Computing Call-Blocking Probabilities in LEO Satellite Networks: The Single-Orbit Case

Abdul Halim Zaim, George N. Rouskas, *Senior Member, IEEE*, and Harry G. Perros, *Senior Member, IEEE*

***Abstract*—We study the problem of carrying voice calls over a low-earth-orbit satellite network and present an analytical model for computing call-blocking probabilities for a single orbit of a satellite constellation. We have devised a method to solve the corresponding Markov process efficiently for orbits of up to five satellites. For orbits consisting of a larger number of satellites, we have developed an approximate decomposition algorithm to compute the call-blocking probabilities by decomposing the system into smaller subsystems and iteratively solving each subsystem in isolation using the exact Markov process. Our approach can capture blocking due to handoffs for both satellite-fixed and earth-fixed constellations. Numerical results demonstrate that our method is accurate for a wide range of traffic patterns and for orbits with a number of satellites that is representative of commercial satellite systems.**

***Index Terms*—Call blocking probability, decomposition algorithms, handoffs, low-earth-orbit (LEO) satellite networks.**

I. Introduction

CURRENTLY, we are witnessing an increase in the demand for a broad range of wireless telephone and Internet services. Satellite-based communication is poised to provide mobile telephony and data transmission services on a worldwide basis in a seamless way with terrestrial networks. Satellite systems are insensitive to location and can be used to extend the reach of networks and applications to anywhere on earth.

Satellites can be launched in different orbits, of which the low earth orbit (LEO), the medium earth orbit (MEO), and the geostationary orbit (GEO) are the most well known. LEO satellites are placed in orbits at an altitude of less than 2000 km above the earth. Their orbit period is about 90 min, and the radius of the footprint area of a LEO satellite is between 3000–4000 km. The duration of a satellite in LEO orbit over the local horizon of an observer on earth is approximately 20 min, and the propagation delay is about 25 ms. A few tens of satellites on several orbits are needed to provide global coverage. MEO satellites are placed in circular orbits at an altitude of around 10 000 km. Their orbit period is about 6 h, and the duration of a satellite in MEO orbit over the local horizon of an observer on earth is a few hours. Fewer satellites on two or three orbits are enough to provide global coverage in a MEO system. Propagation delay in a MEO system is about 125 ms. GEO satellites are also in circular orbits in the equatorial plane at an altitude of 35 786 km, with an orbital period equal to that of the earth. A satellite in GEO orbit appears to be fixed above the earth's surface. The footprint of a GEO satellite covers nearly one-third of the earth's surface (between 75° south to 75° north). Therefore, a near global coverage can be obtained with three satellites, but the propagation delay is 250 ms.

A LEO or MEO satellite system is a set of identical satellites launched in several orbital planes, with the orbits having the same altitude. The satellites move in a synchronized way in trajectories relative to the earth. Such a set of satellites is referred to as a *constellation*. The position of all the satellites in relation to the earth at some instance of time repeats itself after a predetermined period, which is usually several days, while a satellite within an orbit also comes to the same point on the sky relative to the earth after a certain time.

In a LEO or MEO satellite system, satellites can communicate directly with each other by line of sight using intraplane intersatellite links (ISLs), which connect satellites in the same orbital plane, and interplane ISLs, which connect satellites in adjacent planes. ISLs introduce flexibility in routing; they can be used to build redundancy into the network; and they permit two users in different footprints to communicate without the need of a terrestrial system. To improve the bandwidth and frequency efficiency, the satellite footprint area is divided into smaller cells. For each cell within a footprint, a specific beam of the satellite is used. A constellation of satellites may provide either *satellite-fixed* or *earth-fixed* cell coverage. In the first case, the satellite antenna sending the beam is fixed, and as the satellite moves along its orbit, its footprint and the cell move as well. In the case of earth-fixed cell coverage, the earth's surface is divided into cells, as in a terrestrial cellular system, and a cell is serviced continuously by the same beam during the entire time that the cell is within the footprint area of the satellite.

As satellites move, fixed and mobile users hand off from one beam to another (beam handoff) or from one satellite to another (satellite handoff). The velocity of a satellite is much higher than that of objects on earth. Therefore, the number of handoffs during a telephone call depends on the call duration, the beam size, the satellite footprint size, and the satellite speed, while the location and mobility of a user only effects the time a handoff takes place. For example, for a call duration of 3 min, the customer of a LEO satellite constellation with an elevation angle of 10° will experience two handoffs. In an earth-fixed system, each beam is assigned to a fixed cell on the earth within the satellite's footprint. During a satellite handoff, all beams are reassigned to their respective cells in the adjacent footprint area. Therefore, in these systems, both beam and satellite handoffs occur at the same time. In a satellite-fixed system, a user may be handed off to the next beam in the same satellite or the satellite behind, as

Manuscript received July 20, 2000; revised August 2, 2001.

The authors are with the Department of Computer Science, North Carolina State University, Raleigh, NC 27695-7534 USA (e-mail: ahzaim@csc.ncsu.edu; rouskas@csc.ncsu.edu; hp@csc.ncsu.edu).

Publisher Item Identifier S 0018-9545(02)00430-9.

the cell defined by the beam moves away from the user. The newly entered beam may or may not have enough bandwidth to carry the handed-off traffic. In the case of satellite-based telephony, the new beam may not be able to carry a handed-off telephone call, in which case, the call will be dropped. In general, handoffs in satellite systems impose a big problem from the point of quality of service.

Existing and planned LEO/MEO satellite systems for worldwide mobile telephony include Globalstar, Iridium, ICO, Ellipso, Constellation, Courier, and Gonets. These systems differ in many aspects, including the number of orbits and the number of satellites per orbit, the number of beams per satellite, their capacity, the band they operate (S-band, L-band, etc.), and the access method employed (frequency-, time-, code-division multiple access). Also, these systems provide different services, and they may or may not have on-board switching capabilities. For instance, Iridium has on-board digital processing and switching, while other systems, such as the Globalstar, act as a bent pipe. Despite these differences, from the point of view of providing telephony-based services, the principles of operation are very similar, and thus, the analytical techniques to be developed in this work will be applicable to any LEO/MEO satellite system that offers such services.

A. Related Research

Because of the importance of satellite systems, several performance studies have been undertaken. A typical way of modeling a satellite system in the literature is to represent each cell as an $M/M/K/K$ queue. This approach permits the calculation of various useful performance measures, such as the call-blocking probability. However, this type of model does not take into account the fact that the amount of traffic in one cell depends on the amount of traffic in one or more other cells. These types of traffic dependencies are taken into account in our models described in Section II.

In [6], Ganz *et al.* investigated the distribution of the number of handoffs and the average call-drop probability for LEO satellite systems. Both beam-to-beam and satellite-to-satellite handoffs were taken into account. Each cell was modeled as an $M/M/K/K$ queue, where K denotes the number of channels per cell, assuming that the number of handoff calls entering a cell is equal to the number of handoff calls leaving the cell. In [8], Jamalipour *et al.*investigated the traffic characteristics of LEO systems and proposed a probability density function to locate the position of each user. Using this function, the normalized throughput and average delay was calculated. See also [9] and [7]. In [12], Pennoni and Ferroni described an algorithm to improve the performance of LEO systems. They defined two queues for each cell: one for new calls and one for handoff calls. The calls are held in these two queues for a maximum allowed waiting time. That is, they are dropped if they are not served within this time. The queue for new calls has a maximum waiting time equal to 20 s. The queue for handoff calls has a maximum waiting time equal to the crossover time of the overlapping zone of two adjacent cells. The handoff queue has higher priority than the new calls queue. Simulation results showed that this algorithm decreased the call-dropping rate drastically. In [13], Ruiz *et al.* used teletraffic techniques to calculate the blocking and handoff probabilities. Various channel assignment strategies were investigated. In [5], Dosiere *et al.* defined a model for calculating the handoff traffic rate. The authors divided a street of coverage into small pieces, where each piece is equal to the footprint area of a satellite. Given the total arrival distribution for the whole street of coverage, the arrival rate of each satellite was obtained by integrating that distribution for the satellite interval. The handoff rate was then calculated through a second integration. Once the handoff rate has been obtained, the blocking probability can be calculated using the Erlang loss formula.

A number of authors have also dealt with the very interesting problem of routing in a satellite system. In [17] and [18], Werner *et al.* proposed a dynamic routing algorithm for asynchronous transfer mode (ATM)-based LEO and MEO satellite systems. Due to the fact that satellites move in orbits and that orbits slowly rotate around the earth, the network topology can be seen as consisting of a series of topologies that continuously repeat themselves. For each topology, end-to-end routes are calculated. Subsequently, an optimization procedure is carried out over all the network topologies with a view to minimizing the occurrence of handoffs between successive topologies. In [11], Mauger and Rosenberg proposed the *virtual node* routing algorithm for ATM traffic. Users are mapped onto virtual nodes, and each virtual node is served by a satellite. When the satellite passes, the next satellite takes its place and serves the virtual node. Routing is performed according to the topology of the virtual nodes.

Chang *et al.* proposed the finite-state automaton (FSA) model in [3] and [1] to solve the ISL link assignment problem in LEO satellite systems. The total time it takes the position of all the satellites over the earth to repeat itself is divided into equal-length intervals during which the visibility between satellites—that is, the network topology of the satellites—does not change. Given a traffic matrix for each interval, a link assignment algorithm is run with a view to maximizing the residual capacity of the bottleneck links. The result is a table that shows connectivity between satellites for each interval. These tables can be stored in each satellite, and during the real-time operation of the system, the intersatellite links are established according to these tables. Further related research can be found in [4] and [2].

Uzunalioglu *et al.* suggested in [15] and [16] a connection handoff protocol for LEO satellite systems. First, a minimum cost route for a connection between two points on the earth is obtained. This route is used for as long as possible. When a handoff occurs at either end of the connection, the protocol simply adds the new link to the path. This continues for a predetermined amount of time, when the protocol computes a new end-to-end path for the connection. In [14], Uzunalioglu proposed a probabilistic routing protocol based on the above approach. Finally, a new traffic load balancing algorithm was proposed by Kim *et al.*in [10].

B. Contributions and Organization

In this paper, we study the problem of carrying voice calls over a LEO satellite network and present an analytical model for computing call-blocking probabilities for a single orbit of a

satellite constellation. We first derive an exact Markov process, and corresponding queueing network, for a single orbit under the assumption that satellites are fixed in the sky (i.e., there are no handoffs of voice calls). We show that the queueing network has a product-form solution, and we develop a method for computing the normalizing constant. In terms of time complexity, our method represents a significant improvement (which we quantify) over a brute-force calculation; however, it can be applied directly to orbits with at most five satellites. For a system with a larger number of satellites, we then present an approximate decomposition algorithm to compute call-blocking probabilities by decomposing the system into smaller subsystems and solving each subsystem in isolation, using the exact solution described above. This approach leads to an iterative scheme, where the individual subsystems are solved successively until a convergence criterion is satisfied.

Next, we introduce handoffs by considering the system of satellites as they orbit the earth. For an orbit with earth-fixed coverage, we then show that there is no blocking due to handoffs, and thus, the solution (exact or approximate) obtained under the assumption that satellites are fixed in the sky can be used to compute call-blocking probabilities in this case. For an orbit with satellite-fixed coverage, on the other hand, blocking due to handoffs does occur. In this case, we show how the queueing network described above can be extended to model call handoffs by allowing customers to move from one node to another. We derive the rate of such node-to-node transitions in terms of the speed of the satellites and the shape of the footprints. We also show that the new queueing network has a product-form solution similar to that under the no-handoffs assumption. Thus, the exact and approximate algorithms developed above can be applied directly to compute call-blocking probabilities under the presence of handoffs.

This paper is organized as follows. In Section II, we develop an exact Markov process model under the assumption that satellites are fixed in the sky (i.e., no handoffs take place). In Section III, we present an approximate decomposition algorithm for a large number of satellites. In Section IV, we extend our approach to model handoffs for both earth-fixed and satellite-fixed coverage. We present numerical results in Section V. In Section VI, we conclude this paper by discussing possible directions in which this work may be extended in the future.

II. An Exact Model for the No-Handoffs Case

Let us first consider the case where the position of the satellites in the single orbit is fixed in the sky, as in the case of geostationary satellites. The analysis of such a system is simpler, since no calls are lost due to handoffs from one satellite to another, as when the satellites move with respect to the users on the earth. This model will be extended in the following section to account for handoffs in constellations with both earth-fixed and satellite-fixed coverage.

Each up- and downlink of a satellite has the capacity to support up to C_{UDL} calls, while each intersatellite link has capacity equal to C_{ISL} calls. Let us assume that call requests arrive at each satellite according to a Poisson process, and that call holding times are exponentially distributed. We now show how to compute blocking probabilities for the three satellites in the single orbit of Fig. 1. The analysis can be generalized to analyze $k > 3$ satellites in a single orbit. For simplicity, we consider only shortest path routing, although the analysis can be applied to any fixed routing scheme whereby the path taken by a call is fixed and known in advance of the arrival of the call request.

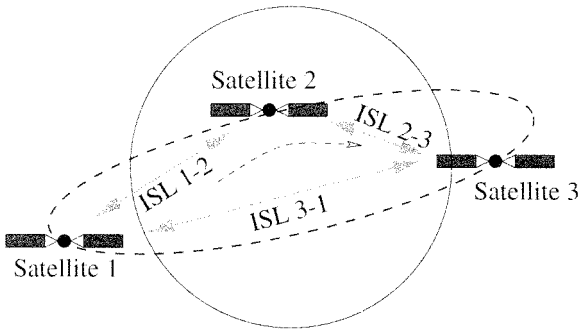

Fig. 1. Three satellites in a single orbit.

Let n_{ij} be a random variable representing the number of active calls between satellite i and satellite $j, 1 \leq i, j \leq 3$, regardless of whether the calls originated at satellite i or j. Let λ_{ij} (respectively, $1/\mu_{ij}$) denote the arrival rate (respectively, mean holding time) of calls between satellites i and j. Then, the evolution of the three-satellite system in Fig. 1 can be described by the six-dimensional Markov process

$$\underline{n} = (n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33}). \tag{1}$$

Also let $\underline{1}_{ij}$ denote a vector with zeros for all random variables except random variable n_{ij}, which is one. The state transition rates for this Markov process are given by

$$r(\underline{n}, \underline{n} + \underline{1}_{ij}) = \lambda_{ij}, \qquad \forall i, j \tag{2}$$

$$r(\underline{n}, \underline{n} - \underline{1}_{ij}) = n_{ij}\mu_{ij}, \qquad \forall i, j,\ n_{ij} > 0. \tag{3}$$

The transition in (2) is due to the arrival of a call between satellites i and j, while the transition in (3) is due to the termination of a call between satellites i and j.

Due to the fact that some of the calls share common up-and-down and intersatellite links, the following constraints are imposed on the state space:

$$2n_{11} + n_{12} + n_{13} \leq C_{\mathrm{UDL}} \tag{4}$$

$$n_{12} + 2n_{22} + n_{23} \leq C_{\mathrm{UDL}} \tag{5}$$

$$n_{13} + n_{23} + 2n_{33} \leq C_{\mathrm{UDL}} \tag{6}$$

$$n_{12} \leq C_{\mathrm{ISL}} \tag{7}$$

$$n_{13} \leq C_{\mathrm{ISL}} \tag{8}$$

$$n_{23} \leq C_{\mathrm{ISL}}. \tag{9}$$

Constraint (4) ensures that the number of calls originating (equivalently, terminating) at satellite 1 is at most equal to the capacity of the up- and downlink of that satellite. Note that a call that originates and terminates within the footprint of satellite 1 captures two channels; thus the term $2n_{11}$ in constraint (4). Constraints (5) and (6) are similar to (4) but correspond to satellites 2 and 3, respectively. Finally, constraints (7)–(9) ensure that the number of calls using the link between two satellites is at most

equal to the capacity of that link. Note that because of (4)–(6), constraints (7)–(9) become redundant when $C_{\rm ISL} \geq C_{\rm UDL}$. In other words, there is no blocking at the intersatellite links when the capacity of the links is at least equal to the capacity of the up- and downlinks at each satellite.[1]

It is straightforward to verify that the Markov process for the three-satellite system shown in Fig. 1 has a closed-form solution, which is given by

$$\begin{aligned} P(\underline{n}) &= P(n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33}) \\ &= \frac{1}{G} \frac{\rho_{11}^{n_{11}}}{n_{11}!} \frac{\rho_{12}^{n_{12}}}{n_{12}!} \frac{\rho_{13}^{n_{13}}}{n_{13}!} \frac{\rho_{22}^{n_{22}}}{n_{22}!} \frac{\rho_{23}^{n_{23}}}{n_{23}!} \frac{\rho_{33}^{n_{33}}}{n_{33}!} \end{aligned} \tag{10}$$

where G is the normalizing constant and $\rho_{ij} = \lambda_{ij}/\mu_{ij}, i, j = 1, 2, 3$, is the offered load of calls from satellite i to satellite j. As we can see, the solution is the product of six terms of the form $\rho_{ij}^{n_{ij}}/n_{ij}!, i, j = 1, 2, 3$, each corresponding to one of the six different types of calls. Therefore, it is easily generalizable to a k-satellite system, $k > 3$.

An alternative way is to regard this Markov process as describing a network of six $M/M/K/K$ queues, one for each type of calls between the three satellites. Since the satellites do not move, there are no handoffs, and as a consequence customers do not move from one queue to another (we will see in Section IV-B that handoffs may be modeled by allowing customers to move between the queues). Now, the probability that there are n customers in an $M/M/K/K$ queue is given by the familiar expression $(\rho^n/n!)/(\sum_{k=0}^{K} \rho^k/k!)$, and therefore, the probability that there are $(n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33})$ customers in the six queues is given by (10). Unlike previous studies reported in the literature, our model takes into account the fact that the six $M/M/K/K$ queues are not independent, since the number of customers accepted in each $M/M/K/K$ queue depends on the number of customers in other queues, as described by (4)–(9).

Of course, the main concern in any product-form solution is the computation of the normalizing constant

$$G = \sum_{\underline{n}} \frac{\rho_{11}^{n_{11}}}{n_{11}!} \frac{\rho_{12}^{n_{12}}}{n_{12}!} \frac{\rho_{13}^{n_{13}}}{n_{13}!} \frac{\rho_{22}^{n_{22}}}{n_{22}!} \frac{\rho_{23}^{n_{23}}}{n_{23}!} \frac{\rho_{33}^{n_{33}}}{n_{33}!} \tag{11}$$

where the sum is taken over all vectors $\underline{n}$ that satisfy (4)–(9). We now show how to compute the normalizing constant G in an efficient manner.

We can write $P(\underline{n})$ as

$$\begin{aligned} &P(n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33}) \\ &\quad = P(n_{11}, n_{22}, n_{33} \,|\, n_{12}, n_{13}, n_{23}) P(n_{12}, n_{13}, n_{23}) \\ &\quad = P(n_{11} \,|\, n_{12}, n_{13}, n_{23}) P(n_{22} \,|\, n_{12}, n_{13}, n_{23}) \\ &\qquad \times P(n_{33} \,|\, n_{12}, n_{13}, n_{23}) P(n_{12}, n_{13}, n_{23}) \\ &\quad = P(n_{11} \,|\, n_{12}, n_{13}) P(n_{22} \,|\, n_{12}, n_{23}) \\ &\qquad \times P(n_{33} \,|\, n_{13}, n_{23}) P(n_{12}, n_{13}, n_{23}) \end{aligned} \tag{12}$$

The second step in (12) is due to the fact that once the values of random variables n_{12}, n_{13}, n_{23}, representing the number of calls in each of the intersatellite links, are fixed, then the random variables n_{11}, n_{22}, and n_{33} are independent of each other (refer also to Fig. 1). The third step in (12) is due to the fact that random variable n_{11} depends on n_{12} and n_{13} and is independent of the random variable n_{23}; similarly for random variables n_{22} and n_{33}.

When we fix the values of the random variables n_{12} and n_{13}, the number of up-and-down calls in satellite 1 is described by an $M/M/K/K$ loss system, and thus

$$P(n_{11} \,|\, n_{12}, n_{13}) = \sum_{0 \leq 2n_{11} \leq C_{\rm UDL} - n_{12} - n_{13}} \frac{\rho_{11}^{n_{11}}}{n_{11}!}. \tag{13}$$

Similar expressions can be obtained for $P(n_{22} \,|\, n_{12}, n_{23})$ and $P(n_{33} \,|\, n_{13}, n_{23})$, corresponding to satellites 2 and 3, respectively. We can now rewrite (11) for the normalizing constant as follows:

$$\begin{aligned} G = &\sum_{0 \leq n_{12}, n_{13}, n_{23} \leq \min\{C_{\rm UDL}, C_{\rm ISL}\}} \frac{\rho_{12}^{n_{12}} \rho_{13}^{n_{13}} \rho_{23}^{n_{23}}}{n_{12}! n_{13}! n_{23}!} \\ &\times \left[\left(\sum_{0 \leq 2n_{11} \leq C_{\rm UDL} - n_{12} - n_{13}} \frac{\rho_{11}^{n_{11}}}{n_{11}!} \right) \right. \\ &\times \left(\sum_{0 \leq 2n_{22} \leq C_{\rm UDL} - n_{12} - n_{23}} \frac{\rho_{22}^{n_{22}}}{n_{22}!} \right) \\ &\left. \times \left(\sum_{0 \leq 2n_{33} \leq C_{\rm UDL} - n_{13} - n_{23}} \frac{\rho_{33}^{n_{33}}}{n_{33}!} \right) \right]. \end{aligned} \tag{14}$$

Let $C = \max\{C_{\rm ISL}, C_{\rm UDL}\}$. Using (14), we can see that the normalizing constant can be computed in $O(C^3)$ time rather than the $O(C^6)$ time required by a brute-force enumeration of all states, a significant improvement in efficiency.

Once the value of the normalizing constant is obtained, we can compute blocking probabilities by summing up all the appropriate blocking states. Consider the three-satellite orbit of Fig. 1. The probability that a call that either originates or terminates at satellite 1 will be blocked on the up- and downlink of that satellite is given by

$$P_{\rm UDL_1} = \sum_{2n_{11} + n_{12} + n_{13} = C_{\rm UDL}} P(\underline{n}) \tag{15}$$

while the probability that a call originating at satellite i (or satellite j) and terminating at satellite j (or i) will be blocked by the intersatellite link (i, j) is

$$P_{{\rm ISL}_{ij}} = \begin{cases} 0, & C_{\rm ISL} > C_{\rm UDL} \\ \sum_{n_{ij} = C_{\rm ISL}} P(\underline{n}), & \text{otherwise.} \end{cases} \tag{16}$$

Once the blocking probabilities on all up-and-down and intersatellite links have been obtained using expressions similar to (15) and (16), the blocking probability of calls between any two satellites can be easily obtained. We note that (15) and (16) explicitly enumerate all relevant blocking states, and thus, they involve summations over appropriate parts of the state space

[1]When there are more than three satellites in an orbit, calls between a number of satellite pairs may share a given intersatellite link. Consequently, the constraints of a k-satellite orbit, $k > 3$, corresponding to (7)–(9) will be similar to constraints (4)–(6), in that the left-hand side will involve a summation over a number of calls. In this case, blocking on intersatellite links may occur even if $C_{\rm ISL} \geq C_{\rm UDL}$.

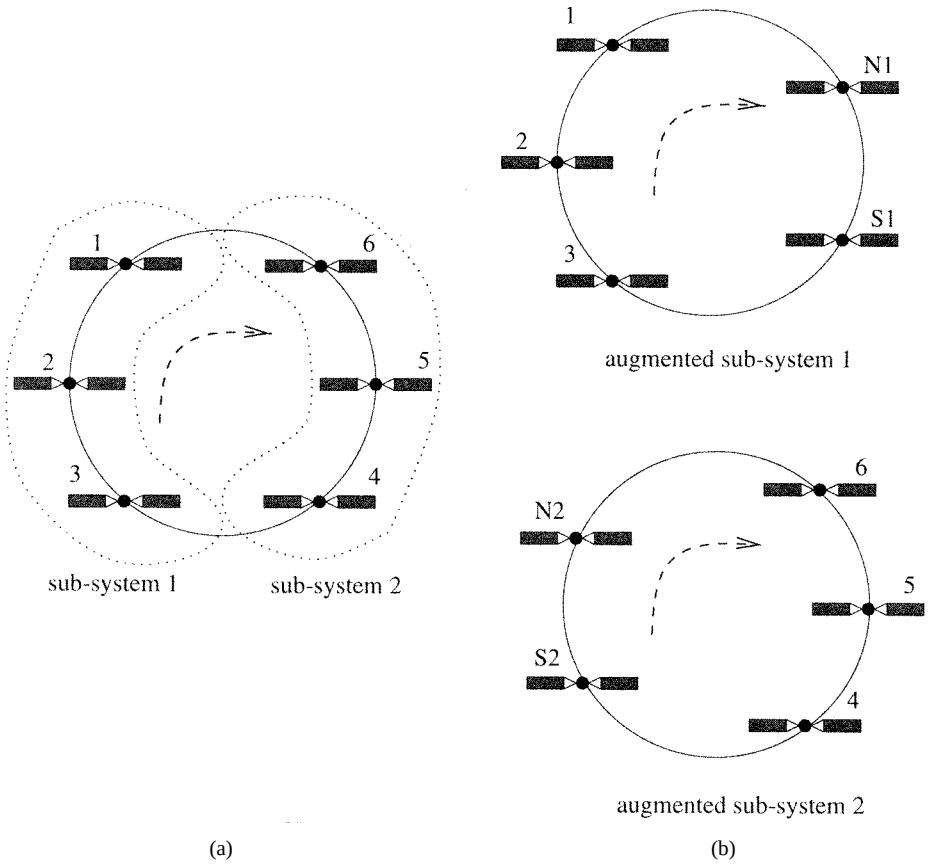

Fig. 2. (a) Original six-satellite orbit and (b) augmented subsystems.

of the Markov process for the satellite orbit. Consequently, direct computation of the link blocking probabilities using these expressions can be computationally expensive. We have been able to express the up-and-down and intersatellite link blocking probabilities in a way that allows us to compute these probabilities as a byproduct of the computation of the normalizing G. As a result, all blocking probabilities in a satellite orbit can be computed in an amount of time that is equal to the time needed to obtain the normalizing constant, plus a constant. The derivation of the expressions for the link blocking probabilities is a straightforward generalization of the technique employed in (12) and is omitted.

III. A Decomposition Algorithm for the No-Handoffs Case

Let k be the number of satellites in a single orbit and N the number of random variables in the state description of the corresponding Markov process, $N = k(k+1)/2$. Using the method described above, we can compute the normalizing constant G in time $O(C^{N-k})$, as opposed to time $O(C^N)$ needed by a brute-force enumeration of all states. Although the *improvement* in the running time provided by our method for computing G increases with k, the value of N will dominate for large values of k. Numerical experiments with the above algorithm indicate that this method is limited to $k = 5$ satellites. That is, it takes an amount of time on the order of a few minutes to compute the normalizing constant G for five satellites. Thus, a different method is needed for analyzing realistic constellations of LEO satellites.

In this section, we present a method to analyze a single orbit with k satellites, $k > 5$, by decomposing the orbit into subsystems of three or fewer satellites. Each subsystem is analyzed separately, and the results obtained by the subsystems are combined using an iterative scheme.

To explain how the decomposition algorithm works, let us consider the case of a six-satellite orbit, as shown in Fig. 2(a). This orbit is divided into two subsystems. Subsystem 1 consists of satellites 1, 2, and 3, and subsystem 2 consists of satellites 4, 5, and 6. To analyze subsystem 1 in isolation, we need to have some information from subsystem 2. Specifically, we need to know the probability that a call originating at a satellite in subsystem 1 and terminating at a satellite in subsystem 2 will be blocked due to lack of capacity in a link in subsystem 2. Also, we need to know the number of calls originating from subsystem 2 and terminating in subsystem 1. Similar information is needed from subsystem 1 in order to analyze subsystem 2.

In view of this, each subsystem is augmented to include two fictitious satellites, which represent the aggregate behavior of the other subsystem. In subsystem 1, we add two new satellites, which we call N1 and S1, as shown in Fig. 2(b). A call originating at a satellite $i, i = 1, 2, 3$, and terminating at a satellite $j, j = 4, 5, 6$, will be represented by a call from i to one of the fictitious satellites (N1 or S1). Depending upon i and j, this call may be routed differently. For instance, let us assume that $i = 2$ and $j = 4$. Then, in our augmented subsystem 1, this

Decomposition Algorithm for A Single Orbit

A 6-satellite orbit is decomposed into two 3-satellite sub-systems as in Figure 2. Sub-system 1 consists of satellites 1 to 3 in the original orbit plus fictitious satellites N1 and S1, while sub-system 2 consists of satellites 4 to 6 of the original orbit plus fictitious satellites N2 and S2.

1. begin
2. $h \leftarrow 0$ //Initialization step
   // $p_{ij}(h)$ is the probability that an inter-sub-system call
   // will be blocked in sub-system 1
   // $q_{ij}(h)$ is the probability that an inter-sub-system call
   // will be blocked in sub-system 2
   $q_{ij}(h) \leftarrow 0, \quad 1 \leq i \leq 3 < j \leq 6$
3. $h \leftarrow h + 1$ //h-th iteration
4. $\lambda_{ij}(h) \leftarrow \lambda_{ij}, \quad 1 \leq i \leq j \leq 3$ //Sub-system 1
   $\lambda_{1,N1} = (1 - q_{16})\lambda_{16} + (1 - q_{15})\lambda_{15}$
   $\lambda_{1,S1} = (1 - q_{14})\lambda_{14}$
   $\lambda_{2,N1} = (1 - q_{26})\lambda_{26} + (1 - q_{25})\lambda_{25}$
   $\lambda_{2,S1} = (1 - q_{24})\lambda_{24}$
   $\lambda_{3,N1} = (1 - q_{36})\lambda_{36}$
   $\lambda_{3,S1} = (1 - q_{34})\lambda_{34} + (1 - q_{35})\lambda_{35}$
   Solve sub-system 1 as described in the text to obtain
   new values for $p_{ij}(h)$
5. $\lambda_{ij}(h) \leftarrow \lambda_{ij}, \quad 4 \leq i \leq j \leq 6$ //Sub-system 2
   $\lambda_{N2,4} = 0$
   $\lambda_{S2,4} = (1 - p_{14})\lambda_{14} + (1 - p_{24})\lambda_{24} + (1 - p_{34})\lambda_{34}$
   $\lambda_{N2,5} = (1 - p_{15})\lambda_{15} + (1 - p_{25})\lambda_{25}$
   $\lambda_{S2,5} = (1 - p_{35})\lambda_{35}$
   $\lambda_{N2,6} = (1 - p_{16})\lambda_{16} + (1 - p_{26})\lambda_{26} + (1 - p_{36})\lambda_{36}$
   $\lambda_{S2,6} = 0$
   Solve sub-system 2 as described in the text to obtain
   new values for $q_{ij}(h)$
6. Repeat from Step 3 until the blocking prob. converge
7. end of the algorithm

Fig. 3. Decomposition algorithm for a single orbit of a satellite constellation.

call will be routed to satellite S1 through satellite 3. However, if $j = 6$, the call will be routed to satellite N1 through satellite 1.[2] In other words, satellite N1 (respectively, S1) in the augmented subsystem 1 is the destination for calls of the original orbit that originate from satellite $i, i = 1, 2, 3$ and are routed to satellite $j, j = 4, 5, 6$ in the clockwise (respectively, counterclockwise) direction in Fig. 2(a). Similarly, calls originating from satellite $j, j = 4, 5, 6$, to satellite $i, i = 1, 2, 3$, and routed in the counterclockwise (respectively, clockwise) direction are represented in subsystem 1 as calls originating from N1 (respectively, S1) to i. Again, the originating satellite (N1 or S1) for the call depends on the values of i and j and on the path that the call follows in the original six-satellite orbit.

[2]While this discussion assumes shortest path routing, our model can handle any fixed-routing scheme.

Subsystem 2 is likewise augmented to include two fictitious satellites N2 and S2 [see Fig. 2(b)], which represent the aggregate behavior of subsystem 1. Satellites N2 and S2 become the origin and destination of calls traveling from subsystem 2 to subsystem 1, and vice versa, in a manner similar to N1 and S1 described above.

A summary of our iterative algorithm is provided in Fig. 3. Below, we describe the decomposition algorithm using the six-satellite orbit shown in Fig. 2(a). Recall that $\lambda_{ij}, 1 \leq i \leq j$, is the arrival rate of calls between satellites i and j. For analyzing the augmented subsystems in Fig. 2(b), we will introduce the new arrival rates $\lambda_{i,N1}, \lambda_{i,S1}, \lambda_{N2,j}$, and $\lambda_{S2,j}, i = 1, 2, 3, j = 4, 5, 6$. Specifically, $\lambda_{i,N1}$ (respectively, $\lambda_{i,S1}$) accounts for all calls between satellite $i, i = 1, 2, 3$, and a satellite in subsystem 2 that are routed in the clockwise (respectively, counterclock-

wise) direction. Similarly, $\lambda_{N2,j}$ (respectively, $\lambda_{S2,j}$) accounts for all calls between subsystem 1 and satellite $j, j = 4, 5, 6$, that are routed in the clockwise (respectively, counterclockwise) direction.

Initially, we solve subsystem 1 in isolation using

$$\lambda_{1,N1} = (1 - q_{16})\lambda_{16} + (1 - q_{15})\lambda_{15} \quad (17)$$
$$\lambda_{1,S1} = (1 - q_{14})\lambda_{14} \quad (18)$$
$$\lambda_{2,N1} = (1 - q_{26})\lambda_{26} + (1 - q_{25})\lambda_{25} \quad (19)$$
$$\lambda_{2,S1} = (1 - q_{24})\lambda_{24} \quad (20)$$
$$\lambda_{3,N1} = (1 - q_{36})\lambda_{36} \quad (21)$$
$$\lambda_{3,S1} = (1 - q_{34})\lambda_{34} + (1 - q_{35})\lambda_{35}. \quad (22)$$

Quantity $q_{ij}, 1 \leq i \leq 3 < j \leq 6$, represents the current estimate of the probability that a call between a satellite i in subsystem 1 and a satellite j in subsystem 2 will be blocked due to the lack of capacity in a link of subsystem 2. For the first iteration, we use $q_{ij} = 0$ for all i and j; how these values are updated in subsequent iterations will be described shortly. Thus, the term $(1-q_{16})\lambda_{16}$ in (17) represents the *effective* arrival rate of calls between satellites 1 and 6, as seen by subsystem 1; similarly for the other terms in (17)–(22).

The solution to the first subsystem yields an initial value for the probability $p_{ij}, 1 \leq i \leq 3 < j \leq 6$, that a call between a satellite i in subsystem 1 and a satellite j in subsystem 2 will be blocked due to lack of capacity in a link of subsystem 1. Therefore, the effective arrival rates of calls between, say, satellite 1 and satellite 4 that are offered to subsystem 2 can be initially estimated as $(1 - p_{14})\lambda_{14}$. We can now solve subsystem 2 in isolation using[3]

$$\lambda_{N2,4} = 0 \quad (23)$$
$$\lambda_{S2,4} = (1 - p_{14})\lambda_{14} + (1 - p_{24})\lambda_{24} + (1 - p_{34})\lambda_{34} \quad (24)$$
$$\lambda_{N2,5} = (1 - p_{15})\lambda_{15} + (1 - p_{25})\lambda_{25} \quad (25)$$
$$\lambda_{S2,5} = (1 - p_{35})\lambda_{35} \quad (26)$$
$$\lambda_{N2,6} = (1 - p_{16})\lambda_{16} + (1 - p_{26})\lambda_{26} + (1 - p_{36})\lambda_{36} \quad (27)$$
$$\lambda_{S2,6} = 0. \quad (28)$$

Based on the above discussion, $\lambda_{S2,4}$ in (24) represents the effective arrival rate of calls between a satellite in subsystem 1 and satellite 4, as seen by subsystem 2. Expressions (23)–(28) can be explained in a similar manner. The solution to the second subsystem provides an estimate of the blocking probabilities $q_{ij}, 1 \leq i \leq 3 < j \leq 6$, that calls between satellites in the two subsystems will be blocked due to lack of capacity in a link of subsystem 2.

The new estimates for q_{ij} are then used in (17) to (22) to update the arrival rates to the two fictitious satellites of augmented subsystem 1. Subsystem 1 is then solved again, and the estimates p_{ij} are updated and used in (23)–(28) to obtain new arrival rates for the fictitious satellites of subsystem 2. This leads to an iterative scheme, where the two subsystems are solved successively until a convergence criterion (e.g., in terms of the values of the call-blocking probabilities) is satisfied.

Orbits consisting of any number $k > 5$ of satellites can be decomposed into a number of subsystems, each consisting of three satellites of the original orbit (the last subsystem may consist of fewer than three satellites). The decomposition method is similar to the one above, in that for subsystem l, the remaining satellites are aggregated to two fictitious satellites. Each subsystem is analyzed in succession as described above. We note that when employing the decomposition algorithm, the selection of the subsystem size will depend on the number of satellites in the original orbit and how efficiently we can calculate the exact solution of the Markov process associated with each subsystem. It is well known in decomposition algorithms that the larger the individual subsystems that have to be analyzed in isolation, the better the accuracy of the decomposition algorithm. Thus, as we mentioned above, we have decided to decompose an orbit into subsystems of the largest size (three of the original satellites plus two fictitious ones) for which we can efficiently analyze the Markov process, plus, possibly, a subsystem of smaller size if the number of satellites is not a multiple of three. Note also that proving convergence for decomposition algorithms of the type presented here is a difficult task, and we have not been able to show that the decomposition algorithm will always converge. However, in our experimentation with a wide range of traffic patterns and orbit sizes, we have found that the blocking probabilities converge to within 10^{-6} in only a handful of iterations using our algorithm.

IV. Modeling Handoffs

A. Earth-Fixed Coverage

Let us now turn to the problem of determining blocking probabilities in a single orbit of satellites with earth-fixed coverage; such a coverage is provided by the Teledesic constellation. Let k denote the number of satellites in the orbit. In this case, we assume that the earth is divided into k fixed cells (footprints) and that time is divided in intervals of length T such that, during a given interval, each satellite serves a certain cell by continuously redirecting its beams. At the end of each interval, i.e., every T time units, all satellites simultaneously redirect their beams to serve the next footprint along their orbit. They also hand off currently served calls to the next satellite in the orbit.

We make the following observations about this system. Handoff events are periodic with a period of T time units, and handoffs take place in bulk at the end of each period. Also, there is no call blocking due to handoffs, since, at each handoff event, a satellite passes its calls to the one following it and simply inherits the calls of the satellite ahead of it. Finally, within each period T, the system can be modeled as one with no handoffs, such as the one described in the previous section. Given that the period T is equal to the orbit period (approximately 90 min) divided by the number of satellites, we can assume that the system reaches steady state within the period, and thus, the initial conditions (i.e., the number of calls inherited by each satellite at the beginning of the period) do not affect its behavior.

[3]In (23), we have that $\lambda_{N2,4} = 0$ because we assume that calls between satellites in subsystem 1 and satellite 4 are routed in the counterclockwise direction; similarly for (28).

Now, since every T units of time, each satellite assumes the traffic carried by the satellite ahead, from the point of view of an observer on the earth, this system appears to be as if the satellites are permanently fixed over their footprints. Hence, we can use the decomposition algorithm presented above to analyze this system.

B. Satellite-Fixed Coverage

Consider now satellite-fixed cell coverage. As a satellite moves, its footprint on the earth (the cell served by the satellite) also moves with it. As customers move out of the footprint area of a satellite, their calls are handed off to the satellite following it from behind. In general, the footprints of adjacent satellites overlap with each other, and customers in an overlapping area may have access to both satellites. We assume that a call is handed off from its current satellite to the satellite following it in the orbit when it is determined that the latter provides higher signal quality. Therefore, for the purposes of the following discussion, we define the footprint of a satellite as the area on the earth in which the signal from this satellite is of higher quality than the signal (if any) received by another satellite.

To model handoffs in this case, we make the assumption that potential customers are uniformly distributed over the part of the earth served by the satellites in the orbit. This assumption has the following two consequences.

1) The arrival rate λ to each satellite remains constant as it moves around the earth. Then, the arrival rate of calls between satellite i and satellite j is given by $\lambda_{ij} = \lambda r_{ij}$, where r_{ij} is the probability that a call originating by a customer served by satellite i is for a customer served by satellite j.
2) The active customers served by a satellite can be assumed to be uniformly distributed over the satellite's footprint. As a result, the rate of handoffs from satellite i to satellite j that is following from behind is proportional to the number of calls at satellite i.

Clearly, the assumption that customers are uniformly distributed (even within an orbit) is an approximation. In Section VI, we will discuss how the results presented in this section can be extended to accurately model the situation when customers are not uniformly distributed.

Let A denote the area of a satellite's footprint and v denote a satellite's speed. As a satellite moves around the earth, within a time interval of length Δt, its footprint will move a distance of ΔL, as shown in Fig. 4. Calls involving customers located in the part of the original footprint of area ΔA (the handoff area) that is no longer served by the satellite are handed off to the satellite following it. Let $\Delta A = A\beta\Delta L$, where β depends on the shape of the footprint. Because of the assumption that active customers are uniformly distributed over the satellite's footprint, the probability q that a customer will be handed off to the next satellite along the sky within a time interval of length Δt is

$$q = \frac{\Delta A}{A} = \beta \Delta L = \beta v \Delta t. \tag{29}$$

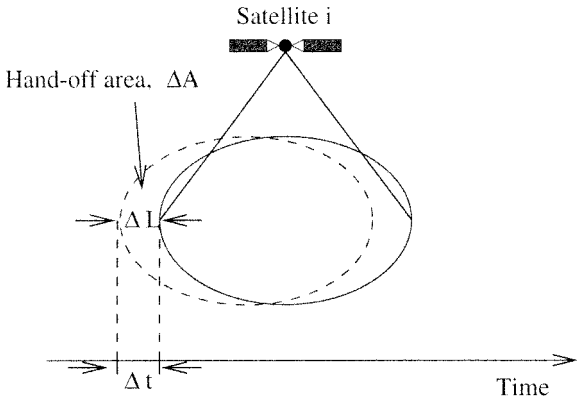

Fig. 4. Calculation of the handoff probability.

Define $\alpha = \beta v$. Then, when there are n customers served by a satellite, the *rate* of handoffs to the satellite following it will be αn.

Let us now return to the three-satellite orbit (see Fig. 1) and introduce handoffs. This system can be described by a continuous-time Markov process with the same number of random variables as the no-handoffs model of Section II (i.e., $n_{11}, \ldots, n_{33}$) and the same transition rates [(2) and (3)], but with a number of additional transition rates to account for handoffs. We will now derive the transition rates due to handoffs.

Consider calls between a customer served by satellite 1 and a customer served by satellite 2. There are n_{12} such calls serving $2n_{12}$ customers: n_{12} customers on the footprint of satellite 1 and n_{12} on the footprint of satellite 2. Consider a call between customer A and customer B served by satellite 1 and 2, respectively. The probability that customer A will be in the handoff area of satellite 1 but customer B will not be in the handoff area of satellite 2 is $q(1-q) = q - q^2$. But, from (29), we have that $\lim_{\Delta t \to 0}(q^2)/(\Delta t) = 0$, so the rate at which these calls experience a handoff from satellite 1 to satellite 3 that follows it is αn_{12}. Let $\underline{n} = (n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33})$, and define $\underline{1}_{ij}$ as a vector of zeroes for all variables except variable n_{ij}, which is one. Based on the above discussion, we thus have

$$r(\underline{n}, \underline{n} - \underline{1}_{12} + \underline{1}_{23}) = \alpha n_{12}, \quad n_{12} > 0. \tag{30}$$

Similarly, the probability that customer B will be in the handoff area of satellite 2 but customer A will not be in the handoff area of satellite 1 is $q(1-q) = q - q^2$. Thus, the rate at which these calls experience a handoff from satellite 2 to satellite 1 that follows it is again αn_{12}

$$r(\underline{n}, \underline{n} - \underline{1}_{12} + \underline{1}_{11}) = \alpha n_{12}, \quad n_{12} > 0. \tag{31}$$

On the other hand, the probability that both customers A and B are in the handoff area of their respective satellites is q^2, which, from (29), is $o(\Delta t)$, and thus simultaneous handoffs are not allowed.

Now consider calls between customers that are both served by the same satellite, say, satellite 1. There are n_{11} such calls serving $2n_{11}$ customers. The probability that exactly one of

the customers of a call is in the handoff area of satellite 1 is $2q(1-q)$, so the rate at which these calls experience handoffs (involving a single customer) to satellite 3 is $2\alpha n_{11}$

$$r(\underline{n}, \underline{n} - \underline{1}_{11} + \underline{1}_{13}) = 2\alpha n_{11}, \quad n_{11} > 0. \tag{32}$$

As before, the probability that both customers of the call are in the handoff area of satellite 1 is q^2, and again, no simultaneous handoffs are allowed.

The transition rates involving the other four random variables in the state description (1) can be derived using similar arguments. For completeness, these transition rates are provided in (33)–(38)

$$r(\underline{n}, \underline{n} - \underline{1}_{13} + \underline{1}_{12}) = \alpha n_{13}, \quad n_{13} > 0 \tag{33}$$
$$r(\underline{n}, \underline{n} - \underline{1}_{13} + \underline{1}_{33}) = \alpha n_{13}, \quad n_{13} > 0 \tag{34}$$
$$r(\underline{n}, \underline{n} - \underline{1}_{22} + \underline{1}_{12}) = 2\alpha n_{22}, \quad n_{22} > 0 \tag{35}$$
$$r(\underline{n}, \underline{n} - \underline{1}_{23} + \underline{1}_{13}) = \alpha n_{23}, \quad n_{23} > 0 \tag{36}$$
$$r(\underline{n}, \underline{n} - \underline{1}_{23} + \underline{1}_{22}) = \alpha n_{23}, \quad n_{23} > 0 \tag{37}$$
$$r(\underline{n}, \underline{n} - \underline{1}_{33} + \underline{1}_{23}) = 2\alpha n_{33}, \quad n_{33} > 0. \tag{38}$$

From the queueing point of view, this system is the queueing network of $M/M/K/K$ queues described in Section II, where customers are allowed to move between queues according to (30)–(38). (Recall that in the queueing model of Section II, customers are not allowed to move from node to node.) This queueing network has a product-form solution similar to (10). Let γ_{ij} denote the total arrival rate of calls between satellites i and j, including new calls (at a rate of λ_{ij}) and handoff calls (arriving at an appropriate rate). The values of γ_{ij} can be obtained by solving the traffic equations for the queueing network. Let also $\nu_{ij} n_{ij}$ be the departure rate when there are n_{ij} of these calls, including call termination (at a rate of $\mu_{ij} n_{ij}$) and call handoff (at a rate of $2\alpha n_{ij}$). Also, define $\rho'_{ij} = \gamma_{ij}/\nu_{ij}$. Then, the solution for this queueing network is given by

$$\begin{aligned} P(\underline{n}) &= P(n_{11}, n_{12}, n_{13}, n_{22}, n_{23}, n_{33}) \\ &= \frac{1}{G} \frac{(\rho'_{11})^{n_{11}}}{n_{11}!} \frac{(\rho'_{12})^{n_{12}}}{n_{12}!} \frac{(\rho'_{13})^{n_{13}}}{n_{13}!} \frac{(\rho'_{22})^{n_{22}}}{n_{22}!} \frac{(\rho'_{23})^{n_{23}}}{n_{23}!} \\ &\quad \times \frac{(\rho'_{33})^{n_{33}}}{n_{33}!} \end{aligned} \tag{39}$$

which is identical to (10) except that ρ_{ij} has been replaced by ρ'_{ij}.

The product-form solution (39) can be generalized in a straightforward manner for any k-satellite orbit, $k > 3$. We can thus use the techniques developed in Section II to solve the system involving handoffs exactly, or we can use the decomposition algorithm presented in Section III to solve orbits with a large number of satellites.

V. Numerical Results

In this section, we validate both the exact model and the decomposition algorithm by comparing to simulation results. In the figures presented, simulation results are plotted along with 95% confidence intervals estimated by the method of replications. The number of replications is 30, with each simulation run lasting until each type of call has at least 15 000 arrivals. For the approximate results, the iterative decomposition algorithm terminates when all call-blocking probability values have converged within 10^{-6}. We implemented our own simulation software for the satellite system; both the simulation and the iterative algorithm were run on a SUN Sparc-10 workstation.

For the results presented here, we consider three different traffic patterns; similar results have been obtained for several other patterns. Let r_{ij} denote the probability that a call originating by a customer served by satellite i is for a customer served by satellite j.[4] The first pattern is a uniform traffic pattern such that

$$r_{ij} = \frac{1}{k} \quad \forall i, j \text{ (uniform pattern)} \tag{40}$$

where k is the number of satellites. The second is a pattern based on the assumption of traffic locality. Specifically, it assumes that most calls originating at a satellite i are to users in satellites $i-1$, i, and $i+1$, where addition and subtraction is modulo-k for a k-satellite orbit

$$r_{ij} = \begin{cases} 0.3, & j = i-1, i, i+1 \\ \dfrac{0.1}{k-3}, & j \neq i-1, i, i+1 \end{cases} \text{ (locality pattern).} \tag{41}$$

The third pattern is such that there are two communities of users, and most traffic is between users within a given community (e.g., satellites over different hemispheres of the earth)

$$r_{ij} = \begin{cases} \dfrac{0.8}{k/2}, & i, j = 1, \ldots, k/2 \\ \dfrac{0.8}{k/2}, & i, j = k/2+1, \ldots, k \\ \dfrac{0.2}{k/3}, & i = 1, \ldots, k/2, \quad j = k/2+1, \ldots, k \\ \dfrac{0.2}{k/3}, & j = 1, \ldots, k/2, \quad i = k/2+1, \ldots, k \end{cases} \tag{42}$$

A. Results With the Exact Model

In this section, we present results with the exact Markov process model for the no-handoffs case developed in Section II. Recall that we can directly compute the normalizing constant G using (14) for orbits of up to five satellites. Thus, we obtain the blocking probability values by solving the exact Markov process for a five-satellite orbit and the three traffic patterns discussed above.

Fig. 5 plots the blocking probability against the capacity C_{UDL} of up- and downlinks, when the arrival rate $\lambda = 10$ and the capacity of intersatellite links $C_{\text{ISL}} = 10$, for the uniform traffic pattern. Three sets of plots are shown: one for calls originating and terminating at the same satellite (referred to as "local calls" in the figure), one for calls traveling over a single intersatellite link, and one for calls traveling over two

[4]Our objective in this section is simply to demonstrate the accuracy of the exact model of Section II and the decomposition algorithm of Section III, both of which assume that satellites are fixed in the sky. Thus, the traffic patterns we consider can be thought of as being "attached" to either the satellites or the areas of the earth covered by each satellite.

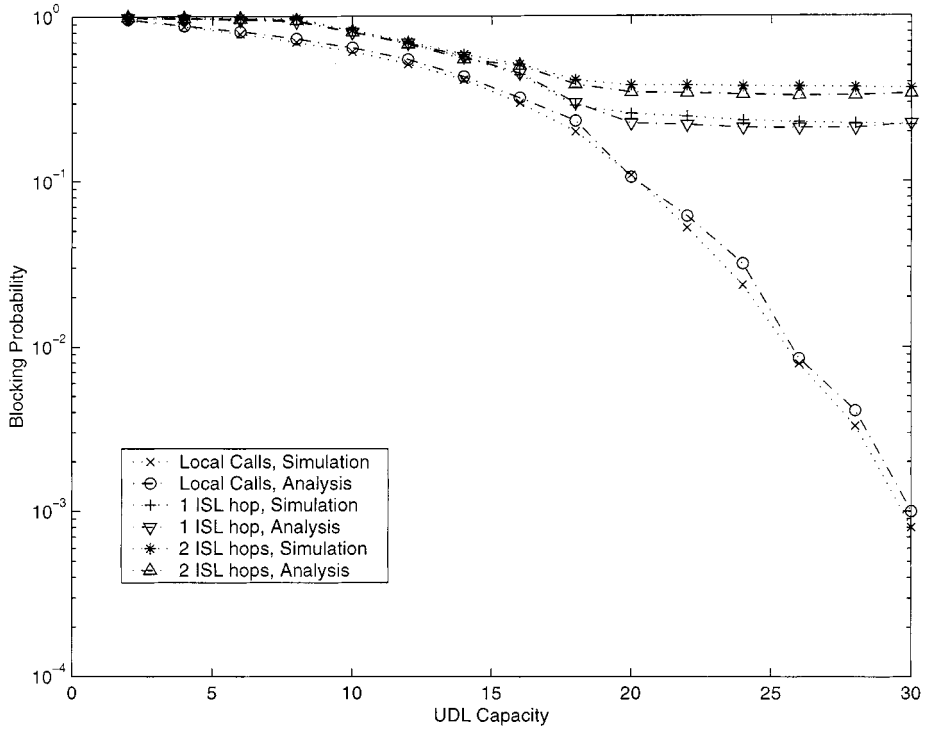

Fig. 5. Call-blocking probabilities for a five-satellite orbit, $\lambda = 10, C_{\rm ISL} = 10$, uniform pattern.

intersatellite links.[5] Each set consists of two plots: one corresponding to blocking probability values obtained by solving the Markov process and one corresponding to simulation results (including simulation results is redundant, since we solve the Markov process exactly; however, we have decided to present them anyway).

From the figure, we observe that as the capacity $C_{\rm UDL}$ of up- and downlinks increases, the blocking probability of all calls decreases. However, for calls traveling over at least one intersatellite link, the blocking probability curve flattens out after an initial drop. This behavior is due to the fact that for small values of $C_{\rm UDL}$, the up- and downlinks represent a bottleneck; thus, increasing $C_{\rm UDL}$ reduces the call-blocking probability significantly. However, once $C_{\rm UDL}$ increases beyond a certain value, the intersatellite links become the bottleneck, and the blocking probability of calls that have to travel over these links is not affected further. On the other hand, the blocking probability of calls not using intersatellite links (i.e., those originating and terminating at the same satellite) decreases rapidly as $C_{\rm UDL}$ increases, dropping to zero for values $C_{\rm UDL} > 30$ (because of the logarithmic scale, values of zero cannot be shown in Fig. 5, so there are no values plotted when $C_{\rm UDL} > 30$ for the curves of local calls).

Fig. 6 plots the blocking probability for the same calls as in Fig. 5, against the capacity $C_{\rm ISL}$ of intersatellite links; for the results presented, we assume that $\lambda = 10$ and $C_{\rm UDL} = 20$. In this figure, we can see that as the value of $C_{\rm ISL}$ increases, the blocking probability of calls using intersatellite links decreases, as expected. However, the blocking probability of local calls (i.e., calls originating and terminating at the same satellite, which do not use intersatellite links) increases with increasing $C_{\rm ISL}$. This behavior can be explained by noting that as $C_{\rm ISL}$ increases, a larger number of nonlocal calls (i.e., calls using intersatellite links) is accepted (since their blocking probability decreases). Since both local and nonlocal calls compete for up- and downlinks, an increase in the number of nonlocal calls accepted will result in higher blocking probability for local calls. But when the value of $C_{\rm ISL}$ exceeds the value of $C_{\rm UDL}$ (which is equal to 20 in this case), the up- and downlinks become the bottleneck, and further increases in $C_{\rm ISL}$ have no effect on blocking probabilities.

Fig. 7 is similar to Fig. 5 except that the arrival rate is $\lambda = 5$ instead of 10 (all other parameters are as in Fig. 5). The behavior of the various curves is similar to that in Fig. 5. The main difference is that the blocking probabilities in Fig. 7 are significantly lower, a result that is expected due to the lower arrival rate.

Finally, Figs. 8 and 9 show results for the same parameters as in Fig. 5 but correspond to the locality and two-community traffic patterns, respectively. Again, the behavior of the curves is similar for all three figures, although the actual blocking probability values depend on the traffic pattern used.

The results in Figs. 5–9 illustrate the fact that the blocking probability values obtained by solving the Markov process match the simulation results; this is expected since the Markov process model we developed is exact. Thus, this model can be used to study the interplay between various system parameters (e.g., $C_{\rm ISL}$, $C_{\rm UDL}$, traffic pattern, etc.), and their effect on the call-blocking probabilities, in an efficient manner. We note that solving the Markov process takes only a few minutes, while

[5]These are the only possible types of calls in a five-satellite orbit and shortest path routing. Furthermore, because of symmetry, the results are the same regardless of the satellite at which the calls originate or terminate.

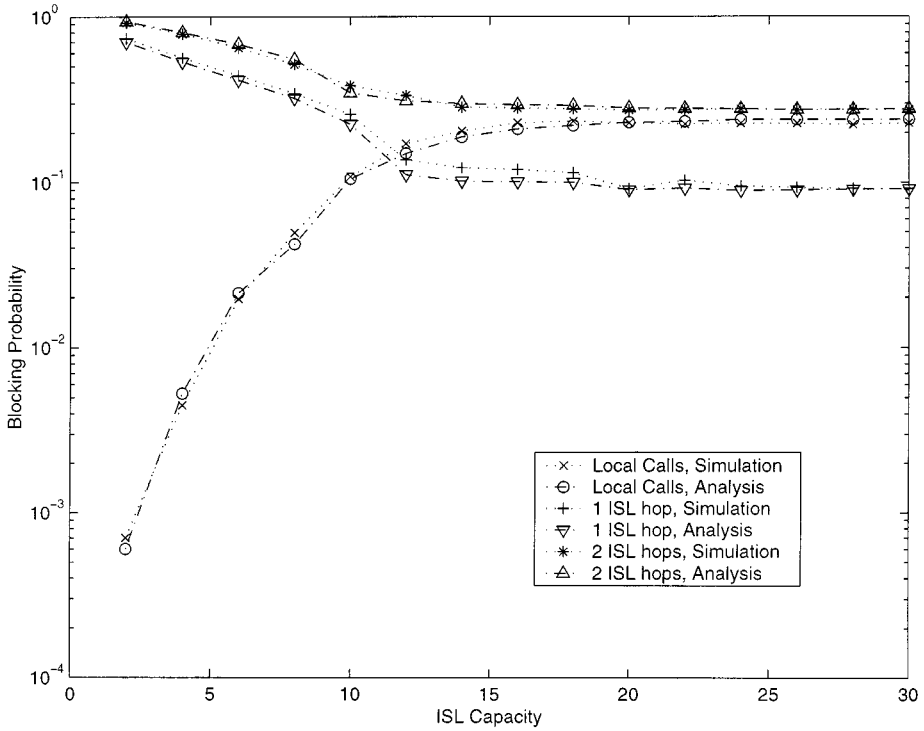

Fig. 6. Call-blocking probabilities for a five-satellite orbit, $\lambda = 10, C_{\mathrm{UDL}} = 20$, uniform pattern.

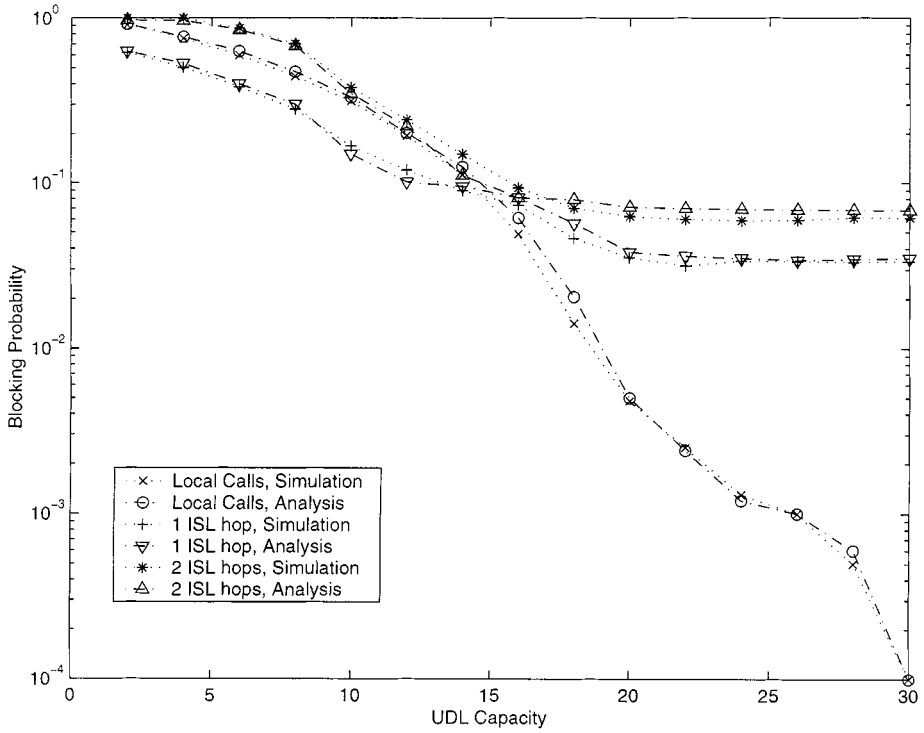

Fig. 7. Call-blocking probabilities for a five-satellite orbit, $\lambda = 5, C_{\mathrm{ISL}} = 10$, uniform pattern.

running the simulation takes anywhere between 30 min and several hours, depending on the value of the arrival rates.

B. Validation of the Decomposition Algorithm

We now validate the decomposition algorithm developed in Section III by comparing the blocking probabilities obtained by running the algorithm to simulation results. We consider a single orbit of a satellite constellation consisting of 12 satellites, a number representative of typical commercial satellite systems. In all cases studied, we have found that the algorithms converge in only a few (less than ten) iterations, taking a few minutes to terminate. On the other hand, simulation of 12-satellite orbits is quite expensive in terms of computation time, taking several hours to complete.

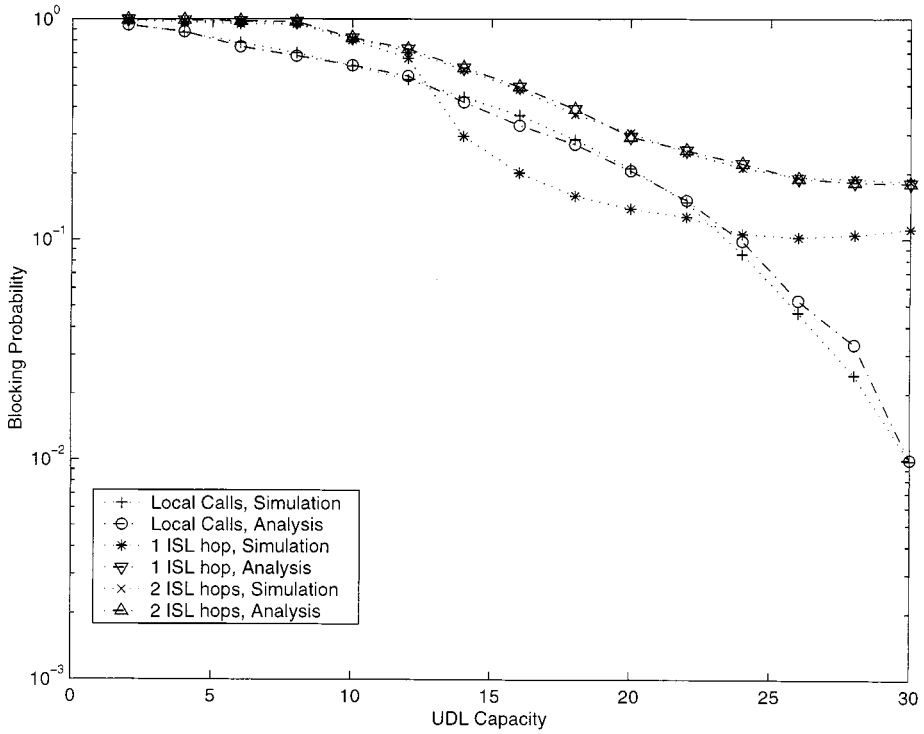

Fig. 8. Call-blocking probabilities for a five-satellite orbit, $\lambda = 10, C_{\mathrm{ISL}} = 10$, locality pattern.

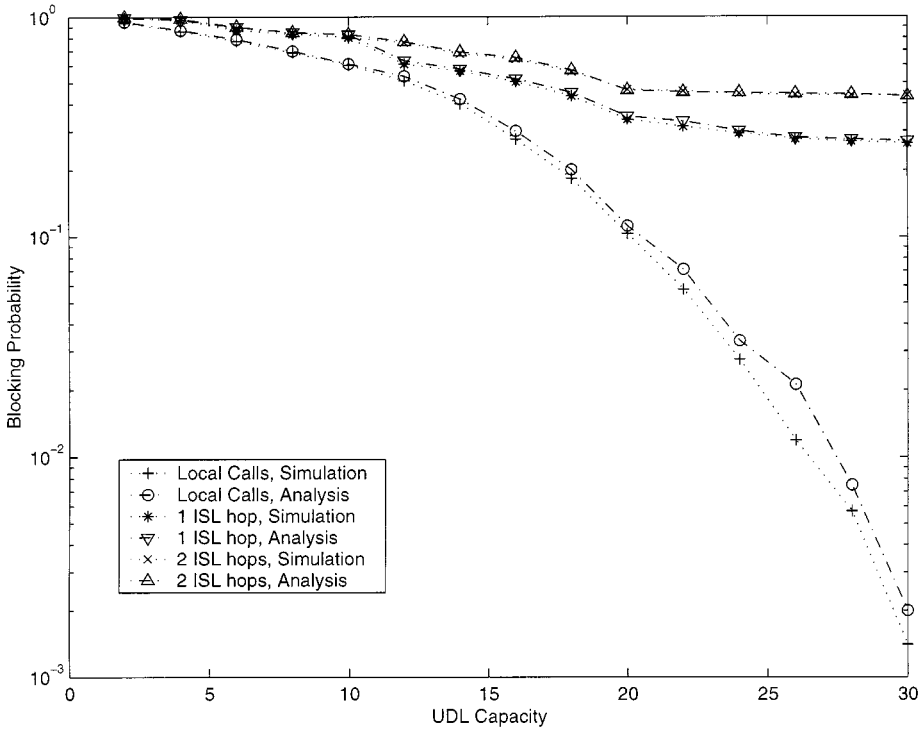

Fig. 9. Call-blocking probabilities for a five-satellite orbit, $\lambda = 10, C_{\mathrm{ISL}} = 10$, two-community pattern.

Fig. 10 plots the blocking probability against the capacity C_{UDL} of up- and downlinks, when the arrival rate $\lambda = 5$ and the capacity of intersatellite links $C_{\mathrm{ISL}} = 20$ for the uniform traffic pattern. Six sets of calls are shown: one for local calls and five for nonlocal calls. Each set consists of two plots: one corresponding to blocking probability values obtained by running the decomposition algorithm of Section III and one corresponding to simulation results. Each nonlocal call for which

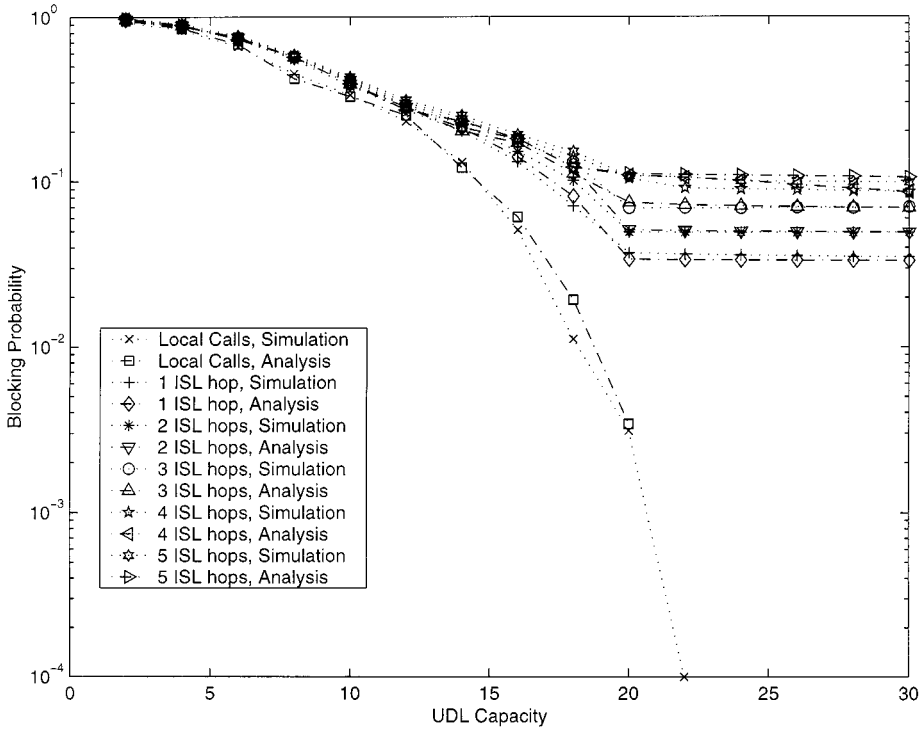

Fig. 10. Call-blocking probabilities for a 12-satellite orbit, $\lambda = 5, C_{\mathrm{ISL}} = 20$, uniform pattern.

results are shown travels over a different number of intersatellite links, from one to five. Thus, the results in Fig. 10 represent calls between all the different subsystems in which the 12-satellite orbit is decomposed by the decomposition algorithm.

From the figure, we observe the excellent agreement between the analytical results and simulation. The behavior of the curves can be explained by noting that when the capacity C_{UDL} of up- and downlinks is less than 20, these links represent a bottleneck. Thus, increasing the up- and downlink capacity results in a significant drop in the blocking probability for all calls. When $C_{\mathrm{UDL}} > 20$, however, the intersatellite links become the bottleneck, and nonlocal calls do not benefit from further increases in the up- and downlink capacity. We also observe that the larger the number of intersatellite links over which a nonlocal call must travel, the higher its blocking probability, as expected. The blocking probability of local calls, on the other hand, drops to zero for $C_{\mathrm{UDL}} > 20$, since they do not have to compete for intersatellite links.

Figs. 11 and 12 are similar to Fig. 10 but show results for the locality and two-community traffic patterns, respectively. For the results presented, we used $\lambda = 5$ and $C_{\mathrm{ISL}} = 10$, and we varied the value of C_{UDL}. We observe that the values of the call-blocking probabilities depend on the actual traffic pattern, but the behavior of the various curves is similar to that in Fig. 10. Finally, in Fig. 13, we fix the value of C_{UDL} to 20, and we plot the call-blocking probabilities for the two-community traffic pattern against the capacity C_{ISL} of the intersatellite links.

Overall, the results in Figs. 10–13 indicate that analytical results are in good agreement with simulation over a wide range of traffic patterns and system parameters. Thus, our decomposition algorithm can be used to estimate call-blocking probabilities in LEO satellite systems in an efficient manner.

VI. Concluding Remarks

We have presented an analytical model for computing blocking probabilities for a single orbit of a LEO satellite constellation. We have devised a method for solving the exact Markov process efficiently for up to five-satellite orbits. For orbits consisting of a larger number of satellites, we have developed an approximate decomposition algorithm to compute the call-blocking probabilities by decomposing the system into smaller subsystems and iteratively solving each subsystem in isolation using the exact Markov process. We have also shown how our approach can capture blocking due to handoffs for both satellite-fixed and earth-fixed orbits.

The analytical model we presented in this paper can be extended in several directions, some of which are the subject of current research. Assuming that a constellation of satellites consists of R orbits, a natural approach is to decompose it into R subsystems, each representing a single orbit, which are then solved using the techniques we developed here. When combining the solutions to the R subsystems, the traffic on interorbit links must also be accounted for. While in this paper we have considered a fixed routing scheme, alternate routing schemes can be modeled using the techniques we developed in [19, Section IV-B]. It is also possible to improve the performance of handoff calls by reserving a set of channels on each link for the exclusive use of these calls. Channel reservation can be modeled by a modified Markov process for the single subsystem studied

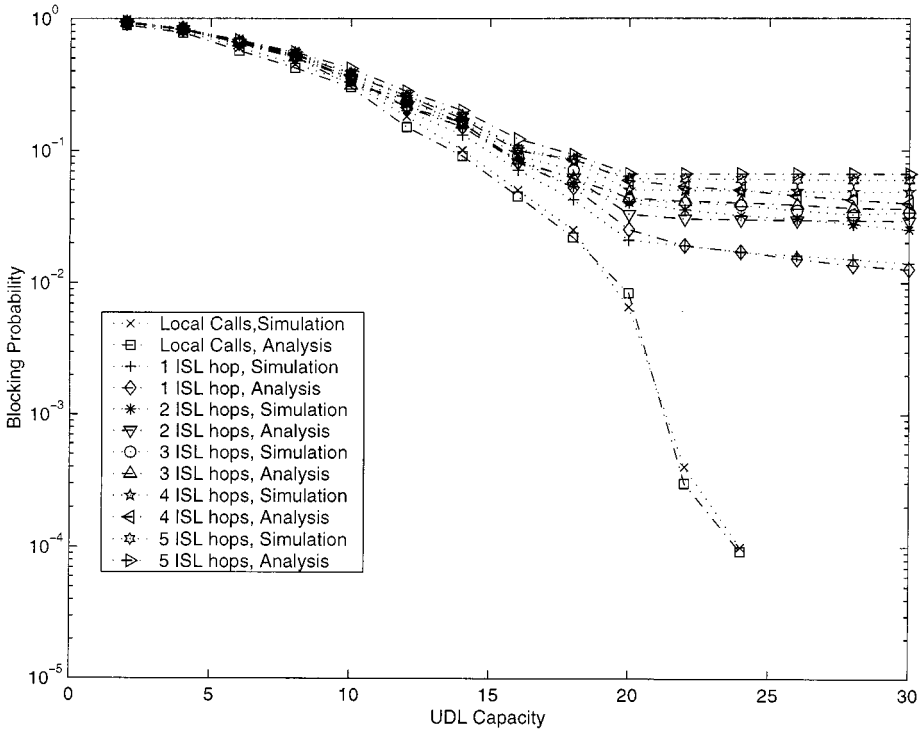

Fig. 11. Call-blocking probabilities for a 12-satellite orbit, $\lambda = 5, C_{\mathrm{ISL}} = 10$, locality pattern.

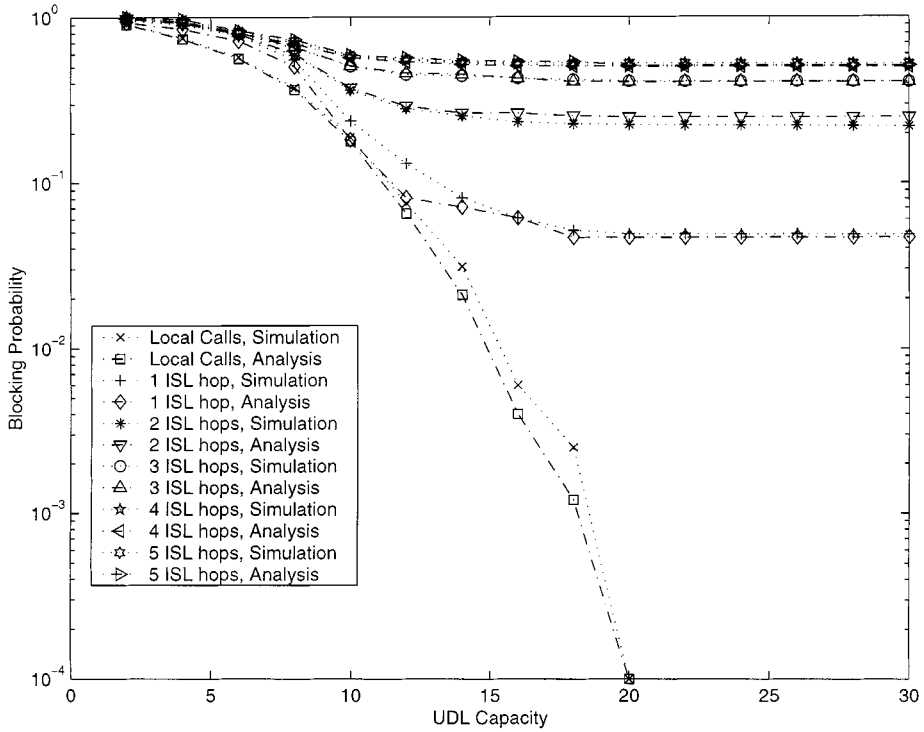

Fig. 12. Call-blocking probabilities for a 12-satellite orbit, $\lambda = 5, C_{\mathrm{ISL}} = 10$, two-community pattern.

in Section II; we believe that a closed-form solution for the modified process can be obtained. Finally, it is possible to extend our approach to analyze the case of heterogeneous traffic (i.e., when customers are not uniformly distributed over the earth, an assumption we made in Section IV). One approach to account for different geographic arrival rates is to segment the band of earth covered by the satellites into fixed regions, each with a different arrival rate of new calls. This approach gives rise to a periodic Markov process model whose special structure can be exploited to solve it efficiently.

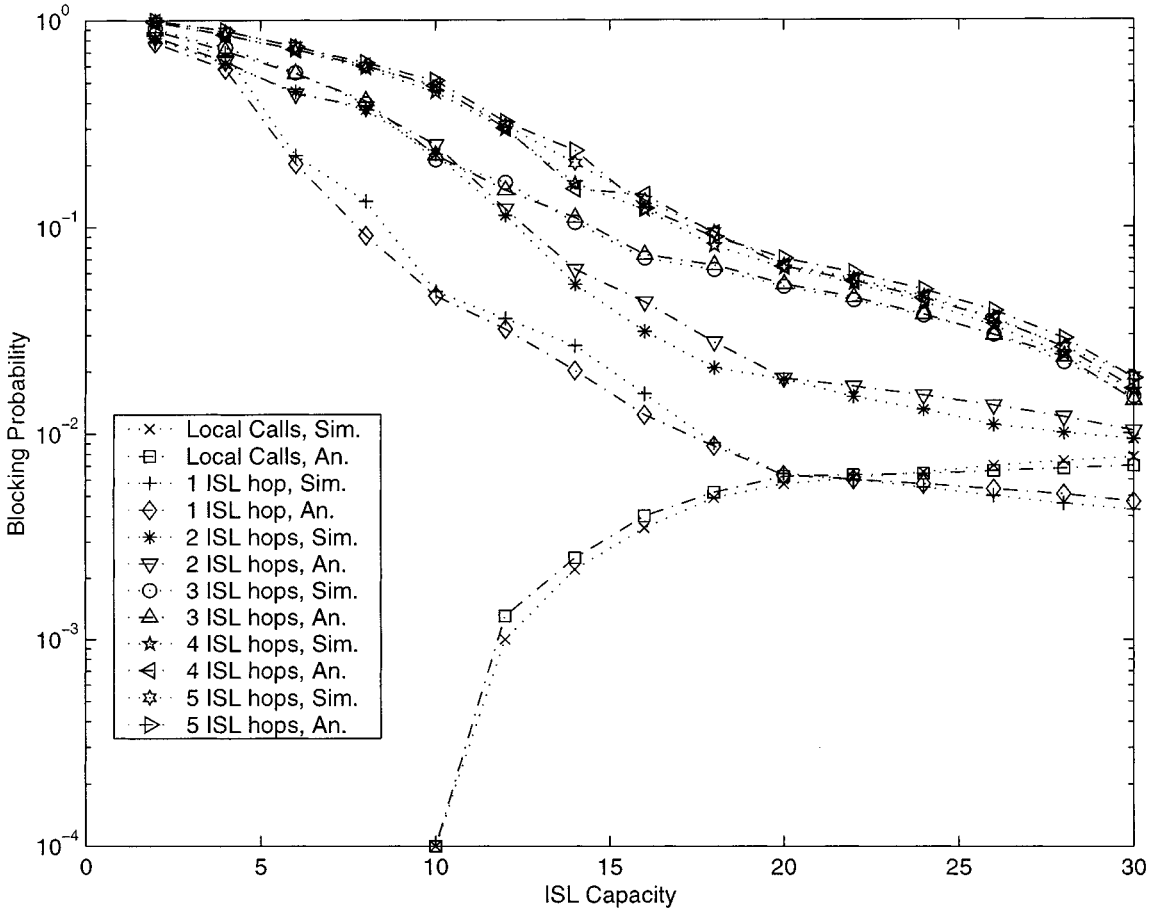

Fig. 13. Call-blocking probabilities for a 12-satellite orbit, $\lambda = 5, C_{\mathrm{UDL}} = 20$, two-community pattern.

Our approach cannot be used directly to model dynamic (or adaptive) routing algorithms, or scenarios in which the traffic intensity varies over time. New techniques are needed to analyze satellite constellations under these assumptions, and we are currently investigating these problems.

REFERENCES

[1] H. S. Chang, B. W. Kim, C. G. Lee, S. L. Min, Y. Choi, H. S. Yang, and C. S. Kim, "Topological design and routing for LEO satellite networks," in *Proc. IEEE GLOBECOM*, 1995, pp. 529–535.

[2] H. S. Chang, B. W. Kim, C. G. Lee, S. L. Min, Y. Choi, H. S. Yang, D. N. Kim, and C. S. Kim, "Performance comparison of optimal routing and dynamic routing in low earth orbit satellite networks," in *Proc. IEEE Vehicular Technology Conf.*, 1996, pp. 1244–1249.

[3] ——, "Performance comparison of static routing and dynamic routing in low earth orbit satellite networks," in *Proc. IEEE Vehicular Technology Conf.*, 1996, pp. 1240–1243.

[4] ——, "FSA-based link assignment and routing in low earth orbit satellite networks," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 1037–1048, Aug. 1998.

[5] F. Dosiere, T. Zein, G. Maral, and J. P. Boutes, "A model for the handover traffic in low earth-orbiting satellite networks for personal communications," in *Proc. IEEE GLOBECOM*, 1993, pp. 574–578.

[6] A. Ganz, Y. Gong, and B. Li, "Performance study of low earth orbit satellite systems," *IEEE Trans. Commun.*, vol. 42, pp. 1866–1871, Feb./Mar./Apr. 1994.

[7] A. Jamalipour, *Low Earth Orbital Satellites for Personal Communication Networks*. Reading, MA: Artech House, 1998.

[8] A. Jamalipour, M. Katayama, and A. Ogawa, "Traffic characteristics of LEOS-based global personal communications networks," *IEEE Commun. Mag.*, vol. 35, pp. 89–93, Feb. 1997.

[9] A. Jamalipour, M. Katayama, T. Yamazato, and A. Derroni, "A performance analysis on the effects of traffic nonuniformity in low earth-orbit satellite communication systems," in *Proc. SITA*, 1993, pp. 203–206.

[10] Y. S. Kim, Y. H. Bae, Y. Kim, and J. P. Boutes, "Traffic load balancing in low earth orbit satellite networks," *Comput. Commun. Networks*, pp. 191–195, 1998.

[11] R. Mauger and C. Rosenberg, "QoS guarantees for multimedia services on a TDMA-based satellite network," *IEEE Commun.*, vol. 35, pp. 56–65, July 1997.

[12] G. Pennoni and A. Ferroni, "Mobility management in LEO/ICO satellite systems: Preliminary simulation results," in *Proc. PIMRC*, 1994, pp. 1323–1329.

[13] G. Ruiz, T. L. Doumi, and J. G. Gardiner, "Teletraffic analysis and simulation of mobile satellite systems," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 311–320, Feb. 1998.

[14] H. Uzunalioglu, "Probabilistic routing protocols for low earth orbit satellite networks," in *Proc. IEEE ICC*, 1998, pp. 89–93.

[15] H. Uzunalioglu and W. Yen, "Managing connection handover in satellite networks," in *Proc. IEEE GLOBECOM*, 1997, pp. 1606–1610.

[16] H. Uzunalioglu, W. Yen, and I. F. Akyildiz, "A connection handover protocol for LEO satellite ATM networks," in *Proc. ACM/IEEE MobiCom'98*, 1998, pp. 204–214.

[17] M. Werner, "A dynamic routing concept for ATM-based satellite personal communication networks," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 1636–1648, Oct. 1997.

[18] M. Werner, C. Delucchi, H. J. Vogel, G. Maral, and J. J. DeRidder, "ATM-based routing in LEO/MEO satellite networks with intersatellite links," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 69–81, Jan. 1997.

[19] Y. Zhu, G. N. Rouskas, and H. G. Perros, "A path decomposition algorithm for computing blocking probabilities in wavelength routing networks," *IEEE/ACM Trans. Networking*, vol. 8, pp. 747–762, Dec. 2000.

Abdul Halim Zaim received the M.Sc. degree from the Computer Engineering Department, Bogazici University, Turkey, in 1996 and the Ph.D. degree from the Electrical and Computer Engineering Department, North Carolina State University, Raleigh, in 2001.

As a Teaching Assistant, he taught several courses at Istanbul University between 1993–1997. In 1998–1999, he was with Alcatel, Raleigh, NC. He is currently a Postdoctoral Teaching Associate at North Carolina State University. His research interests include computer performance evaluation, satellite and high-speed networks, and computer network design.

George N. Rouskas (S'92–M'95–SM'01) received the diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1989 and the M.S. and Ph.D. degrees in computer science from the College of Computing, Georgia Institute of Technology, Atlanta, in 1991 and 1994, respectively.

He joined the Department of Computer Science, North Carolina State University, Raleigh, in 1994, where he has been an Associate Professor since 1999. During the 2000–2001 academic year, he spent a sabbatical term at Vitesse Semiconductor, Morrisville, NC. In May and June 2000, he was an Invited Professor at the University of Evry, France. His research interests include network architectures and protocols, optical networks, multicast communication, and performance evaluation.

Prof. Rouskas is a member of the ACM and the Technical Chamber of Greece. He received a 1997 NSF Faculty Early Career Development (CAREER) Award. He was coauthor of a paper that received the Best Paper Award at the 1998 SPIE Conference on All-Optical Networking. He received the 1995 Outstanding New Teacher Award from the Department of Computer Science, North Carolina State University, and the 1994 Graduate Research Assistant Award from the College of Computing, Georgia Institute of Technology. He was a Coguest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Protocols and Architectures for Next Generation Optical WDM Networks (October 2000). He is on the editorial boards of the IEEE/ACM TRANSACTIONS ON NETWORKING and *Optical Networks Magazine.*

Harry G. Perros (M'87–SM'97) received the B.Sc. degree in mathematics from Athens University, Athens, Greece, in 1970, the M.Sc. degree in operational research with computing from Leeds University, U.K., in 1971, and the Ph.D. degree in operations research from Trinity College, Dublin, Ireland, in 1975.

From 1976 to 1982, he was an Assistant Professor in the Department of Quantitative Methods, University of Illinois at Chicago. In 1979, he spent a sabbatical term at INRIA, Rocquencourt, France. In 1982, he joined the Department of Computer Science, North Carolina State University, Raleigh, as an Associate Professor, where since 1988 he has been a Professor. During academic year 1988–1989, he was on a sabbatical leave of absence, first at BNR, Research Triangle Park, NC, and subsequently at the University of Paris VI, France. During academic year 1995–1996, he was on a sabbatical leave of absence at Nortel, Research Triangle Park. He has published extensively in the area of performance modeling of computer and communication systems. He has organized several national and international conferences. He is the author of *Queueing Networks with Blocking: Exact and Approximate Solutions* (Oxford, U.K.: Oxford University Press, 1994). He is Chairman of the IFIP Working Group 6.3 on the Performance of Communication Systems. His current research interests are in the areas of optical networks and their performance, as well as software performance evaluation.