

Satellite-UMTS Traffic Dimensioning and Resource Management Technique Analysis

Yim Fun Hu, Ray E. Sheriff, Enrico Del Re, *Senior Member, IEEE*,
Romano Fantacci, *Senior Member, IEEE*, and Giovanni Giambene, *Member, IEEE*

Abstract— The universal mobile telecommunications system (UMTS) will consist of space UMTS (S-UMTS) and terrestrial UMTS (T-UMTS) components. An algorithm for predicting the traffic capacity in terms of the number of subscribers for the satellite component of UMTS is presented. The algorithm takes into account the take-up characteristics of new products, the growth of gross domestic product (GDP), the projection of population, the tariff of the service, and price fall over the forecast period. The predicted traffic is used to generate a traffic grid in terms of Erlang of dimension 36×72 in steps of 5° in both the latitude and longitude directions.

The traffic grid is used to evaluate the performance of a dynamic channel allocation (DCA) technique as well as a fixed channel allocation (FCA) technique. Both channel allocation techniques have been considered with the queuing of handover (QH) requests. In order to compare the respective techniques' performance, a low-earth orbit mobile satellite system (LEO-MSS) mobility model is developed to take into account the effect of satellites' motion during interbeam handovers. A theoretical model for obtaining the values of blocking probabilities for low-traffic loads is presented. Finally, the performance of the DCA-QH technique is compared with the FCA-QH technique under suitably defined traffic and mobility conditions.

Index Terms— Dynamic channel allocation, low-earth orbit, mobile satellite communications, resource management, traffic capacity, UMTS.

I. INTRODUCTION

THE DEFINITION of third-generation mobile telecommunication systems is now in progress. In Europe, activities are concentrating on the standardization of the universal mobile telecommunications system (UMTS). UMTS will consist of both space UMTS (S-UMTS) and terrestrial UMTS (T-UMTS) segments. This paper considers two key issues required to characterize the satellite component of UMTS: traffic dimensioning and resource management strategies.

The work presented in this paper has been carried out within the framework of the European Union-funded RACE II SAINT (satellite integration in the future mobile network) project.

Manuscript received October 12, 1995; revised January 10, 1997. This work was supported in part by the European Commission which partially funded the RACE II SAINT Project, under which the traffic model and resource management technique for S-UMTS were developed.

Y. F. Hu and R. E. Sheriff are with the Department of Electronic and Electrical Engineering, University of Bradford, West Yorkshire, BD7 1DP, U.K.

E. Del Re, R. Fantacci, and G. Giambene are with the Department of Electronic Engineering, University of Florence, Florence, 50139, Italy.

Publisher Item Identifier S 0018-9545(98)07848-7.

Satellites are expected to play complementary rather than competitive roles in UMTS. The satellite's role is largely dependent on the degree of coverage offered by the terrestrial network. The following roles have been identified for the satellite component [1]–[3].

- *Coverage Completion*: Implementation of terrestrial mobile systems is a continuous process; initially mobile services are made available in densely populated urban areas, followed by extension to suburban areas, airports, motorways, etc., and, finally, rural areas. Progression down this implementation path is largely controlled by the economics and demographics of the region concerned. Due to the large coverage area offered by a satellite beam, which provides equal priority coverage for all areas, the satellite component can be used to complete the coverage of the terrestrial network. Furthermore, aeronautical and maritime users may be solely dependent on the satellite component for the provision of services.
- *Coverage Extension*: In this instance, the satellite system can be used to extend coverage boundaries of the terrestrial network.
- *Disaster Proof Availability*: Satellite systems can provide a backup service if some form of natural or man-made disaster reduces the effectiveness of the terrestrial network.
- *Rapid Deployment*: Satellite systems can be used to rapidly extend the coverage of the terrestrial network.
- *Global Roaming*: Satellite systems can provide users with global roaming capability.
- *Dynamic Traffic Management*: The satellite resource can be used to off-load some of the traffic from the terrestrial network. For example, a mobile moving from one terrestrial cell to another where no channels are available could be rerouted over the satellite, according to an internetwork handover (provided that the user accepts greater charges).

The first five of the above roles can be considered as being complementary to the terrestrial service. The final role is supportive, which can lead to a decrease in the blocking probability of the terrestrial network, or alternatively an increase in the network capacity for the same grade of service.

Based on publicly available market predictive studies and taking into account the complementary roles of the space segment identified above, an algorithm for traffic volume estimation for satellite UMTS has been developed. This algorithm makes use of worldwide databases containing demographic and economic data, as shown in Section II-C. In order to

provide for such an estimation, the effects of competing services have been considered so that likely market demand can be forecast.

Since the satellite channels are more scarcely available (and, then, more costly) than the terrestrial one (i.e., satellites are power and bandwidth limited), satellite resources in an integrated system have to be carefully utilized.

The effectiveness of the satellite's supportive role will largely be determined by the resource assignment strategy adopted by the network. The performances of radio resource management techniques have in the past been evaluated according to some presumed traffic volume distribution and a suitable model for the generation of handover requests during call lifetime. While this approach provides valuable information on the relative merits of one technique against another, optimal resource allocation strategies cannot be designed unless realistic conditions for both traffic distributions and user mobility are considered.

The main objective of this paper is to provide a performance evaluation of a dynamic channel allocation technique with queuing of handover (DCA-QH) requests, taking into account: 1) a predicted traffic volume distribution for future S-UMTS based on the historical market development of telecommunication services and 2) a realistic user mobility model according to which handover requests are generated during call lifetime. To the knowledge of the authors, an analysis of this kind, in combining both the commercial and the technical aspects associated with the provision of nongeostationary mobile satellite services, has not been carried out before.

This paper consists of two main parts: the first part concentrates on the traffic volume prediction in S-UMTS, and the second part analyzes the performance of the DCA-QH technique proposed for S-UMTS in an LEO-MSS scenario on the basis of the traffic distribution obtained in the first part. In the analysis, the space segment is assumed to consist of 66 LEO satellites, equally distributed over 11 orbital planes; each satellite has multispot beam capability. As part of the analysis, a mobility model for interbeam handovers has been derived. For the sake of completeness, the performance of the proposed DCA-QH technique has been also compared with fixed channel assignment with queuing of handover requests (FCA-QH).

II. MOBILE SATELLITE TRAFFIC MODEL

A. Background

There is very little publicly available information on mobile satellite traffic modeling and what is available tends to be sparse in content. One source of available information can be obtained from the European Commission's study "Satellite Personal Communications" [4], part of which described a method for predicting the market for future satellite services. This method has been used as a basis for the development of the traffic model presented in this paper.

A basic assumption of this study is that, at least until around the year 2005, demand for voice will outweigh that for data and

paging services. To estimate the voice market, the following assumptions were made [4].

- Areas served by existing telecommunications technologies, terrestrial wireless or cellular, were excluded.
- The size of the satellite market in the remaining areas was assumed to depend upon the same underlying relationship between the penetration and the affordability to pay for the services as that experienced in the terrestrial market.
- The satellite market in the remaining areas was assumed to develop at the same pace as the terrestrial wireless/cellular, provided that both the tariff and the quality of the satellite services are set at the same level as the terrestrial services.

As far as voice applications are concerned, the key markets for S-UMTS services that were identified came from three user groups.

- 1) *Cellular In-Fill Users*: Users in areas where there is no cellular coverage or in countries where cellular telephony may be used as a substitute for unavailable fixed telephony services.
- 2) *Rural Fixed Users*: Users in areas of low-population density or where fixed telephony services are not available.
- 3) *International Business Travellers*: Business travellers and travellers to areas of the world which currently have poor or no terrestrial cellular services.

The roles of the satellite component of UMTS identified in [1] and [2] are in general agreement with the provision of services to the above user categories.

The potential size of each user group's market depends on several factors:

- 1) economy of the region;
- 2) market penetration of the service;
- 3) tariff of the provided service;
- 4) quality of the provided service;
- 5) takeup rate of the service.

The takeup rate of a service can be characterized by four phases [5]: the introductory phase, growth phase, maturity phase and, finally, decline phase. During the introductory phase, the growth in penetration develops slowly, typically due to: 1) the high introductory tariff of the service and 2) skepticism of potential customers. Following the introductory phase, the growth in demand tends to rise rapidly due to many factors, including: fall in price, market acceptability, service options, etc. This growth continues until the market reaches maturity. The length of the maturity phase generally will be sustained before declining due, mainly, to competition from new services. The characteristics of a takeup curve are shown in Fig. 1.

This characteristic curve can be adequately described by the logistic model [6] which has the following form:

$$p(t) = \frac{M_S}{1 + ae^{bt}} \quad (1a)$$

where

- $p(t)$ penetration at time t ;
- a and b parameters of the model;

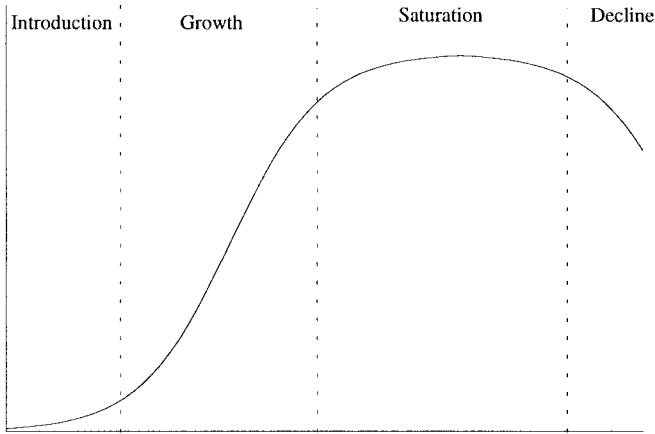


Fig. 1. Takeup curve characteristics.

M_S parameter describing the saturation level.

Transformation of (1a) gives

$$X_t = \alpha + bt \quad (1b)$$

where

$$X_t = \ln \left[\frac{P(t)}{M} - 1 \right] \quad (1c)$$

$$\alpha = \ln a$$

where M is estimated by derivation of relationship between the penetration of the service and the GDP at the year of saturation using the least-square method. With this approach, the number of variables is reduced to two, namely, X_t and α , which can be estimated. This approach has been adopted throughout the paper for the derivation of the takeup rate.

B. Algorithms for the Estimation of Traffic Capacity

Three main steps are involved in predicting the traffic volume for each user group:

- determination of the profitability for provision of a service via satellite;
- determination of the gross potential market (GPM);
- prediction of penetration of the service (Pr), taking into account the takeup rate of the service (Tp).

The three parameters are combined together to give the traffic volume in terms of the number of subscribers N

$$N = \text{GPM} \times \text{Pr} \times \text{Tp}. \quad (2)$$

The following sections describe the algorithm for the prediction of the traffic volume for each user group.

1) *Cellular In-Fill Services*: The first step in sizing the market for the cellular in-fill services is to determine the profitability of the provision of such services via satellite. To determine the profitability, a population density threshold, $\text{DEN}_{\text{thres}}$, has been introduced. This is defined as the threshold of the population density above which the implementation of a terrestrial cellular system is profitable. Conversely, the use of satellites for the provision of mobile services is preferable. According to [4], $\text{DEN}_{\text{thres}}$ for terrestrial rollout is related to the gross domestic product (GDP) per capita. For low GDP

regions, a population threshold $\text{POP}_{\text{thres}}$ is used instead of $\text{DEN}_{\text{thres}}$. The thresholds defined in Table I have been used in the S-UMTS traffic model.

For regions where the provision of cellular services via satellite is profitable, the GPM has to be determined. The GPM for the cellular in-fill market, $\text{GPM}_{\text{cellular}}$, is defined as the number of people residing outside of the terrestrial cellular coverage. If in a region, $p\%$ of the population is inside the terrestrial coverage, then

$$\text{GPM}_{\text{cellular}} = \text{Pop}_{\text{region}}(1 - p/100) \quad (3)$$

where $\text{Pop}_{\text{region}}$ is defined as the population in the region. The penetration of the cellular in-fill services, $\text{Pr}_{\text{cellular}}$, is defined as the percentage of people subscribing to the services. $\text{Pr}_{\text{cellular}}$ depends on the relative tariff of the services. Relative tariff is defined as the ratio of GDP per capita to the tariff of the service.

By gathering data for the income from the cellular services and the number of cellular mobile communications subscribers as well as the GDP per capita for countries over the globe, a relationship between $\text{Pr}_{\text{cellular}}$ and the relative tariff has been developed and is shown in Fig. 2, the derivation of which is given in Appendix A. It can be seen from Fig. 2 that the cellular market penetration can be seen to level off at around about 10%.

In order to estimate the value of M_S in (1a), the saturation level, the growth in GDP and population, as well as the annual tariff, taking into account annual price fall [4], are forecast at the year of saturation. This is carried out on a country by country basis. A saturation period of 15 years is assumed [7]. With the GDP, population, and annual tariff, the relative tariff is calculated according to Appendix A. Based on the historic growth in cellular mobile radio services and on the estimated value of M_S , the values of a and b in (1a) are determined according to the method outlined in [8]. The yearly penetration of cellular in-fill services, $P_{\text{cellular}}(t)$, can then be obtained by applying (1a). Finally, the number of subscribers for cellular in-fill services, N_{cellular} , in year t is given by

$$N_{\text{cellular}} = \text{Pop}_{\text{region}}(1 - p/100) \times P_{\text{cellular}}(t). \quad (4)$$

2) *Rural Fixed Services*: The first step in estimating the number of users of rural fixed services is to determine the profitability threshold for the provision of such services via satellite. This is achieved in the same way as for cellular in-fill services. The second step is to identify the GPM. The GPM of the rural fixed services, $\text{GPM}_{\text{fixed}}$, is defined as follows:

$$\text{GPM}_{\text{fixed}} = N_{\text{household}} - T_{\text{residential}} \quad (5)$$

where $N_{\text{household}}$ is the number of households and $T_{\text{residential}}$ is the number of telephone mainlines in residential areas.

Having determined the GPM, the penetration of the fixed rural services has to be determined. The penetration, Pr_{fixed} , is defined as the expected number of telephone lines per household. Fig. 3 shows the relationship between Pr_{fixed} and the relative tariff, which is defined as the ratio of GDP per household to the tariff of the telephony service, the derivation of which is given in Appendix A.

TABLE I
THRESHOLDS FOR TERRESTRIAL ROLLOUT FOR THE PROVISION OF MOBILE SERVICES

Country GDP Group	Year 2010
High GDP (> 24 kECU/capita)	$DEN_{thres} > 3 \text{ people/km}^2$
Mid GDP (between 7 and 24 kECU/capita)	$DEN_{thres} > 30 \text{ people/km}^2$
Low GDP (< 7 kECU/capita)	$POP_{thres} > 1 \text{ million people}$

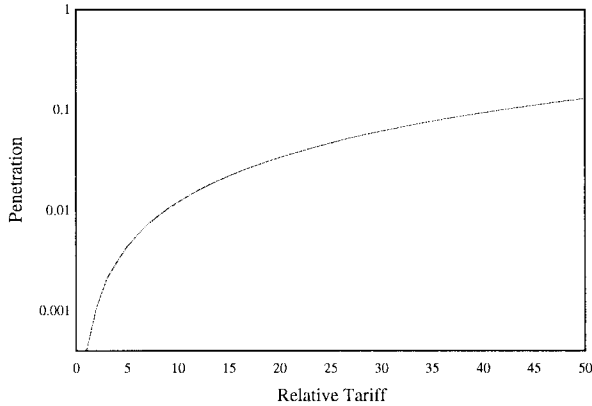


Fig. 2. Relationship between the penetration of cellular services and relative tariff.

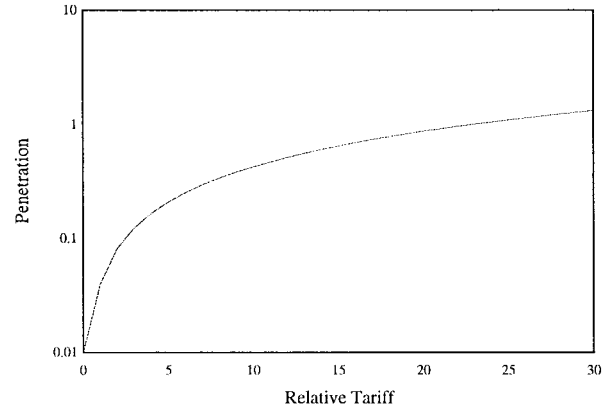


Fig. 3. Relationship between the penetration of rural fixed services and relative tariff.

Following similar procedures as in the cellular in-fill services and by applying the logistic model, the yearly penetration of rural fixed services, $P_{fixed}(t)$, is determined on a country by country basis, based on the historic growth in the telephony services and on the penetration curve, as shown in Fig. 3. Hence, the total number of subscribers, N_{fixed} , for the rural fixed service is (assuming that it is profitable to provide services via satellite)

$$N_{fixed} = (N_{household} - T_{residential}) \times P_{fixed}(t). \quad (6)$$

3) *International Business Travellers*: The methodology adopted in deriving the traffic model for international business travellers is the same as in the cellular in-fill model. However, the gross potential market for this group of users is defined as the number of outbound airline travellers, $N_{air\ travellers}$, per annum, which is related to the GDP of the region. In determining the penetration of provision of S-UMTS services to this group of users, the subscriber payment per year was set three times higher than that of the cellular in-fill model. It has been assumed that the service takeup rate for this group of users is the same as the cellular in-fill users, but with a lower penetration due to higher tariff. Hence, the number of subscribers N_{IBT} for this user group can be predicted as follows:

$$N_{IBT} = N_{air\ travellers} \times P_{IBT}(t) \quad (7)$$

where $P_{IBT}(t)$ is the yearly penetration of S-UMTS services for international business travellers.

C. Data Sources

Any forecast model is subject to uncertainty. One of the major problems in predicting traffic capacity is data acquisition.

Most of the data have been obtained from the United Nations [9], [10], the ITU [11], as well as from the EU's databases [12]. Table II identifies the data used for the prediction of traffic for each user group.

D. Traffic Distribution

On the basis of the market model developed in the previous sections, we are able to evaluate the distribution of subscribers over the world. Fig. 4(a)–(c) shows the distribution of number of subscribers for each user group over the seven regions of the globe, while Fig. 4(d) shows the total distribution in the year 2010. In generating these results, the costing figures shown in Table III were used. It can be seen that for cellular in-fill services, Asia, the former USSR, and Europe have the biggest potential market. For international business travellers, Europe and North America have the biggest potential market. As for rural fixed services, the market is dominated by Asia (mainly China and India), as well as by other developing regions such as South America, the former USSR, and Africa. In general, Asia has the biggest potential market for S-UMTS services.

It has been predicted that by 2010, of the total market, 61% will be due to fixed users, 36% will be due to cellular in-fill, and 3% will be due to IBT's.

Figs. 5 and 6 show the growth in the number of subscribers over the period 1998–2010 with lower and higher tariffs, respectively. In generating these results, the prices [in European currency units (ECU's)] in the following table have been assumed.

E. Generation of Traffic Grid

The prediction of traffic volume by the year 2010 in terms of number of subscribers enables the prediction of traffic in

TABLE II
TRAFFIC MODEL DATA

User group	Data for prediction of the number of potential subscribers	Number of Countries Used
Cellular in-fill	1) Population	225
	2) Rural area population %	225
	3) Gross Domestic Product (GDP)	225
	4) Surface area of the region	225
	5) Number of mobile communications subscribers	216
	6) Population of cities over a million	58
	7) Income from telecommunication services	194
Rural fixed services	1) Number of telephone mainlines	218
	2) % of residential telephone lines	199
	3) Household size	225
	4) GDP	225
	5) Income from telephony services	183
International business travellers	1) Number of outbound airline travellers per annum	146
	2) GDP	225

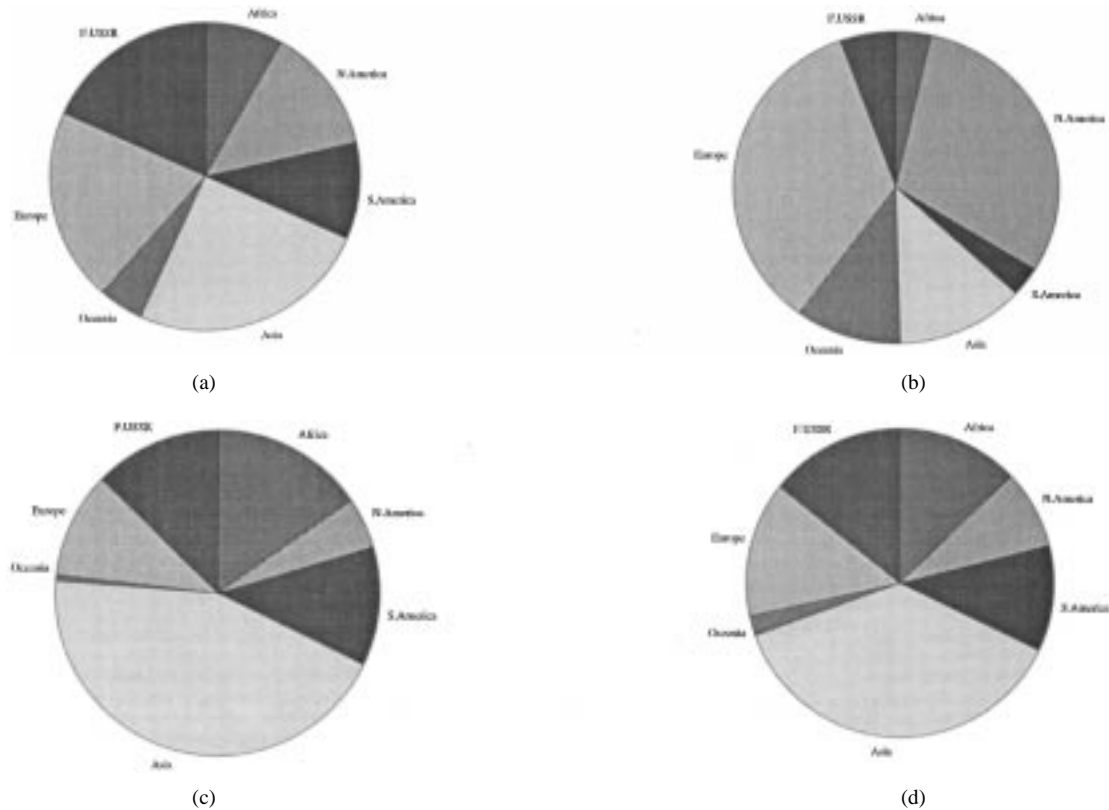


Fig. 4. Distribution of (a) cellular in-fill users, (b) international business travellers, (c) rural fixed users, and (d) S-UMTS users.

Erlang. A traffic grid of dimension 36×72 has been generated in steps of 5° in both latitude and longitude directions. In generating the grid, it is assumed that subscribers are evenly distributed over a country. The traffic generated by a user depends on the hour of the day; here, the peak traffic (rush hour) generated by a user is considered: each mobile subscriber

(both international business travellers and cellular in-fill users) generates 0.01 call per hour; fixed subscribers are assumed to generate 0.04 call per hour.¹ The average call duration τ_m is assumed to be 2 min.

¹ Values derived following correspondence during the RACE II SAINT project activities.

TABLE III
ASSUMED TARIFF AND ANNUAL PRICE FALL FOR S-UMTS SERVICES

Charge in ECU (at system launch)	Fig. 5 (lower tariff)	Fig. 6 (higher tariff)	Annual price fall
<i>Handset/Terminal</i>	700	1000	0.7%
<i>Subscription/year</i>	130	200	0.7%
<i>Usage per call</i>	0.2	0.5	0.5%

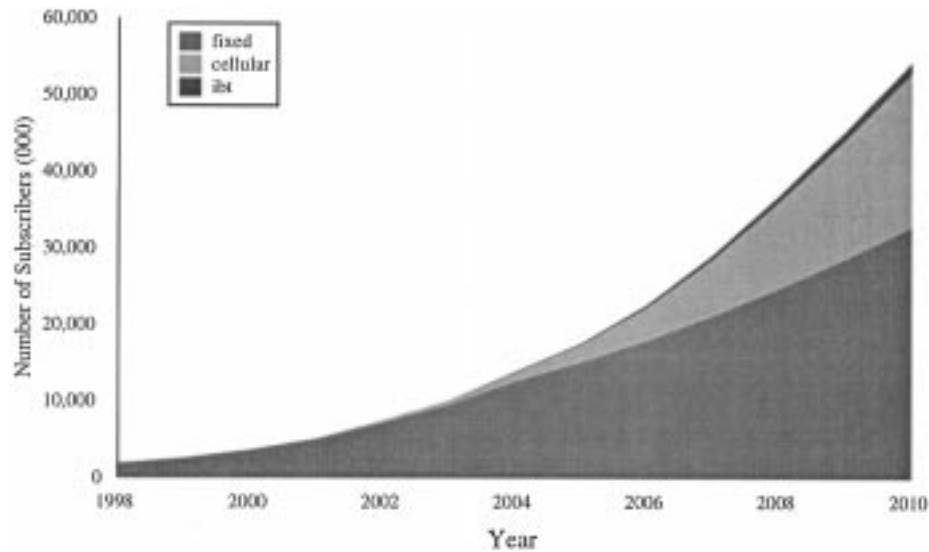


Fig. 5. World market forecast with lower tariff.

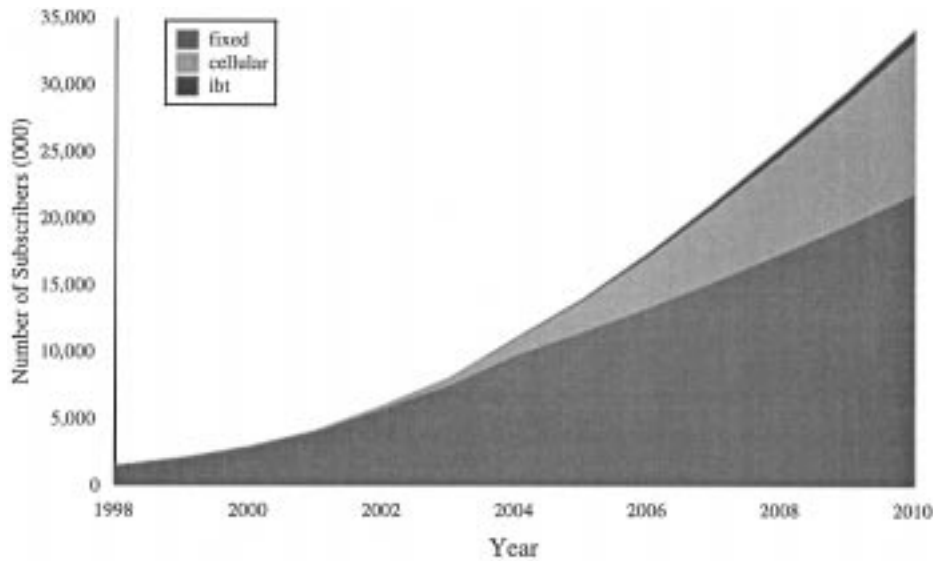


Fig. 6. World market forecast with higher tariff.

Each grid point represents the center of an area of 5° in both the latitude and longitude directions. By assuming that the earth is a perfect sphere of radius R_T , the area of each grid square $A(i)$ centered at (i, j) with latitude $\alpha(i)$, where i and j denote the indexes for the latitude and longitude, respectively,

is given by

$$A(i) \cong \left(R_T \frac{\pi 5^\circ}{180^\circ} \right)^2 \cos(\alpha(i)) \tag{8}$$

where $R_T = 6378$ km. The traffic intensity $\rho(i, j)$ inside the

traffic grid is given by

$$\rho(i, j) = \lambda_1 \tau_m \sum_{k \in S(i, j)} N_{1k} p_k + \lambda_2 \tau_m \sum_{k \in S(i, j)} N_{2k} p_k \quad (9)$$

where

- λ_1 = the mean call arrival rate per mobile subscriber (=0.01 call/hour);
- λ_2 = the mean call arrival rate per fixed user (=0.04 call/hour);
- τ_m = the average unencumbered call duration assumed here equal to 2 min (S-UMTS case);
- $S(i, j)$ = the set of countries inside the grid square (i, j) ;
- N_{1k} = number of S-UMTS mobile subscribers in country k ;
- N_{2k} = number of S-UMTS fixed users in country k ;
- p_k = area percentage of the country k in grid square (i, j) .

It is more meaningful to consider the traffic density within a grid square. For given $\rho(i, j)$ and $A(i)$, the traffic density within the grid square (i, j) is given by

$$\sigma(i, j) = \frac{\rho(i, j)}{A(i)}. \quad (10)$$

III. PERFORMANCE EVALUATION OF DCA-QH TECHNIQUE IN S-UMTS

A. Study Assumptions from the Resource Management Standpoint

The traffic characteristics, i.e., the traffic density and the traffic intensity, for S-UMTS obtained in Section II have been used to evaluate the performance of DCA-QH requests. This technique is based on the cost-function DCA algorithm described in [13] for terrestrial mobile networks that has already been applied to LEO-MSS's [14].

For the cost-function DCA technique, the allocation of channels to new call arrivals and the release of channels at call terminations are based on different cost functions. In order to extend its application to S-UMTS networks, it is assumed that the beamforming technique enables each spot-beam footprint to be equivalent to a cell in the terrestrial cellular network: a hexagonal regular cellular layout is obtained, where parameter R denotes the length of each cell side. The centers of two adjacent cells are separated by a distance of $\sqrt{3}R$. With these assumptions, the cost functions for new call arrivals and call terminations in the satellite network remain the same as those defined in [13].

As an improvement to the model developed in [14], users are also allowed to move across the seam of the satellite cellular network, where the handover occurrence is greatly increased with respect to a perfectly centered motion. This represents a more realistic condition for user mobility. Accordingly, a mobility model suitable for the satellite network under consideration has been defined in Section III-B.

In this analysis, the working orbital configuration for S-UMTS is assumed to be that of the LEO type (Iridium-like),

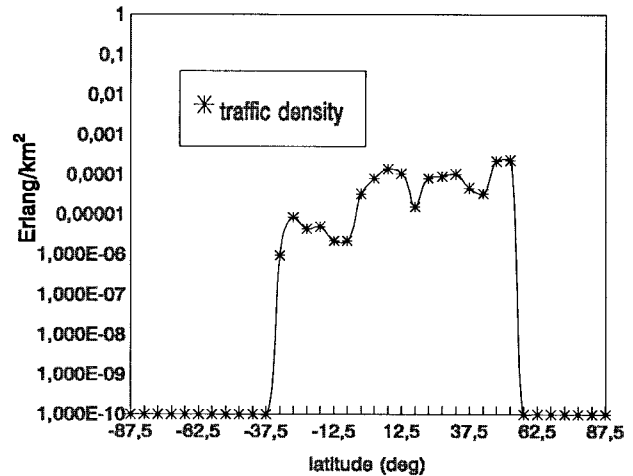


Fig. 8. Peak S-UMTS traffic density found by the cells irradiated by an LEO satellite on a polar plane at a longitude of 10°E.

with 66 satellites equally distributed over 11 polar orbits at an altitude of 780 km. The orbital period is equal to 102 min. We assume that each satellite has a multispot beam antenna that irradiates on the earth a network of cells which is parallelogram shaped with seven cells per side.

It has been considered that the region of the earth covered by the satellites is on a hypothetical polar plane centered at a longitude of 10°E. The satellites on this plane cover a central European region. The method developed to estimate the blocking probability values can be directly applied to other regions as well. The satellites on this plane are assumed to manage the traffic of a belt on the earth with a width of approximately 15° at the equator, as shown in Fig. 7.

The networks of cells managed by the satellites of this plane cross regions with different traffic conditions on the earth and at different hours. These are approximately the regions which are considered in the three columns of the traffic grid for longitudes 5°, 10°, and 15°E, respectively.

For simplicity, it has been assumed that, for a given latitude, the traffic density crossed by the cells of a satellite is constant and equal to the average of the density values of the longitudes 5°, 10°, and 15°E. In Fig. 8, the average traffic density crossed by a network of cells, as derived according to (10), is shown for satellites placed on a polar plane at a longitude of 10°E.

Furthermore, it has been assumed that three different mobile satellite systems (MSS's) will share equally the S-UMTS traffic. For the LEO-MSS under consideration, the traffic intensity per cell has been derived according to (10) and is shown in Fig. 9.

Finally, it has been assumed that in each cell the traffic due to new call attempts is produced according to a Poisson process.

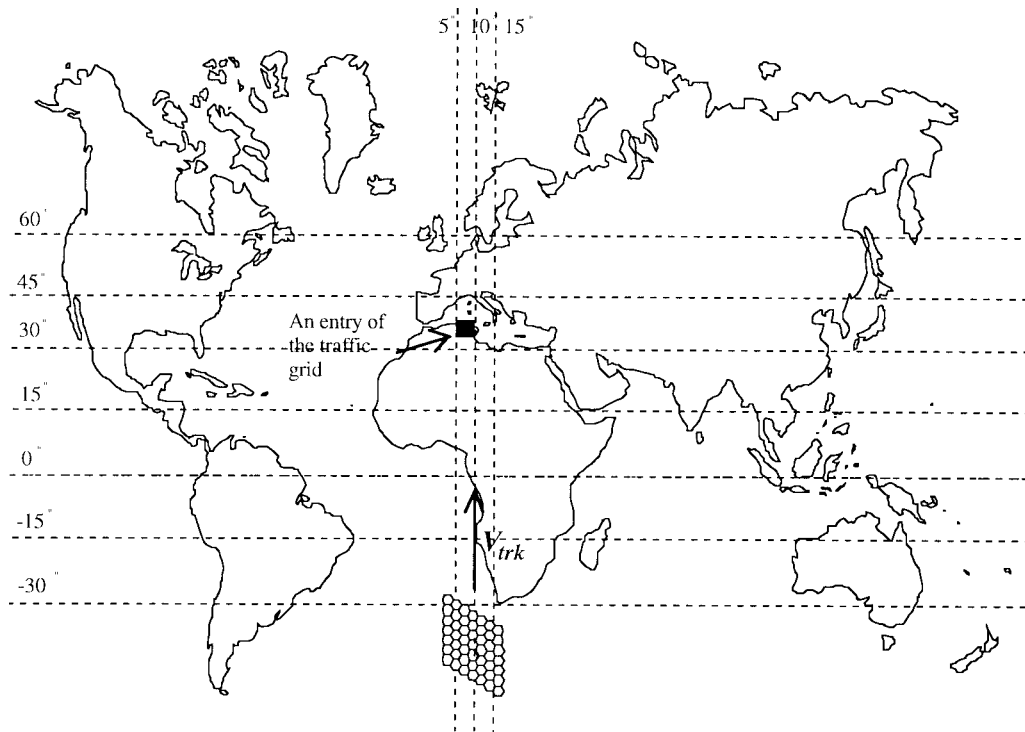


Fig. 7. A network of cells managed by an LEO satellite on a polar plane at a longitude of 10°E using a Mercator projection of the earth.

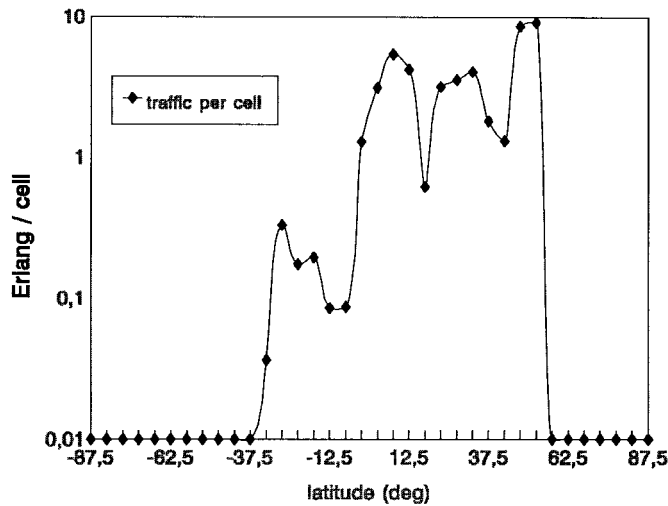


Fig. 9. Peak traffic intensity per cell for the MSS under consideration.

B. The LEO-MSS Mobility Model for the Generation of Interbeam Handover Requests

When a user terminal (UT) with a call in progress crosses the border of an adjacent cell, a new channel must be assigned to the call in the new cell (destination cell) to avoid a forced call termination. This procedure is called *interbeam handover*. In an MSS, the handover procedure may occur between different spot beams of the same satellite or of adjacent satellites. In this paper, only the former case will be considered.

Unlike geostationary satellites, the main problem in LEO-MSS's is due to the satellite's motion with respect to a fixed point on the earth. A convenient representation of the

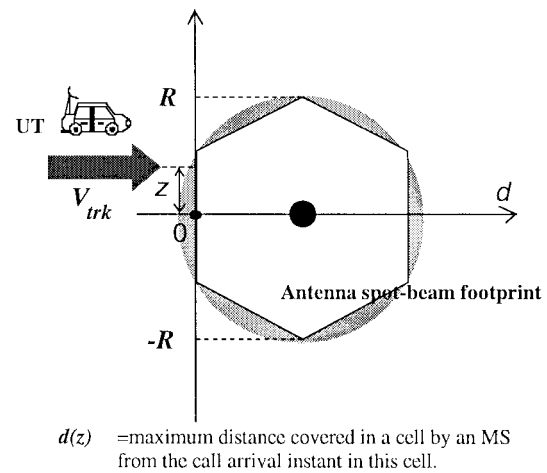


Fig. 10. The LEO mobility model.

LEO mobility scenario is that of considering any UT moving relative to an LEO satellite with a speed equal to the satellite ground-track speed V_{trk} . This approximation is acceptable due to the relatively high value of V_{trk} (about 26000 km/h) in comparison with the other motion component speeds such as the earth rotation around its axis and the UT motion relative to the earth [15], [16].

In order to define the mobility model for interbeam handovers in LEO-MSS's, the following assumptions have been made.

- 1) UT's cross the cellular network with a relative velocity "orthogonal" to the side of the cells (see Fig. 10).
- 2) A UT crosses the cell at a random height z , which varies from $-R$ to R . (see Fig. 10).
- 3) UT's cross cells at a constant velocity (relative to the

- satellite) equal to the satellite ground-track speed V_{trk} .
- 4) When a handover occurs, the destination cell is the neighboring cell in the direction of the satellite motion.
 - 5) From the call outset in a cell, a UT travels a distance (depending on z) defined as:
 - a) uniformly distributed between zero and $d(z)$, if the cell is the source cell for the call;
 - b) deterministically equal to $d(z)$, if the call is in a transit cell.

Note that the source cell refers to the cell where the UT call begins. Any cell reached by the associated user during the call lifetime is referred to as the transit cell.

According to Fig. 10, the expression of the distance $d(z)$ is obtained as follows:

$$d(z) = \begin{cases} \sqrt{3}R, & \text{if } |z| \leq \frac{R}{2} \\ 2\sqrt{3}(R - |z|), & \text{if } R \geq |z| > \frac{R}{2}. \end{cases} \quad (11)$$

The user mobility can be characterized by the following dimensionless parameter:

$$\alpha = \frac{2R}{V_{\text{trk}}\tau_m} \quad (12)$$

where τ_m is the average unencumbered call duration as defined in Section II-E.

For the LEO-MSS under consideration (i.e., Iridium-like), $R = 212.5$ km, $V_{\text{trk}} = 26\,600$ km/h, and $\tau_m = 2$ min, which gives $\alpha = 0.48$.

Due to the high value of V_{trk} , this mobility model is valid for any type of S-UMTS subscriber, regardless of whether they are stationary or travelling on a train, aircraft, or ship.

In order to evaluate the performance of channel allocation techniques, simulations of the satellite cellular system have been based on the previous mobility model and according to the following conditions.

- The simulated cellular network is parallelogram shaped (seven cells per side), and it is folded onto itself in order to naturally produce handover requests for border cells and in order to ensure that each cell in the network has a complete belt of interfering cells [17].
- A maximum queue capacity of ten has been used for the interbeam handover requests.
- The reuse distance is $\sqrt{21}R$ (i.e., the belt of interfering cells is formed by two tiers of cells) [18].
- The number of channels available to the system is $M = 70$ (in the FCA case, according to the selected reuse distance, we have a seven-cell cluster: each cell has ten channels permanently assigned).

On the basis of the uniform distribution of the height z according to which a UT crosses a cell (see the mobility assumptions), the average distance crossed in a cell by a UT, L , is obtained as follows:

$$L = \frac{1}{2R} \int_{-R}^R d(z) dz = \frac{3\sqrt{3}}{4}R. \quad (13)$$

The relative UT motion is characterized by velocity V_{trk} , and then the average UT sojourn time in a cell² (T_s) results in

$$T_s = \frac{L}{V_{\text{trk}}} = \frac{3\sqrt{3}}{8}\alpha\tau_m. \quad (14)$$

By substituting $\alpha = 0.48$ for the Iridium-like mobility case and $\tau_m = 2$ min, (14) gives $T_s = 38$ s. Parameter T_s can be considered as the average time between interbeam handover requests during call lifetime. This result has been verified by simulations according to the previously mentioned simulation model [19] and the above quoted values of system parameters.

Let us denote by n_h the average number of handover requests per call arrival. In [19] and [20], it has been demonstrated that parameter n_h in no blocking conditions, $n_{h|\text{no blocking}}$ (this is the maximum possible value for n_h ; n_h is close to $n_{h|\text{no blocking}}$ for low-traffic loads; see also the next section), only depends on the geometry of the network and on the mobility assumptions: it is given by the ratio between the average unencumbered call duration (τ_m) and the average UT sojourn time in a cell (T_s). Therefore, in the mobility model developed in this paper, $n_{h|\text{no blocking}}$ is given by

$$n_{h|\text{no blocking}} = \frac{\tau_m}{T_s} = \frac{8}{3\sqrt{3}\alpha} \frac{\text{handovers}}{\text{call}}. \quad (15)$$

With $\alpha = 0.48$ (Iridium mobility case), $n_{h|\text{no blocking}}$ is equal to approximately 3.3 handovers per call. This result will be used in the theoretical study on the DCA performance made in Section III-C.

Due to the extremely high-interbeam handover rate, a suitable policy is required to manage the interbeam handover requests. Assume that an active UT is approaching the border of its cell x and is going into an adjacent cell y . In this new cell, the UT must be provided with a new channel to continue with the communication; subsequently, the channel used in cell x must be released by the UT (i.e., *seamless handover, make-before-break arrangement*).

It has been assumed that a good policy to manage the interbeam handover requests is to allow the QH requests if there is no available channel in the destination cell of the user.

The queuing of handover requests is made possible by a certain degree of overlap among the footprints on the earth of adjacent beams. Let us assume that, due to beamforming, spot-beam footprints disposed on the earth according to a hexagonal regular layout (side R) have a circular coverage with radius R' . Within an overlap area, a user can receive the signal from at least two beams. Possible values of the ratio R'/R range from 1 to 1.5 [21]. Obviously, the greater this ratio is, the greater the overlap area and, subsequently, the better the performance achieved by the queuing technique. A conservative choice has been made: $R'/R = 1$.

Consider that a UT sends a handover request as soon as it enters the overlap area between cell x and cell y , where cell x is the originating cell.

The value of the maximum queuing time for any handover request, $t_{w\text{max}}$, can be derived as the time spent by a UT

²The UT sojourn time in a cell is a random variable which represents the time spent from the instant when the UT enters a cell to the instant when the UT reaches the borders with an adjacent cell.

to cross the overlap area between adjacent cells at a given height z with a speed V_{trk} . In general, the distance crossed in the overlap area and $t_{w\text{max}}$ are random variables that depend on several parameters, such as the direction of UT (relative to the spot-beam footprint) motion, the satellite antenna characteristics, the propagation conditions, and the satellite constellation dynamics. For the sake of simplicity and on the basis of the mobility model, the randomness of $t_{w\text{max}}$ only depends on the height z according to which a UT crosses the cellular layout. According to simulations, the average value for $t_{w\text{max}}$ is approximately equal to 7 s [19].

We have considered a first-in-first-out (FIFO) service discipline for the queued handover requests. Let $\Lambda(y)$ be the set of available channels in the cell y at the handover instant, the handover request sent by a UT is served according to the following steps.

- If $\Lambda(y) \neq \emptyset$, an arrival due to handover is served in cell y and a call termination is performed in cell x .
- If $\Lambda(y) = \emptyset$, the handover request is queued waiting for an available channel in cell y . In the meantime, the communication is served by the source cell. If within $t_{w\text{max}}$ a channel becomes available in cell y (and the call related to the handover request is still in progress), the handover request is successfully served, otherwise, the handover fails and the call is dropped.

C. Performance Evaluation

In order to evaluate the performance of the DCA-QH technique, the following parameters are defined:

- P_{b1} blocking probability for new call attempts;
- P_{b2} interbeam handover failure probability;
- P_{drop} call-dropping probability due to an unsuccessful handover;
- P_{ns} probability of an incompletely served call due to either the initial blocking or a subsequent handover failure;
- λ_{na} mean arrival rate per cell of new call attempts;
- λ_h mean arrival rate per cell of calls due to interbeam handovers.

From the above definitions, it is straightforward to note that [13], [14]

$$n_h = \frac{\lambda_n}{\lambda_{na}}. \quad (16)$$

In evaluating the performance of the DCA-QH technique, the assumptions made in Sections III-A and B are used. Furthermore, the blocking values have been estimated in the presence of peak traffic generated by each user (conservative assumption) according to the study made in Section II. In particular, let us refer to Fig. 9, where the behavior of the traffic intensity per cell as a function of the latitude of the related satellite on a polar orbit at a longitude 10° E is shown. As it can be seen from Fig. 9, the traffic intensity ranges from zero to nine Erlangs/cell. For traffic loads greater than

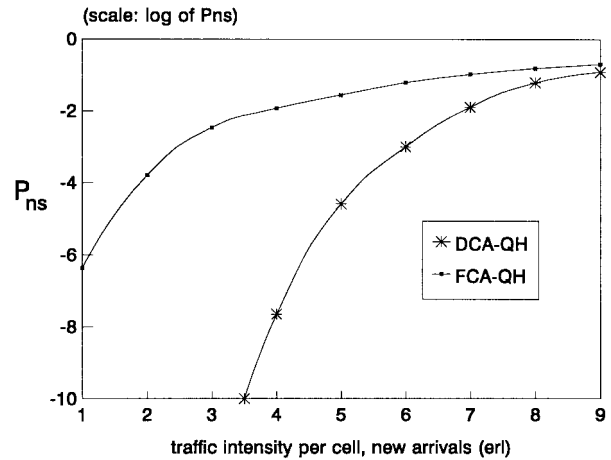


Fig. 11. Behavior of P_{ns} as a function of traffic intensity of new call arrivals ρ_{na} .

five Erlangs/cell (high-traffic loads³), the P_{ns} performance has been obtained through simulations. Instead of simulating the entire satellite dynamics (i.e., a satellite which orbits around the world and crosses regions with different traffic loads, see Fig. 9), it has been considered that an acceptable estimate of the P_{ns} performance could simply be obtained by assuming a uniform traffic distribution for the cells of the simulated network and by varying the traffic intensity for each simulation according to the traffic intensity values shown in Fig. 9 (i.e., from zero to nine Erlangs/cell).

The behavior of P_{ns} as a function of the traffic intensity per cell is shown in Fig. 11 for traffic loads greater than five Erlangs/cell. Such an evaluation may approximate the steady value of P_{ns} obtained in correspondence with each region on the earth in the real case with satellite dynamics and then with temporarily varying conditions of the traffic offered to the cells managed by satellites.

For low-traffic loads per cell (<5 Erlangs/cell), the relative blocking probabilities become too low to be reliably estimated through simulations. Thus, a theoretical evaluation of the P_{ns} performance for low-traffic loads is here proposed. Under low-traffic load conditions, the handover queuing has a decreasing impact on system performance since only few handovers are queued. Therefore, we may assume (with a good approximation) that the DCA-QH technique has the same P_{ns} performance of a simple DCA technique, as proved by simulation results in [14]; the same consideration is valid for FCA-QH versus FCA. In the following analysis, for the sake of simplicity, we will evaluate the performance of the DCA technique and we will extend these results to DCA-QH under the hypothesis of low-traffic loads per cell. The same assumption also holds for FCA as regards FCA-QH.

The theoretical evaluation of the performance of DCA techniques has been carried out in the literature [22] for the case of "fixed users" (i.e., $\alpha \rightarrow \infty$). The users are fixed when they do not change cells during call lifetime (for instance, this is the case of geostationary satellites if the intrinsic user

³This value of traffic load per cell depends on both the number of satellite channels M that has been assumed equal to 70 and on the reuse distance $= \sqrt{21}R$.

mobility can be neglected). Denoting the ERLANG-B formula by $E(\rho, S)$ for a traffic intensity ρ and a number of resources S , then

$$E(\rho, S) = \frac{\rho^S}{S! \sum_{k=0}^S \frac{\rho^k}{k!}}. \quad (17)$$

For fixed users the call blocking probability (i.e., $P_{b1} = P_{ns}$) has been theoretically evaluated in [19] as follows:

$$P_{b1} = P_{ns} = E(7\rho, M') \quad (18)$$

where

$$M' = \delta M, \quad \rho = \text{traffic intensity per cell.}$$

M' is the equivalent number of channels available for a seven-cell cluster. δ is a reduction factor of the available channels per cluster with respect to the system channels M due to the interference from adjacent cells. M' will be rounded to the next highest integer if necessary. In [22], the value of δ has been estimated by simulations for different DCA strategies from the cost function considered here; in particular, “*timid DCA*” and “*aggressive DCA*.” The *timid DCA* algorithm chooses a channel among those available in a cell without making an optimized choice; no channel reconfiguration is performed for a call in progress. The *aggressive DCA* allows the reuse distance to be temporarily violated; calls are rearranged as soon as possible. Our cost-function DCA performs a locally optimized choice in channel selection without violating the reuse distance. Channel reconfigurations may be performed at the end of a call in a cell in order to compact the use of resources. Then, cost-function DCA can be considered as an improved version of *timid DCA*. In [22], the authors show that the value of δ for the *aggressive DCA* is greater than *timid DCA*; in particular, simulation results in [22] give $\delta \approx 0.658$ for *timid DCA* with a planar cellular system and a seven-cell cluster. We expect that the cost-function DCA allows a δ value greater than 0.658, under homogeneous assumptions. According to [22], δ has been estimated by simulations: we have obtained $\delta \approx 0.8$; this value has been validated through the fitting between the blocking results obtained by simulations and those analytically predicted with the method described below.

One of the aims of this paper is to extend the previous study to evaluate P_{ns} for a DCA technique in the case of mobile users under examination. Toward this end, we consider the performance parameter called grade of service (GOS) which takes into account both P_{b1} and P_{b2} values, but with different weights according to the different arrival rates per cell due to new call attempts, λ_{na} , or to handovers, λ_h [23]. GOS can be expressed as follows:

$$\text{GOS} = \frac{\lambda_{na}}{\lambda_h + \lambda_{na}} P_{b1} + \frac{\lambda_h}{\lambda_h + \lambda_{na}} P_{b2} \quad (19)$$

where GOS is the equivalent of call blocking probability (P_{b1}) in the case of fixed users.

The parameter GOS is used to express P_{ns} in the case of mobile users. By definition [14]

$$P_{ns} = P_{b1} + (1 - P_{b1})P_{\text{drop}}. \quad (20)$$

For low-traffic intensity, $P_{\text{drop}} \cong n_h P_{b2}$ [19]. Then, (20) can be expressed as follows by neglecting the product $n_h P_{b2} P_{b1}$ since this is infinitesimal of a higher order than P_{b1} and P_{b2} :

$$P_{ns} \cong P_{b1} + n_h P_{b2}. \quad (21)$$

Reexpressing (19) with n_h gives $\text{GOS} = (P_{b1} + n_h P_{b2}) / (1 + n_h)$. By comparing this expression of GOS with (21), the following result arises:

$$P_{ns} \cong (1 + n_h) \text{GOS}. \quad (22)$$

In (22), the parameter GOS is evaluated by adapting (18), as explained below. First of all, two additional approximations have to be considered, with respect to the modeling assumptions made in Sections III-A and B.

- The handover arrival process toward a cell is Poisson with average rate λ_h independent of the Poisson process which models the arrival of new call attempts per cell (average rate λ).
- The channel holding time due to new call arrivals and handovers is assumed to be exponentially distributed with an average value T_H equal to

$$\frac{\lambda_{na}}{\lambda_{na} + \lambda_h} T_{H1} + \frac{\lambda_h}{\lambda_{na} + \lambda_h} T_{H2}$$

where

T_{H1} average channel holding time of a call in its source cell;

T_{H2} average channel holding time of a call in a transit cell.

Note that both approximations have been already introduced in [14] and validated by simulations. In particular, as to the Poisson approximation for the arrival process of handover requests toward a cell, we can note that this is the output process from the loss queuing system which models adjacent cells. It is a fact that the output process from a loss queuing system with a Poisson input process and exponentially distributed service times is not a Poisson traffic, but a smoothed one. However, such an output process can be approximated by a Poisson process under the condition of low-traffic loads.

Therefore, in adapting the ERLANG-B formula (18) to the present case with mobile users, we have to consider that there are two types of arrivals per cell (i.e., new call attempts and handover requests) which globally are characterized by a Poisson process with a gross arrival rate $\lambda_h + \lambda_{na}$ and have an exponential channel holding time with average value T_H . Then, formula (18) can be reused to express GOS in this case of mobile users, provided that we substitute for ρ a new expression for the traffic intensity per cell, ρ_{eq} , which is obtained as follows:

$$\rho_{\text{eq}} = (\lambda_h + \lambda_{na}) T_H = \lambda_{na} T_H + \lambda_h T_{H2}. \quad (23)$$

From (16) and (23), we have that $\rho_{\text{eq}} = \lambda_{na} (T_{H1} + n_h T_{H2})$.

In general, $T_{H1} + n_h T_{H2} \leq \tau_m$, due to the blocking conditions that affect the value of n_h . Therefore, if GOS is computed by using in (18) a traffic load $\lambda_{na} \tau_m = \rho_{na}$ (the traffic intensity per cell due to new arrivals [23]) instead of ρ_{eq} , then an upper bound is obtained.

Under the assumption of low-traffic loads per cell (i.e., $\lambda_{na} \rightarrow 0$), P_{b1} and P_{b2} are extremely low and the value of n_h is very close to $n_{h|no\ blocking} = 8/(3\sqrt{3}\alpha)$ [see (15)]: in the Iridium-like case ($\alpha = 0.48$), $n_{h|no\ blocking} \cong 3.3$ handovers/call; simulations have confirmed that n_h tends to saturate to a value almost equal to 3.3 handovers/call when $\lambda_{na} \rightarrow 0$ for $\alpha = 0.48$ (i.e., $\tau_m = 2$ min, $V_{trk} = 26,600$ km/h and $R = 212.5$ km). Therefore, under the assumption of low-traffic loads per cell, $n_{h|no\ blocking}$ will be considered instead of n_h and $T_{H1} + n_{h|no\ blocking}T_{H2} \cong \tau_m$; then, $\rho_{eq} \cong \lambda_{na}\tau_m = \rho_{na}$ is used in this case to obtain GOS from (18). Therefore, we obtain from (22) the following expression for P_{ns} :

$$P_{ns} = (1 + n_h)E(7\rho_{na}, \delta M). \quad (24)$$

In (24), the only user mobility-dependent parameter is n_h that, for low-traffic loads, is considered equal to $8/(3\sqrt{3}\alpha)$ [see (15)]. As already explained, we have estimated that $\delta \approx 0.8$ has to be used in (24) for our cost-function DCA technique with a seven-cell cluster; this value of δ allows a good agreement between theoretical and simulation results for the P_{ns} value around five Erlangs/cell (see Fig. 11).

Analogous considerations to those previously described lead to the following formula to evaluate P_{ns} in the FCA-QH case:

$$P_{ns} \cong (1 + n_h)E(\rho_{na}, M/7). \quad (25)$$

Equations (24) and (25) have been used to compare the performance of DCA-QH and FCA-QH for low-traffic loads. The resulting behavior obtained for P_{ns} has been shown in Fig. 11 both for DCA-QH and FCA-QH for low-traffic loads ($\rho_{na} \leq 5$ Erlangs/cell, theoretical study) and medium-high-traffic loads per cell ($\rho_{na} > 5$ Erlangs/cell, simulations). Let us remark that in Fig. 5, for very low-traffic loads per cell, the performance analysis for DCA-QH gives extremely low P_{ns} values which could not be reliably estimated by simulations. This is the reason why we resort to the use of an analytical evaluation which, in these extreme conditions, can anyhow give useful indications about the behavior of DCA-QH and its strong advantages with respect to FCA-QH.

According to the behaviors shown in Fig. 11, the P_{ns} values have been computed as a function of the latitude for the traffic characteristics presented in Fig. 9. The results are presented in Fig. 12. It is straightforward to note the significant reduction of P_{ns} obtained by DCA-QH with respect to FCA-QH for all the traffic conditions of Fig. 9. It can be concluded that DCA-QH allows a better channel utilization than FCA-QH in a real situation for mobility and traffic.

IV. CONCLUSION

Although in the past some considerations on the comparison between the performance of FCA-QH and DCA-QH were made on LEO-MSS systems, this paper represents an important contribution and effort to compare the two techniques under realistic conditions, both in terms of user mobility and traffic distribution.

First of all, an algorithm to predict the number of subscribers for the satellite component of UMTS has been developed;

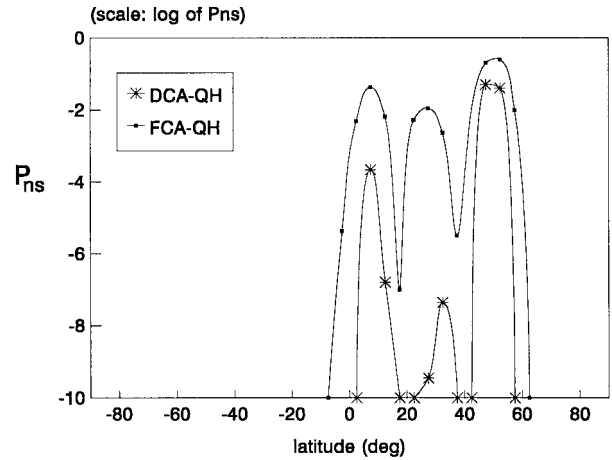


Fig. 12. Comparison of DCA-QH and FCA-QH in terms of P_{ns} for the traffic situation around the meridian of 10° E longitude.

accordingly, we have evaluated a traffic grid for S-UMTS all over the globe in the year 2010. Moreover, a suitable user mobility model has been defined and analyzed for LEO-MSS's. A performance analysis has been carried out for both DCA-QH and FCA-QH under the assumption of low-traffic loads per cell.

Finally, the performance of both DCA-QH and FCA-QH has been derived for S-UMTS according to the above traffic model and mobility model. We have highlighted that DCA-QH significantly increases the quality of service perceived by users. Both the traffic and mobility models represent very general tools that can be applied to any channel allocation strategy in order to evaluate its performance in a real scenario: they may significantly help the network designer to predict the channel capacity needed for an LEO-MSS.

APPENDIX A

DERIVATION OF RELATIONSHIP BETWEEN PENETRATION OF S-UMTS SERVICES AND RELATIVE TARIFF

A. Cellular In-Fill Services

Let

B_i average annual bill of cellular mobile services per subscriber in country i ;

I_i total income from the cellular mobile services in country i ;

N_i total number of cellular mobile subscribers in country i ;

X_i relative tariff for cellular in-fill services in country i ;

Y_i penetration (%) of the cellular in-fill services in country i ;

Pop_i total population in country i .

Then

$$B_i = \frac{I_i}{N_i}$$

$$X_i = \frac{GDP/capita}{B_i}$$

$$Y_i = \frac{N_i}{\text{Pop}_i} \times 100\%.$$

A regression line is then fitted to X_i and Y_i to obtain the relationship shown in Fig. 2.

B. Fixed Services

Let

B_i average annual bill of fixed telephony services per household in country i ;

I_i total income from the fixed telephony services in country i ;

N_i total number of main telephone lines in country i ;

M_i total number of households in country i ;

X_i relative tariff for fixed telephony services in country i ;

Y_i penetration (%) of the fixed telephony services in country i .

Then

$$B_i = \frac{I_i M_i}{N_i}$$

$$X_i = \frac{\text{GDP/household}}{B_i}$$

$$Y_i = \frac{N_i}{M_i} \times 100\%.$$

A regression line is then fitted to X_i and Y_i to obtain the relationship as shown in Fig. 3.

REFERENCES

- [1] "List of operational requirements for UMTS satellite integration," RACE II SAINT Deliverable, R2117 TNO DR 1102-A1, May 1994.
- [2] "Framework for satellite integration within the universal mobile telecommunications system (UMTS)," ETSI Tech. Rep., D-ETR/SMG-51201, Sept. 1993.
- [3] E. Del Re, "Objectives and research activities of COST 227 project-integrated space/terrestrial mobile networks," in *RACE Mobile Telecommunications Workshop*, Amsterdam, The Netherlands, May 17–19, 1994, pp. 590–597.
- [4] "Satellite personal communications and their consequences for European telecommunications trade and industry," Rep. to European Commission, KPMG, Mar. 1994.
- [5] M. J. Baker, *Marketing New Industrial Products*. New York: Macmillan, 1975.
- [6] CCITT, "Telephone network and ISDN—Quality of service, network management and traffic engineering," in *IXth Plenary Assembly*, Melbourne, Australia, vol. II, Nov. 14–25, 1988.
- [7] "Traffic model and characteristics," RACE II SAINT Deliverable, R2117 BRU EEE DR P210-B1, Apr. 1995.
- [8] CCITT GAS 10 Handbook, "Planning data and forecasting methods," Geneva, Switzerland, vol. 1, 1987.
- [9] *Statistical Yearbook*, no. 38, UNESCO.
- [10] *Demographic Yearbook*, no. 42, UNESCO.

- [11] *ITU Statistical Yearbook*, 1994.
- [12] Eurostat Data, European Commission.
- [13] E. Del Re, R. Fantacci, and G. Giambene, "Handover and dynamic channel allocation techniques in mobile cellular networks," *IEEE Trans. Veh. Technol.*, vol. 44, no. 2, May 1995.
- [14] ———, "Efficient dynamic channel allocation techniques with handover queuing for mobile satellite networks," *IEEE J. Select. Areas Commun.*, vol. 13, no. 2, pp. 397–405, 1995.
- [15] ———, "Performance analysis of a dynamic channel allocation technique for satellite mobile cellular network," *Int. J. Satell. Commun.*, vol. 12, pp. 25–32, 1994.
- [16] A. Granz, Y. Gong, and B. Li, "Performance study of low earth orbit satellite systems," *IEEE Trans. Commun.*, vol. 42, nos. 2/3/4, pp. 1866–1871, 1994.
- [17] E. Del Re, R. Fantacci, and G. Giambene, "An efficient technique for dynamically allocating channels in satellite cellular networks," in *Proc. IEEE GLOBECOM'95*, Singapore, Nov. 13–17, 1995, pp. 1624–1628.
- [18] Z. Hass, "Mobile communication networks," in *GLOBECOM'94 Tutorial no. 4*, San Francisco, CA, Nov. 28–Dec. 2, 1994.
- [19] E. Del Re, R. Fantacci, and G. Giambene, "Handover requests queuing in low earth orbit mobile satellite systems," in *Proc. 2nd European Workshop on Mobile Personal Satcoms*, Rome, Italy, Oct. 9–11, 1996, pp. 213–232.
- [20] S. Nanda, "Teletraffic models for urban and suburban microcells: Cell sizes and handoff rates," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 673–682, 1993.
- [21] T. P. Chu and S. S. Rappaport, "Overlapping coverage and channel rearrangement in microcellular communication systems," in *Proc. GLOBECOM'94*, San Francisco, CA, Nov. 28–Dec. 2, 1994, pp. 1674–1678.
- [22] L. J. Cimini, G. J. Goschini, I. Chin-Lin, and Z. Miljanic, "Call blocking performance of distributed algorithm for dynamic channel allocation in microcells," *IEEE Trans. Commun.*, vol. 42, no. 8, pp. 2600–2607, 1994.
- [23] "Radio resource management," RACE II SAINT Deliverable, R2117 FIU DIE DR P211-B1, Sept. 1995.

Yim Fun Hu, for a biography, see this issue, p. 1199.

Ray E. Sheriff, for a biography, see this issue, p. 1199.

Enrico Del Re (M'78–SM'84), for a biography, see this issue, p. 1214.

Romano Fantacci (S'82–M'87–SM'91), for a photograph and biography, see this issue, p. 1214.

Giovanni Giambene (S'94–M'97), for a photograph and biography, see this issue, p. 1214.