

# ADAPTIVE HANDOVER MANAGEMENT FOR MULTISERVICE NGSO SATELLITE SYSTEMS

Evangelos Papapetrou  
Department of Computer Science  
University of Ioannina  
Ioannina, Greece

Petia Todorova  
Fraunhofer Institute FOKUS  
Smart Environment Center  
Berlin, Germany

Stylianos Karapantazis, Niovi Pavlidou  
Electrical and Computer Engineering Dpt  
Aristotle University of Thessaloniki  
Thessaloniki, Greece

## ABSTRACT

In this paper we propose a handover method for multiservice non-geo satellite systems. The proposed algorithm can be used for satellite handover as well as for cell handover when a fixed amount of resources is allocated to each cell. Specifically, we consider the case in which the satellite system provides services to users of different QoS requirements. The first type of users require the minimization and, if possible, the elimination of the probability of forcing a connection in progress to termination during a handover. On the other hand, the users of the second category do not have any specific handover requirements. The new method is evaluated through extensive simulations and is proved to meet the requirements of the first type of users while at the same time maximizes the system utilization for both types of users.

## I. INTRODUCTION

Over the last years, the interest in Non-GeoStationary Orbit (NGSO) satellite systems has declined due to their failure to provide services beyond plain voice communication. However, recently the interest in NGSO systems has been revived in the light of newly planned multiservice systems [1],[2]. The inherent features of NGSO systems such as low propagation delays, high capacities, large coverage areas and on-demand connectivity, constitute an appealing set of advantages in the future era of high speed interactive data services. However, the complexity of an NGSO satellite and the mobility of satellites with respect to the Earth's surface, spawn several management issues that must be thoroughly confronted. Such a well-known issue is the management of users which must be handed over to different satellites (or cells within a satellite) during the lifetime of their communication because of the rapid satellite movement. Forced termination of communication must be averted and at the same time the system resources must be efficiently utilized to support as many users as possible. Connection admission and handover management hold a key role in NGSO satellite systems, especially in the context of all IP multiservice satellite systems in which the problem of forced termination is aggravated. Although connectionless communication is used at the network layer, future high quality services require stable transport layer connections that can be guaranteed to receive specific communication quality.

During the last years, a lot of studies regarding the issue of handover have been conducted [3]. The most appealing, because of their simplicity and effectiveness, are those concerned with systems in which the satellite or a cell in the satellite footprint is assigned a fixed amount of bandwidth. Another ad-

vantage of those methods is that they can be used to address both handover types, namely *satellite* and *cell handover*. The common ground of all proposed techniques is to reserve resources in the forthcoming cell or satellite in order to be used by the transit user when the handover occurs. For example, in [4], the proposed Guaranteed Handover (GH) method tackles the problem of cell handover by reserving resources at the next satellite cell when the user enters to a new cell. By this approach, the method is proved to eliminate the probability of forcing a connection in progress into termination. However, the reservation (also known as *locking*) of resources comes at the cost of low system utilization. That is because the reserved resources are not available to new connections. To reduce the effect of locking, the common idea to a number of methods is to defer the reservation of resources. The time to delay the reservation is decided in a different way, according to the proposed method. In [5], the time of resource locking is decided based on the forced termination requirements, by using a complex Markovian model. The Time-based Channel Allocation Algorithm (TCRA) algorithm, presented in [6], improves on GH by taking advantage of the user positions to delay resource locking. Finally, in [7], [8], the Dynamic Doppler-Based Handover Prioritization (DDBHP) algorithm, which was proposed for both cell and satellite handovers, takes advantage of user locations and system characteristics to decide on the time to delay resource locking. Both TCRA and DDBHP manage to eliminate the forced termination probability.

Despite the plethora of studies on handover management [3], the issue is rarely addressed in multiservice systems. In this case the management of users with different requirements regarding the forced termination probability, is of great interest. In [4] a system with two classes of users is addressed and the problem of cell handover is addressed. The first class of users requires the elimination of forced termination during handover, while the second class has no specific requirements. For users of class 2 no specific reservation mechanism is proposed, raising concerns about fairness. On the contrary, both the method for cell handovers proposed in [9] and the method for satellite handovers proposed in [10] reserve resources in the next cell/satellite when the handover occurs for both classes of users. Furthermore, the treatment of new connections is different (resources in the next cell/satellite are not required for the admission of a new connection) in order to reduce the cost of resource locking.

In this paper, we propose a new method for managing users of no specific requirements when coexisting with users that require small values of forced termination probability. The new algorithm, called Multiservice Adaptive Handover (MAH)

aims at meeting the user requirements and at the same time efficiently utilizing the system resources. In this way MAH also achieves a certain fairness for users of no specific requirements. MAH can be used for satellite handovers and for cell handovers in Fixed Channel Allocation (FCA) systems. However, in the context of this study we will focus on satellite handovers. Finally, MAH is simple in implementation, minimizing the computational load on the satellite.

The rest of this paper is structured as follows. In Section II. the problem of handover in multiservice systems is defined and the proposed method is presented in detail. Then, in Section III. the details of the simulation model, used for evaluating the new method, are provided and the evaluation metrics are discussed, while in Section IV. the obtained results are presented and analyzed. Finally, in Section V. a summary of the performed work is presented and useful conclusions are drawn.

## II. MULTISERVICE ADAPTIVE HANDOVER (MAH)

In order to present the details of MAH, we first have to describe the considered system model.

### A. System Model and Problem Description

The most important ingredient of the system model is the mobility model. Several mobility models have been proposed in the literature for NGSO satellite systems. Mobility models comprise the set of rules that describe the size, the shape and the movement of satellite footprints. In this study we employ a two dimensional model according to which the satellite footprint is modelled as a rectangle. Two-dimensional models are frequently used in the literature ([3],[4],[7],[8],[9],[10],[11]). According to the mobility model the satellite (i.e., the satellite footprints) is considered to move with a speed  $V_{sat}$ . Footprint overlapping or satellite diversity are not considered in the context of this study. The mobility of the satellite is the one causing the need for a handover since it is much higher than the speed of users in fast vehicles and the velocity that corresponds to the rotation of the earth. Let us consider the model presented in fig. 1. Users have to be handed over to the next satellite at time  $t_0 + t_{fp}$ , where  $t_{fp}$  is the time needed by the users to cross the footprint due to the movement of the satellite. As far

the first class (*class 1* users) correspond to a highly demanding interactive service, such as telephony. The basic service requirements for class 1 users are: a) allocation of a predefined amount of bandwidth and, b) the minimization of the probability of forcing an ongoing connection to termination (*forced termination probability* -  $P_f$ ). On the contrary, for users of class 2 there exists a minimum and maximum bandwidth that can be allocated and there is no specific requirement as far as  $P_f$  is concerned. Class 2 service refers to a service similar to web browsing in which the bitrate may be variable. Additionally, since the forced termination of a TCP connection used for web browsing has a smaller effect than in the case of an interactive service, there is no strict relevant requirement. Furthermore, users are considered to be uniformly distributed in the footprint of the satellite.

To manage different classes of users two approaches have been proposed so far. The first one (although proposed in [4] for cell handover, could also be used for satellite handovers) provides a guaranteed handover (GH) service to class 1 users. To achieve this, each user of class 1, entering footprint  $i$ , is required to request at time  $t_0$  (see fig. 1) resources in footprint  $i + 1$ . If the resources exist then they are reserved, otherwise the request is queued. In this way all queued requests are guaranteed to be serviced until time  $t_0 + t_{fp}$ . On the contrary users of class 2 can not request resources. Therefore the successful handover of a user entering footprint  $i$  depends on the existence of resources at this footprint at time  $t_0$ . It is well understood that the reservation of resources by class 1 users raises an issue of fairness for class 2 users. Specifically, class 2 users face significantly greater probability of being blocked before entering the systems (*blocking probability* -  $P_b$ ). Furthermore, users of class 2 suffer high forced termination probabilities since no resource reservation is allowed. To overcome this problem resource reservation by class 2 users is allowed in [10]. However, to preserve priorities between the two types of users, two separate queues are maintained where reservation requests are stored until they are serviced. When resources are made available, requests of class 1 users are serviced with absolute priority. Allowing requests from class 2 users ameliorates the fairness issue but aggravates the system utilization since for more users, resources are reserved in more than one satellites. The proposed method for handover modifies the method described in [10] in order to preserve fairness between the two classes of users and at the same time maximize the system utilization.

### B. The proposed solution

The basic concept of MAH is to take advantage of the fact that class 2 users have no specific requirement regarding  $P_f$ . To this end, we propose delaying the handover requests of class 2 users. Let us consider fig.1. Class 1 users that enter footprint  $i$ , request resources in footprint  $i + 1$  at time  $t_0$ , while users of class 2 request resources at time  $t_0 + t_D$ . The advantages of this strategy are twofold. Obviously, the method economizes on resources by reducing the period of time that resources are reserved for class 2 users. As a result, smaller blocking probability may be achieved for both classes. The second advantage relates to the handover priority of the two classes. By delaying

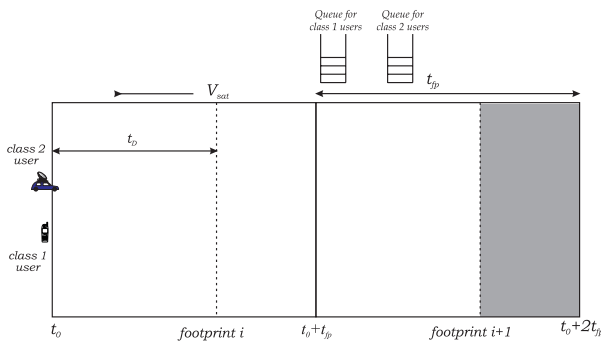


Figure 1: System Model

as users are concerned, we consider two classes. The users of

the handover requests of class 2 users, the method provides increased priority to class 1 users. To clarify this we refer to fig.1. Suppose that a user of class 2 enters footprint  $i$  at  $t_0$ . Its request is delayed for  $t_D$  seconds. It is clear that all requests from class 1 users that arrive at footprint  $i$  between  $t_0$  and  $t_0 + t_D$  are either serviced or are added to the respective queue and will be serviced earlier than the request from the class 2 user. The result of the increased priority of class 1 handover requests is the minimization of  $P_f$  for class 1 users.

The impact of time  $t_D$  in class 2 users is the increased probability of forcing an ongoing call into termination. However, recall that class 2 users do not have any specific requirement regarding  $P_f$ . An appropriate choice of  $t_D$  can provide relatively small  $P_f$  while at the same time boost system utilization. It is clear that if  $t_D$  is set to zero then MAH matches the method proposed in [10]. Clearly  $t_D$  depends on the footprint geometry and on the traffic conditions in the footprint. Our approach is to derive a simple and flexible formula that could provide a rough estimation of  $t_D$  in order to minimize  $P_f$  for class 2 users. To this end, we make the following observation based on fig.1. Suppose a class 2 user enters footprint  $i$  at  $t_0$  and somehow defines  $t_D$ . If its request is queued at time  $t_0 + t_D$ , the time remaining for that request to be serviced is  $t_{fp} - t_D$  and should be sufficient. Suppose that at this point the bandwidth requests in both queues are  $B_Q$ . Those requests should be less than the capacity of the footprint that corresponds in the grey area in fig.1 ( $C_{fraction}$ ) because until time  $t_0 + t_{fp}$  all users in this area will have left the footprint, releasing the corresponding capacity. Recall also that users from footprint  $i$  that will enter footprint  $i + 1$  have already reserve capacity or their requests reside in one of the queues. Considering uniform footprint load  $C_{fraction}$  can be expressed as:

$$C_{fraction} = C \cdot \frac{t_{fp} - t_D}{t_{fp}} \quad (1)$$

where  $C$  is the capacity of a footprint. As a result it holds:

$$B_Q \leq C \cdot \frac{t_{fp} - t_D}{t_{fp}} \quad (2)$$

resulting in:

$$t_D \leq t_{fp} \cdot \left(1 - \frac{B_Q}{C}\right) \quad (3)$$

To preserve simplicity in the previous equations we neglect the impact of class 1 users that may enter footprint  $i$  in the time interval  $(t_0 + t_D, t_0 + t_{fp})$  and can affect previous equations. In any case the previous equations serve only as a rough estimation of  $t_D$ . The basic concern is the simplicity of calculations such that the above equations can be carried out by the users, alleviating in this way the amount of operations needed by the satellite.

### III. SIMULATION MODEL

In order to evaluate the proposed method a simulation tool was developed in C++. Four orbital planes with four satellites in

Table 1: Traffic Parameters

	<b>Class 1</b>	<b>Class 2</b>
Maximum bandwidth(kbps)	30	512
Minimum bandwidth(kbps)	30	64
$\lambda$ (connections/min/footprint)	60-120	9-18
Mean duration( $T_d$ )	180	180
Users/footprint ( $N$ )	1000	400

each one were simulated. The coverage area of a satellite footprint was  $1720 \times 1720 km^2$ , which is roughly equal to the area that a satellite of the Teledesic system (the Boeing design with the 288 satellites) covers. Additionally, the velocity of each satellite was set to  $5.89 km/sec$ . The capacity of the satellite was considered to be  $32 Mbps$ .

In the simulated scenarios we considered two classes of users. The users of class 1 require a fixed amount of  $30 kbps$  while class 2 users may be allocated from  $64$  to  $512 kbps$  depending on the traffic conditions in the footprint. Specifically, a class 2 user is always allocated  $512 kbps$  when such capacity is available. In any other case the user is allocated the remaining capacity provided that it is at least  $64 kbps$ . Otherwise the connection is dropped. New connection requests are generated according to a Poisson process with mean arrival rate  $\lambda$ . The mean duration of connections is  $T_d$ . In each footprint there are  $N$  users of each class, uniformly distributed. In Table 1 the simulation parameters concerning traffic are summarized.

We compare the proposed method with the algorithm presented in [10]. As mentioned previously the algorithm mandates that users of both class should request resources for footprint  $i + 1$  when entering footprint  $i$ . Hereinafter this scheme will be referred to as "Early Reservation". As far as admission control is concerned, we implemented for MAH the same strategy as in [10]. That is, new connections are admitted into the network as long as the minimum required bandwidth is available. Finally, in both schemes we used the FIFO policy for both queues where requests are stored.

#### A. Evaluation metrics

For the evaluation of the two schemes we used the following metrics: 1) Blocking Probability ( $P_b$ ), i.e., the probability that a new user is blocked, 2) Forced Termination Probability ( $P_f$ ), i.e. the probability of forcing an ongoing call to termination, 3) Grade of Service (GoS), 4) Mean capacity allocated to users of class 2 ( $C_{mean}^2$ ), and 5) Mean footprint capacity allocated to users that were successfully serviced ( $C_{ef}$ ). While the usefulness of  $P_b$  and  $P_f$  is straightforward, GoS [12] expresses the combined impact of both  $P_b$  and  $P_f$  experienced by users. It is defined by:

$$GoS = 0.1P_b + 0.9P_f \quad (4)$$

in order to emphasize on the impact of  $P_f$ . Another important metric is  $C_{mean}^2$  which expresses the mean capacity allocated to users of class 2. As mentioned previously, class 2 users may be allocated  $64$  to  $512 kbps$  depending on the traffic conditions. An efficient handover algorithm should economize on

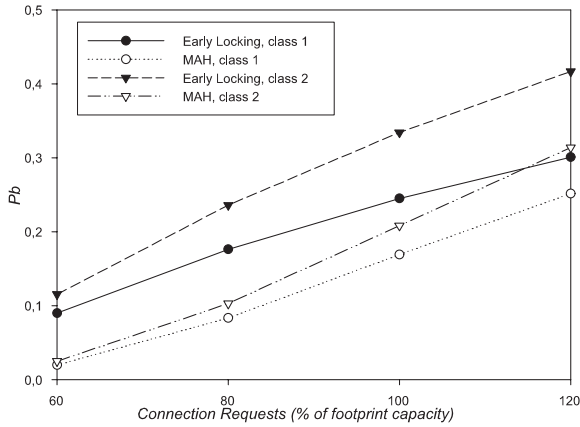


Figure 2: Blocking probability for the two classes of users

the satellite resources and therefore being able to provide increased bandwidth to users of class 2. Finally,  $C_{eff}$  provides the fraction of footprint capacity that was used to service users and therefore measures the ability of an algorithm to minimize the reserved for handovers bandwidth.

#### IV. PERFORMANCE RESULTS

In the experiments that were conducted the two algorithms are compared under variable traffic load. Specifically, we varied the connection arrival rate for class 1 from 60-120 connections/min/footprint with steps of 20 connections/min/footprint and the arrival rate for class 2 from 9-18 connections/min/footprint with steps of 3 connections/min/footprint. In all cases the ratio of class 1 and class 2 traffic was kept constant. The resulting offered load ranges from 60%-120% of the total footprint capacity if class 2 connections are considered to be of 512kbps. However, the actual bandwidth of class 2 connections is less than 512kbps and therefore the actual offered load never exceeds the 100% of the total footprint capacity.

In fig.2  $P_b$  is presented for the two algorithms and for both traffic classes. It is clear that MAH achieves better performance for both traffic classes over the entire range of offered load. The improvement compared to the performance of “Early Locking” ranges from 16% to 78% for users of class 1 and from 24.5% to 78.5% for class 2 users. It is clear that the improvement is evenly applied to users of both classes. Furthermore, MAH preserves the priority between the two classes of traffic and performs better than “Early Locking” does for class 1 users in almost the entire range. This result is owed to the ability of MAH to reduce the bandwidth reserved by class 2 users.

Indeed this is proved in fig.3, where the mean bandwidth allocated to successfully served users is presented. MAH maximizes  $C_{eff}$  over the entire range of offered load and manages an improvement that ranges from 21% to 33.5%. As the offered load increases, MAH manages to allocate increasing bandwidth to users. On the contrary when “Early Locking” is used, the increment of user allocated bandwidth is impeded by increas-

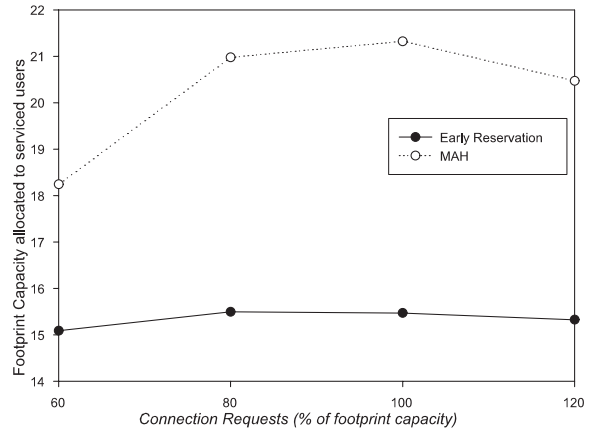


Figure 3: Mean capacity allocated to successfully serviced users for classes of users

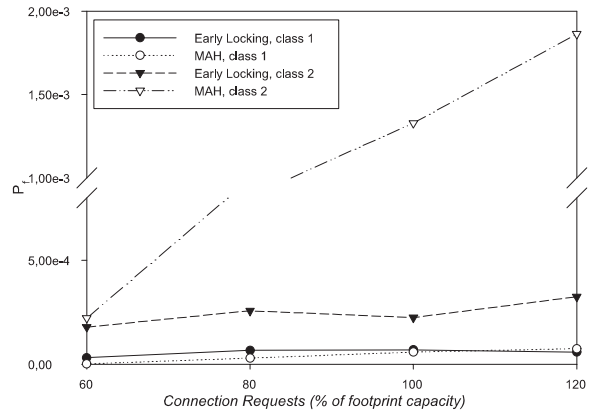


Figure 4: Forced termination probability for the two classes of users

ing demands for resource reservations. Finally, for high traffic loads MAH’s performance slightly deteriorates. This can be explained by Eq.3, according to which the reservation of resources is in general longer (increased  $t_D$ ) when the traffic load ( $B_Q$ ) is high. However, even in this case MAH manages an improvement of approximately 30% compared to the performance of “Early Locking”.

The price to pay for the improved performance of MAH is the increased  $P_f$  for the users of class 2, since their reservation requests are intentionally delayed. Fig. 4, which presents  $P_f$  for both handover methods and traffic classes, confirms the expected result. However, the actual value of  $P_f$  is kept at very low levels, which in combination with the fact that users of class 2 have no specific requirements can be considered an acceptable performance. Furthermore, MAH preserves and even slightly improves the performance for users of class 1.

The actual impact of the increased  $P_f$  for users of class 2 is depicted in fig.5, where the GoS, perceived by users of both classes, is presented. The results prove that the levels of  $P_f$

can be considered low, especially when compared to those of  $P_b$ . As a result, GoS is dominated by  $P_b$  and MAH presents an improved performance compared to “Early Locking” for both classes of users and over the entire range of traffic load. The improvement ranges from 16% to 78% for class 1 and from 21% to 78% for class 2 users.

Finally, the mean capacity allocated to class 2 users is depicted in fig.6. MAH achieves an impressive increase of 35% compared to the capacity allocated when “Early Locking” is utilized. This result must be interpreted in combination with the fact that, as was presented previously, MAH also admits more users. In other words, MAH not only services more users but also provides them more bandwidth. Concluding, it is worth to be noted that the improvement achieved with MAH, is preserved even when the satellite is heavily loaded.

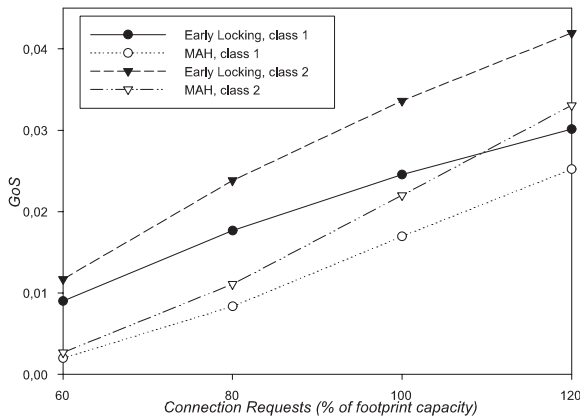


Figure 5: Grade of Service perceived by users of both classes

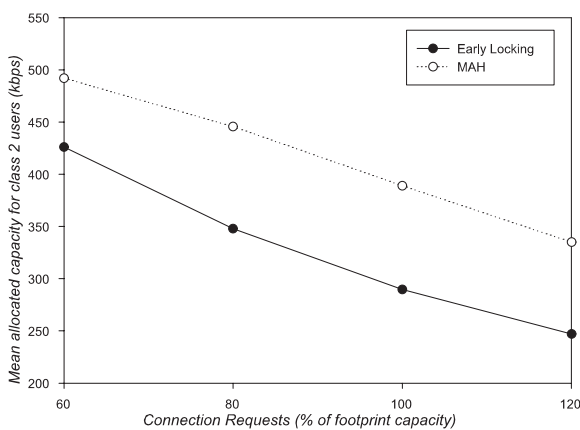


Figure 6: Mean capacity allocated to class 2 users

## V. CONCLUSIONS

In this paper we proposed a new handover method for multi-service non-geo satellite systems, called Multiservice Adaptive

Handover (MAH), which can be used both for satellite and cell handovers. MAH makes the observation that users with no QoS requirements may defer their reservation requests in order to improve the utilization of the satellite's capacity. To this end, MAH takes advantage of the traffic conditions in the satellite to calculate the time period for which a user may defer its requests. The calculation is made at the user side, alleviating in this way the computational burden on the satellite. The proposed method was evaluated by means of extensive simulation studies and was proved to have an improved performance compared to previously proposed methods. MAH not only keeps forced termination probability low but also manages to service more users, allocating at the same time more bandwidth to each user.

## VI. ACKNOWLEDGEMENT

This work was carried out in the framework of the EU funded Network of Excellence “SatNex II” (contract No. IST-027393).

## REFERENCES

- [1] (2006) Globalstar press releases. [Online]. Available: <http://www.globalstar.com>
- [2] T. Taleb, N. Kato, and Y. Nemoto, “Recent Trends in IP/NGEO Satellite Communication Systems: Transport, Routing, and Mobility Management Concerns,” *IEEE Wireless Communications*, vol. 12, no. 5, pp. 63–69, October 2005.
- [3] P. K. Chowdhury, M. Atiquzzaman, and W. Ivancic, “Handover Schemes in Satellite Networks: State-of-the-Art and Future Research Directions,” *IEEE Communications Surveys*, vol. 8, no. 4, 4th Quarter 2006.
- [4] G. Maral, J. Restrepo, E. del Re, R. Fantacci, and G. Giambene, “Performance analysis for a guaranteed handover service in an LEO constellation with satellite-fixed cell,” *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, pp. 1200–1214, 1998.
- [5] Y. Xu, Q. Ding, and C. Ko, “Elastic handover scheme for leo satellite mobile communication systems,” in *IEEE Global Telecommun. Conf., San Francisco, CA*, pp. 1161–1165.
- [6] L. Boukhatem, A.-L. Beylot, D. Gaïti, and G. Pujolle, “TCRA: A Time-based Channel Reservation Scheme for Handover Requests in LEO Satellite Systems,” *International Journal of Satellite Communications and Networking*, vol. 21, no. 3, pp. 227–240, 2003.
- [7] E. Papapetrou and F.-N. Pavlidou, “Analytic Study of Doppler-based handover management in LEO satellite systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 3, pp. 830–839, 2005.
- [8] —, “QoS handover management in LEO/MEO satellite systems,” *Wireless Personal Communications*, vol. 24, no. 2, pp. 189–204, January 2002, Special Issue on Broadband Mobile Terrestrial-Satellite Integrated Systems.
- [9] P. Todorova, S. Olariu, and H. N. Nguyen, “A two-cell lookahead call admission and handoff management scheme for multimedia leo satellite systems,” in *Proc. of the 36th Hawaii International Conference on System Sciences (HICSS'03)*, 2003.
- [10] S. Karapantazis, P. Todorova, and F.-N. Pavlidou, “On bandwidth and inter-satellite handover management in multimedia leo satellite systems,” in *Proc. of the Advanced Satellite Mobile Systems (ASMS) Conference 2006, Herrsching am Ammersee, Germany*, 2006.
- [11] E. del Re, R. Fantacci, and G. Giambene, “Efficient dynamic channel allocation techniques with handover queuing for mobile satellite networks,” *IEEE J. Select. Areas Commun.*, vol. 13, no. 2, p. 397 405, February 1995.
- [12] —, “Characterization of user mobility in low earth orbit mobile satellite systems,” *Wireless Networks*, vol. 6, no. 3, p. 165179, May 2000.