with a period of the order of minutes), the hub also computes the value of a threshold, to be used for stream traffic call admission control, based upon the estimation of an average fading level. Though further analysis and comparisons are necessary, and despite the various approximations used in the model for the optimization, the adaptive control exhibits the capacity to follow variations in the traffic rates, as well as changes in fade attenuation at a "macroscopic" level. In particular, the network manager can act on the performance indexes by means of a weight parameter in any system condition, rather than by fixing a threshold that may be optimized only for a specific situation.

Acknowledgements

¹The authors wish to thank Dr. Eng. Claudio Balestrino for his precious support in the simulation programs.

References

- [1] N. Celandroni, E. Ferro, N. James, F. Potorti, "FODA/IBEA: a flexible fade countermeasure system in user oriented networks", *Internat. J. Satellite Commun.*, vol. 10, no. 6, pp. 309-323, Nov.-Dec. 1992.
- [2] N. Celandroni, E. Ferro, F. Potorti, "Experimental results of a demand-assignment thin route TDMA system", *Internat. J. Satellite Commun.*, vol. 14, no. 2, pp. 113-126, March-April 1996.
- [3] R. Bolla, F. Davoli, M. Marchese, "A bandwidth allocation scheme in a satellite environment to support multimedia communications", *Proc. 6th Ka-Band Utilization Conf.*, Cleveland, OH, May 2000, pp. 613-620.
- [4] R. Bolla, F. Davoli, M. Marchese, "A bandwidth allocation strategy for multimedia traffic in a satellite network", *Proc. IEEE Globecom 2000*, San Francisco, CA, Nov. 2000, pp. 1130-1134.
- [5] B. Tsybakov, N.D. Georganas, "On self-similar traffic in ATM queues: definition, overflow probability bound, and cell delay distribution", *IEEE/ACM Trans. Networking*, vol. 5, no. 3, pp. 397-409, June 1997.
- [6] B. Tsybakov, N.D. Georganas, "Overflow probability in an ATM queue with self-similar input traffic", *Proc. IEEE Internat. Conf. Commun. (ICC'97)*, Montreal, Canada, 1997.
- [7] K. W. Ross, *Multiservice Loss Models for Broadband Telecommunication Networks*, Springer-Verlag, London, 1995.
- [8] F. Carducci, M. Francesi, "The Italsat satellite system", *Internat. J. Satellite Commun.*, vol. 13, pp. 49-81, 1995.
- [9] ITU-T, Recommendation G.821, vol. III - fasc. III.3, Malaga-Torremolinos Plenary Assembly, Oct. 1984.

¹ This work was partially supported by the Italian National Research Council (C.N.R.), under the "5%" Multimedia Program.