COMSAT

Technical Review

Volume 23 Number 2, Fall 1993

COMSAT TECHNICAL REVIEW is published by COMSAT Corporation. Subscriptions are: one year, $25 U.S.; two years, $45; three years, $60; single copies, $15; article reprints, $3. Overseas airmail delivery is available at an additional cost of $18 per year. Make checks payable to COMSAT and address to: Records Department, COMSAT Corporation, 22300 COMSAT Drive, Clarksburg, MD 20871-9475, U.S.A.

ISSN 0095-9669

COMSAT TECHNICAL REVIEW

Volume 23 Number 2, Fall 1993

Index: coding; digital transmission; mobile communications; models; modulation, demodulation, modems

Performance of M-ary FSK modulation in a shadowed land mobile satellite communications channel

R. A. KHALONA

(Manuscript received March 4, 1994)

Abstract

M-ary frequency shift keying (MFSK) is a power-efficient modulation method that is currently being studied for low-power and low-data-rate applications via satellite. The power efficiency of this method increases as the signal alphabet increases, at the cost of increased complexity and reduced bandwidth efficiency. The uncoded performance of MFSK in a land mobile satellite communications channel is analyzed using a channel model that includes the effects of shadowing. Results for coded performance using Reed-Solomon coding with hard-decision decoding are derived and used to determine the codes that require the smallest signal-to-noise ratio for a given bit error probability and for varying degrees of shadowing. These results should be useful in determining link margins, and as a reference for validating the findings of simulation studies.

Introduction

M-ary frequency shift keying (MFSK) is a power-efficient modulation scheme in which efficiency improves as the number of frequencies employed, M, increases, at the expense of additional complexity and reduced bandwidth efficiency. This scheme has been found to be advantageous in low-rate, low-power applications via satellite. In a previous paper [1], the performance of this modulation scheme was analyzed in a unified manner

that provided results for slow, frequency-nonselective Rayleigh and Rician fading channels, as well as in the additive white Gaussian noise (AWGN) channel. This was accomplished by using Rician statistics to model the amplitude of the received (faded) signal. The validity of this model has been verified for maritime and aeronautical environments by comparing measured distributions for the amplitude of the received signal with Rician distributions having varying carrier-to-multipath ratios, *C/M* [2]–[4]. For land mobile satellite communications, however, the Rician model is generally not suitable due to the presence of signal blockage (*e.g.*, shadowing from buildings and other obstacles). Instead, various channel models have been proposed that characterize the amplitude distribution of the received signal in the presence of shadowing [5],[6].

In this paper, the performance of uncoded MFSK modulation with optimum noncoherent detection is determined for a land mobile satellite communications channel, using the channel model proposed by Loo [5]. This model was selected because of its analytical tractability and ease of implementation for software simulation [7]. The use of noncoherent detection can be justified because it is robust in a fading environment and has less implementation complexity than coherent detection. In addition, in the absence of fading when the received signal-to-noise ratio, *S/N*, is large, the performance of optimum noncoherent detection approaches that of coherent detection, as shown in Figure 1 [1].

First, the channel model proposed in Reference 5 is briefly reviewed. The performance of uncoded MFSK using this channel model is then derived, and results are presented for M = 2, 4, 8, 16, and 32. These results are used to select the optimum Reed-Solomon code for 32-FSK in terms of minimum required energy-per-bit to noise-power density ratio, E_b/N_o, for a given bit error probability.

A channel model for land mobile satellite communications

The channel model considered here is obtained by assuming that the line-of-sight (LOS) signal component in the presence of shadowing is lognormally distributed in amplitude and uniformly distributed in phase, while the multipath signal component is Rayleigh distributed in amplitude and uniformly distributed in phase [5]. Denoting the amplitude and phase of the composite received signal by α and θ, respectively, this can be expressed as

$$ae^{j\theta} = ze^{j\alpha} + we^{j\beta} \quad (a, z, w > 0) \tag{1}$$

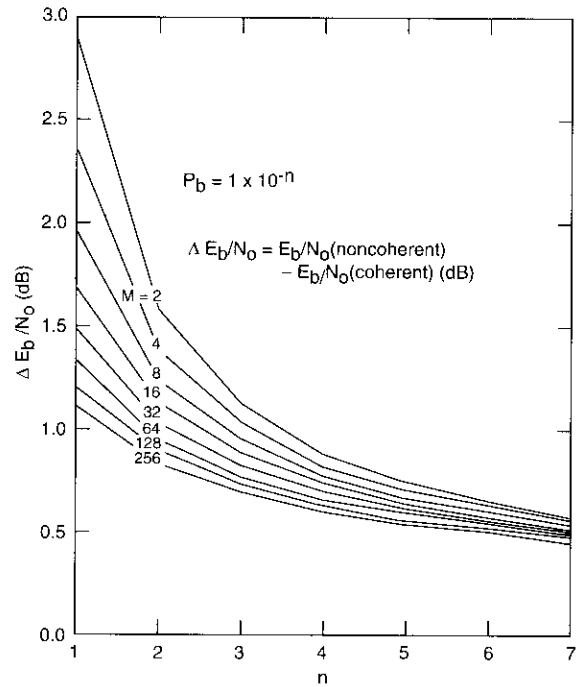


Figure 1. E_b/N_o Penalty (dB) in Using Noncoherent Detection With Respect to Coherent Detection as a Function of P_b and M for the AWGN Channel [1]

where z and w are lognormal and Rayleigh distributed, respectively, and the phases α and β are both uniformly distributed. When noncoherent detection is used and fading is sufficiently slow, the performance of a modulation scheme can be established based on the amplitude statistics of the received signal (LOS plus multipath signal components). With the above assumptions, the probability density of the amplitude of the received signal is given by [5]

$$p_A(a) = \frac{ae^{-\frac{a^2}{2b_0}}}{b_0\sqrt{2\pi d_0}} \int_0^\infty \frac{1}{z} e^{-\frac{z^2}{2b_0}} I_0\left(\frac{az}{b_0}\right) e^{-\frac{(\ln z - \mu)^2}{2d_0}} \, dz \tag{2}$$

where b_0 is the average multipath power relative to the unfaded LOS signal component, and μ and d_0 are the mean and variance, respectively, of the

lognormally distributed shadowed LOS signal component. The numerical values of these parameters, as a function of the degree of shadowing (obtained in propagation measurements at L-band, performed in Canada [8]) are given in Table 1 [9].

In the next section, this channel model is used to determine the performance of uncoded MFSK modulation with optimum noncoherent detection in a land mobile satellite link.

TABLE 1. PARAMETERS IN LOO'S CHANNEL MODEL*

DEGREE OF SHADOWING	$\sqrt{d_0}$	μ	b_0
Light	0.115	0.115	0.158
Average	0.161	-0.115	0.126
Heavy	0.806	-3.91	0.0631

*The parameters used are those in Table 1 of Reference 9. The discrepancy between these parameters and those of Reference 5 was kindly communicated to the author by C. Loo of CRC.

Uncoded MFSK performance

In MFSK modulation, one of $M = 2^k$ signals is transmitted every T seconds, where T is the time allotted for the transmission of each M-ary symbol and k is the number of information bits per symbol. Neglecting possible random phase shifts for each MFSK tone, which do not affect the performance of the noncoherent detector, each signal can be described by

$$s_i(t) = \sqrt{\frac{2E_s}{T}} \cos(2\pi f_i t), \quad 0 \le t \le T, \quad i = 1,2, \ldots, M \tag{3}$$

where $E_s = k\,E_b$ is the energy per symbol, and E_b is the energy per information bit. The received signal, $y(t)$, can be expressed as

$$y(t) = a(t) \cdot s_i(t) + n(t) \tag{4}$$

where $a(t)$ is a stationary and slowly varying channel amplitude (*i.e.*, slow fading) with amplitude statistics given by equation (2) above, and $n(t)$ is a zero mean Gaussian noise process with one-sided power spectral density, N_0 W/Hz. In what follows, it is mathematically convenient to characterize the fading process by defining $r(t) = a^2(t)$. That is, fading is dealt with on a power basis. Since $a(t) \ge 0$, the probability density function for $r(t)$ can be shown to be

$$p_R(r) = \frac{p_A(\sqrt{r})}{2\sqrt{r}} = \frac{e^{-\frac{r}{2b_0}}}{2b_0\sqrt{2\pi d_0}} \int_0^\infty \frac{1}{z}\, e^{-\frac{z^2}{2b_0}} I_0\left(\frac{\sqrt{r}\,z}{b_0}\right) e^{-\frac{(\ln z - \mu)^2}{2d_0}}\, dz \tag{5}$$

When fading is slow (*i.e.*, the bandwidth of the fading process, $r(t)$, is significantly smaller than the symbol rate), the symbol error probability for uncoded MFSK with ideal noncoherent detection can be found by averaging its unfaded performance over the statistics of the fading process, as

$$P_s = \int_0^\infty P_{s/r}(r, \gamma_s) p_R(r)\, dr \tag{6}$$

where $\gamma_s = k\,E_b/N_o$ is the S/N per MFSK symbol, and $P_{s/r}(r,\gamma_s)$ is the symbol error probability conditioned on a time-varying channel amplitude given by [1]

$$P_{s/r}(r,\gamma_s) = \int_0^\infty u\, I_0\left(\sqrt{2r\gamma_s}\, u\right) \exp\left[-\frac{u^2 + 2r\gamma_s}{2}\right] \left[1 - \left(1 - e^{-\frac{u^2}{2}}\right)^{M-1}\right] du \tag{7}$$

A well-known alternative form of equation (7), which is easier to compute when M is not too large, is given by

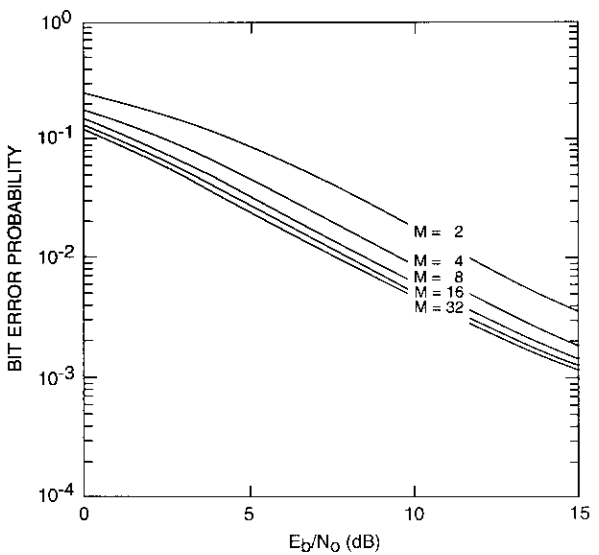
$$P_{s/r}(r,\gamma_s) = \sum_{i=1}^{M-1} C_i^{M-1} \frac{(-1)^{i+1}}{i+1} \exp\left(-\frac{i}{i+1} r\gamma_s\right) \tag{8}$$

where $C_i^M = M!/[(M-i)!\,i!]$. Substituting equations (8) and (5) into equation (6) and performing the integral with respect to r results in the following analytical expression for the average symbol error probability:

$$P_s = \frac{1}{\sqrt{2\pi d_0}} \sum_{i=1}^{M-1} C_i^{M-1} \frac{(-1)^{i+1}}{i + 2b_0\gamma_s i} \int_0^\infty \frac{1}{z}\, e^{-\frac{(\ln z - \mu)^2}{2d_0}}\, e^{-\frac{z^2\gamma_s i}{i+1+2b_0\gamma_s i}}\, dz \tag{9}$$

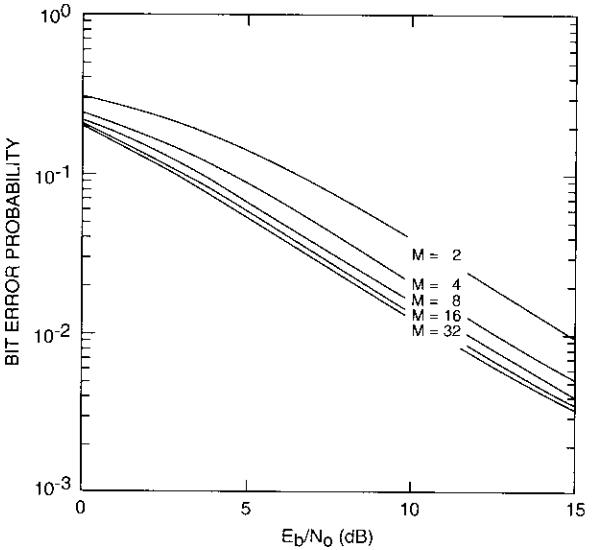
which has been integrated numerically to obtain the results given here. The integral in equation (9) was computed by using the substitution $x = e^{-z}$ and evaluating the resulting integral using a combination of Simpson's rule and the extended trapezoidal rule [10]. The uncoded bit error probability of MFSK for $M = 2, 4, 8, 16$, and 32 and for light, average, and heavy shadowing (as quantified by the parameters in Table 1) is shown in Figure 2. The uncoded bit error ratio (BER) ranges for these graphs were chosen to correspond to decoded BER ranges of interest when Reed-Solomon coding and hard-decision decoding are used, as discussed in the next section.

For orthogonal MFSK, the bit error probability can be related to the symbol error probability through $P_b = [M/2(M-1)]P_s$. The results given in Figure 2 for binary FSK agree with those presented previously by Loo [9]. They can also be useful in determining the performance of differential phase shift keying (DPSK), since for AWGN and slow fading channels the performance of DPSK is 3 dB better than that of noncoherent binary FSK. For fast fading, however, the latter may outperform the former [11].

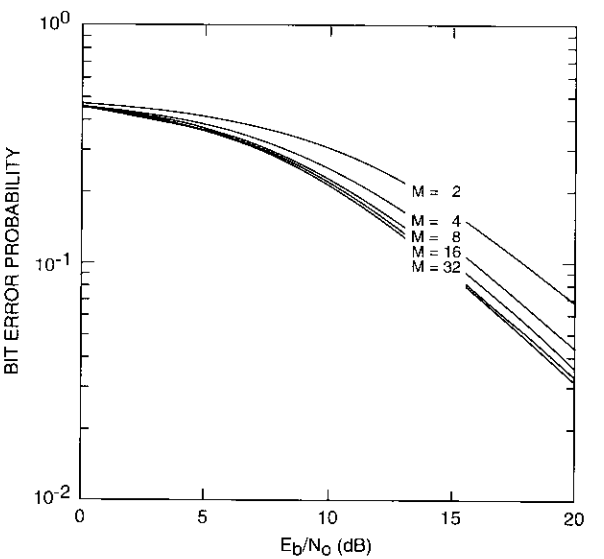


(a) Light Shadowing

Figure 2. *Uncoded Performance of Noncoherent MFSK Modulation in a Land Mobile Satellite Link With Shadowing*



(b) Average Shadowing



(c) Heavy Shadowing

Figure 2. *Uncoded Performance of Noncoherent MFSK Modulation in a Land Mobile Satellite Link With Shadowing (Cont'd)*

Close inspection of the results in Figure 2 reveals that, for moderate to high *S/N*, approximately a 2-dB improvement in performance is obtained as *M* increases from 2 to 4, and approximately 1-dB improvement results as *M* increases from 4 to 8, with diminishing performance returns as *M* increases further. These results will now be used to optimize Reed-Solomon (RS) code selection for a particular MFSK modulation scheme.

Reed-Solomon code optimization for 32-FSK

The results given in Figures 1 and 2 for uncoded performance reveal that MFSK performance does improve as *M* increases; however, for $M > 32$ the improvement is rather small. This indicates a tradeoff between improved power performance on the one hand and increased complexity plus reduced bandwidth efficiency on the other. The results presented in the previous section are used here to optimize code selection for 32-FSK, which is judged to offer a reasonable compromise between increased complexity and bandwidth efficiency for a target bit error probability, assuming RS coding with hard-decision decoding. Some of the merits of this modulation/coding scheme in terms of performance and implementation complexity have already been discussed in Reference 1.

When MFSK is used in conjunction with RS coding and hard-decision decoding, and when sufficiently long interleaving is used, decoded performance can be specified in terms of the uncoded symbol error probability as

$$P_b = \frac{M}{2(M-1)}\left[\frac{1}{N}\sum_{i=t+1}^{N} i\, C_i^N P_s^i (1-P_s)^{N-i}\right] \qquad (10)$$

where an (N, K) RS code is assumed. Such a code is capable of correcting up to $t = \text{int}\,[(N-K)/2]$ RS symbol errors. In evaluating equation (10), the *S/N* per MFSK coded symbol $\gamma_{sc} = r_c k\, E_b/N_o$ is substituted for γ_s in equation (9), where $r_c = K/N$ is the code rate of the RS code.

Optimizing the RS code for MFSK entails finding the value of *K* for which the required E_b/N_0 for a given bit error probability is minimized, assuming a direct mapping from RS symbols to MFSK tones (*i.e.*, $N = M - 1$), as discussed in Reference 1. The results of this optimization for 32-FSK are given in Figure 3, which shows the required E_b/N_0 for a bit error probability of 10^{-4} for light, average, and heavy shadowing. These results are summarized in Table 2, which gives the optimum value of *K* as a function of the degree of shadowing, and the required value of E_b/N_0 for the above bit error probability.

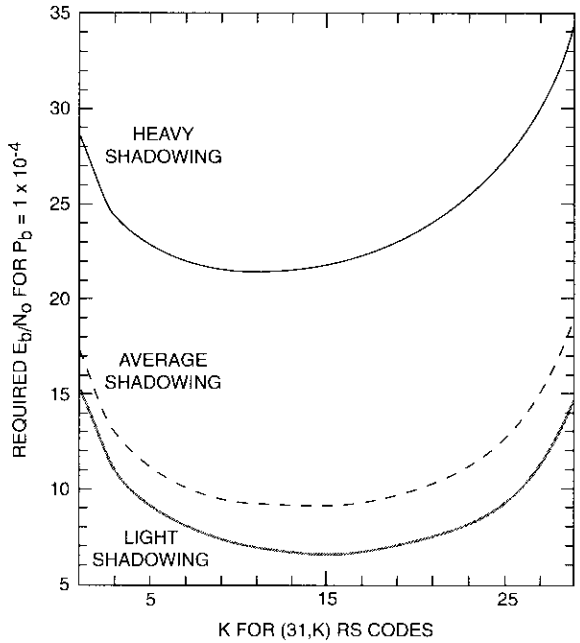


Figure 3. *Determination of Optimum RS Codes for 32-FSK in a Shadowed Land Mobile Satellite Link*

TABLE 2. OPTIMUM (31, K) RS CODE(S) FOR 32-FSK MODULATION IN A SHADOWED LAND MOBILE SATELLITE LINK

DEGREE OF SHADOWING	OPTIMUM *K*	E_b/N_0 (dB)
Light	15	6.6
Average	13 or 15	9.1
Heavy	11	21.4

The above results assume that very long interleaving is used. It is expected that limited interleaving will increase the E_b/N_0 requirement over that given in Table 2. Quantifying this interleaving degradation does not appear to be feasible without resorting to time domain simulations. The performance of 32-FSK with the optimum RS codes of Table 2 and hard-decision decoding is shown in Figure 4 for the three degrees of shadowing considered above.

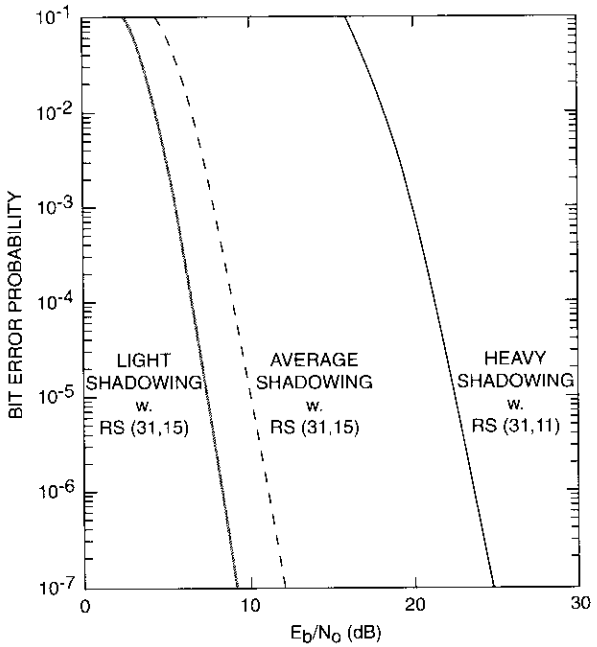


Figure 4. *Performance of 32-FSK With Optimum RS Codes of Table 2 in a Shadowed Land Mobile Satellite Link*

Conclusions

An analytic method for determining the uncoded performance of MFSK in a land mobile satellite communications channel has been developed. This method uses Loo's channel model [5] to account for the effects of signal shadowing. Channel amplitude variations are assumed to be slow with respect to signal variations. For low-data-rate applications via satellite, this assumption is probably valid for slowly moving users, even for small data rates. For stationary users, especially those who do not have a clear LOS to the satellite, the lognormal shadowing model may not be adequate, since fading may be too slow. In this case, performance might be more accurately determined by assuming a fixed power loss in the link. For nonstationary users, depending on the data rate, the validity of this assumption may be questionable, since fast fading can cause significant additional degradation in uncoded performance.

Results for coded performance using RS coding with hard-decision decoding were derived and used to determine the codes that yield the smallest *S/N* requirement for a given bit error probability and for varying degrees of shadowing. For coded performance, the degradation due to slow fading (*e.g.*, for stationary users) will strongly depend on the depth of interleaving employed, since slow fading generally requires that data be interleaved to a greater extent to disperse errors caused by long fades.

For those interested in other classes of codes, Lutz [13] gives rough guidelines for selecting the best interleaving depth for Bose-Chaudhuri-Hocquenghem (BCH) block codes and convolutional codes as a function of the data rate and maximum expected fading bandwidth.

References

[1] R. Khalona, "Unified Performance Analysis of Reed-Solomon Coded M-ary FSK Modulation in AWGN, Rician and Rayleigh Fading Channels," 2nd International Conference on Universal Personal Communications, Ottawa, October 1993, *Proc.*, pp. 662–668.

[2] W. A. Sandrin and D. J. Fang, "Multipath Fading Characterization of L-Band Maritime Mobile Satellite Links," *COMSAT Technical Review*, Vol. 16, No. 2, Fall 1986, pp. 319–338.

[3] J. Hagenauer *et al.*, "The Maritime Satellite Communication Channel—Channel Model, Performance of Modulation and Coding," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-5, No. 4, May 1987, pp. 701–713.

[4] A. Neul *et al.*, "Propagation Measurements for the Aeronautical Satellite Channel," IEEE 37th Vehicular Technology Conference, Tampa, Florida, June 1987, *Proc.*, pp. 90–97.

[5] C. Loo, "A Statistical Model for a Land Mobile Satellite Link," *IEEE Transactions on Vehicular Technology*, Vol. VT-34, No. 3, August 1985, pp. 122–127.

[6] E. Lutz *et al.*, "The Land Mobile Satellite Communication Channel—Recording, Statistics and Channel Model," *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 2, May 1991, pp. 375–386.

[7] C. Loo and N. Secord, "Computer Models for Fading Channels With Applications to Digital Transmission," *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 4, November 1991, pp. 700–707.

[8] J. Butterworth, "Propagation Measurements for Land Mobile Satellite Systems at 1542 MHz," Communications Research Center (Canada), CRC Technical Note No. 723, August 1984.

[9] C. Loo, "Digital Transmission Through a Land Mobile Satellite Channel," *IEEE Transactions on Communications*, Vol. 38, No. 5, May 1990, pp. 693–697.

[10] W. H. Press *et al.*, *Numerical Recipes*, Cambridge: Cambridge University Press, 1986, Section 4.2.

[11] M. Schwartz, W. Bennett, and S. Stein, *Communication Systems and Techniques*, New York: McGraw-Hill, 1966, Sections 11-2 and 11-3.
[12] J. Proakis, *Digital Communications*, 2nd ed., New York: McGraw-Hill, 1989.
[13] E. Lutz, "Code and Interleaver Design for Data Transmission Over Fading Channels," IEEE Global Telecommunications Conference, Atlanta, November 1984, *Proc.*, pp. 381–386.

Ramon A. Khalona received the BS, MS, and PhD degrees, all in electrical engineering, from the Illinois Institute of Technology (IIT) in 1982, 1984, and 1990, respectively. From 1984 to 1986, he performed earth station engineering work at the Managua Earth Station in Nicaragua, and then returned to IIT as an instructor in electrical engineering, teaching courses in communication systems and radio frequency design while earning a PhD with emphasis in telecommunication systems.

From 1990 to 1994, Dr. Khalona was a Member of the Technical Staff in the Communications Technology Division at COMSAT Laboratories, where he was engaged in system studies related to mobile satellite communications, including propagation transmission impairment analysis, channel modeling, and simulations. He subsequently joined Nippondenso America, Inc., LA Laboratories, in Carlsbad, California, where he is currently a Senior Member of the Technical Staff conducting R&D on cellular and personal communications. Dr. Khalona is a member of the IEEE Communications Society and the IEEE Information Theory Society.

Index: mobile communications, non-geosynchronous orbits, propagation, simulation, transmission

Availability prediction for handheld communications systems using non-geosynchronous satellites

W. A. SANDRIN AND D. V. HASCHART

(Manuscript received August 16, 1993)

Abstract

A variety of satellite constellations have been proposed for providing handheld-terminal telephone service via satellite. The constellations under consideration can be categorized as low earth orbit (LEO), intermediate circular orbit (ICO), or geosynchronous earth orbit (GEO). It is particularly important that a constellation allow calls to be established and maintained when blockage is present in the user's local environment. This issue is addressed for both LEO and ICO systems, and quantitative examples are given for selected LEO/ICO constellations, including the probability of a user completing a call of arbitrary duration in an environment with a precisely defined skyline that determines the directions in which communications links can be maintained. The examples indicate the quality of service that can be expected (given certain assumptions regarding user environment, link margin, and handover) and provide a basis for comparing the candidate constellations.

Introduction

A number of different satellite system designs are currently under consideration for providing voice service to handheld terminals. These systems can be categorized as low earth orbit (LEO), intermediate circular orbit (ICO), or geosynchronous earth orbit (GEO) satellite systems.

While many factors—including launch economics, spacecraft technology, earth segment complexity, network and handover issues, spectrum allocation

and sharing, and capacity—influence the selection of a preferred orbital configuration, one very important consideration is service availability. The customer environment is a major factor affecting availability for any type of satellite system; however, there is a fundamental difference between GEO and LEO/ICO systems in terms of link reliability if the customer is located in a non-ideal environment where service may be interrupted due to blockage or shadowing.

In a GEO system, link quality to a customer using a handheld terminal is usually not time-sensitive (neglecting multipath considerations), so that once a customer has established a good-quality link, the link is likely to remain good for the duration of the call. In contrast, link quality in a LEO/ICO system is time-sensitive because the elevation angle and azimuth to the satellite are constantly changing. Furthermore, abrupt changes in direction occur during satellite handover. The study described here was conducted to quantify availability for LEO and ICO systems. It examined the ability of these systems to provide reliable service to a customer who either receives or initiates a telephone call.

The term *availability,* when applied to communications systems, has a number of different formal definitions. Usually, availability is a measure of the fraction of time that a system is usable in some sense, although in some analyses of mobile systems availability has been associated with the fraction of geographical area for which communications links can be maintained. Here, *availability* is used in the context of a measure of time, although its specific definition is not the commonly used fraction (or percentage) of time that a link is usable, but rather the probability that a user can successfully complete a call lasting N minutes (where N is an arbitrary length of time) given a specific user skyline and a specific LEO or ICO satellite constellation. This definition provides a meaningful basis for comparing various non-geosynchronous satellite systems.

LEO/ICO system availability methodology and assumptions

In a LEO or ICO system, a user with a handheld terminal will experience uninterrupted service provided the network always hands over (reroutes) a call in progress to a satellite visible to the user. LEO/ICO system constellations are designed so that at least one satellite is always visible above a specified minimum elevation angle. Therefore, as long as the user has an unblocked view above the minimum angle, over 360° of azimuth, uninterrupted service is possible. (Satellites in LEO/ICO constellations are sometimes visible at elevation angles below the specified minimum, a factor that is important in computing availability when the user has an irregular skyline.)

The higher the minimum elevation angle for which satellite constellations are designed, the better their availability will be, since the chances of satellite blockage are reduced. However, LEO/ICO satellite constellations which provide higher minimum elevation angles are more expensive than those providing lower minimum angles, since more satellites are required for a given orbit altitude. Constellations with minimum elevation angles ranging from 10° to 30° have been proposed, and example constellations covering this range are compared herein.

For many potential user environments, the local topography (or "skyline") will extend above the minimum elevation angle in certain azimuthal directions. As a result, there may be some periods of time when no satellite is visible to the user. During these intervals, calls in progress will be terminated and new calls cannot be initiated or received. As the caller's skyline becomes increasingly filled with high-elevation obstacles (such as buildings, hilly terrain, or dense vegetation), service interruptions can be expected to increase.

The analyses that follow will quantify the relationship between the user's skyline and the quality of service (link availability) that can be expected from different LEO/ICO constellations. The analyses are based on two idealizing assumptions:

- That handover is perfect; *i.e.*, that the call in progress is always handed over to a visible satellite, provided one is available.
- That service interruption occurs whenever the user does not have a direct line-of-sight path to any satellite.

These assumptions permit the problem to be analyzed directly and unambiguously. Effects such as diffraction, shadowing, scattering, and multipath must also be considered in quantifying the relationship between availability and link margin. However, as discussed later, the amount of required margin can be inferred based on other factors. For link designs where the margin is relatively small (approximately 6 dB or less), the results presented here can give a reasonable estimate of link availability, provided the first assumption regarding perfect handover applies.

Figure 1 gives an overview of the process that was used to estimate link availability. The first step is to define the user environment and the satellite constellation. Information on the user environment is then used to model the user skyline. In the examples shown here, the user's skyline (in terms of elevation angle *vs* azimuth) was obtained in a number of ways. In the first

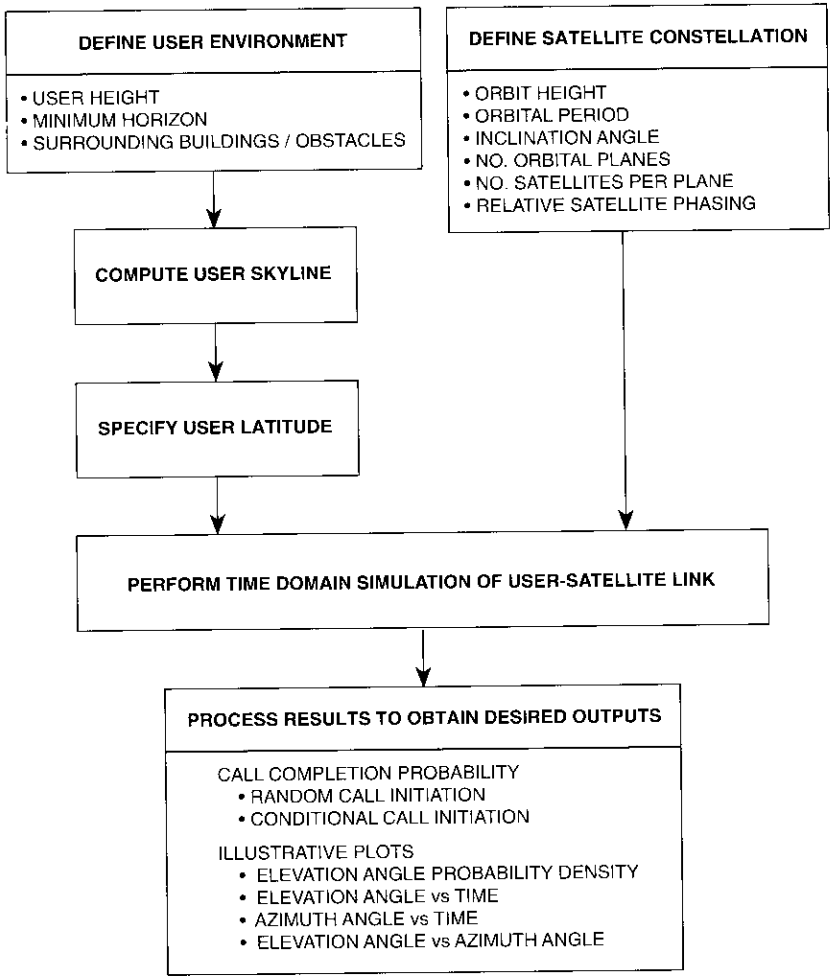


Figure 1. Overview of LEO/ICO Availability Calculation Methodology

approach, a "low-rise" urban setting was modeled by a cluster of buildings of specific sizes, with the user located at a specific point. This model was then used to compute the user's skyline. In subsequent examples, more-general skylines were modeled using different clusters of wall arrangements, each configured to produce the same average elevation angle. A simpler skyline model consisting of a constant elevation angle for all values of azimuth was also used. This latter model places the user in the center of a large circular pit—an unrealistic case, but one that is useful for calibration and comparison purposes.

In addition to the user environment, the user's latitude must be specified, since coverage for LEO/ICO constellations is sensitive to this parameter. (The long-term statistics of interest in this study are not sensitive to user longitude.) Figure 1 identifies the types of parameters that must be specified for the selected satellite constellation.

Given the satellite constellation data and the user skyline, a time domain simulation is run which computes, as a function of time, the elevation angle and azimuth to all satellites above the unblocked horizon, and then determines which satellites are visible to the user. The computed time record data are then processed to determine the relationship between the length of a call and the probability of successfully completing the call.

Examples given later provide other processed output, including plots of elevation angle or azimuth angle *vs* time, and elevation angle probability density functions. These intermediate relationships, which are generated in the process of computing availability, provide insight into the nature of the blockage problem.

The principal output of the LEO/ICO analysis is the relationship between the length of a call and the probability of successfully completing the call, given a specific satellite constellation, user skyline, and user location. Although this relationship is very sensitive to the specific user skyline assumed, the primary utility of such a relationship is twofold. First, it provides a quantitative indication of availability for environments that are representative of those for typical users. Second, it offers a mechanism for quantitative comparisons among different candidate constellations.

The software developed to perform the analysis illustrated in Figure 1 is described in an Appendix to this paper.

Comparison of four LEO/ICO systems

The example comparisons presented here are designed specifically to illustrate the comparison technique used and the differences among four example

LEO/ICO constellations, each having a minimum elevation angle between 10° and 30°. The examples also compare a LEO constellation and an ICO constellation having minimum elevation angles of approximately 20°.

The characteristics of the four LEO/ICO systems being compared are given in Table 1. Each constellation has been considered as a candidate for supporting a handheld communications system, with attention at the time of this study being focused particularly on the LEO-54 and ICO-12 constellations. The minimum visibility elevation angle for a constellation varies with user latitude. A latitude of 30° was selected for the examples given because, at this latitude, the minimum elevation angle is about the same (approximately 20°) for the LEO-54 and ICO-12 systems.

Availability results for a "low-rise" urban setting

A practical means of comparing the four candidate constellations is to compute the probability of completing a call lasting N minutes for specific local topographies. Figure 2 illustrates the basic user environment assumed for the examples given here. Two orientations are used: one in which the streets are aligned north-south and east-west (Figure 2), and another in which the same street layout is rotated 45° counterclockwise. This user setting is designated a "low-rise" urban environment. Note that the user in these examples is located 17.7 ft ($12.5\sqrt{2}$) from the corner of a 96-ft-high building at the street intersection. The user's optical skyline (in terms of elevation angle *vs* azimuth angle) which results from this environment is shown in Figure 3, where North is the 0° azimuth value. The skyline for the other example environment, in which the entire setting is rotated 45° clockwise, is the same as shown in Figure 3 except that 45° is subtracted from the azimuth scale.

Figures 4 and 5 plot the probability of completing a call lasting N minutes in the two example environments. The graphs are generated using the methodology shown in Figure 1, and the following assumptions apply:

- The user's environment is that shown in Figure 2.
- Handover is always perfect.
- The call is maintained as long as there is at least one satellite visible to the user. That is, as a satellite dips below the user's skyline, handover to another satellite is assumed, provided another satellite is visible to the user at that instant.
- The call occurs at a random time, with the user either initiating or receiving the call.
- For azimuth angles where the user sees the horizon instead of a building, a 5° minimum usable elevation angle is assumed. Hence,

TABLE 1. CHARACTERISTICS OF LEO/ICO SYSTEMS COMPARED

DESIGNATION	NO. SATELLITES	SYSTEM TYPE	NO. PLANES	NO. SATELLITES/ PLANE	INCLINATION ANGLE (deg)	ORBIT HEIGHT ABOVE EARTH'S SURFACE (km)	ORBITAL PERIOD (min)	MIN. ELEVATION ANGLE AT 30° LATITUDE (deg)	APPROX. MAX. VIEW TIME FOR 5° HORIZON (min)
LEO-35	35	LEO	5	7	90.0	1,584.9	117.9	10	21.0
LEO-54	54	LEO	9	6	55.0	1,800.0	122.7	20	26.5
ICO-12	12	ICO	3	4	50.7	10,355.0	359.0	20	144.8
ICO-15	15	ICO	3	5	53.5	11,622.2	400.6	30	164.5

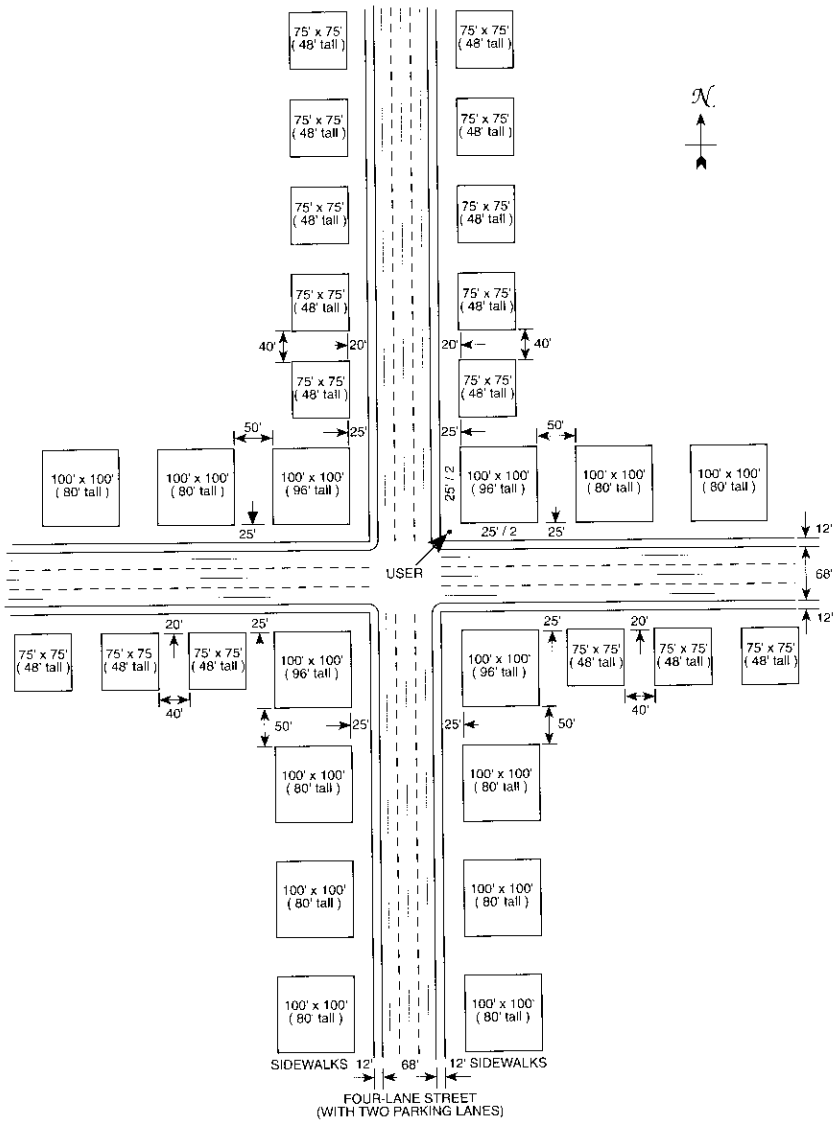


Figure 2. Simulated Low-Rise Urban Intersection

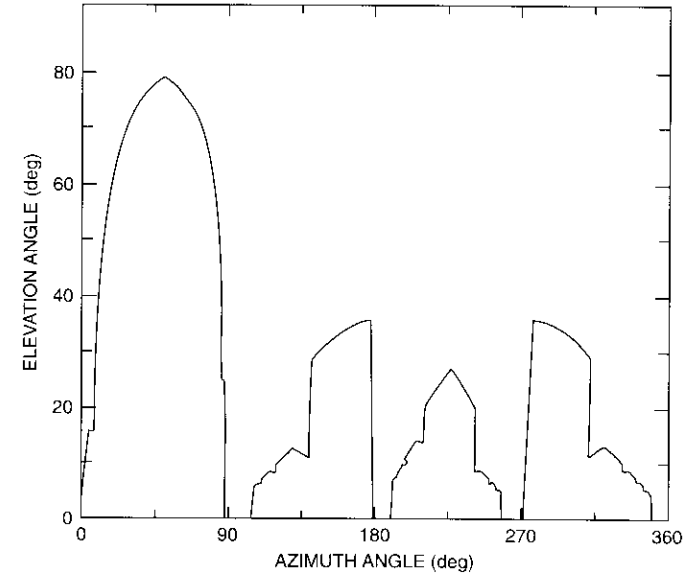


Figure 3. Low-Rise Urban Skyline (avg. elev. = 26.82°)

when the user has an unblocked view of a satellite at a 5° elevation angle, communication is maintained if all other satellites having higher elevation angles are blocked by the user's skyline at that instant.

- The link is broken when line-of-sight (optical) visibility to all spacecraft is lost.

As shown in Figure 4, the two ICO constellations provide better availability than the two LEO configurations, except for short call durations (less than 7 minutes), when LEO-54 is superior to ICO-12. It should be noted that, while most commercial telephony systems achieve availabilities well in excess of 0.95, the results shown in Figure 4 would, by comparison, be unacceptable for long call durations, especially for the LEO systems.

A 45° rotation of the street alignment (Figure 5) resulted in perfect availability for both ICO systems. The LEO-54 system was essentially unchanged (only slightly worse), and the LEO-35 system, which was poor in Figure 4, was significantly poorer. To assist in explaining these differences, a series of plots were made of the azimuth and elevation angles for the satellites having the highest and second highest elevation angles. The plots showed that, for the rotated street intersection model, the backup satellite is unblocked for both ICO constellations.

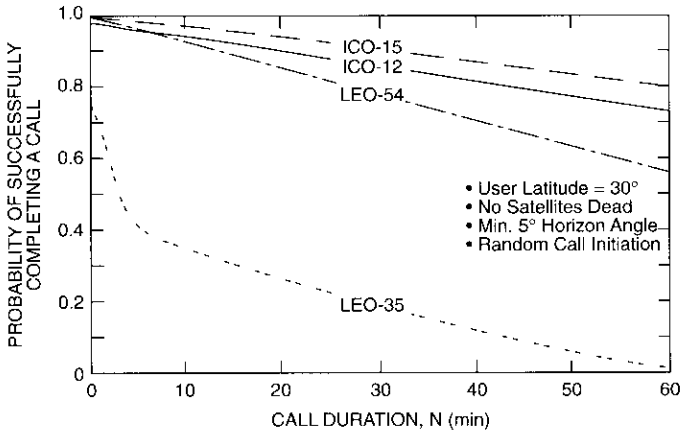


Figure 4. Probability of Completing a Call of N Minutes Assuming a City Intersection Skyline

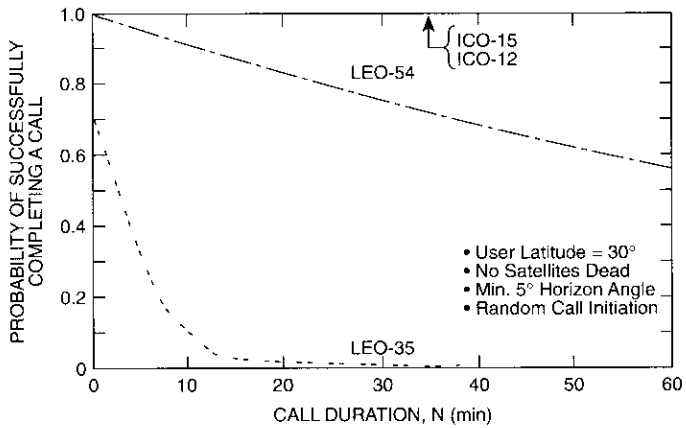


Figure 5. Probability of Completing a Call of N Minutes Assuming a City Intersection Skyline Rotated 45°

The results shown in Figure 5 for the ICO constellations, with the 45° street alignment, indicate that a satellite is always visible to the user despite the proximity of large obstacles, whereas for the original north-south alignment, there are occasions when no satellite is visible. This phenomenon is also illustrated in Figures 6a and 6b, which plot elevation angle (to the highest visible satellite) *vs* time for the respective street alignments. In Figure 6a,

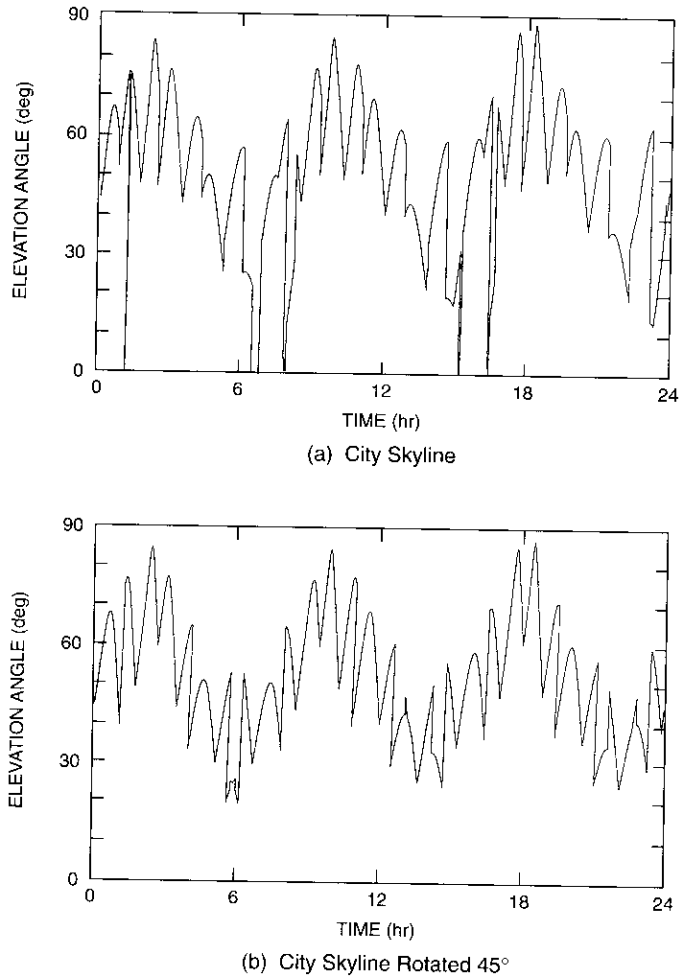


Figure 6. Elevation Angle vs Time for an ICO-12 System at 30° User Latitude

when no satellite is visible, the elevation angle is plotted as zero, and it can be seen that there are a few intervals over a 24-hr period during which no satellite is visible (thus making the probability of completing a randomly initiated call less than 1). When the street alignment is rotated 45° clockwise (Figure 6b), a satellite is always visible to the user.

Comparison of constellation elevation angle distributions

To provide insight into the nature of the four satellite constellations being compared, Figure 7 shows the probability density of the elevation angle of the highest satellite for each constellation. In each case, the probability density is zero below the minimum elevation angle for that constellation, as given in Table 1. Also note that this plot is specific to a 30° latitude.

The plot in Figure 7 indicates (but does not conclusively prove, for reasons discussed below) which constellations will have the best availability. For example, because the mean elevation angle of the LEO-35 constellation is lower than that of the other constellations, LEO-35 is likely to have the worst availability in a blocked environment. The fact that the LEO-54 constellation is skewed toward low elevation angles while the ICO-12 constellation is skewed toward high elevation angles indicates that, for many blockage environments, the ICO constellation will provide better availability than the LEO constellation, although both have a minimum elevation angle of about 20°.

It should be kept in mind, however, that the probability densities shown in Figure 7 provide no information regarding other satellites with lower elevation angles which may also be visible to the user. Thus, it cannot be concluded that the constellation with the highest average elevation angle is always the best for all skylines. (As noted for the example depicted in Figure 4, the LEO-54 constellation has better availability than the ICO-12 constellation for calls of less than 7 minutes duration.)

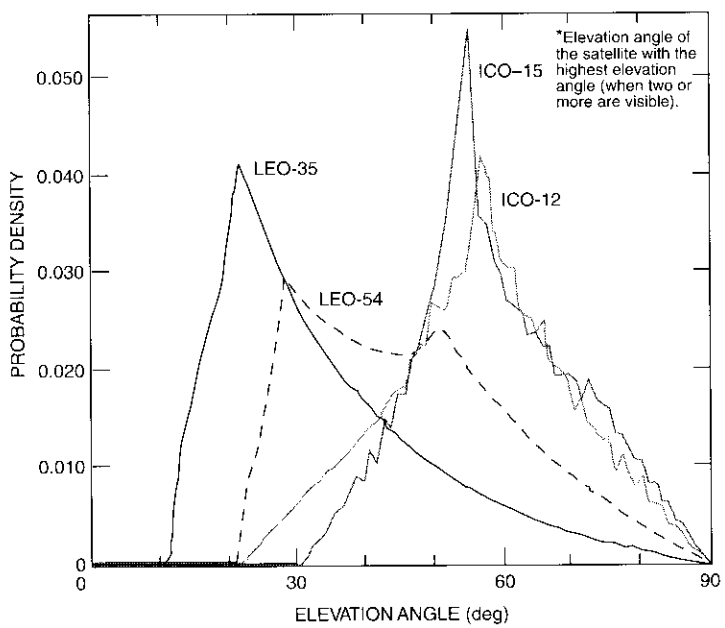


Figure 7. Elevation Angle Probability Density for a User at 30° Latitude*

Availabilities assuming constant elevation angle horizons

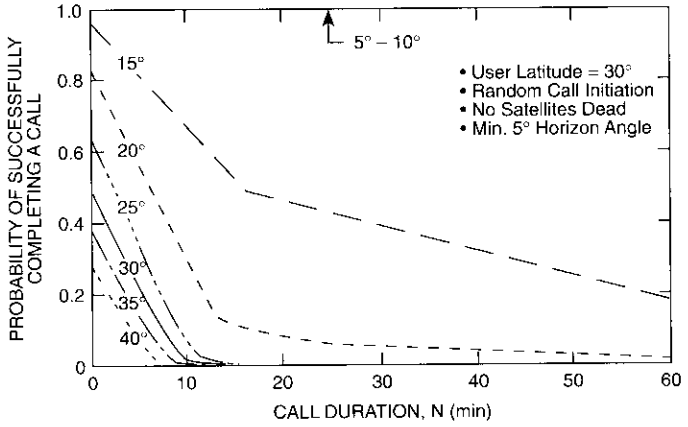
An alternative to comparing constellations on the basis of specific user environments is to make comparisons assuming a user horizon that is fixed at a constant elevation angle for all values of azimuth. This models the case where the user is located at the center of a large circular pit. While this is not very representative of most user environments, it is easily specified and unambiguous. Consequently, it provides useful information regarding the relative performance of different constellations. As will be discussed later, a constant-elevation-angle horizon is not the worst-case horizon. Other horizons exist that have the same average elevation angle but result in significantly poorer availability.

Figure 8 presents the results of the availability analysis for the four constellations being compared. As can be seen, when the horizon is below the minimum elevation angle guaranteed by the particular constellation, availability is perfect; when the user's horizon is higher than the minimum elevation angle, availability is not perfect. Thus, it is not surprising that the LEO-35 constellation, with its minimum elevation angle of 10° (Figure 8a), is the worst of the four, while the ICO-15 constellation, with its minimum elevation angle of 30°, is the best (Figure 8d).

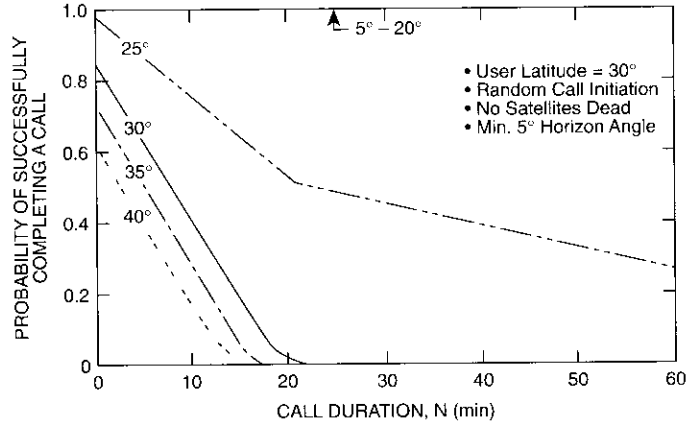
It is noteworthy that there is a significant difference in availability between the LEO-54 constellation and the ICO-12 constellation (Figures 8b and 8c), even though both have approximately the same minimum elevation angle of 20°. Comparing these two cases reveals that the ICO-12 constellation has significantly better availability than the LEO-54 constellation. This difference is attributed to the fact that the satellites of the ICO-12 constellation are visible to the user at relatively high elevation angles for a greater amount of time, whereas the LEO-54 satellites spend more time at lower elevation angles. This explanation is supported by the elevation angle probability density functions given in Figure 7.

Example of availability computed at a higher latitude

In the examples shown in Figures 4 and 5 a latitude of 30° was selected, since the minimum elevation angles for that latitude for both the ICO-12 and

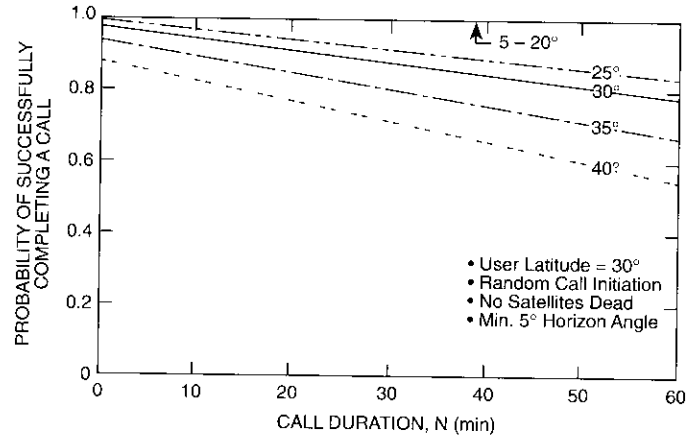


(a) LEO-35

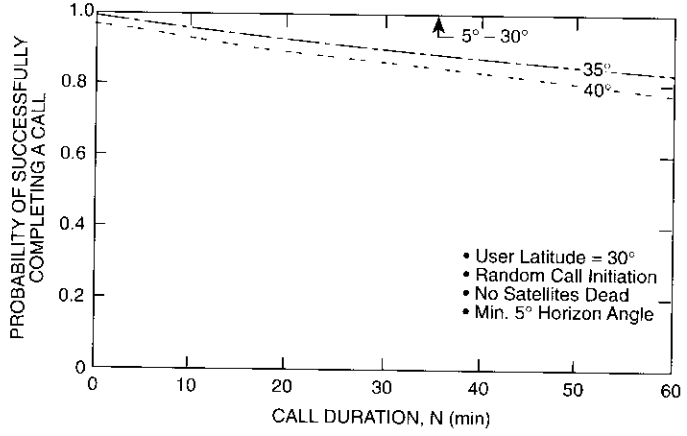


(b) LEO-54

Figure 8. Probability of Completing a Call of N Minutes for a Constant-Elevation-Angle User Horizon



(c) ICO-12



(d) ICO-15

Figure 8. Probability of Completing a Call of N Minutes for a Constant-Elevation-Angle User Horizon (Cont'd)

the LEO-54 constellations were about 20°. However, since elevation angle distributions change with latitude, it can be expected that availability will also change with user latitude.

Figure 9 shows the elevation angle probability density distributions at a latitude of 38° for the four constellations. The minimum elevation angle for the LEO-54 constellation has now become about 30°, whereas (from Figure 7) at a latitude of 30°, the minimum elevation angle was about 20°. Other differences in the elevation angle probability density functions can be seen by comparing Figures 7 and 9.

Figure 10 gives the results of the availability analysis for the site shown in Figure 2 at a latitude of 38°. Comparing the results in Figure 10 with those in Figure 4 reveals a dramatic improvement for the LEO-54 constellation, while the other three constellations show only slight changes in availability, with the ICO-15 and LEO-35 constellations improving slightly and the ICO-12 constellation becoming slightly worse.

Although the purpose of this example was to demonstrate that availability for LEO/ICO constellations is sensitive to latitude, the results for this one

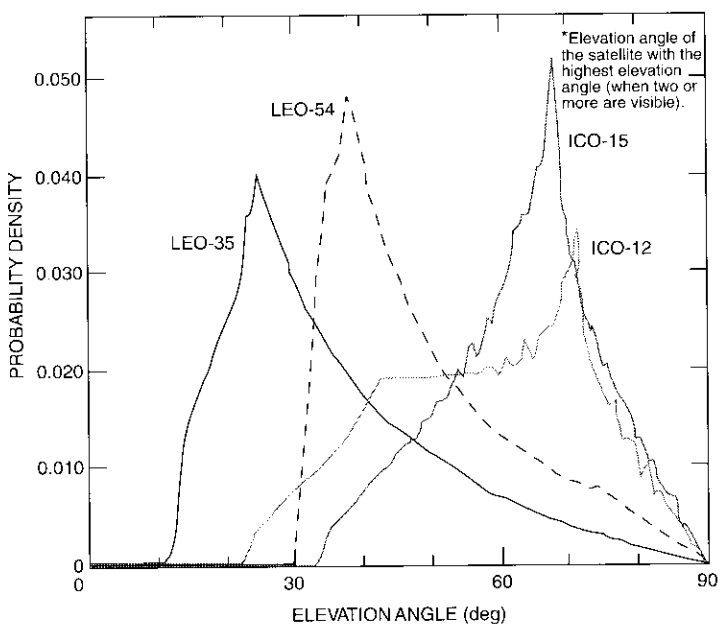


Figure 9. Elevation Angle Probability Density for a User at 38° Latitude*

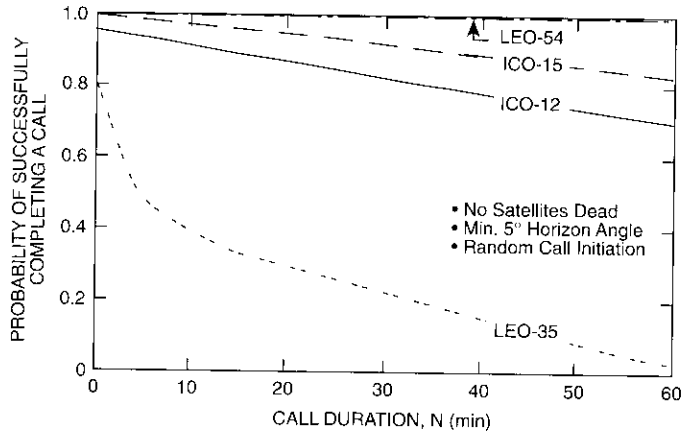


Figure 10. Probability of Completing a Call of N Minutes Assuming a City Intersection Skyline at 38° Latitude

example are not sufficient for generalizations regarding the relationship between availability and latitude. A better approach would be to compare the elevation angle probability density functions at many latitudes for constellations of interest and to include a series of availability plots for various latitudes, for different user environments.

Conditioned call completion probabilities

All of the probability-of-call-completion examples presented above assume that the call is initiated at a random time, which is the usual assumption made in teletraffic analyses. Although this is a reasonable assumption for competition-sensitive systems, especially when received calls are considered, availability for user-placed calls would be expected to improve if the user had a "green light" display on the handset which indicated that at least one satellite was visible. The probability of completing a call of N minutes would be enhanced, at the expense of having to wait for a satellite to become visible.

Figure 11 is an example of such a conditioned-availability relationship. This is the same as the example shown in Figure 4, except that calls are only initiated during periods when at least one satellite is visible to the user. Figure 12 complements Figure 11 by showing the time a user must wait for a visible satellite when a call is initiated randomly. Comparing Figures 4 and 11 reveals that imposing this condition makes little difference in the computed availabilities for the three best systems, and only modestly improves the availability of the poorest system. This is because, for the three best systems,

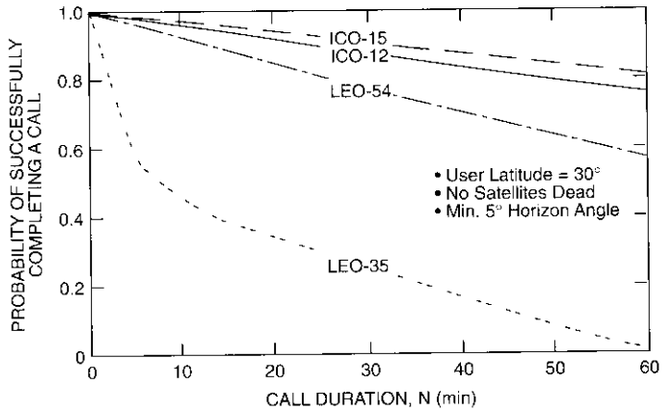


Figure 11. Probability of Completing a Call of N Minutes Assuming a City Intersection Skyline and a Call Initiated When a Satellite is Visible

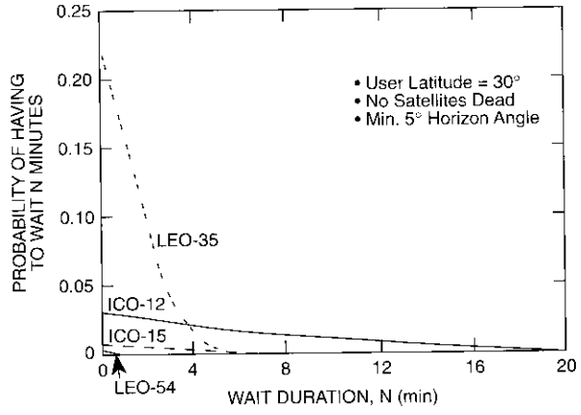


Figure 12. Probability of Having to Wait N Minutes Before a Satellite Becomes Visible, Assuming a City Intersection Skyline and Random Call Initiation Time

the probability of not having a visible satellite at any given time is very small. Therefore, this type of conditioned call access will only be of value when system availability is relatively poor to begin with.

Summary of LEO/ICO comparisons

The relationships described above were selected to compare four satellite constellations in terms of their service availability. In general, the higher a constellation's minimum elevation angle, the better its availability.

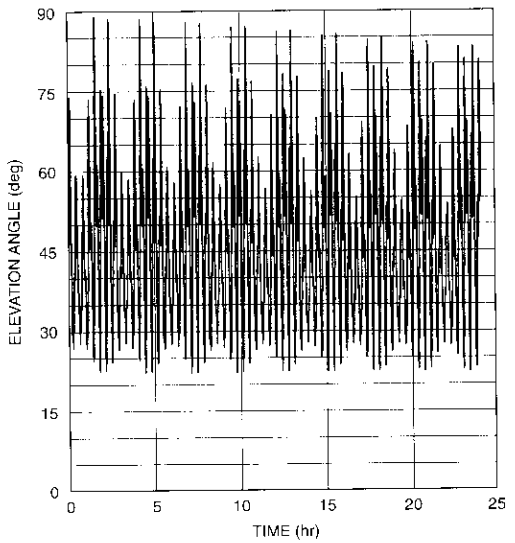
Two of the constellations examined (LEO-54 and ICO-12) have similar minimum elevation angles at a latitude of 30°. From the results presented here, it appears that the ICO-12 constellation would offer superior availability over a wider range of environments, although this observation is not conclusive, and other examples in which the LEO-54 constellation is the better of the two are also given (*i.e.*, Figure 4 for short calls, and Figure 10). More extensive comparisons using a greater number of sample environments and more values of latitude would be required in order to select one specific constellation over another. Also, these comparisons were made solely on the basis of an idealized availability analysis; other factors (identified in the Introduction) must be considered when selecting a preferred constellation.

Elevation angle and azimuth angle plots

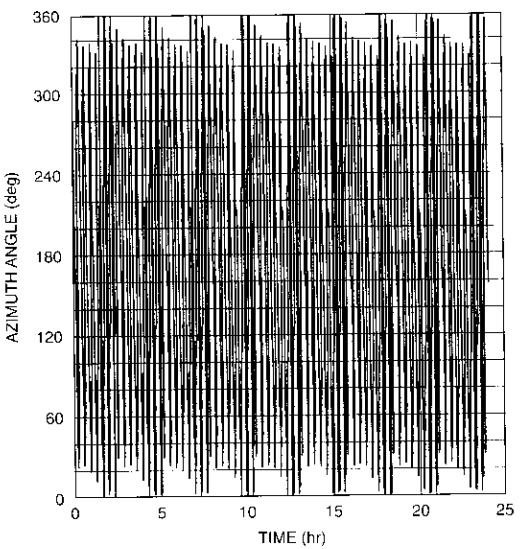
In the preceding section, the four LEO/ICO constellations described in Table 1 were compared, primarily on the basis of computed availability estimates for users located in various environments. An intermediate step in producing the availability plots is the computation of elevation and azimuth angles for the satellites as seen from the user's location.

Plots of elevation and azimuth angles for LEO/ICO constellations provide insight into their nature. As an example, the LEO-54 constellation is illustrated through plots of elevation angle *vs* time (Figure 13a), azimuth angle *vs* time (Figure 13b), and elevation angle *vs* azimuth angle (Figure 13c). In all of these plots, the values of elevation angle and azimuth angle are for the satellite with the highest elevation angle as seen by a user at 30° latitude. It is also assumed that there are no blockages between the user's location and the satellite (*i.e.*, that the user's skyline is a 0° horizon). The plots illustrate the following features, which may help provide an intuitive understanding of the nature of LEO/ICO systems:

- Satellite handover, which is identified by abrupt changes in the plotted functions.

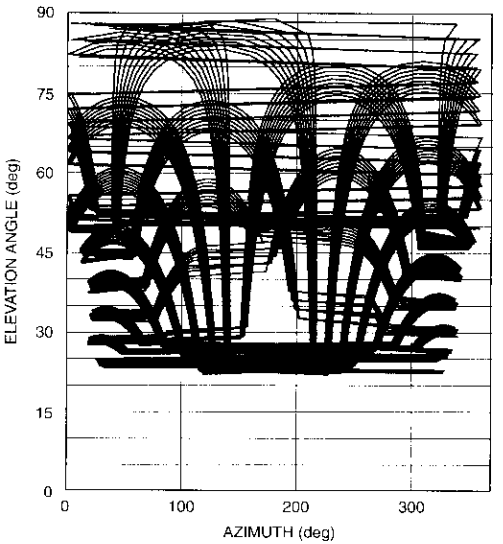


(a) Elevation Angle vs Time



(b) Azimuth Angle vs Time

Figure 13. Plots of Elevation and Azimuth Angle for a LEO-54 System at 30° User Latitude



(c) Elevation Angle vs Azimuth

Figure 13. Plots of Elevation and Azimuth Angle for a LEO-54 System at 30° User Latitude (Cont'd)

- Time variations in both elevation and azimuth angles, and the difference in these variations between LEO and ICO systems.
- Periodic patterns in the elevation and azimuth angles.

Effect of different skyline shapes on availability

Although it is readily apparent that the availability of a LEO/ICO system will degrade as the average elevation angle of the user's skyline increases, it is also of interest to know how the specific shape of a user's skyline affects availability. This relationship is illustrated here by assuming 13 different skyline shapes (Figure 14), each with the same average elevation angle. For each skyline, availability is computed in terms of the probability of completing a call lasting N minutes. The results are then compared, both with each other and with availability results obtained using skylines of constant elevation angle.

In the availability calculations, the heights of the obstacles shown in the figure have been adjusted so that the average skyline elevation angles are 16° and 25°. The examples discussed here assume a LEO-35 constellation (as described in Table 1) and a user latitude of 38.6° (Washington, D.C.).

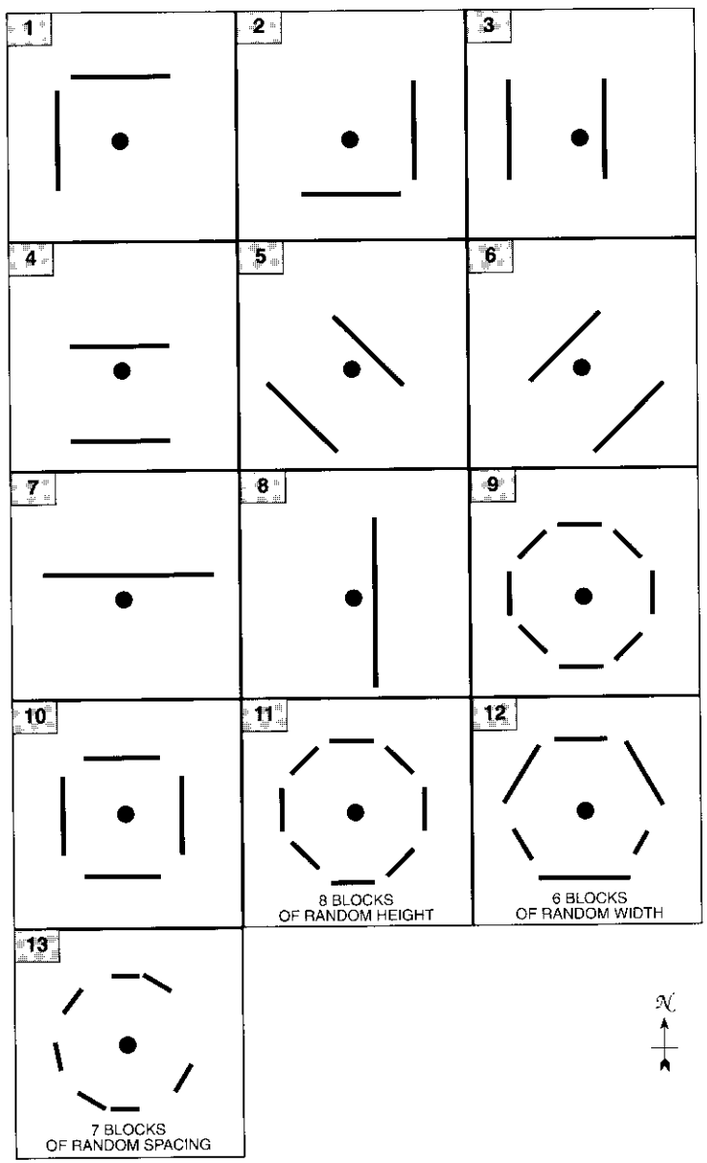


Figure 14. Skyline Blockage Scenarios

Figures 15a and 15b present the results of availability calculations for 16° and 25° average elevation angle skylines, respectively. Overlaid on these figures are the availabilities that result from uniform skylines of constant elevation angle. The same assumptions discussed previously were made in obtaining these availability results. That is, that handover is always made to a satellite visible to the user, that the link is broken when there are no satellites visible to a user, and that line-of-sight visibility exists.

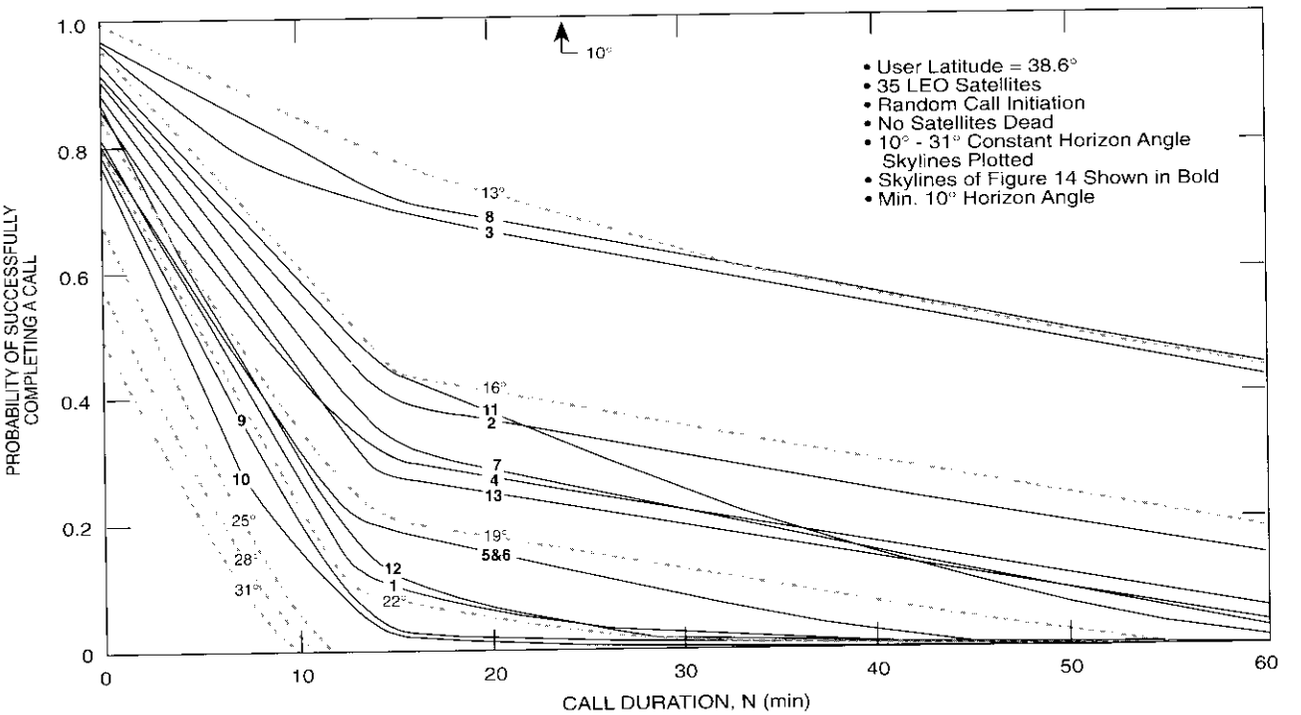
Examination of Figure 15 reveals that availability is highly sensitive to the specific shape of the skyline, and that there is a large variance among the examples. For example, for the 16° average skylines shown in Figure 15a, the availability performance for individual skylines ranges from being nearly equivalent to a uniform skyline of 13° (skyline 8) to being worse than a uniform skyline of 22° (skyline 10). Skylines with good visibility in the north-south direction (such as skylines 3 and 8) have the best availability characteristics. This is a consequence of the polar orbits of the satellite constellation used in this example.

Discussion of idealizing assumptions

Line of sight

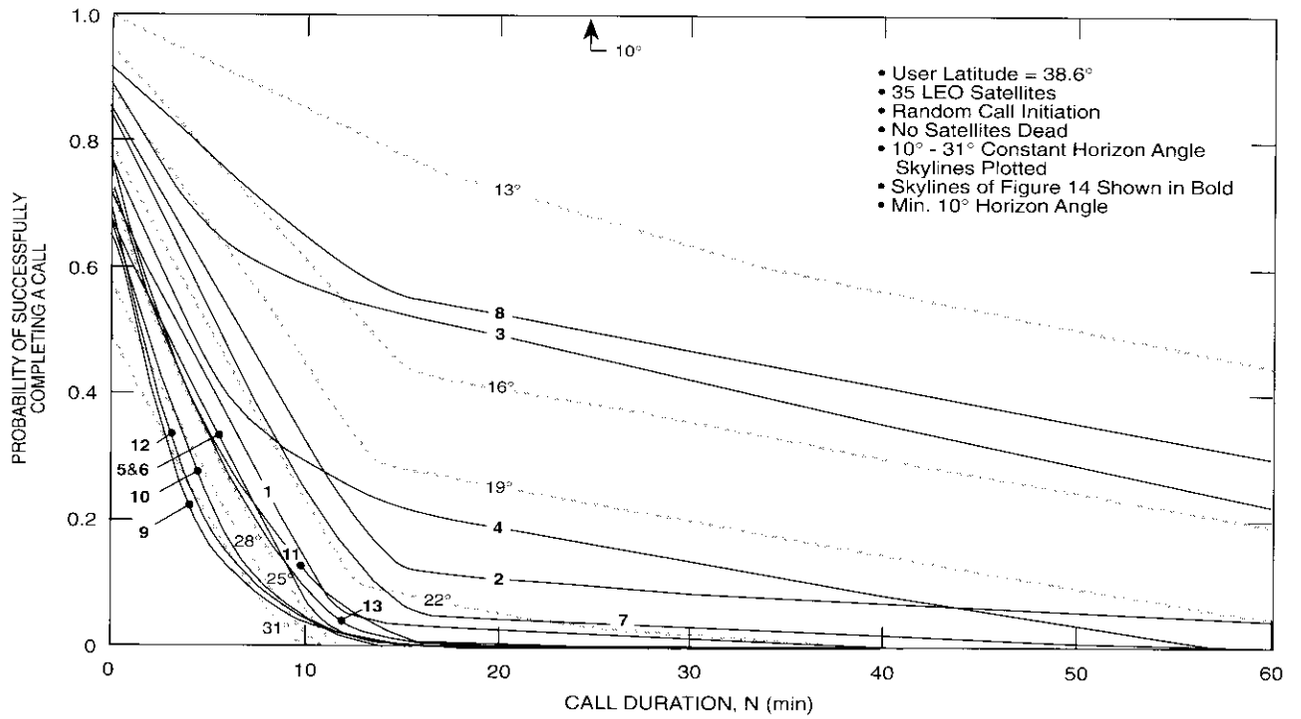
As noted previously, the examples given in this paper are idealized in that they assume that a link is broken if the user loses line-of-sight (optical) visibility to all satellites. Because blockages do not affect radio waves in the same way as light waves, and because RF links between the user and satellite will be designed with excess power margins, the applicability of these line-of-sight visibility assumptions must be examined.

Actual links at L-band differ from ideal line-of-sight paths due to effects such as diffraction, multipath, scattering, and attenuation caused by vegetation, buildings, and other obstacles. Therefore, for links with very large power margins, availability will be significantly better than the idealized examples shown here. Larger margins permit operation in shadowed regions because of diffraction, operation with scattered components, and operation with signals that are partially attenuated due to vegetation and losses. However, it can be argued that the line-of-sight model used here is appropriate for links having smaller power margins, and therefore that the availability estimates based on this assumption are reasonable for such cases. For reasons discussed below, an appropriate value for a "smaller margin" in this case is about 6 dB, although it is recognized that, in terms of conventional satellite links, 6 dB is not a small margin.



(a) Average Skyline Blockage Elevation Angle for the 13 Skylines of Figure 14 = 16°

Figure 15. Probability of Completing a Call of N Minutes Assuming Various Skyline Blockages or Constant Horizon Angle Skylines



(b) Average Skyline Blockage Elevation Angle for the 13 Skylines of Figure 14 = 25°

Figure 15. Probability of Completing a Call of N Minutes Assuming Various Skyline Blockages or Constant Horizon Angle Skylines (Cont'd)

Theoretically, diffraction from an ideal straight edge will produce an attenuation of 6 dB relative to an unblocked transmission when a source is positioned at the point where the straight edge intersects the line between the source and the receiver (*i.e.*, the point where optical blockage begins) [1],[2]. A recent experiment [3] reported a reduction of 5.5 dB in the level of a received L-band satellite beacon at the point where the edge of the roof of an office building intersected the line-of-sight path between the satellite and the receiver, where the elevation angle was 15°. Even though this is a single experimental result (in which the corner of the roof was rounded), it seems to confirm that the theoretical loss of 6 dB may also be a practical assumption at the elevation angle where the satellite is directly in line with the edge of an essentially opaque obstruction, such as a large building.

When the skyline consists of vegetation, the loss at elevation angles below the skyline depends on the nature of the vegetation (*i.e.*, the type of tree, distribution of tree and shrub sizes and heights, and the length of the vegetation path in the direction of the satellite). Although the loss due to vegetation is variable, measurements described in References 3 and 4 indicate that losses are substantial even for small amounts of vegetation. For example, data collected by Goldhirsh and Vogel [4] predict a loss of 24 dB at 1.6 GHz, at an elevation angle of 20°, for a pear tree canopy in full foliage, while data given in Reference 3 show losses ranging from 13 to 22 dB due to a tree line with an average height of 80 ft at an elevation angle of 15° and a frequency of 1.6 GHz. From these and similar data (summarized in Reference 4) it can be concluded that link margins will have to be substantial (on the order of 20 dB) if transmissions are to penetrate foliage. For smaller link margins (on the order of 6 dB), a skyline of vegetation will have approximately the same effect on availability as an opaque skyline of the same shape.

Measurements made inside small buildings indicate losses ranging from about 10 to 20 dB [5] at L-band. Hence, using the same reasoning as for vegetation, it can be concluded that, unless large margins are used, nearly all buildings will appear essentially opaque to satellite links with margins on the order of 6 dB.

Given sufficient link margin, communications can be maintained in shadowed regions via reflected and scattered components. This type of link (called a Rayleigh channel) is characteristic of terrestrial mobile links, where margins are well in excess of 20 dB (typically by tens of dBs) and line-of-sight user locations are rare.

Based on the above discussion, it can be concluded that the availability results given previously should be good approximations of the actual availabilities that can be expected for LEO/ICO links with link margins on the order of 6 dB, where this 6 dB is in excess of the margin allowances for head blockage (including the antenna distortion due to head proximity) and for multipath. Hence, if the total link margin is on the order of 6 dB, the results predicted here will be optimistic (*i.e.*, will correspond to skylines that are somewhat lower).

On the other hand, if the link margin is 20 dB or greater, the availability results presented here would be pessimistic. That is, availability will probably not improve substantially over the results given here until the link margins are strong enough to penetrate vegetation and buildings, or until they produce enough scattered and reflected components so that the link can operate under either Rician conditions (with low carrier-to-multipath ratios) or Rayleigh conditions.

Perfect handover

The second idealizing assumption anticipates perfect (*i.e.*, seamless) handover to another visible satellite whenever an established link to a satellite is broken due to shadowing. Highly reliable handover is essential for any LEO/ICO system, and it is assumed that the necessary techniques can be developed. Since these techniques will obviously require sufficient margin on the links to the satellite(s) to which the link will be switched when blockage occurs, it is further assumed that blockage considerations will be the primary determinant of availability, and that handover will not be sensitive to the user skyline. Thus, in order for a LEO/ICO system to be successful, reliable handover will be required in any user environment. Also, as the user's environment becomes increasingly shadowed and handover becomes relatively frequent, problems should not arise if handover is designed to be reliable, as is anticipated in the initial system design. (Handover algorithms must preclude excessively frequent handovers, which might, for example, be initiated by channel fading.) Hence, it is reasonable to conclude that the availability results given here are not significantly affected by the assumption of ideal handover. However, the results presented here are optimistic compared to systems designed to hand over calls at preset times, that is, those that do not perform handover when blockage occurs.

Conclusions

Four satellite systems that use non-geosynchronous orbits to provide handheld communications have been compared in terms of their service availability. In addition to demonstrating the analysis techniques used, the quantitative LEO/ICO comparisons have shown how availability for various environments

varies with the minimum elevation angle guaranteed by a constellation, as well as with orbital altitude. The examples given showed the importance of using a high minimum elevation angle. They also showed that an ICO system will generally have better availability than a LEO system, given the same minimum elevation angle, because ICO satellites tend to remain at higher elevation angles for a greater percentage of time. For many of the cases selected, availability (in terms of a caller being able to successfully complete a call) is not comparable with that usually associated with commercial service. Substantial link margins (on the order or 20 dB) may be required for some constellations (particularly LEO constellations) in order to bring availability up to a commercially acceptable level.

Acknowledgments

The work described in this paper was performed under the sponsorship of COMSAT Mobile Communications as part of an effort to evaluate candidate systems designed to support future handheld communications services. The authors wish to thank B. Hutchinson and C. Baxter for the assistance and suggestions they provided in formulating the relationships used in this study, and L. White and D. J. Shyy for providing an independent check of the azimuth and elevation angle vs *time relationships for certain constellations, which helped to validate the software developed by the authors.*

References

[1] H. R. Reed and C. M. Russell, *Ultra High Frequency Propagation*, Lexington, Massachusetts: Boston Technical Publishers, 1964, p. 121.

[2] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, New York: John Wiley & Sons, 1981, p. 462.

[3] R. M. Allnutt, Private communication, January 1993.

[4] J. Goldhirsh and W. Vogel, "Propagation Effects of Land Mobile Satellite Systems Overview of Experimental and Modeling Results," NASA Reference Publication 1274, February 1992.

[5] W. Vogel, G. Torrence, and N. Golshan, "Satellite Sound Broadcast Propagation Measurements and System Impairments," AIAA 14th International Communication Satellite Systems Conference,Washington, D.C., March 1992, *A Collection of Technical Papers*, Pt. 3, pp. 1461–1470. AIAA Paper No. 92-2002-CP.

Appendix. LEO/ICO availability analysis software

The following is a brief description of the software developed to provide the results presented in this paper. Included are the preparation of skyline data, time domain simulation of the user-satellite link, and processing of the resulting time domain simulation data to produce the desired outputs. The basic methodology used in the LEO/ICO availability analysis was illustrated in Figure 1.

Skyline computation

A skyline is defined by the minimum (unblocked) elevation angle in every azimuthal direction. This information is used in the time domain simulation of the user-satellite link to decide whether a specific satellite is blocked. Two different methods were used to produce the skyline data.

The first method assumes that all obstacles are oriented so that a line drawn from the user to the plane of the face of the obstacle is perpendicular, and that the obstacle's face has a rectangular outline. This allows an elevation angle calculation to be performed on the outline of each obstacle, and requires that the height, width, distance from the user, and azimuthal direction of each obstacle be specified, as well as the user's height. This type of skyline calculation was used for the computations depicted in Figures 12 and 13.

The second method of producing skyline data assumes right-angled, hexahedron-shaped obstacles (*i.e.*, boxlike objects with four sides, a top, and a bottom), and allows for a more practical orientation of the obstacles. Defining the skyline involves specifying the X and Y coordinates of each object relative to the user's position, the dimensions of each object (height, width, and depth), and the user's height. Then, while stepping around the user in azimuth, using 0.5° steps, the lowest unblocked elevation angle is calculated to define the skyline. The possibility of higher buildings being behind other buildings is also addressed in this process. This methodology was used to produce the skyline data of Figure 3 from the simulated "low-rise" urban intersection model illustrated in Figure 2. Although this method assumes that the obstacles are aligned with the X and Y-axes, the assumption is not restrictive because the entire scene can be rotated by any azimuthal angle and the user can be placed anywhere, except inside an obstacle.

Time domain simulation

The first step in the time domain simulation methodology is to place the user and satellites in their initial positions. The program employs a modified inertial spherical coordinate system with its origin at the center of the earth.

The angular coordinates are referred to the earth's latitude and longitude values at the beginning of the simulation (when $t = 0$), and the radial coordinate is specified by the distance from the origin. This coordinate system is well-suited to examining the dynamic situation of a stationary user on the earth, which is rotating within the spherically shaped "cage" produced by the LEO/ICO satellite orbit tracks. Next, the maximum slant range is calculated to eliminate "unseen" satellites (those that are blocked from view by the earth). Dead (inactive) satellites are also removed from consideration. The elevation and azimuth angles are then calculated for the remaining satellites, and this information, along with the elevation angle/azimuth angle skyline "mask" previously generated, is used to eliminate the blocked satellites. The remaining satellites (if any) are ranked according to their elevation angles. If the number of visible satellites is two or more, the "best" satellite is defined as the one with the highest elevation angle, or the one that remains visible for the longest period of time. Depending on the specified output, the pertinent data for the best satellite are either preprocessed and stored for further processing (when call duration/waiting time results are desired), stored without preprocessing (for the elevation angle probability calculations), or dumped to a file (for the time-plot outputs). An option also exists to select the "second best" satellite for availability analysis, if desired.

Preprocessing is performed in one of two ways: either for irregular skylines, as in Figures 4 and 5, or for skylines of constant elevation angle, as in Figure 7. For the irregular skyline models (with discrete obstacles and a minimum horizon angle of 5° or 10°), only one call duration histogram is required. This histogram represents the number of times the best satellite existed above the 5° or 10° minimum horizon angle for a certain duration. The histogram is composed of 4,320 bins corresponding to durations of from 10 s to 12 hr, in 10-s steps. For skylines of constant elevation angle, a separate call duration histogram is generated for each of eight constant elevation angle horizons. Each histogram then represents the number of times the best satellite existed above a different minimum horizon angle for a certain duration. The following discussion of the preprocessing methodology addresses both single- and multiple-histogram cases.

For the irregular skyline case where only one histogram is used, preprocessing compares the best satellite's elevation angle with its previous value to determine if the minimum elevation angle horizon has been crossed. If a positive transition occurs (*i.e.*, the previous elevation angle was less than the current elevation angle), a counter is started. If no transition occurs and the counter has been started previously, the counter is incremented. If a negative transition occurs (*i.e.*, the previous elevation angle was greater than the current elevation angle), the value of the counter is used to update the duration histogram, and the counter is cleared. Incomplete durations (with no transitions at the beginning and end of the duration) are logged as "lost time." Durations over a specified maximum (typically 12 hours) are logged as "overflow" durations. This accounts for elevation angle durations which outlast the entire 12-hr simulation.

For the case where multiple histograms are generated, each for a different value of constant elevation angle skyline, preprocessing compares the best satellite's elevation angle with its previous value to determine if one of the eight minimum elevation angle horizons has been crossed. If a positive transition has occurred, a counter is started for that minimum elevation angle horizon, and the counters for all lesser minimum elevation angle horizons are incremented. If no transition occurs, then all the previously started counters are incremented. If a negative transition occurs the value of the counter for that minimum elevation angle horizon is used to update its duration histogram, the counter is cleared, and the counters for all lesser minimum elevation angle horizons are incremented. Incomplete and overflow durations are handled in the same manner as for the single-histogram case.

The logged data for the elevation angle probability calculations are stored as a histogram representing the number of times the elevation angle of the best satellite was at certain values. Specifically, the elevation angle of the best satellite is used to select which of the 90 histogram bins (corresponding to elevation angles of from 1° to 90°) is to be incremented.

After the data for this time slot have been logged, the time is incremented by a specified value and the satellite positions and user's position are updated. Note that the satellite positions change due to their orbital motion, while the user's position changes due to the rotation of the earth. The entire satellite selection and data logging process repeats until the simulation time reaches a specified endpoint, which is typically 12 hours, except when the time-plot output option is desired, in which case the endpoint can be any value.

After the endpoint is reached, the simulation is terminated and the user's initial longitude is incremented by a specified value. If this new initial longitude does not exceed a specified overall range, then the entire simulation repeats from the beginning, while retaining the previously logged data. Combining the logged data from the many simulations representing different initial user longitude values produces the effect of simulating the satellite system over a very long time, and these data produce long-term statistics for call duration, waiting time, and elevation angle probabilities. Since the data for the time-plot output option are not processed for statistics, its simulation ends when the simulation time first reaches its endpoint. Once the longitude

has exceeded its specified overall range, the time domain simulation is complete and processing of the logged data into the desired outputs begins.

Output processing

The logged data are processed into the desired outputs by following one of two paths, depending on the output options selected. (The time-plot outputs require no processing.) The elevation angle probability output option calculates the probability density of the elevation angle based on the elevation angle histogram data collected during the time domain simulation. These data represent the number of times that the "best" satellite occurred at each elevation angle (from 1° to 90°, in 1° steps). Specifically, the histogram data are processed into a probability density by dividing the contents of each histogram bin by the total number of possible occurrences. The resultant probability density values are then summed to verify the calculations.

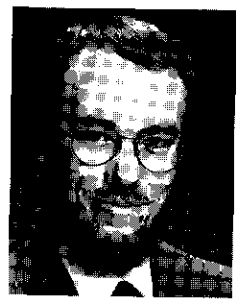
The second output option calculates call duration based on the call duration histogram data. These data result from preprocessing the elevation angle data, and represent the number of times that the "best" satellite existed above the 5° (or 10°) minimum elevation angle horizon for a certain duration. The processing methodology described below applies for single-histogram cases, which represent irregular skylines. For cases with multiple histograms (*i.e.*, skylines of constant elevation angles), the single-histogram processing is simply repeated for each histogram.

The call duration probabilities for randomly initiated calls are based on calculating the probability that a certain call duration could exist, given the histogram data for the entire simulation. Specifically, the probability of a certain call duration is calculated by summing the products of two probabilities for all measured durations. The first probability represents the likelihood that the specific call duration could exist within that measured duration (given a random starting position within that measured duration). For example, a call duration of X would have a 50-percent chance of a successful completion within a measured duration of 2X, and a 90-percent chance within a measured duration of 10X. The second probability represents the likelihood that the randomly initiated call would fall within that measured duration. This depends on the length of the measured duration, the number of occurrences of the measured duration, and the duration of the entire simulation. This calculation is repeated for all desired call duration values. The software also has the ability to produce call waiting time statistics that express the probability that a user must wait N minutes to establish a call, and call duration statistics when a call initiation is attempted only when at least one satellite is known to be visible.

Software implementation

The LEO/ICO Availability Analysis Program and the skyline calculation programs were written in FORTRAN and implemented on a Hewlett-Packard (HP) Series 700 workstation. The resulting skyline data (elevation angle *vs* azimuth angle) were plotted using Xmgr, a plotting package on the HP workstation. The resulting call duration/waiting time probability data were transferred to a Macintosh computer and plotted using Microsoft Excel. The time-plot and elevation probability data were plotted using Xmgr on the HP workstation. Quick-look plots for some of the results were made using Xgraph, an alternate plotting package on the HP workstation.

William A. Sandrin received a BSEE from the University of Vermont in 1963, and an MEE from Rensselaer Polytechnic Institute in 1966. From 1966 to 1971, he worked for the Boeing Company, Aerospace Group, where he was involved in space communications system design and in the development of phased-array antennas. In 1971, he joined COMSAT Laboratories, where he has been engaged in multibeam antenna studies; transponder linearizer development; transmission impairments analysis; studies of domestic, maritime, and aeronautical satellite systems; studies of adaptive phased-array antennas for spacecraft and small-terminal applications; frequency planning studies; and transmission software development. He is currently a Senior Scientist in the Communications Technology Division. Mr. Sandrin is a Senior Member of IEEE.



David V. Haschart received a BSEE from Ohio University in 1986 and a MSEE from the University of Toledo in 1991. He is currently a Member of the Technical Staff in the Communications Systems Studies and Design Department of the Communications Technology Division at COMSAT Laboratories.

Since joining COMSAT in 1991, Mr. Haschart has worked primarily in the areas of system studies and computer modeling. In the systems area, he has been involved with comprehensive transmission analysis/tradeoffs and

system architecture definition for various fixed and mobile communications scenarios (including a satellite-based digital-audio distribution system, a Ka/L-band hybrid satellite-based personal communications concept, an advanced satellite newsgathering system, and various orbital configurations of the Inmarsat-P system). In the area of computer modeling, he has been responsible for a number of computer simulations of non-geosynchronous satellite systems—examining various interference scenarios, the dynamics of system coverage, and other measures of system suitability.

Index: amplifiers, integrated circuits/MMICs, models, solid-state devices

*1-W Ku-band MMIC SSPAs for communications satellite phased-array antenna applications**

J. I. UPSHUR, R. K. GUPTA, G. C. ESTEP, AND R. T. KROLL

(Manuscript received May 2, 1994)

Abstract

The multiple-beam, Ku-band active phased-array antennas currently being developed for communications satellites require broadband, reproducible, and power-efficient solid-state power amplifiers (SSPAs). Since these amplifiers will operate in a multicarrier environment, they must be designed to meet a stringent linearity specification while functioning at high efficiency. The design and development of a fully monolithic 1-W SSPA, and the production of SSPA modules for a 64-element phased-array antenna application, are described. A summary of phased-array antenna systems considerations is included. Monolithic microwave integrated circuit (MMIC) design approaches and the thermal analysis of the SSPA module are also reviewed.

The measured RF performance of the SSPA shows 22.5-dB small-signal gain, output power of +30 dBm, and two-tone third-order intermodulation, C/I_3, greater than 16 dB. The power-added efficiency of the overall 1-W amplifier is greater than 25 percent, with a power chip efficiency greater than 28 percent. The measured RF performance of 80 SSPA modules is summarized in the form of histograms. The excellent performance uniformity obtained over all modules is the result of process-insensitive MMIC designs, chip selection procedures using on-wafer DC and RF probe data, and a repeatable module assembly approach.

*Substantial portions of this paper appeared in the Proceedings of the AIAA 15th International Communications Satellite Systems Conference, San Diego, California, February/March 1994, *A Collection of Technical Papers*, Pt. 1, pp. 316–322. AIAA Paper No. 94-0952-CP.

Introduction

Modern phased-array antennas for communications satellites require large numbers of solid-state power amplifiers (SSPAs). These components must be miniaturized and need to perform with high efficiency due to the limited DC power available on board the satellite. The amplifiers must also be linear in order to operate in a multicarrier communications environment. In addition, all modules in the array must operate with a high degree of gain and phase uniformity for good beam isolation.

The development of SSPA technology has seen continuing reductions in both the size and mass of these amplifiers. The hybrid microwave integrated circuit technology was based on discrete power field-effect transistors (FETs) fabricated on low-loss ceramic substrates. This allowed a reduction in physical size compared to packaged FET devices with Duroid substrates, although assembly required more complex steps and bonding operations. In recent years, work has focused on monolithic and quasi-monolithic implementations of the power amplifiers to achieve high reliability, uniformity, and low cost [1]. In the fully monolithic approach, the power FETs are integrated on the same gallium arsenide (GaAs) chip as the matching circuitry, allowing the greatest reduction in size and weight achieved thus far. This approach also offers enhanced reliability, with relatively few parts to be assembled into the overall module, making the technology suitable for phased-array antenna applications [2],[3].

The objective of the work reported herein was to develop and produce a fully monolithic SSPA module for use in Ku-band direct radiating array antennas. The modular design approach incorporates the necessary tradeoffs to simultaneously achieve good efficiency and linearity, while meeting gain and power requirements over the full 10.7- to 12.75-GHz band. Other requirements for monolithic microwave integrated circuit (MMIC) components include low sensitivity to process variations and a module design based on a repeatable assembly process. The three-stage SSPA is designed as a single-stage driver chip followed by a two-stage power chip. This approach maximizes overall fabrication yield while maintaining simplicity in SSPA assembly.

System considerations

Future high-capacity communications satellite systems will require phased-array antennas that can generate a number of narrow "pencil" beams. These beams may be dynamically reconfigured so that their shapes and/or locations, as well as their relative beam powers, can be adjusted as needed.

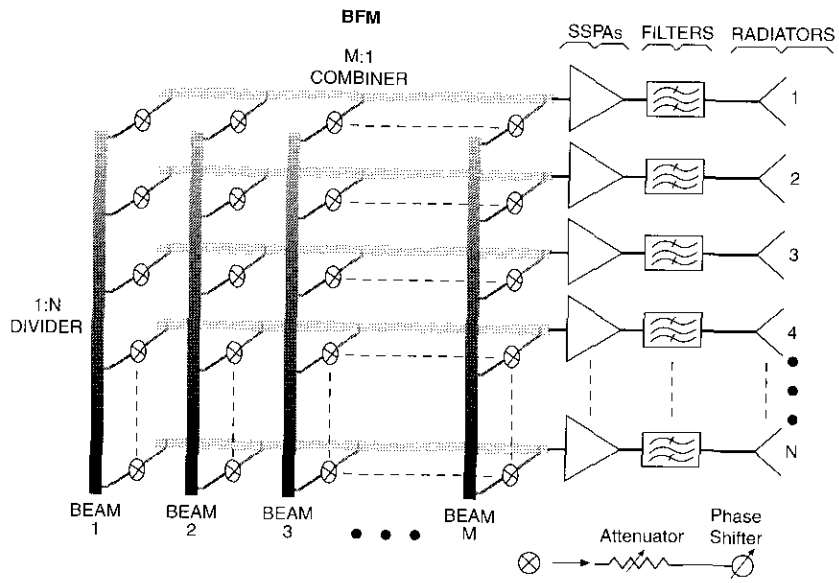


Figure 1. *Multibeam Active Antenna Block Diagram*

Figure 1 is a system block diagram of an M-beam antenna. The beam inputs are individually routed to N-way in-phase power dividers, where the N outputs have equal amplitude. Each output is connected to an attenuator and phase shifter, as shown in the diagram. MMIC components are typically used for the attenuation and phase shift functions to minimize size and mass, and to maintain uniformity among the $M \times N$ units. The outputs of the attenuator/phase shift blocks are then connected to a series of N in-phase, equal-amplitude M-way combiners. Each of the N combiner outputs feeds an SSPA, filter, and radiating element, as shown. The beam-routing section of the antenna system has an $M \times N$ matrix configuration and is referred to as the beam-forming matrix (BFM). A unique feature of these systems is the use of one SSPA for each radiating element, resulting in a total of N amplifiers in this example [4].

Other sections of the antenna system include an array of heat pipes located close to the SSPAs for efficient thermal dissipation, and a means of distributing bias and control voltages to the SSPAs. The beam-forming capability is provided by a digital control section, in conjunction with the phase shifters and attenuators. The signals feeding each radiating element must have amplitude and phase relationships within close tolerance for proper beam formation [5].

In addition to SSPA uniformity, the BFM must be designed for correct amplitude and phase response on all paths. These antenna systems have the potential to dramatically increase satellite communications capacity and flexibility while maintaining the low mass and size necessary to minimize payload launch cost [6].

The power-added efficiency (PAE) of the SSPA becomes a critical performance parameter due to the large number of radiating elements required and the fact that the SSPA is the primary consumer of DC power on board the spacecraft. Figure 2 shows the relationship between the total DC power requirement for the SSPAs of an active array and the PAE of each amplifier. It is clear that high SSPA efficiency is necessary, especially in large arrays, because of the limited DC power available.

An added benefit of high-efficiency operation is that the junction temperature of the power FETs will be lower for a given RF output power. This improves device reliability and increases design lifetime for the communications payload. The 1-W Ku-band SSPAs were developed for a 64-element array, which will require roughly 250 W of DC power.

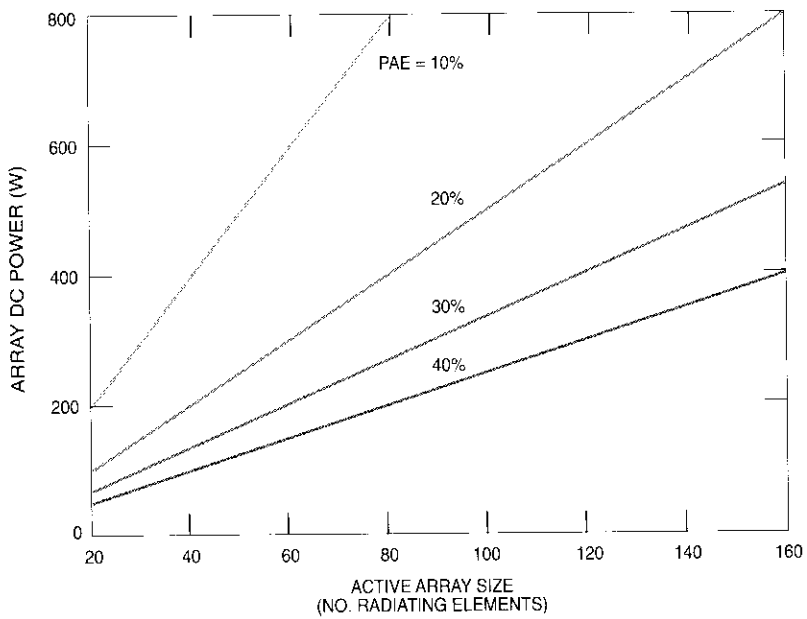


Figure 2. *Impact of SSPA Efficiency on Antenna DC Power Requirement*

SSPA MMIC designs

To obtain an overall gain of 22.5-dB, the SSPA was designed with three stages partitioned as a single-stage driver chip followed by a two-stage power amplifier chip. The size of the output-stage FET was chosen based on the output power requirement. Because the individual stages must satisfy different linearity, efficiency, and gain requirements, their matching techniques, modes of operation, and device widths are different. In the first stage, high gain and high two-tone third-order intermodulation, C/I_3, are more important than high efficiency. In the middle stage, high C/I_3 and high efficiency are equally important, while in the output stage, high efficiency becomes more important than high C/I_3.

In the design phase of this work, the performance of the individual stages was examined to determine the impact on SSPA efficiency of losses incurred in lumped-element matching networks. Where losses were too high, distributed-element matching approaches were used, despite their generally larger size.

The other important factor addressed in choosing the matching network implementation was circuit sensitivity to process variations, especially for the phased-array application where a large number of SSPA units must maintain phase and amplitude performance within specified windows. For circuit uniformity, all shunt capacitor tuning elements in both the driver and power amplifiers are realized as open-circuited transmission lines, although a metal-insulator-metal (MIM) capacitor would require less area. In both circuits, all MIM shunt capacitors use relatively large (8-pF) values for RF bypass, where capacitance variations up to ±20 percent have a negligible impact on circuit performance. This exceeds the allowable capacitance tolerance in the fabrication process. The series DC-blocking capacitors were also chosen to have a similar low sensitivity to fabrication variations.

Driver chip

The driver chip employs both lumped- and distributed-element matching networks, since the effects of loss are negligible in terms of the overall performance of the SSPA. To obtain optimum gain from the driver, the input and output matching circuits provide a conjugate match to the FET. As shown in the driver schematic of Figure 3, the input matching network employs shunt inductance near the FET gate. The series resistance placed in the upper inductive element improves the input match and gain flatness over the operating band. The shunt inductance is used to resonate the input capacitive reactance of the FET, thus forming a real impedance value which is transformed to 50 Ω by the remaining series inductance and shunt capacitance. The output matching

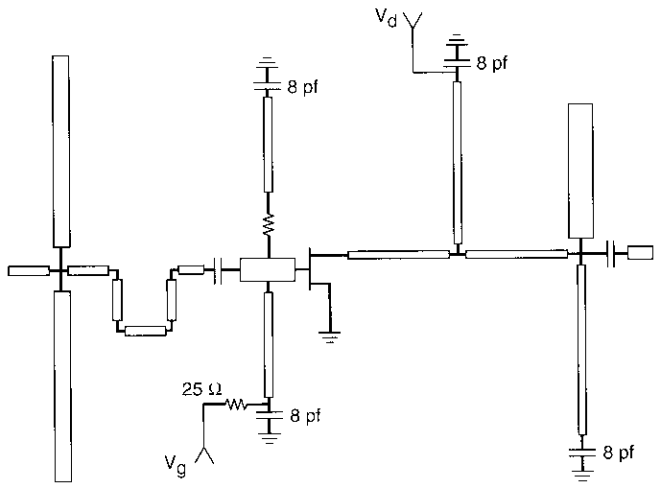


Figure 3. *Driver Amplifier Schematic*

circuit incorporates an inductive T-network with a shunt L-C section to achieve a broadband match between the FET drain and the 50-Ω system impedance.

The driver FET has eight gate fingers, each 120 μm wide, for a total gate periphery of 0.96 mm. The gate length is 0.4 μm, instead of the 0.5-μm used in the power chip. The shorter gate increases gain by about 1 dB, with a negligible effect on overall SSPA efficiency. As a result of this gate length reduction, the driver FET has 15-percent lower normalized gate-source capacitance and 8-percent higher normalized transconductance. The driver FET size produces input and output impedance levels sufficiently high to enable relatively simple matching networks to provide a good match across the band of interest. The FET device width was chosen so that the output power density at 5- to 6-dB backoff in the conjugately matched condition would be sufficient to drive the power chip. In this mode of operation, the driver will contribute only minimally to overall nonlinearity.

Biasing for the gate and drain is done via the shunt stubs, using 8-pf RF bypass capacitors. The gate bias section includes a 25-Ω resistor in the RF decoupling network to improve broadband circuit stability. The resistor is located adjacent to the RF bypass capacitor (Figure 3) so that in-band gain performance is not affected. The driver amplifier chip dimensions are 1.35 × 2.60 mm on 90-μm-thick gallium arsenide (GaAs). Figure 4 is a photograph of the chip showing the on-wafer probe pads at the input and output. An example of on-wafer measured driver amplifier gain is given in Figure 5. The gain

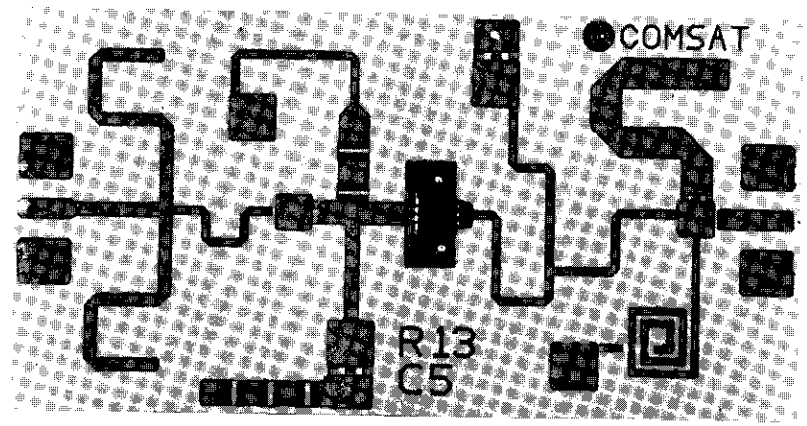


Figure 4. *Ku-Band Driver Chip*

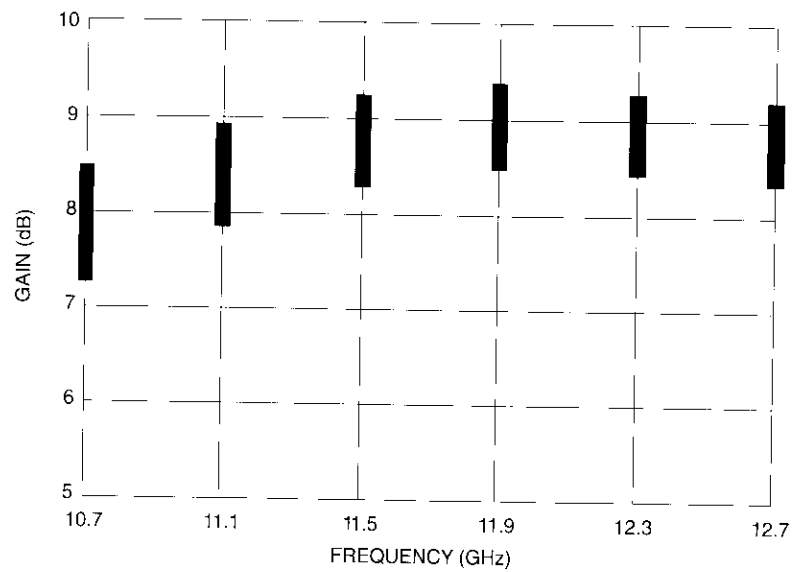


Figure 5. *Gain Distribution of MMIC Drivers From On-Wafer Probe Data*

uniformity for the chips from this wafer is within ±0.5 dB over most of the band. Figure 6 compares the modeled and measured gain response of a single driver chip using +4.5-V drain bias. The correlation over the 10.7- to 12.75-GHz band is excellent.

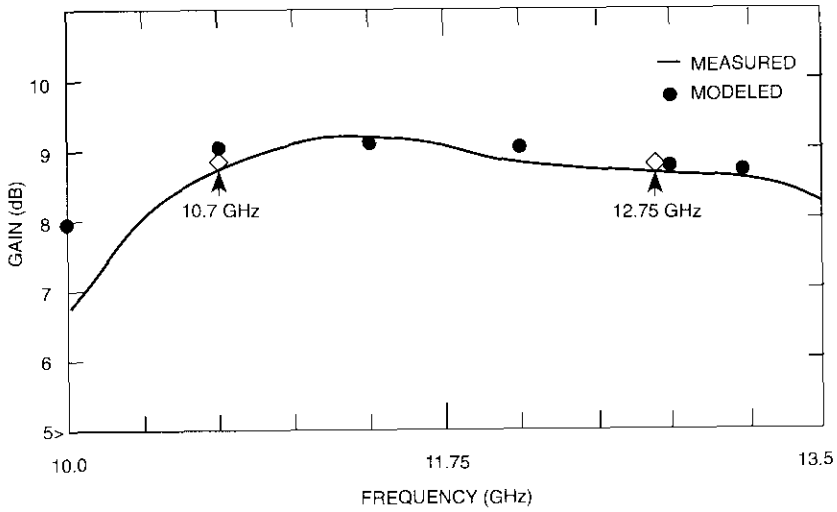


Figure 6. *Modeled vs Measured Gain of Driver Chip*

Power amplifier chip

The power amplifier chip uses distributed-element matching in the form of printed metallization to minimize sensitivity to process variations and reduce circuit losses. The two-stage configuration of the power chip is shown in Figure 7. The first stage has a significant effect on overall SSPA efficiency, since it consumes more DC power than the driver stage. This stage was designed to operate efficiently and at a reasonable C/I_3 level, since it is followed by the additional gain of the output stage. To obtain the desired efficiency and C/I_3, the first stage is operated in Class-AB mode and under 2- to 3-dB output backoff.

The input matching network has shunt inductive elements similar to those of the driver chip. The resulting real impedance is transformed to 50 Ω by the *k*-inverter section, consisting of two series inductors (L) with a shunt capacitor (C) at the midpoint. This structure simulates a quarter-wave transformer but requires less chip area. The shunt L-C section at the input, and the shunt resistors at the FET gate, improve the bandwidth of the input match.

While the input of the second-stage device is conjugately matched for high gain, the output of the FET is matched for high power and efficiency. The efficiency of the second stage has the greatest impact on overall SSPA efficiency. The output stage is operated in Class-AB mode, with the output-matching circuit designed to obtain maximum power and efficiency from the

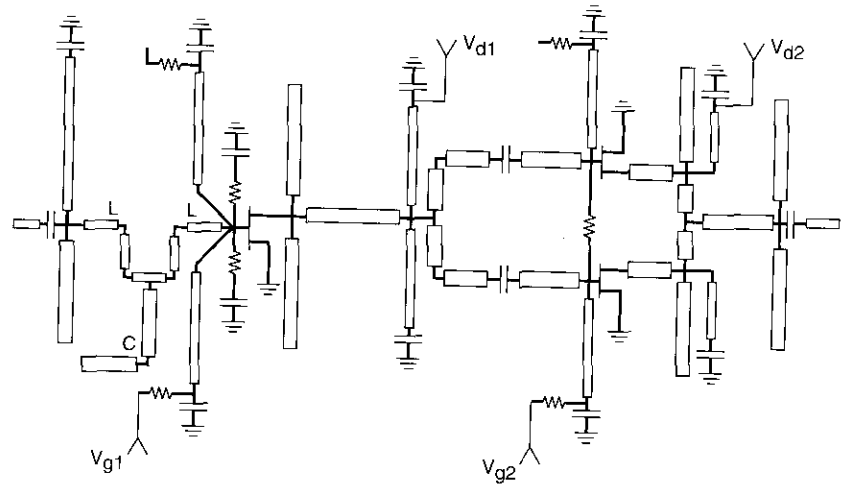


Figure 7. *Power Amplifier Schematic*

FET. The optimum load impedances for large-signal operation were determined from experimental load-pull measurements on test devices having the same gate length and a known gate width. The load impedance values for the first and second stages were then calculated by scaling the measured value based on the actual device gate widths used in the amplifier. The resulting optimum load impedance for the second-stage device is plotted in Figure 8 over the 10.7- to 12.75-GHz band. Also shown is the modeled load impedance presented to the second-stage device by the output matching network, which was calculated using the computer model of the MMIC. The modeled load impedance is within about 10 percent of the required optimum load impedance over the operating band.

The FET of the second stage is partitioned as two cells, which are partially matched and then combined at a point in the interstage matching network to achieve the required bandwidth. The remainder of the interstage network completes the transformation to the optimum load impedance for the first-stage device. Similar techniques are used in the output matching network to transform the 50-Ω system impedance to the optimum load for the output device.

Both stages contain harmonic tuning elements at the FET drains to improve the PAE. A broadband, low-impedance reactive termination at the second harmonic band is achieved with an open-circuited stub one-eighth-wavelength long at the fundamental frequency. This becomes a quarter-wave stub at the

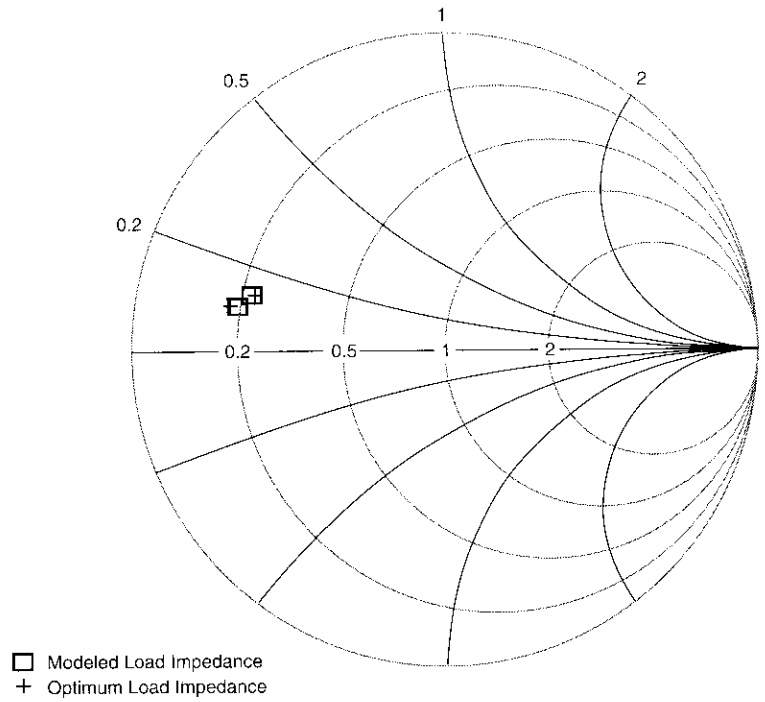


Figure 8. *Modeled Output Network Impedance vs Optimum Load Impedance for Output Device*

harmonic frequency, which transforms the open circuit to low impedance at the drain side of each FET. The capacitive reactance created at the fundamental frequency is incorporated into the matching networks as a tuning element. The resistor placed between the gate terminals of the two FET cells of the second stage improves the odd-mode stability of the circuit. Figure 9 is a photograph of the power MMIC. As in the driver chip, the on-wafer probe pads are visible at the input and output bonding areas. The power chip dimensions are 2.55 × 3.65 mm on 90-μm-thick GaAs.

The on-wafer probe data for the MMIC chips of a power amplifier wafer are shown in Figure 10. Gain variation slopes from approximately 15 dB at 10.7 GHz to 11 dB at 12.75 GHz. The majority of chips produce gain within a ±1-dB window. The observed gain slope is a consequence of the tuning pads designed into the power chip, which could not be connected into the circuit at the time of on-wafer probing. These pads can be seen in Figure 9, primarily in

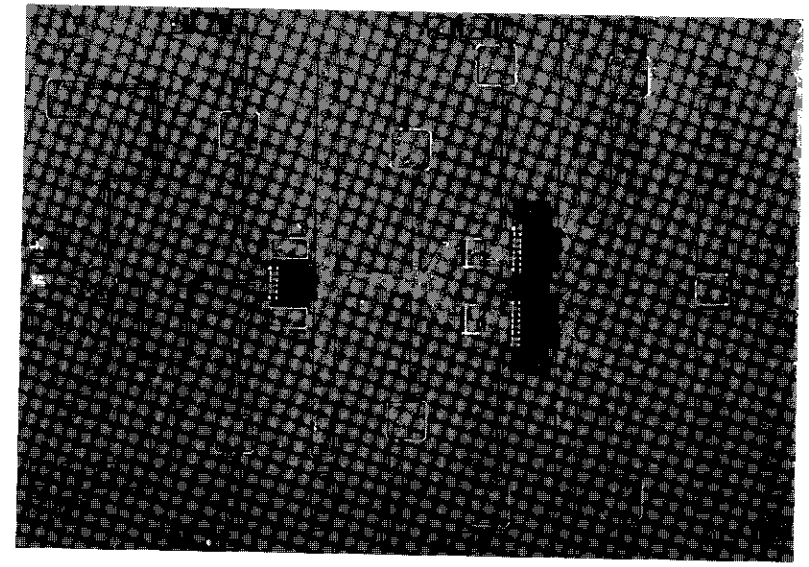


Figure 9. *Ku-Band Power Amplifier Chip*

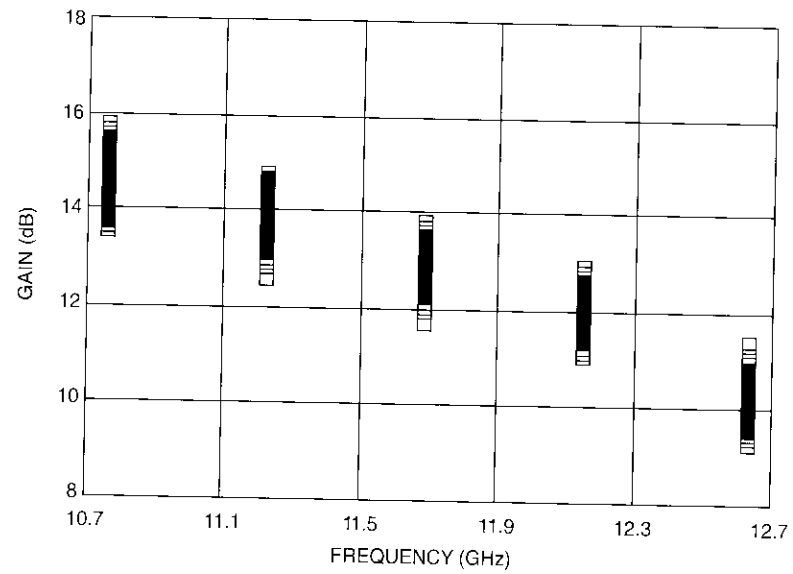


Figure 10. *On-Wafer Gain Measurements of Power Chips*

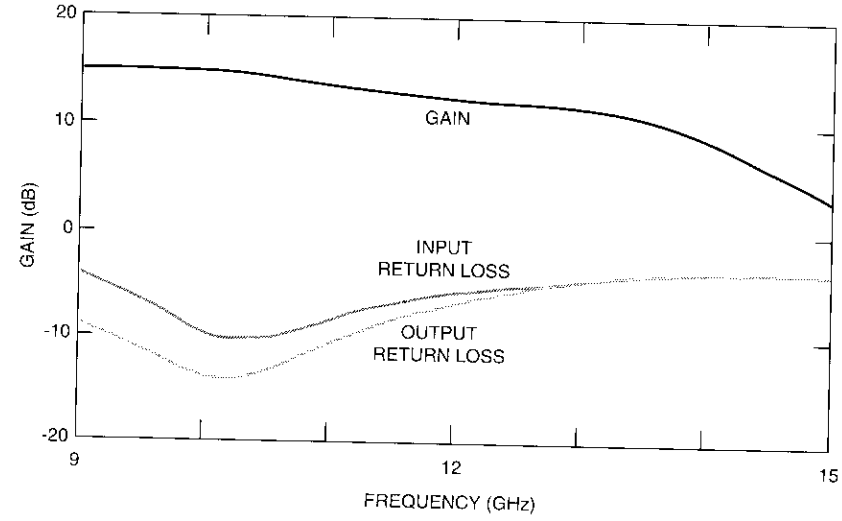
the input network and to a lesser extent in the interstage and output networks. The tuning pads provide a means of compensating for uncertainties in the FET model and the optimum load condition for power operation. As a first check of the on-wafer probe results, the computer model for the MMIC chip was analyzed under the assumption that no tuning pads were connected. The predicted result is shown in Figure 11a as the simulated power chip performance (without tuning). The on-wafer measured gain for a representative MMIC chip in the untuned state (Figure 11b) demonstrates a similar gain variation relative to the predicted curve. When the discrete chips were measured after wafer dicing, the small-signal RF results responded to the tuning conditions as predicted by the computer model. Following tuning pad connection, the designed small-signal gain response was obtained, as shown in Figure 12. The repeatability of the circuit tuning procedure made it possible to use the on-wafer gain data to accurately select the power MMIC before module assembly.

Circuit fabrication

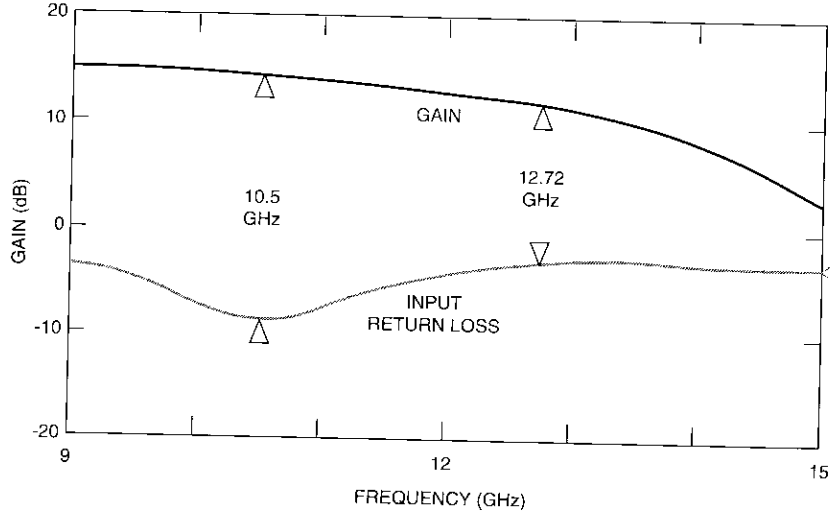
The MMIC fabrication method used to produce the Ku-band circuit is characterized by mesa isolation and a self-aligned gate process that incorporates direct-write electron-beam (e-beam) lithography to form submicrometer gate structures. A molecular beam epitaxy (MBE) machine was used to grow the active layers. Both the driver and power chips are fabricated on 90-μm-thick material. The FET devices of both chips use through-substrate via-holes to provide low-impedance source grounding and additional heat sinking for the FET. The devices have an active-layer peak doping concentration of 2×10^{17} cm^{-3}, and a unit gate width of 125 μm [7].

The mesa pattern, which includes active areas for FETs and resistors, and alignment marks for the ohmic mask, is exposed and developed by a photolithographic process. Unwanted active layer material is removed with an etchant, resulting in mesas with gentle slopes to ensure the continuity of metal layers that cross the mesa edge. This process step provides electrical isolation between the various active devices of the circuit. Ohmic contact metallization is patterned by a liftoff process, and then alloyed by rapid thermal annealing. A separate baseplate mask is used to form the MIM capacitors. The baseplate metallization is also patterned by liftoff, and the wafer is then coated with plasma-enhanced chemical vapor deposition (PECVD) silicon nitride.

Gate-level patterns are defined by direct-write e-beam lithography, and the gate recess is etched through the openings in the e-beam resist. The gate metallization is then deposited by e-beam evaporation, and gate passivation is performed after liftoff with PECVD silicon nitride.



(a) Predicted



(b) Measured On-Wafer

Figure 11. *Power Chip Performance (without tuning)*

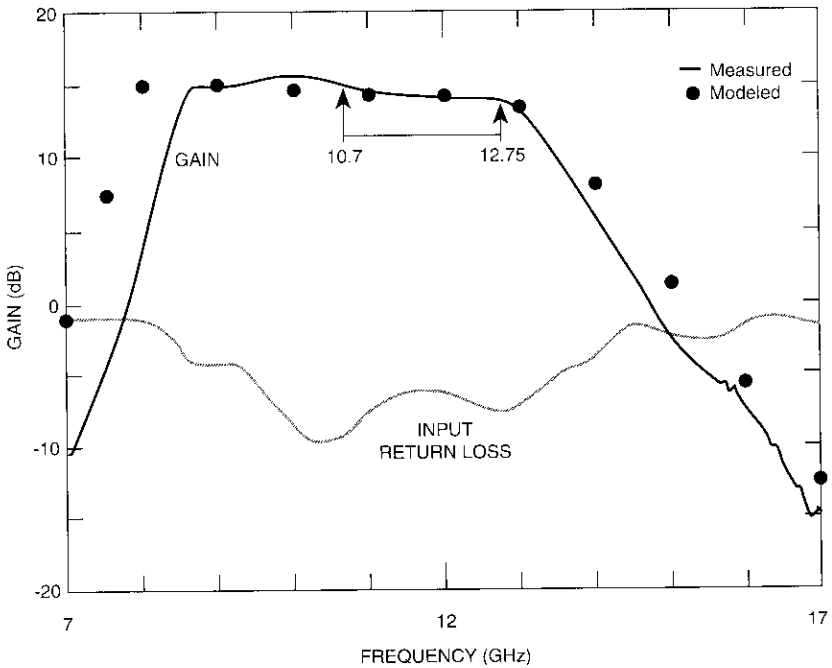


Figure 12. *Modeled vs Measured Gain of Power Chip*

SSPA module design and assembly

Figure 13 is a photograph of the fully assembled SSPA module, which consists of a copper-tungsten carrier to which the driver and power chips are attached with 80/20 gold-tin solder. RF input/output lines on alumina, and the gate and drain bias boards, are mounted to the carrier using conductive epoxy. Both bias boards were fabricated on 0.375-mm-thick alumina.

The drain bias board contains patterned gold bias lines to transfer three independent drain biases to the MMIC chips. The board has a shunt capacitor on each line for additional RF decoupling. During testing, as in the active phased-array application, the drain bias voltages were fixed at +4.5, +7.5, and +7.5 V for the three stages, respectively.

The gate bias circuit operates from a single -5 V supply and has three emitter-followers and variable resistors to supply each FET gate with an independently controllable, constant voltage source. In addition to the chip selection procedures described below, the gate voltage adjustment capability allowed the gain level to be equalized among all modules during the final RF test.

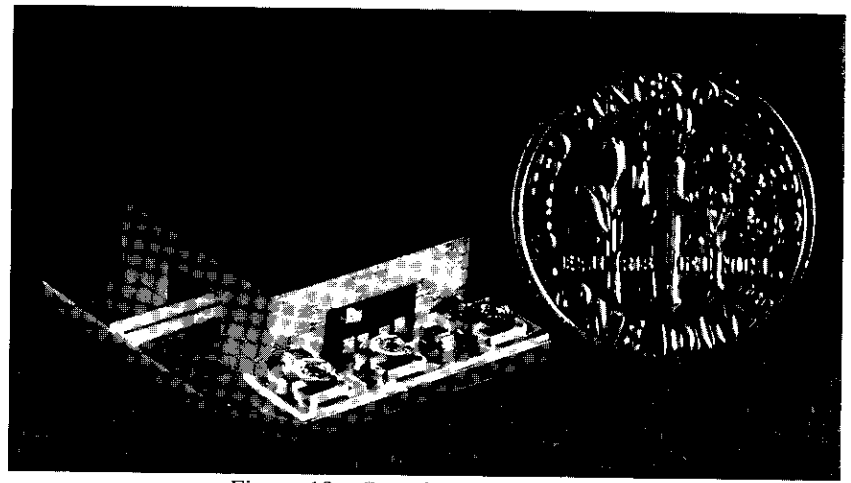


Figure 13. *Complete SSPA Module*

The metal lid provides physical protection for the MMICs and forms an RF enclosure to minimize radiation effects. The interior dimensions of the lid cavity are selected such that the cutoff frequency is well above the band of operation, thus guarding against potential oscillation problems when the lid is in place. The overall dimensions of the SSPA module are approximately $2.5 \times 2.5 \times 0.63$ cm.

A multinode finite-element thermal model of the MMIC chips and copper-tungsten carrier was used to determine the thermal resistance between the FET channel area and the bottom of the carrier. The GaAs chip was divided into six layers, each 15-μm thick, for more accurate calculation of thermal spreading resistance. The model also takes into account the reduction of GaAs thermal conductivity with increasing temperature. Assuming a baseplate temperature of 60°C, the maximum predicted FET channel temperature is 107°C—well below the typical reliability upper limit of 140°C. This thermal model conservatively neglects the additional heat conduction provided by the FET via-holes.

As described previously, the driver and power MMIC chips contain provisions for on-wafer RF testing. As a final step before chip dicing, the circuits were evaluated for small-signal gain and return loss behavior over the 10.70- to 12.75-GHz band. A typical on-wafer measurement procedure created a database of chip S-parameters at five frequencies uniformly spaced across the full band. Results for each chip were stored with the chip row and column identifier as a reference for unique correlation. In addition to the RF measurements, each amplifier FET was measured for DC characteristics, including

saturation current density, transconductance, pinch-off voltage, and breakdown voltage. These results were also stored in a database, with a row and column identifier for each chip.

The DC and RF data were used in a detailed chip selection procedure prior to module assembly, to ensure optimum uniformity and RF yield. The primary performance indicators for the driver chip were RF gain and return loss. The DC FET measurements were available as a secondary indicator; however, the gain level and flatness of the driver are more critical in determining overall module performance compared to power and efficiency. This information is available from the RF small-signal measurement.

Power chip selection relies on both DC and RF on-wafer data. The RF data are used as an initial screen for small-signal performance. The DC data (especially the saturation current density and breakdown voltage) indicate the large-signal performance of the chip.

Measured SSPA RF performance

Test fixtures were designed and fabricated to allow complete RF testing of the SSPA modules. A computer-controlled test facility was implemented for efficient testing of the units. Basic small-signal testing included frequency-swept gain and input/output return loss measurements. These were performed at narrowband (10.5 to 13 GHz), with the option of broadband (0.05- to 20-GHz) measurement. Another setup under the same computer control characterized SSPA output power, PAE, and C/I_3 as a function of input power. Transfer phase variation as a function of input drive was measured with a vector network analyzer. The measurements were generally performed at 25°C, although measurements at 50°C were conducted on a sample basis. Noise-power ratio measurements were also performed on the sample units.

The small-signal swept gain of the complete three-stage module is shown in Figure 14. The gain is nominally 22.5 dB over the band of interest, with approximately ±0.5-dB gain variation. At 50°C, the small-signal gain decreases by approximately 2 dB.

Figure 15 is a plot of the SSPA output power *vs* input drive at the five standard measurement frequencies across the band. The module achieves greater than 31-dBm saturated output power, and about 30 dBm at 2-dB gain compression. The somewhat lower power capability at frequency f_2 has been attributed to a test-fixture-related discontinuity and its mismatching effect on the output of the SSPA at this frequency. A similar plot of efficiency *vs* input drive (Figure 16) shows peak efficiencies of approximately 30 percent, and 22 to 26 percent at 2-dB compression.

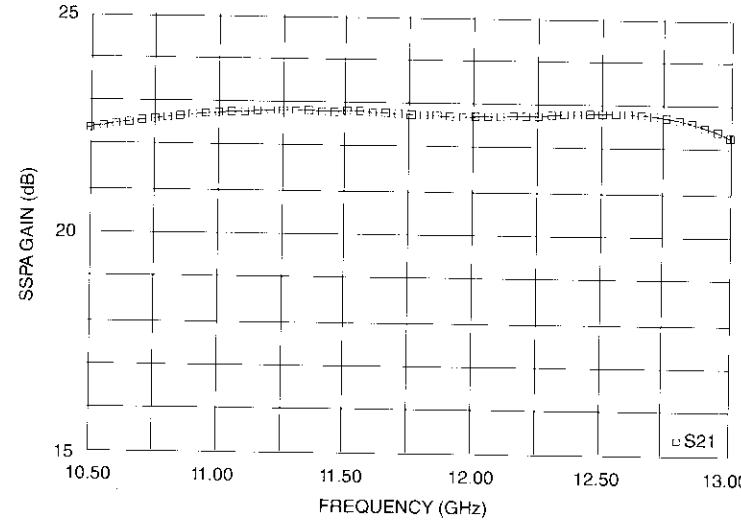


Figure 14. *Swept Small-Signal Gain of SSPA Unit*

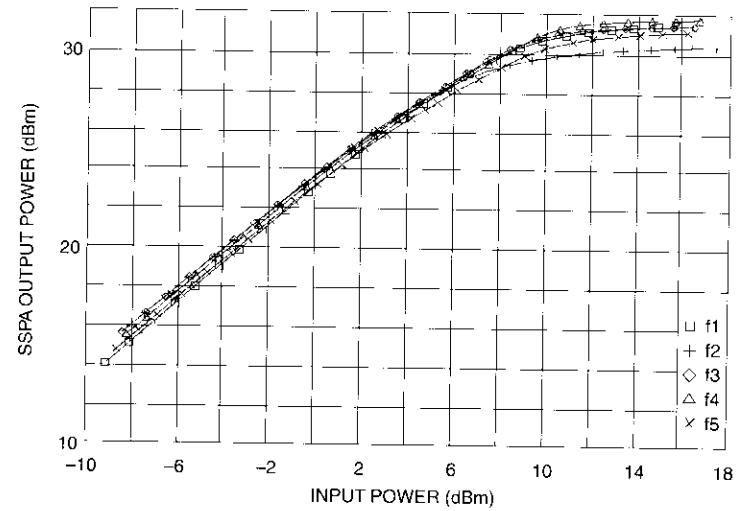


Figure 15. *Output Power vs Input Power of SSPA*

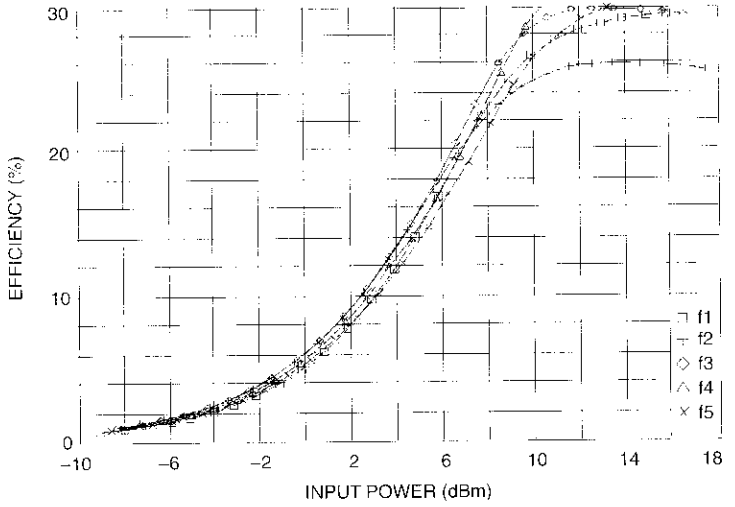


Figure 16. *SSPA PAE vs Input Drive*

The C/I_3 measurements are shown in Figure 17. Frequency separation is 10 MHz. The results were generally better than 16 dB at 2-dB compression, except at frequency f_2 (which is again consistent with the test fixture discontinuity effect noted previously).

In the active transmit phased-array antenna application, each radiating element of the antenna is driven by a transmit module partitioned as the SSPA section, the electronic power conditioner (EPC) section, and the output filter. The SSPA section is located adjacent to the heat pipes for optimum thermal dissipation. The SSPA, in cascade with the pre-driver stages, produces greater than 65-dB overall small-signal gain.

The excellent uniformity of RF performance attained is summarized in a composite swept gain plot of more than 75 units (Figure 18). A total of 80 modules were produced, including 64 for the actual antenna array, plus a number of spare units. The histogram in Figure 19 represents the small-signal gain at band edges and center for all 80 modules. Approximately 75 of the modules have gain within a 2-dB window across the full frequency band. Similarly, the output power averages 30 dBm, with almost all units lying between 29.5 and 30.5 dBm (Figure 20). All output powers here are referenced to an input power of +10 dBm. The C/I_3 performance is generally better than 16 dB at the 2-dB gain compression point (Figure 21).

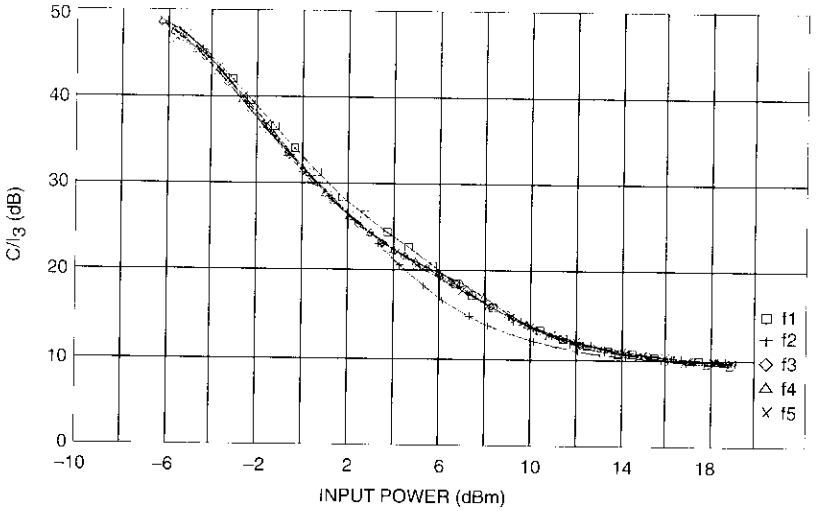


Figure 17. *SSPA C/I_3 vs Input Drive*

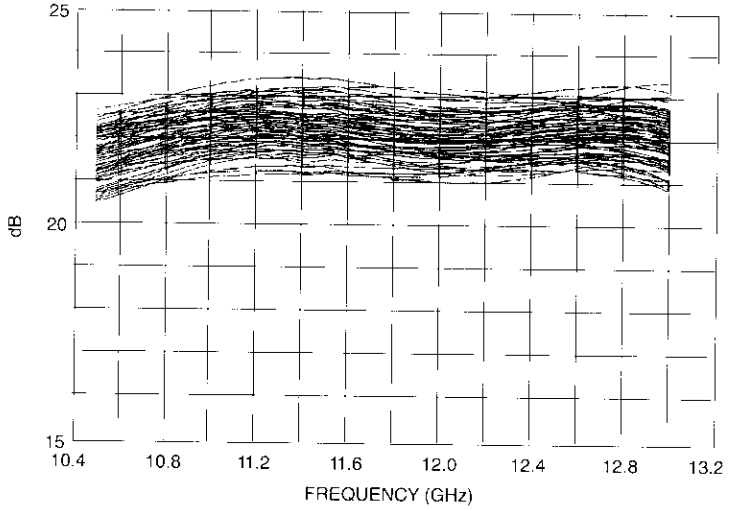


Figure 18. *Composite Plot of Small-Signal Gain for More Than 75 SSPAs*

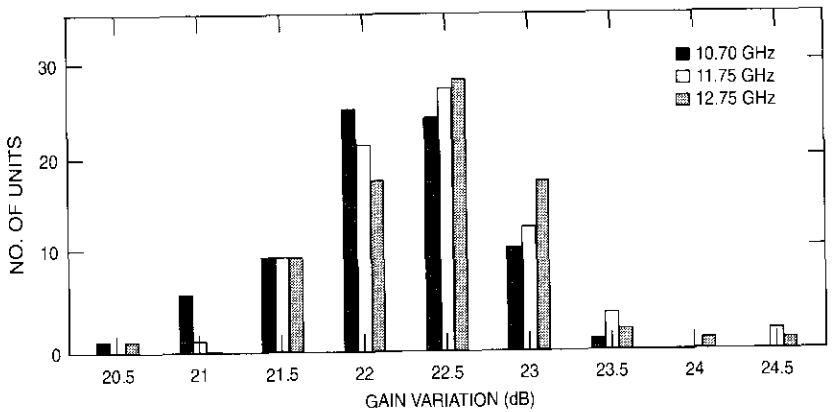


Figure 19. *Gain Distribution of 80 SSPA Modules*

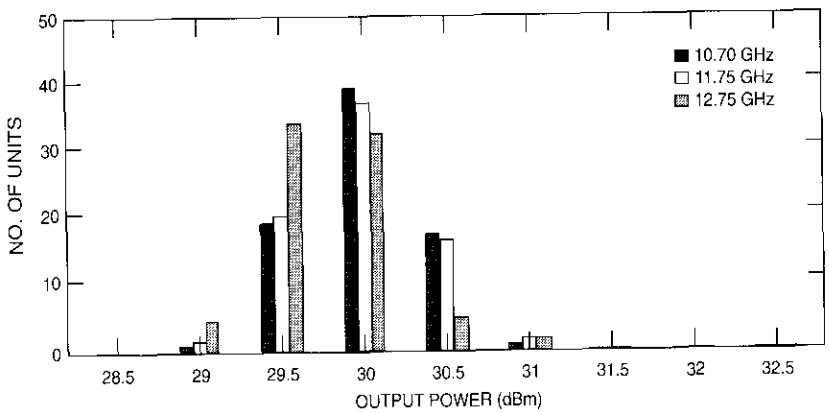


Figure 20. *Output Power Distribution of 80 SSPA Modules*

Conclusions

A fully monolithic 1-W Ku-band SSPA module consisting of a single-stage driver amplifier and a two-stage power amplifier has been developed for a communications satellite phased-array antenna. This work has verified that GaAs MMIC technology can be used to produce 1-W SSPAs for phased-array applications. Efficient assembly and RF performance uniformity were demonstrated for 80 production units. In general, the small- and large-signal performance of the modules was consistent with the designs. Assembly of gate- and drain-bias supplies on the MMIC chip carriers simplified module assembly and testing. In the future, similar design and production methods should be extendable to the development of flight-qualified SSPAs for space communications applications.

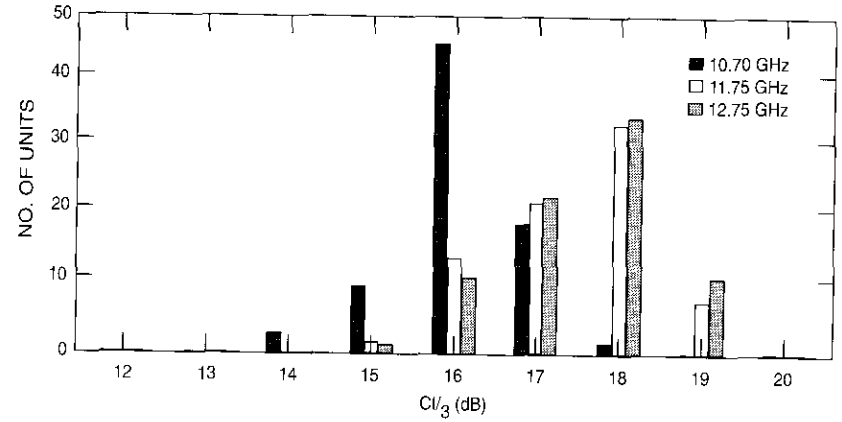


Figure 21. *C/I_3 Uniformity for 80 SSPA Modules*

Acknowledgments

The authors acknowledge the support of D. Roques and B. Cogo of Alcatel Espace in the Ku-band SSPA design and production. They wish to thank F. Phelleps and A. Cornfeld for MMIC fabrication, and J. Tyler, T. Morgan, and S. Haynes for assembling the SSPA modules. They are also grateful to F. Assal and J. Evans for technical and management support of this project.

References

[1] B. D. Geller *et al.*, "Production of Miniaturized Multistage SSPAs for a Ku-Band Array Antenna," AIAA 14th International Communication Satellite Systems Conference, Washington, DC, March 1992, *A Collection of Technical Papers*, Pt. 1, pp. 659–664. AIAA Paper No. 92-1897-CP.

[2] R. Gupta and F. Assal, "Transponder RF Technologies Using MMICs for Communications Satellites," AIAA 14th International Communication Satellite Systems Conference, Washington, DC, March 1992, *A Collection of Technical Papers*, Pt. 1, pp. 446–455. AIAA Paper No. 92-1859-CP.

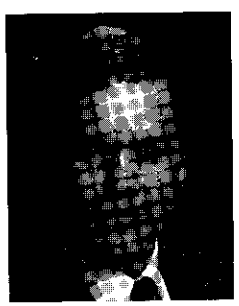
[3] J. R. Potukuchi *et al.*, "MMIC Insertion in a Ku-Band Active Phased Array for Communications Satellites," IEEE MTT-S International Microwave Symposium, Dallas, Texas, May 1990, *Digest*, Vol. 2, pp. 881–884.

[4] A. I. Zaghloul *et al.*, "Active Phased-Array Development at C- and Ku-Band," AIAA 15th International Communications Satellite Systems Conference, San Diego, CA, February/March 1994, *A Collection of Technical Papers*, Pt. 3. AIAA Paper No. 94-1069-CP.

[5] A. I. Zaghloul, "Statistical Analysis of EIRP Degradation in Antenna Arrays," *IEEE Transactions on Antennas and Propagation*, Vol. AP-33, No. 2, February 1985, pp. 217–221.

[6] R. M. Sorbello, "Advanced Satellite Antenna Developments for the 1990s," AIAA 12th International Communication Satellite Systems Conference, Arlington, Virginia, March 1988, *A Collection of Technical Papers*, pp. 652–659. AIAA Paper No. 88-0873.

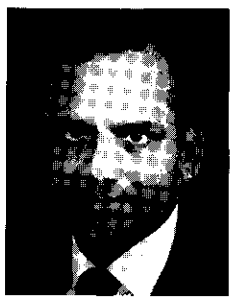
[7] J. I. Upshur *et al.*, "A Fully Monolithic Broadband Power Amplifier for Ku-Band Communications Satellite Applications," 35th Midwest Symposium on Circuits and Systems, Washington, DC, August 1992, *Proc.*, pp. 1020–1023.



John I. Upshur received a BS in physics from Davidson College in 1976, and MS and PhD degrees in electrical engineering from the University of Virginia, in 1978 and 1983, respectively. Prior to joining COMSAT in 1985, he was with M/A-COM Advanced Semiconductor Operations, where he was involved with gallium arsenide (GaAs) MMIC design and circuit fabrication. He developed fabrication processes for 44-GHz MMIC PIN-diode phase shifters and millimeter-wave mixer diodes.

Since joining COMSAT in 1985, Dr. Upshur has been involved in a variety of MIC and MMIC circuit design and measurement programs. He has worked extensively with development of 2.5-W MMIC power amplifiers, as well as 5- and 10-W SSPA modules. Tasks in other areas have included the design and characterization of Ku-band HEMT-based LNAs, MIC and MMIC X-band phase shifter circuits, and MMIC dielectric resonator oscillators for down-converter applications.

Dr. Upshur is currently a Senior Member of the Technical Staff in the Microwave Components Department of the Satellite and Systems Technologies Division. His current work focuses on the design and testing of X- and Ku-band high-efficiency SSPAs for space communications applications. He served as Vice-Chairman (1993-1994) and Chairman (1994-1995) of the Washington, D. C./Northern Virginia Chapter of the IEEE Microwave Theory and Techniques (MTT) Society. He is a member of Sigma Xi, Tau Beta Pi, and IEEE.

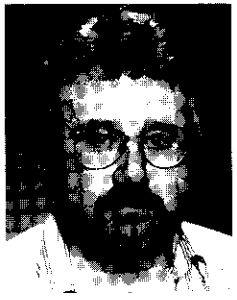


Ramesh K. Gupta received a BSc (honors) in electronics and communications engineering from Punjab University, Chandigarh, India, in 1974, and an MSc and PhD in electrical engineering from the University of Alberta, Edmonton, Canada, in 1976 and 1980, respectively. He also received an MBA from the Wharton School of Business, University of Pennsylvania, in 1989. He was Punjab University Merit Scholar (1970–74) and was awarded the Alberta Government Telephones Centennial Fellowship (1976–79) for graduate research in telecommunications.

In 1980, Dr. Gupta joined COMSAT Laboratories, where he is a Department Manager in the Satellite and Systems Technology Division. He has contributed to the application of hybrid MIC and GaAs MMIC technology in advanced satellite subsystems, and has been responsible for development of wideband microwave switch matrix (MSM) arrays, components for a 120-Mbit/s CQPSK onboard modem, and active beam-forming networks for single- and multibeam Ku-band phased-array antennas using MIC and MMIC technologies. He has also contributed to the spacecraft and ground network study and support effort for the INTELSAT VI satellite program.

Dr. Gupta has served as Vice-Chairman (1987–88) and Chairman (1988–89) of the Washington, DC/Northern Virginia Chapter of the IEEE Microwave Theory and Techniques (MTT) Society. He has authored/coauthored more than 45 papers on solid-state devices and circuits, GaAs MMICs, advanced microwave subsystems, and satellite systems. He holds two patents and was a co-recipient of the ICDSC-8 Best Paper Award for his contributions to a paper documenting the development and testing of the SSTDMA subsystem for INTELSAT VI. He is a Senior Member of IEEE.

George C. Estep received a BSEE from Virginia Tech in 1986. He joined COMSAT in 1988 and is currently a Senior Member of the Technical Staff in the Microwave Components Department, where he is responsible for hardware and software design and development for various satellite communications applications. These include a multibeam beam-forming matrix for spaceborne phased-array antennas, MMIC designs, frequency synthesizers, and microprocessor and analog controller designs.



Robert Kroll received an AA in electronics technology from Montgomery College in 1976, and a BS in electronics technology from Capitol Institute of Technology in 1979. Prior to joining COMSAT, he was a customer engineer with IBM, where he maintained various mainframe subsystems, including 3330 and 3333 hard disk drives, 2540 reader/punch, and 3420 tape drives.

Since joining COMSAT in 1979, Mr. Kroll has been extensively involved in the RF characterization of microwave components using TRL, SOL, and time-domain calibration techniques. These components have included MMIC attenuators, phase shifters, passive networks, and SSPAs. He has worked on CAD layouts of MICs, and on the design and fabrication of MIC and MMIC test fixtures. He has also designed specialized software code for commanding and controlling an earth station terminal, and assisted in logic design and assembly of the IOTT Test Unit Command Generator and Telemetry Receiver.

Mr. Kroll is currently an assistant staff member in the Microwave Components Department of the Satellite and Systems Technologies Division. His primary focus is on the development and maintenance of automated microwave measurements systems and the analysis of measured data for archival and presentation graphics.

Index: monitoring and control, key management, frequency planning, resource allocation, orbit prediction

Interim system planning computer

W. L. COOK

(Manuscript received February 22, 1994)

Abstract

The Interim System Planning Computer (ISPC) will be used to plan and manage networks operating on military communications satellites. The ISPC architecture supports a hierarchical division of responsibilities between the system manager and the managers of the individual networks that make up the system. The system- and network-level functions may be contained within a single processor or distributed among processors located at various sites. The basic ISPC planning and operations functions include scenario definition, resource allocation, ephemeris propagation, key management, and system monitoring.

The ISPC operates on engineering workstations with dual-screen displays, and the design incorporates a user-friendly graphical user interface and a distributed relational database. To support the distributed nature of the ISPC architecture, ISPCs can communicate with each other over both local area networks and wide area networks, via X.25 connections, or via satellite control circuits using a link-level protocol designed to ensure high data throughput.

Introduction

The Interim System Planning Computer (ISPC) is a distributed computer-based facility that will be used to plan, monitor, and reconfigure satellite communications networks operating within the Universal Modem System (UMS). The UMS will provide fixed-ground, transportable-ground, airborne, and ship users with survivable antijam, anti-scintillation, low-probability-of-exploitation interoperable digital data communications. The system will use

nonprocessing transponders on U.S., British, French, and North Atlantic Treaty Organization (NATO) satellites, including DSCS II, DSCS III, SKYNET 4, SYRACUSE, NATO III, and NATO IV.

The mission of the ISPC is to provide scenario definition, resource allocation, ephemeris propagation, key management, system monitoring, data dissemination, handover control, computer operations, and general support functions at both the system and network levels. This paper describes the system architecture of the ISPC, with emphasis on the graphical user interface (GUI), relational database, communications functions, and the system management functions it performs.

ISPC overview

The ISPC supports hierarchical management functions operating at both the system and network levels. Each ISPC can operate in one or more of the following modes:

- *System Operations Controller (SOC) Mode.* In this mode, the ISPC manages a system consisting of up to 32 networks, operating on as many as five satellites. The SOC plans system scenarios, allocates satellite resources to networks, plans interoperable circuits and system control circuits, performs ephemeris propagation, manages cryptographic keys, monitors overall system status, and directs system-wide reconfigurations.
- *Network Operations Controller (NOC) Mode.* In the NOC mode, the ISPC manages a network of up to 128 universal modems (UMS). The NOC plans network scenarios, allocates power and bandwidth to user circuits, manages keys for intranetwork use, generates and distributes network implementation plans (NIPS) and NIP updates, monitors circuit and terminal performance, and directs network reconfigurations.
- *Alternate SOC (ASOC) Mode.* In this mode, the ISPC serves as a hot standby for the SOC. The ASOC monitors overall system status and maintains a system database concurrent with that of the SOC in order to assume the SOC role in the event of a handover.
- *Alternate NOC (ANOC) Mode.* In the ANOC mode, the ISPC serves as a hot standby for the NOC. The ANOC monitors circuit and terminal performance and maintains a network database concurrent with that of the NOC in order to assume the NOC role in the event of a handover.

The following functions are performed by all ISPCs, regardless of mode:

- Maintain database concurrence.
- Manage planned or unplanned handover of control functions to other ISPCs to ensure UMS survival.
- Manage ISPC startup, recovery after brief outages, and shutdown.
- Provide general support, including database management, operator interface, alarm and message handling, and logging and retrieval.
- Emulate the operator interface units for collocated UMs.

The ISPC source code is written in the Ada programming language. The software development process conforms to the documentation, configuration management, quality assurance, and testing requirements of DOD-STD-2167A.

The ISPC hardware consists of a Digital Equipment Corporation (DEC) VAXstation 4000 Model 60 with dual screens controlled from a single keyboard and mouse, running the VMS operating system. The ISPC takes full advantage of advanced computer hardware and software technologies, including interactive GUIs, distributed relational databases, and computer communications.

ISPC operator interface

The ISPC provides a user-friendly graphical operator interface which generates displays, processes operator inputs, and provides for efficient operator control of all ISPC functions. It satisfies stringent human engineering performance requirements and is compatible with the skill levels of the ISPC operators. The design employs techniques and features known to improve operator comprehension and minimize operator overload.

The operator interface is based on the X Window system—a network-transparent windowing system that has become an industry standard. Use of this technology permits applications to operate across networks in a heterogeneous and geographically distributed environment. OSF/Motif, a tool kit offered by the Open Software Foundation, provides the operator interface "look and feel." The AXIS Toolkit, a COMSAT-proprietary user interface development environment, permits software developers to generate the operator interface interactively, thus minimizing the software development effort.

The ISPC operator interface provides full support for dual screens at each operator workstation. The software allows the operator to select the screen on which individual application windows will appear, or to use a default screen layout. If one screen fails, the operator can continue to work on the remaining

screen. The ISPC operator interface screen hierarchy is presented in an Appendix to this paper.

ISPC database

The data necessary to define a UMS configuration are stored in a relational database, which is maintained in a consistent state at all ISPC sites. The database contains both scenario-dependent and scenario-independent data. A scenario defines the specific system or network resources required in order to respond to a specific threat environment. The ISPC stores up to 10 system scenarios and 10 network scenarios for each network supported. An operator may create "official" scenarios, which are automatically distributed to all other ISPCs, as well as "unofficial" scenarios for "what-if" planning.

The ISPC database is controlled by DEC's Rdb/VMS relational database management system. Rdb/VMS uses the Structured Query Language (SQL), which is the industry-standard relational database language.

Data specified at the system level include the locations of all ISPCs and UMs, as well as satellite configurations, terminal parameters, network membership, control circuit and interoperable (*i.e.*, internetwork) circuit descriptions, ephemeris parameters, TRANSEC and COMSEC keys, and the power and bandwidth allocations to each network. Network-level data include modem parameters, group membership, intranetwork circuit descriptions, and circuit power allocations.

The ISPC automatically disseminates system scenario-independent data and official system scenario data specified at the SOC to all other ISPCs in the system. It also disseminates network scenario-independent data and official network scenario data specified at the NOC to all ISPCs that serve as ANOCs. Automatic dissemination of these data ensures database concurrency throughout the system, which facilitates handover of control in the event of ISPC failure. The operator may also manually transfer selected unofficial data to other ISPCs.

Inter-ISPC communications

The SOC operator has complete flexibility in specifying the architecture of the UMS—in particular, the geographical distribution of ISPCs and UMs, the roles assumed by each ISPC, and the communications links between them. Figure 1 shows the ISPC operational environment, which consists of a number of ISPCs connected via local area networks (LANs) to collocated UMs acting as either network control terminals (NCTs) or network terminals (NTs).

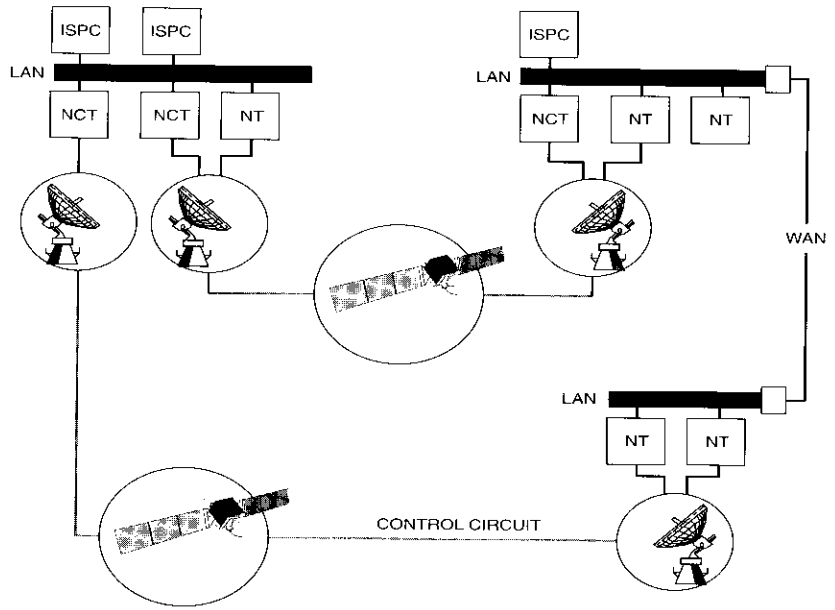


Figure 1. *ISPC Operational Environment*

Interconnection between sites is provided by wide area networks (WANs), packet-switched networks (using X.25), or satellite control circuits established through the UMs.

The ISPC software consists of multiple application processes that use the services of the LAN message router (LMR) to send messages to UMs and to applications in other ISPCs. The LMR employs the LAN message protocol (LAMP) format for all messages transmitted over the LAN. For communications with other ISPCs, it uses the satellite communications protocol (SCP), a selective-repeat transport-level protocol that ensures efficient transmission over satellite links characterized by high bit error ratios (BERs) and/or long transmission delays. Reliable delivery and flow control of packets between the ISPC and its collocated UMs is provided by the telecommunications protocol/internetworking protocol (TCP/IP).

Figure 2 shows the protocol layers used for ISPC-to-ISPC communications. The UM accepts packets from the ISPC and sends them over an appropriate satellite control circuit to a remote UM, which forwards them to the destination ISPC. If an ISPC needs to send messages over a satellite control circuit provided

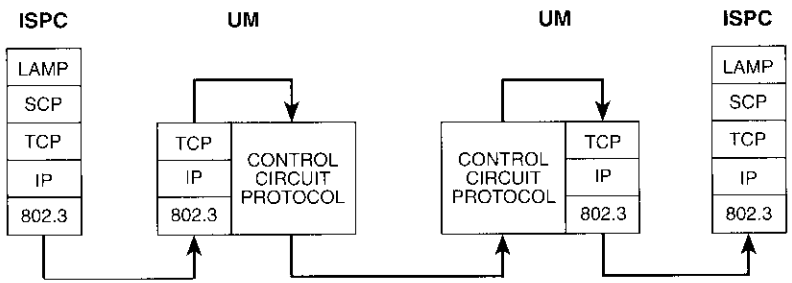


Figure 2. *Protocol Usage*

by a UM to which it is not directly connected, it sends those messages first to an ISPC which is connected to that UM, which will perform the appropriate routing.

Key features of the communications software are the dynamic routing protocol and the SCP protocol, which are described below.

Dynamic routing protocol

The ISPC employs a dynamic routing protocol to transfer messages between ISPCs using the LAN and control circuits. This protocol automatically selects the shortest available path to each destination ISPC. When one or more circuits fail, the routing algorithm automatically switches to alternate predefined routes, without operator intervention. The operator is notified if no alternate route is available.

This protocol is based on the routing information protocol (RIP) [1], which has become a standard for the exchange of routing information between gateways and hosts. The ISPC routing protocol, UMS RIP, uses many of the concepts and algorithms of the standard RIP, but differs from the standard RIP in that it:

- Performs routing based on ISPC IDs rather than Internet addresses.
- Uses a unique message format.
- Does not support the routing table request message.
- Always sends complete routing table information, never partial updates.
- Does not delete a destination from the routing table when it becomes unreachable, rather the destination remains with a distance of infinity.
- Supports multiple physical links between two ISPCs.

Like the standard RIP, the UMS RIP uses a distance vector algorithm to determine the total "distance" from a source to a destination. The ISPC associates a distance with each link—with LAN connections having the lowest values of distance, and satellite control circuits having higher values. The total distance is the sum of all distances from the source to the destination.

The general UMS RIP algorithm can be summarized as follows. The algorithm maintains a table of routing information with an entry for each destination ISPC. Each entry contains the total distance to the destination and an indication of the physical link of choice. The algorithm also maintains a table of physical links to neighboring ISPCs. The status of each link is altered based on connection updates and the presence or absence of periodic routing status information on the link. Changes in link status may result in the selection of an alternate physical link to the same ISPC, and/or the sending of route status messages to neighbors.

Route status messages containing the total distance to each ISPC are sent periodically to every neighbor. When routing status information is received from a neighbor, the distance associated with the link on which the status was received is added to each distance in the status table. The resulting distances are compared with current routing table entries and, if a new distance is less, the entry is updated to adopt the new route. Each routing table entry that specified a route through the sending neighbor is updated to reflect the new distance, even if the distance is greater. If the routing information was changed, routing status information is sent to neighbors.

Satellite communications protocol

The SCP [2] is a highly reliable host-to-host protocol for use between ISPCs connected via the LAN or satellite control circuits. It provides very high throughput performance, even under stressed conditions, and can operate efficiently over satellite links with delays varying from 270 ms to 60 s and link rates from 300 to 4,800 bits/s (and beyond) without reconfiguration for each specific link. The SCP employs a selective retransmission strategy, as opposed to the go-back-*N* strategy commonly used by other protocols. The protocol performs the following functions:

- *Connection Management.* The SCP provides multiple independent logical connections between any pair of ISPCs. Multiple SCP connections may share a satellite control circuit.
- *Data Transfer Functions.* Data are delivered reliably and in sequence to the destination. Packets with errors are recovered by retransmissions.

An alarm is generated for messages that cannot be delivered after repeated attempts.

- *Segmenting and Reassembly.* Messages larger than 128 octets are fragmented into multiple packets to fit into 128-octet UMS frames (the last frame may be 32 octets) and reassembled into messages prior to delivery to the destination application.
- *Priority.* Messages are associated with one of three priority levels: low, medium, or high. Between any two ISPCs, a separate SCP connection will be set up for each priority level, with the higher level connection having priority access to a satellite control circuit.
- *Flow Control.* The SCP provides flow control between a source and destination so that the source will send data no faster than the destination can receive.

Figure 3 shows the throughput performance of the SCP, based on simulation results. The figure also depicts the performance of an ideal go-back-N (nonselective retransmission) type protocol and an ideal selective retransmission protocol with infinite receive buffer size. The latter curves are based on analytical results and represent an upper bound on performance for the respective class of protocols.

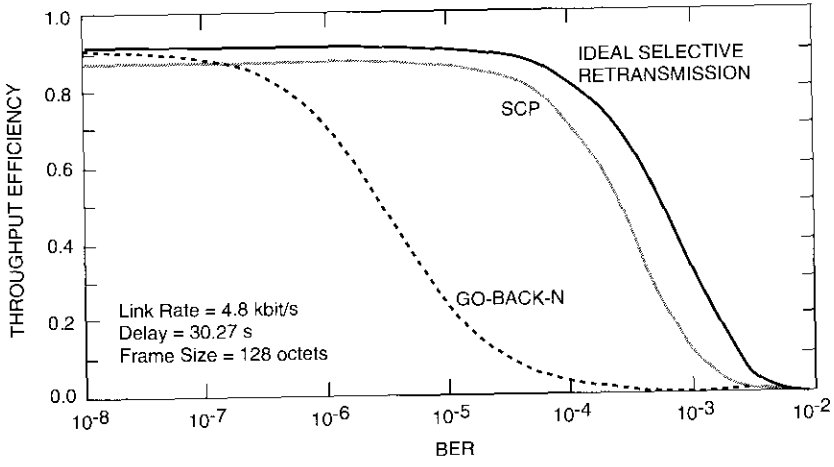


Figure 3. *SCP Performance*

Planning

Planning for the UMS system is performed at both the system and network levels, as shown in Figure 4. It consists of three steps: scenario definition, resource allocation, and NIP generation.

Scenario definition

The ISPC operator plans system and network configurations (*i.e.*, scenarios) to accommodate anticipated threat and environmental conditions. At the system level, the SOC operator defines the ISPCs, networks, modems, satellites, terminal types, and terminals in the system. The operator may also define up to 10 system scenarios, each associated with a specific mission requirement or anticipated threat. For each system scenario, the SOC operator identifies the satellites used, jamming threat, transponder connectivities, and antenna pointing angles, as well as the interoperable (between networks) circuits and control circuits used in that scenario. Figure 5 shows a typical dialog for entering system-level data, in this case terminal types.

At the network level, the NOC operator specifies the modem operational parameters, interface equipment, and modem groups pertaining to the operator's own specific network (or networks). The operator may also define up to 10 network scenarios, each based on a specific system scenario. For each network scenario, the NOC operator identifies the group memberships, modem devolution rank, and alarm thresholds, as well as the intranetwork circuits used in that scenario.

Resource allocation

Using the resource allocation function, the ISPC operator can compute the terminal and satellite resources required to satisfy each scenario. The responsibility for allocating resources is shared between the SOC and NOC, with the SOC allocating resources to networks, and the NOC assigning power to individual circuits within the network.

Based on a high-level definition of communications requirements, the SOC generates an estimate of the total resources (transponder power and bandwidth) required to support each network. The SOC then checks these values against satellite and terminal parameters and operational constraints and forwards them to the NOCs.

Using these inputs and the definition of required connectivity for a network configuration, the NOC determines the power level required for each circuit and the predicted performance and margin of each link. This analysis accounts

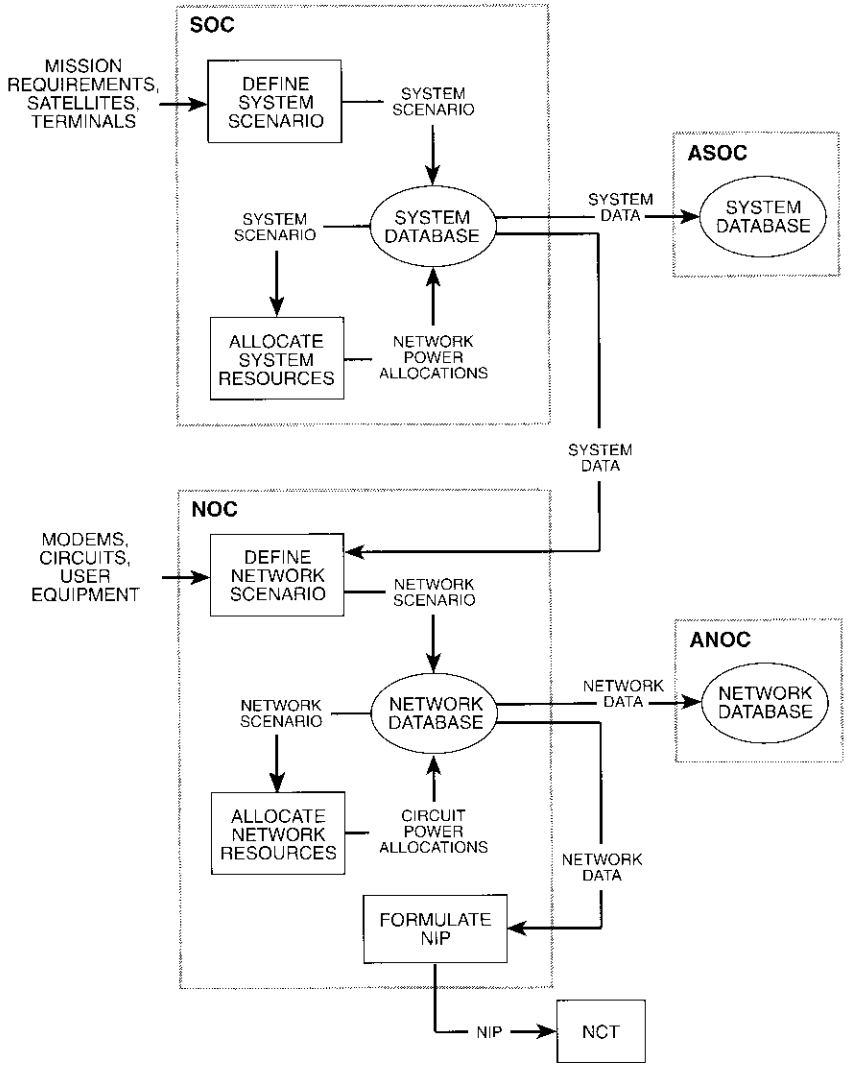


Figure 4. *ISPC Planning Function*

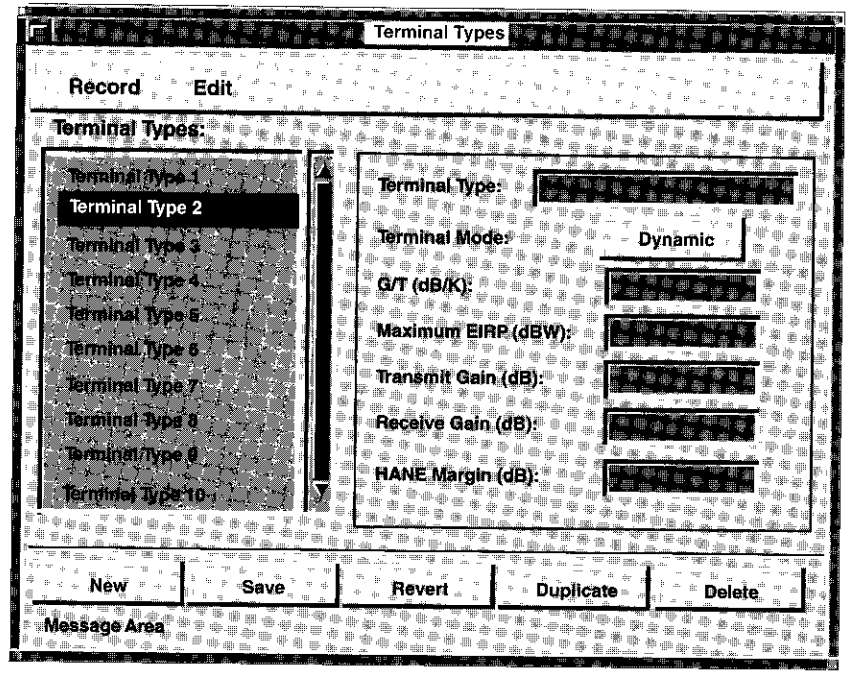


Figure 5. *Typical Data Input Dialog*

for the effects of thermal noise, interference, intermodulation, jamming, high-altitude nuclear explosions (HANEs), and atmospheric fading. Figure 6 is a simplified diagram of the satellite communications link.

The nonlinear characteristics of the satellite transponder are designed to be frequency-independent (or "memoryless") and are represented by a complex Fourier-Bessel series representing the complex output carrier level to the input carrier level for a single carrier [3]. This representation is used to calculate the output power levels of individual signals, given their input power levels (thereby accounting for small-carrier suppression effects), and to determine the intermodulation noise spectrum.

For the purpose of determining intermodulation noise only, the inputs to the transponder are represented by a composite Gaussian signal having the same spectral shape as the composite of the UMS signals, broadband noise jammers, Gaussian jammers, nonconstant envelope jammers, signals from other UMS networks, and non-UMS signals sharing the same transponder. (Other UMS networks sharing the same transponder are assumed to use the full power

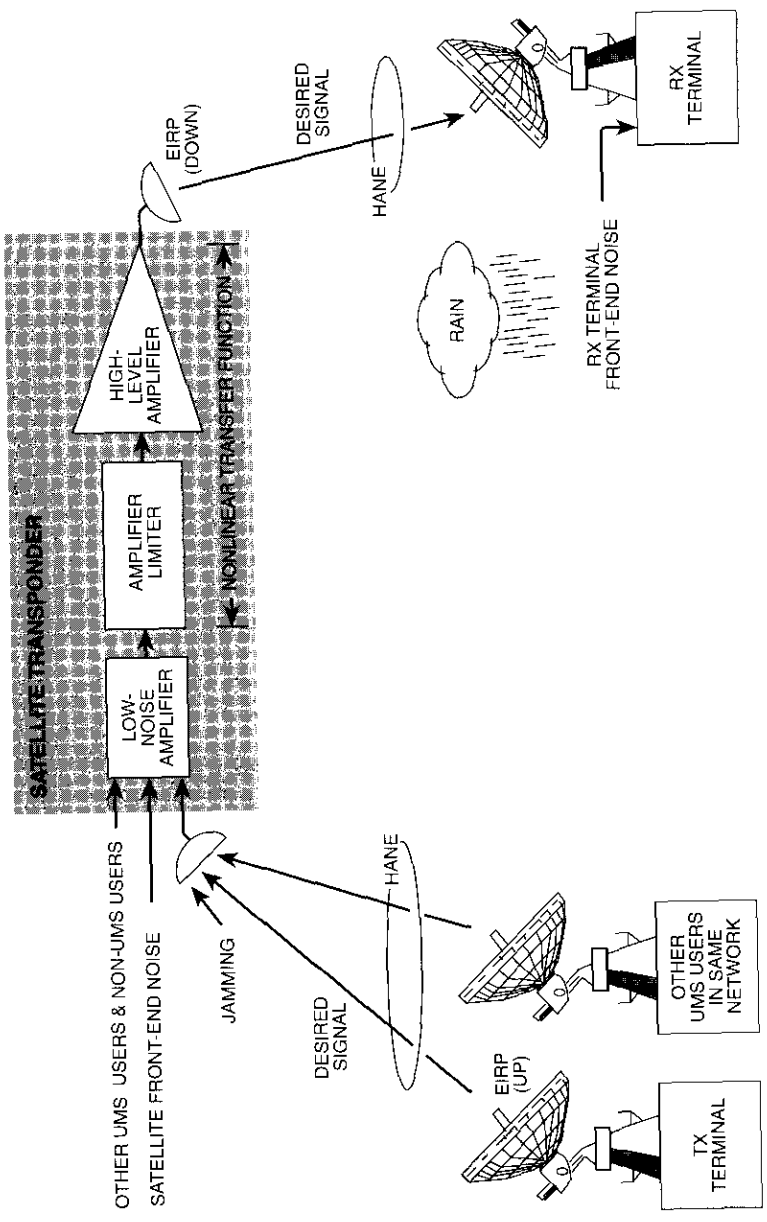


Figure 6. *Communications Link*

and bandwidth resources allocated to them by the SOC.) The validity of the Gaussian model is based on the assumption that frequency-hopped signals have random, equiprobable frequency and phase, and furthermore that even for constant-envelope signals, the spectral properties become Gaussian when the number of signals is greater than 10. The resulting intermodulation spectrum is determined by convolving the composite spectrum with itself the requisite number of times to account for both third- and fifth-order intermod products. The principal advantage of this approach over calculating individual intermod products is that the execution time is (essentially) independent of the number of input signals. The operator may trade off execution time *vs* accuracy by selecting the spectral resolution to be used, and by indicating whether fifth-order intermod products are to be accounted for.

Given the required BER for each link, the ISPC determines the minimum energy-per-bit to noise-power density ratio, E_b/N_o, required at the receive terminal. A simple and efficient optimization procedure determines the satellite transponder input power needed for each circuit in order to achieve the required E_b/N_o. The link delta is the difference between the actual and minimum required E_b/N_o. The link deltas for each circuit are computed and compared with operator-specified minimum margins to determine the adjusted value of transmitted power for the next iteration. The iteration process continues until every link meets its prescribed BER, or until the process diverges, indicating that the objective for one or more links is unachievable.

In performing link calculations, the ISPC computes the satellite antenna gain values in the direction of individual earth station locations. The satellite antenna gain patterns are represented either analytically (using a model for a circular antenna beam) or by means of measured gain values in tabular form. For multiple-beam antennas, the individual feeds are represented by a grid of measured gain values for each feed, and by an amplitude and phase weighting associated with each feed. Antennas with fixed noncircular beam shapes can be modeled as several circular beams combined with fixed amplitude and phase weightings.

NIP generation

Upon completion of the resource allocation process, the NOC operator may perform a validity check to ensure that all information in the scenario is complete and consistent, and then generate a NIP and distribute it to all modems in the network. Each NIP contains up to 10 configurations, where each configuration corresponds to a valid network scenario. An ISPC is capable

of creating, saving, retrieving, and editing up to 10 configurations for each NIP in up to 32 networks. Subsequent modifications to network scenarios are distributed to the modems in the form of NIP updates.

Ephemeris calculations

The ISPC performs calculations to predict satellite position and velocity values, which in turn are used by the modems to predict earth-terminal-to-satellite range and Doppler effects for up to 36 days into the future (see Figure 7).

The SOC operator initiates ephemeris propagation by entering an orbit vector, which defines the state of the satellite at a single known point in time. A sophisticated four-body orbit model predicts the precise position and velocity

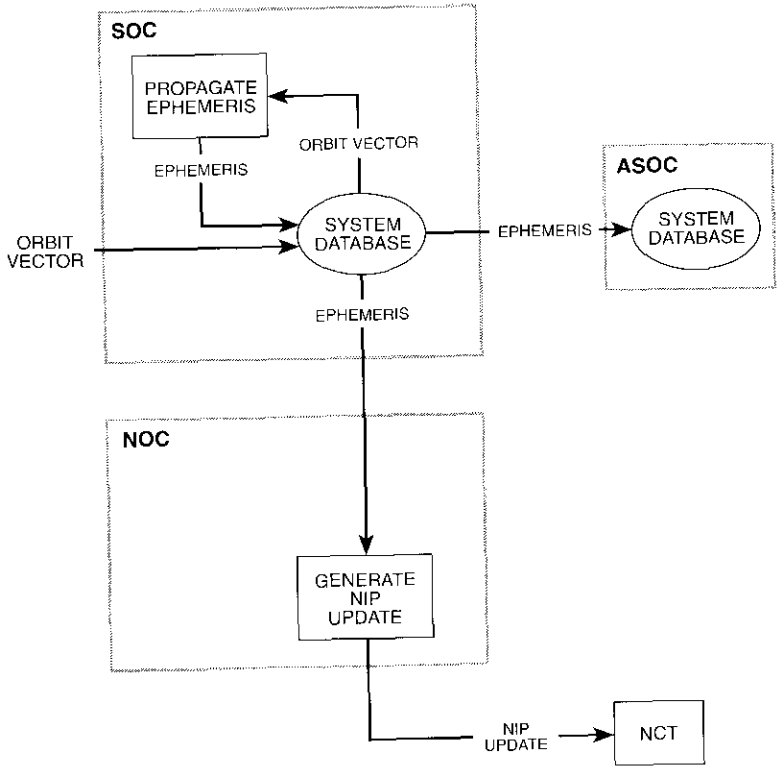


Figure 7. *Ephemeris Propagation*

of the satellite at hourly intervals over an operator-defined schedule. A Gaussian "variation of parameters" formulation of the equations of motion is used to express the acceleration of the satellite as a function of the significant forces acting upon it. The relevant forces modeled are the point-mass gravitational forces of the earth, sun, and moon; the nonspherical component of the earth's gravitational potential; and the solar radiation pressure. The function numerically integrates the equations of motion using the Bulirshe-Stoer method to produce predicted precision state vectors at intervals of 1 hour for each day over the 36-day period.

The output of this process is a table of Keplerian orbital parameter sets for the period—one set for each day. Each parameter set defines an elliptical Keplerian orbit obtained by minimizing the least-square error between the precision state vectors generated for that day and the simpler two-body model of the orbit. These parameter sets are then distributed to the NOCs, where they are included in an NIP update distributed to the modems. Each parameter set will allow the modem to run the simpler orbit model for a day to predict earth-terminal-to-satellite range and Doppler effects to the desired accuracy, and to aid the modem in acquiring and tracking signal timing and frequency.

Key management

Data security in the UMS depends on the secrecy of the cryptographic keys. Key management (Figure 8) includes all of the procedures employed for generating, storing, assigning, distributing, and terminating the keys.

The ISPC Key Manager performs the following functions:

- Sends control messages to the external rekey equipment.
- Receives status messages and keys from the rekey equipment.
- Stores keys and assigns them to appropriate modems.
- Supports the key distribution protocol, which includes transmitting keys to modems, receiving status messages, and servicing alarms.
- Provides rekeying to zero the keys in selected modems.

The Key Manager offers over-the-air key distribution on a scheduled basis, at least once per traffic key cryptoroliover period. In addition, the modem operator may direct the Key Manager to perform unscheduled over-the-air rekeying to recover from the compromise of one or more modem keys.

The ISPC monitors the current status of all keys based on status information included by the NCT in the Resource Utilization Database (RUD) reports, which are transmitted periodically to the NOC.

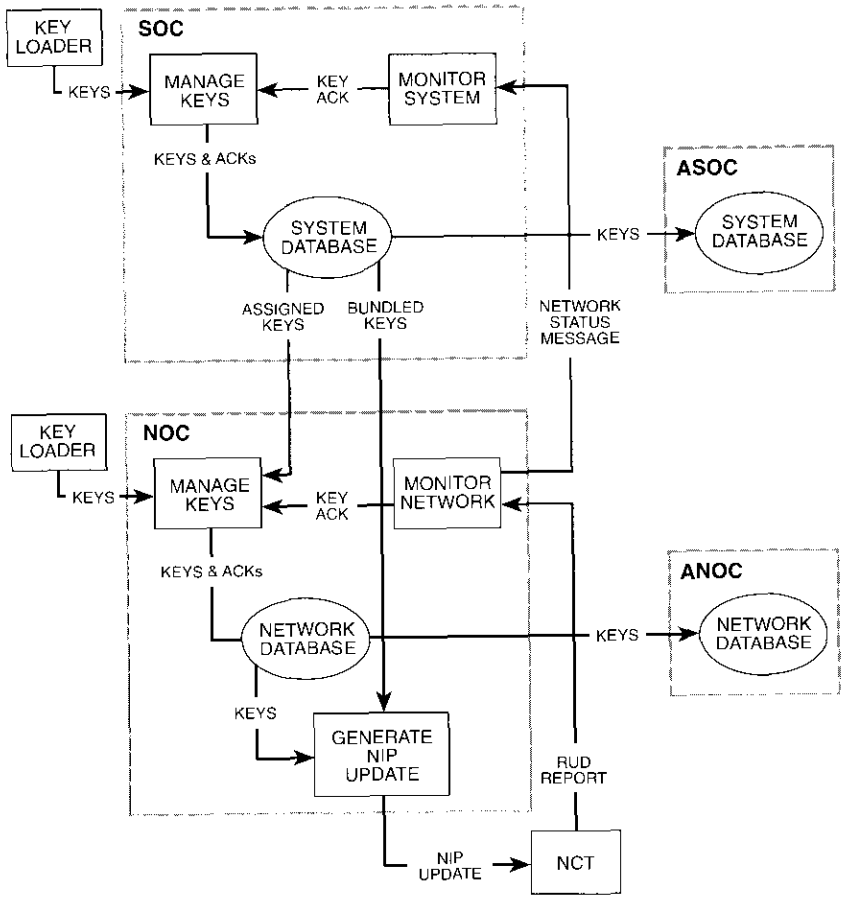


Figure 8. *Key Management*

Monitoring

Both the SOC and NOC perform monitoring functions, as shown in Figure 9. The NOC receives RUD data from the NCT for every modem in the network at minimum intervals of 1 minute. The RUD data reported to the NOC include the E_b/N_o and BER measured at the modems, the status of equipment used for active circuits, and the modem transmit power. Based on this information, the NOC generates operational status displays and sends network status reports to

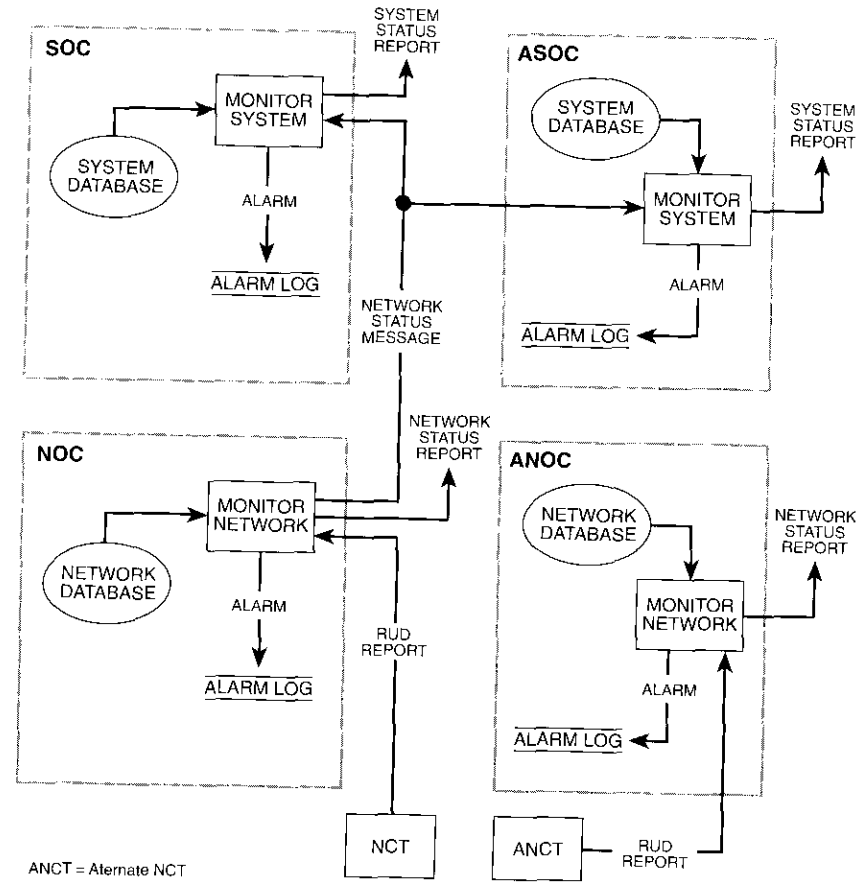


Figure 9. *ISPC Monitoring Function*

the SOC. Both the SOC and NOC perform analyses to support operator troubleshooting.

The NOC generates an operational status display of the following information for each active circuit in the network:

- Link performance deviation (planned *vs* reported E_b/N_o and BER)
- Terminal power deviation (planned *vs* reported transmit power)
- Link measurement error (computed *vs* measured E_b/N_o)

- Reported *vs* planned network power utilization
- Alarms, if deviations are outside operator-specified limits
- Operational status (Available, Busy, Failed, or Planned Maintenance) of all UM ports
- Alarms, if circuits cannot be supported due to failed or unavailable equipment.

Figure 10 shows a typical link status monitoring dialog.

The SOC accepts network status reports as input and generates a system status report that can be displayed, printed, or logged, at the option of the SOC operator. An alarm summary indicates the highest priority alarm(s) reported in each network status report.

Based on the monitoring information displayed at both the network and system levels, ISPC operators can determine the status of all major system components and the degree to which network performance conforms to the plan. When anomalies are encountered, the operator can quickly identify the source of the problem and take action to either reconfigure to another scenario

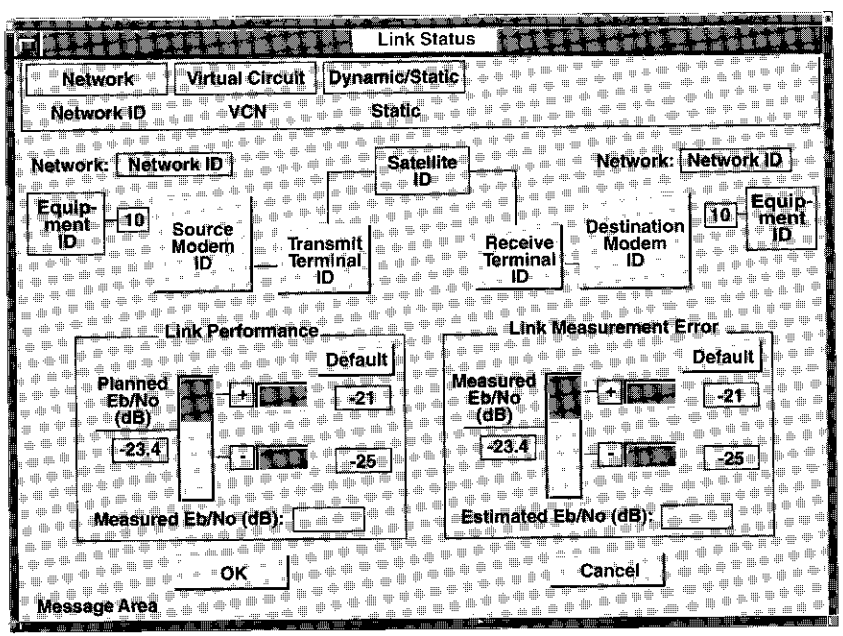


Figure 10. *Link Status Dialog*

or to generate a NIP update to modify an existing scenario. All monitoring data are logged for subsequent retrieval and review.

Conclusions

The ISPC represents a significant step toward integrating diverse satellite system management functions within a single architectural framework and computing platform. Recent major advances in computer system technology, such as graphical user interfaces, relational databases, and computer communications, have resulted in a system that is efficient, flexible, and user friendly, and can be readily expanded to meet future system management needs.

Acknowledgments

This paper is based on work performed at COMSAT Laboratories, under the sponsorship of the U.S. Army. The views, opinions, and/or findings contained in this paper are those of the author and should not be construed as an official Department of the Army position, policy, or decision, unless designated by other documentation.

The ISPC was developed at COMSAT Laboratories in support of Magnavox Electronic Systems Company, the prime contractor to the U.S. Army for the full-scale engineering development phase of the UMS program. The ISPC was delivered to Magnavox in October 1994 for integration with the Universal Modem.

Contributors to this effort from COMSAT Laboratories include A. Agarwal, P. Articola, R. Buhrman, R. Clawson, J. Fleming, T. Hazelwood, E. Homan, L. Karl, G. C. Lee, G. Y. Lee, M. Miller, G. Moxley, C. Nunan, D. Opiekun, J. Opiekun, L. Parker, S. Ramaswamy, R. Regner, K. Saralkar, B. Scherzinger, W. Schmidt, C. Shiplett, J. Snyder, M. Thakkar, and J. Wernimont.

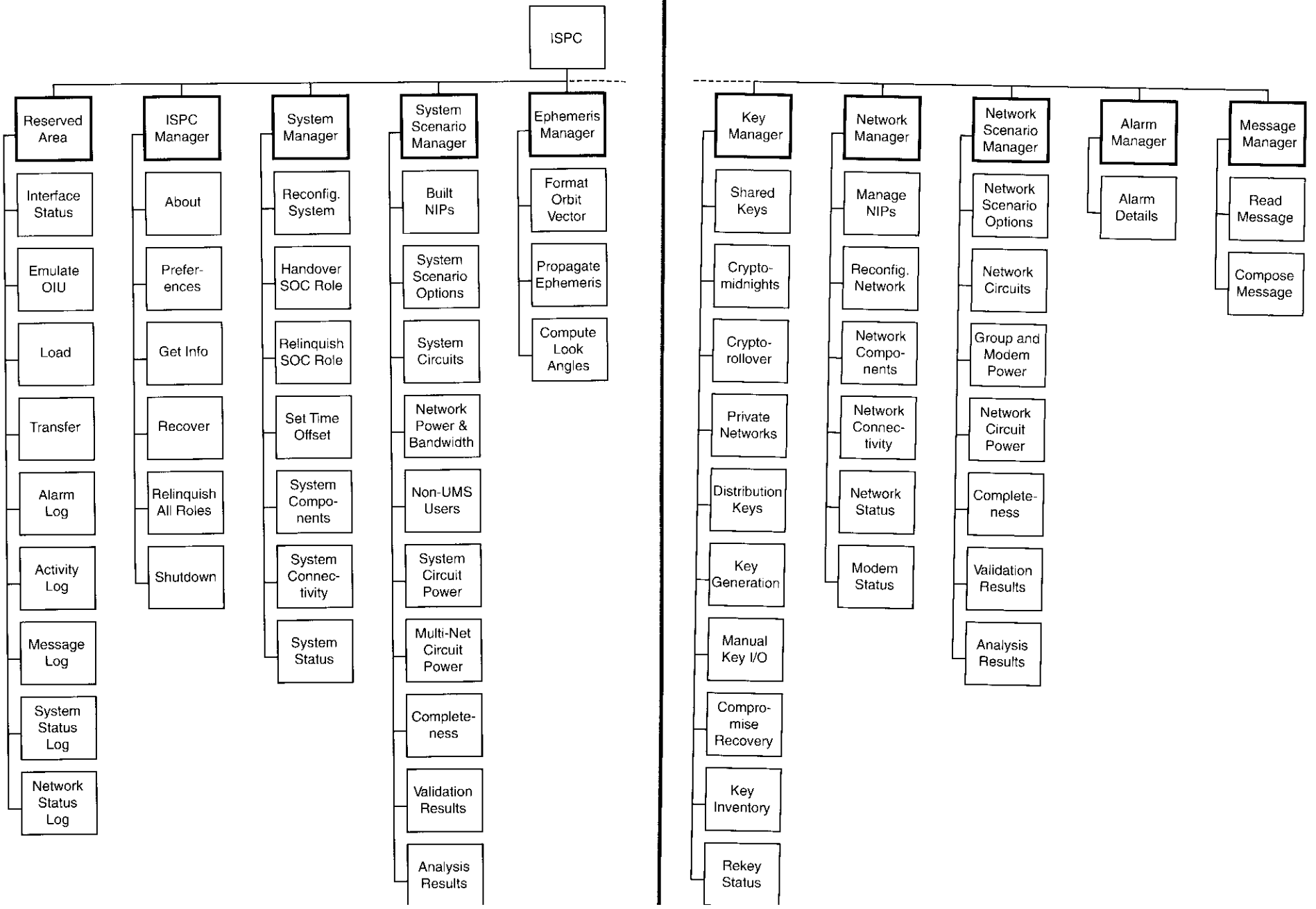
References

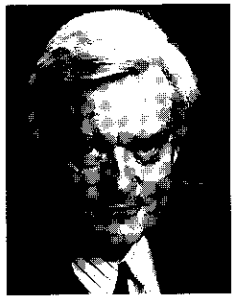
[1] "Routing Information Protocol," Internet Network Working Group, Request for Comments 1058, C. Hedrick, Rutgers University, June 1988.

[2] D. M. Chitre, "Analysis of Throughput Efficiency and Delay in ARQ Systems," *COMSAT Technical Review*, Vol. 11, No. 2, Fall 1981, pp. 345–368.

[3] J. C. Fuenzalida, O. Shimbo, and W. L. Cook, "Time-Domain Analysis of Intermodulation Effects Caused by Nonlinear Amplifiers," *COMSAT Technical Review*, Vol. 3, No. 1, Spring 1973, pp. 89–143.

Appendix: Hierarchy of ISPC windows





William L. Cook received a BS in engineering mechanics from Lehigh University, an MS in engineering sciences from Purdue University, and a DSc in computer sciences from George Washington University. He is currently Vice President of the System Development Division (SDD) at COMSAT Laboratories, where he is responsible for the design and implementation of satellite system management facilities, including satellite system planning tools and earth station and network monitoring and control facilities. SDD also conducts research and development in the areas of graphical user interfaces, database technology, and distributed computing environments.

Since joining COMSAT Laboratories in 1969, Dr. Cook has been responsible for the development of analysis techniques and computer software used in the evaluation and optimization of communications satellite systems and subsystems. These software tools include the General Antenna Program (GAP), the Long Range Planning Model (LRPM), the COMSAT Intermodulation Analyzer (CIA), the Satellite Transmission Impairments Program (STRIP), the Satellite Performance Model (SPM), and the Channel Modeling Program (CHAMP). He has managed experimental programs in high-speed computer communications via satellite, and has directed research activities in computer communications, local area networks, and interface processor development.

Dr. Cook was responsible for a number of major development projects at COMSAT for both commercial and government customers, including the INTELSAT Operations Center TDMA Facility (IOCTF), the TDMA System Monitor (TSM), and the distributed processing network architecture used in the Geostar data processing facility. He was also responsible for development of the Interim System Planning Computer (ISPC); a distributed planning, monitoring and control facility for the Universal Modem System (UMS), which is being developed for the U.S. Army; and new versions of INTELSAT's STRIP7 frequency planning program and (SS)BTPGEN TDMA burst time plan generation system.

Prior to joining COMSAT, Dr. Cook was a Structural Engineer at the Goddard Space Flight Center, and a member of the group responsible for developing the NASA Structural Analysis Program (NASTRAN).

Translations of Abstracts

Qualité de la MDF à "M" fréquences dans une voie occultée de communication par satellite du service mobile terrestre

R. A. KHALONA

Sommaire

La modulation par déplacement de fréquence à "M" fréquences (MDFM) est une méthode de modulation caractérisée par une grande efficacité énergétique. Elle est actuellement à l'étude pour des applications à puissance réduite et à faible débit binaire telles que le téléappel mondial par satellite. Son efficacité énergétique augmente à mesure que s'étend l'alphabet de transmission, au prix d'une complexité accrue et d'une perte d'efficacité de la largeur de bande. La qualité de la MDFM sans codage dans une voie de communication par satellite du service mobile terrestre est analysée au moyen d'un modèle de voie qui tient compte des effets de l'occultation. Les résultats de l'analyse de la qualité obtenue avec le codage Reed-Solomon et décodage à décisions fermes sont utilisés pour déterminer les codes nécissitant le plus petit rapport signal à bruit pour une probabilité donnée d'erreur sur les bits et pour divers degrés d'occultation. Ces résultats devraient faciliter la détermination des marges de liaison et devraient servir de reference pour la validation des conclusions des études de simulation.

Prévision de la disponibilité de systèmes portatifs de télécommunications par satellites non géosynchrones

W. A. SANDRIN ET D. V. HASCHART

Sommaire

Diverses constellations de satellites ont été proposées pour un service de téléphonie par satellite à appareils portatifs. Les constellations à l'étude peuvent se diverser en trois catégories: systèmes à orbite basse, systèmes à orbite circulaire intermédiare et systèmes à orbite de géosynchronisme. Il est particulièrement important qu'une constellation permette d'établir et de maintenir des communications lorsqu'un blocage se produit dans le secteur de l'usager. Cette question est examinée dans le cas des systèmes à orbite basse et à orbite circulaire intermédiaire, et des exemples quantitatifs sont donnés pour certaines de ces constellations, y compris la probabilité du maintien par un usager d'une communication d'une durée arbitraire dans un secteur ayant une ligne d'horizon précisément définie que détermine les directions dans lesquelles il est