

A Multiple Subset Sum Formulation for Feedback Implosion Suppression over Satellite Networks

Gun Akkor, John S. Baras, and Michael Hadjitheodosiou
Electrical and Computer Engineering Department, and
Institute for Systems Research,
University of Maryland, College Park, MD 20742, USA.
e-mail: {akkor, baras, michalis}@isr.umd.edu

Abstract—In this paper, we present a *feedback implosion suppression* (FIS) algorithm that reduces the volume of feedback information transmitted through the network without relying on any collaboration between users, or on any infrastructure other than the satellite network. Next generation satellite systems that utilize the Ka frequency band are likely to rely on various fade mitigation techniques, in order to guarantee a service quality that is comparable to other broadband technologies. User feedback would be a valuable input for a number of such components, however, collecting periodic feedback from a large number of users would result in the well-known *feedback implosion* problem. Feedback implosion is identified as a major problem when a large number of users try to transmit their feedback messages through the network, holding up a significant portion of the uplink resources and clogging the shared uplink medium. In this paper, we look at a system where uplink channel access is organized in time-slots. The goal of the FIS algorithm is to reduce the number of uplink time-slots hold up for the purpose of feedback transmission. Our analysis show that the FIS algorithm effectively suppresses the feedback messages of 95% of all active users, but still achieves acceptable performance results when the ratio of available time-slots to number of users is equal to or higher than 5%.

I. INTRODUCTION

The role satellite systems play in today's communication infrastructure is changing rapidly as a result of several technological advances in the design of these systems. These new technologies are making satellite communication systems more competitive against other broadband solutions (e.g. cable, ADSL) [1]. One of the key advances comes in the utilization of the Ka frequency band (20-30 GHz). Ka-band is very desirable, particularly for multimedia communication, because it offers wider bandwidth segments, which are unavailable at lower frequency bands. However, rain and atmospheric attenuation present a significant challenge to transmission of signals at Ka-band frequencies. In order to guarantee a service quality that is comparable to other broadband technologies, next generation systems will likely rely on various fade mitigation techniques [2].

User feedback would be a valuable input for a number of such components, however, collecting periodic feedback from a large number of users would result in the well-known *feedback implosion* problem. Feedback implosion is identified

This material is based upon work supported by NASA under award number NCC8235.

as a major problem when a large number of users try to transmit their feedback messages through the network, causing a high traffic concentration and backlog. Although not unique to satellite based networks, feedback implosion phenomenon has additional side effects in this context, since a significant portion of the uplink resources may be hold up to transmit these feedback messages instead of useful data traffic, and the access to the shared medium may be clogged [3].

In some of the cases, it maybe possible to reduce the volume of user feedback by exploiting the nature of the information and by considering that (i) the feedback may contain redundant information (due to correlations among the users), and (ii) the protocol may need to track the behavior of only a subset of users, e.g. the group of users with worst case channel conditions [4]. In this paper, we assume that a protocol performs between a central satellite gateway and a group of direct users that are located inside the footprint of the satellite, in an attempt to collect periodic channel state information (CSI). CSI collected from users may be used as an input to adaptive coding and power control components at the satellite. We present a *feedback implosion suppression* (FIS) algorithm to complement such a protocol. The algorithm reduces the volume of feedback information transmitted through the network without relying on any collaboration between users, or on any infrastructure other than the satellite network. In [4], authors propose a feedback suppression algorithm for a similar satellite system, however, our work significantly differs from their work in the way (i) priority feedback is determined and, (ii) feedback is transmitted over the return channels.

The rest of the paper is organized as follows. In the next section, we describe our target satellite network architecture. In Section III, we present an overview of the underlying CSI collection protocol. Section IV discusses our FIS algorithm and its components. In Section V, we present numerical performance results. Last section concludes the paper.

II. SYSTEM DESCRIPTION

The system we consider is a star topology, multiple spot-beam satellite network, where a Ka-band, geo-synchronous satellite provides broadband services to a large number of users located inside its footprint. In this scenario, users access the terrestrial backbone network through a gateway node referred to as the network operations center (NOC) (Fig. 1). In each

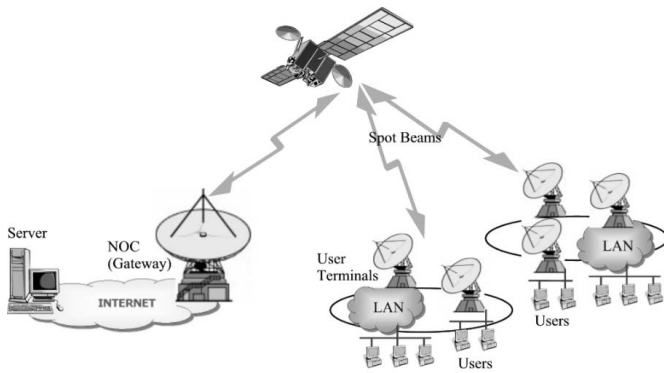


Fig. 1. Satellite communication system architecture. The satellite provides broadband access to users across multiple spot-beam locations.

spot-beam, the users access the uplink in multiple-frequency time-division multiple access (MF-TDMA) mode. The users acquire access by requesting a number of time-slots from the NOC, which invokes a resource allocation algorithm to share the uplink capacity (available time-slots) of the spot-beam among all active users. The result of the resource allocation is broadcast to all active users.

In the rest of the paper, we assume that a feedback message can be transmitted in a single uplink time-slot, and our FIS algorithm runs at the NOC in conjunction with the resource allocation algorithm in assigning time-slots to users. Therefore, the goal of the algorithm is to limit the total number of uplink time-slots used for the purpose of feedback transmission in every spot-beam.

III. CSI COLLECTION PROTOCOL

In this section, we present the CSI collection protocol behavior that will be relevant to the operation of our feedback implosion suppression policy. The CSI of interest, in this case, is the signal attenuation level due to atmospheric and rain fading as measured at the input of the user receiver equipment. The goal of the protocol is to calculate the *maximum* signal attenuation level by collecting periodic reports on the attenuation levels of individual users.

It is assumed that the set of all active users are readily available at the NOC at all times, since this information is also required for login, bandwidth allocation, and possibly billing purposes. Therefore, when a user terminal u_i becomes active, the protocol initializes and keeps a state variable \hat{s}_i at the NOC for recording the attenuation level of that user. The protocol updates the state variables and calculates a new maximum at every collection period by a two-step process:

- 1) Let \mathcal{U}_k denote the set of active users at the start of the CSI collection period k at time $t = kT$, where T is the collection period, and let $s_i(kT)$ denote the attenuation level sample measured by user u_i at the start of this collection period. Users send their measurements to the NOC, which collects and updates the values:

$$\hat{s}_i[k] \leftarrow s_i(kT) \quad \forall i \in \mathcal{U}_k. \quad (1)$$

- 2) After the collection, the NOC calculates the new maximum:

$$s_{\max}[k] \stackrel{a}{=} \max_{i \in \mathcal{U}_k} \{\hat{s}_i[k]\} \stackrel{b}{=} \max_{i \in \mathcal{U}_k} \{s_i(kT)\}. \quad (2)$$

The update operation in (1) requires all active users to acquire access to uplink time-slots to transmit their feedback messages. However, observe that feedback volume would be minimized if only the user with the maximum attenuation level responded at every collection period. In order to reduce the volume of feedback information that is transmitted through the network, our FIS algorithm modifies the behavior of the CSI collection protocol such that, with FIS in place, only a subset $\hat{\mathcal{U}}_k \subseteq \mathcal{U}_k$ of active users report their measurements using at most M_k uplink time-slots for the collection period. Therefore, not all state variables are updated as in (1) at the end of collection period k :

$$\hat{s}_i[k] \leftarrow s_i(kT) \quad \forall i \in \hat{\mathcal{U}}_k, \quad (3a)$$

$$\hat{s}_i[k] \leftarrow \hat{s}_i[k-1] \quad \forall i \notin \hat{\mathcal{U}}_k. \quad (3b)$$

Consequently, the maximum calculated using the state variables may not be equal to the maximum of the attenuation samples measured at time $t = kT$, — i.e. equality (b) in (2) may not hold in general. Therefore, we denote this maximum calculated when FIS algorithm is in place by:

$$\hat{s}_{\max}[k] = \max_{i \in \mathcal{U}_k} \{\hat{s}_i[k]\}. \quad (4)$$

The goal of the FIS algorithm is, therefore, to limit the error between the actual maximum in (2) and the partial maximum in (4) by determining the users in set $\hat{\mathcal{U}}_k$ and their assignments to a given number of uplink time-slots at every collection period. In the next section, we describe the FIS policy for determining set $\hat{\mathcal{U}}_k$.

IV. FIS POLICY

A. Policy Formulation

Let M_k be the maximum number of time-slots that can be allocated to transmission of feedback messages for collection period k . The system may reserve a fix number of time-slots for feedback transmission, or the value of M_k be determined by the remaining time-slots after the user requests for data transmissions are accommodated. The FIS policy selects users and assigns them to one of the M_k available uplink time-slots by solving a *multiple subset sum* problem (MSSP) [5]. Multiple subset sum problem is a variant of the bin packing problem in which the number of bins is given. Mathematically, the MSSP can be described as follows:

Definition 1: Given a set $\{u_i\}_{i=1}^N$ of items, each item u_i having a positive weight w_i , and a set $\{b_j\}_{j=1}^M$ of identical bins with positive capacity C , what is the assignment that maximizes:

$$\max \sum_{j=1}^M \sum_{i=1}^N w_i x_{ij} \quad (5a)$$

$$\text{subject to } \sum_{i=1}^N w_i x_{ij} \leq C, \quad \forall j, \quad (5b)$$

$$\sum_{j=1}^M x_{ij} \leq 1, \quad \forall i, \quad (5c)$$

$$x_{ij} \in \{0, 1\}, \forall i, j. \quad (5d)$$

Without loss of generality, we assume that $N \geq M$, otherwise the problem is trivially solved.

At the start of the CSI collection period k , using the set of active users in \mathcal{U}_k , and the state variables $\{\hat{s}_i[k-1]\}_{i \in \mathcal{U}_k}$, the FIS protocol solves an instance of the MSSP where,

$$N = |\mathcal{U}_k|, \text{ cardinality of the set } \mathcal{U}_k, \quad (6a)$$

$$M = M_k, \text{ number of available time-slots,} \quad (6b)$$

$$w_i = \min\{\hat{s}_i[k-1], A_{\max}\}, \forall i \in \mathcal{U}_k, \quad (6c)$$

$$C = A_{\max}, \quad (6d)$$

where A_{\max} is the maximum attenuation level that can be compensated by the fade countermeasures implemented by the system. Equation (6c) allows all users with a reported attenuation level larger than A_{\max} be treated as equals since the system does not have the means to compensate for their attenuation level above this limit.

Let \mathbf{x}^k be the solution vector that represents the assignment of items to bins. The assignment of items to bins represents the assignment of users to uplink time-slots, such that $x_{ij}^k = 1$ if user u_i is selected and assigned to time-slot j for $\forall i \in \mathcal{U}_k$, and $j = 1, 2, \dots, M_k$, else $x_{ij}^k = 0$.

The NOC broadcasts the solution vector over the network as a part of the existing on-demand bandwidth allocation procedure and reserves the necessary uplink time-slots for the collection period. The MSSP formulation provides a procedure for selecting and assigning users to available uplink time-slots, however, it may assign more than one user to a particular time-slot. Therefore, the FIS algorithm must have a collision avoidance strategy to limit the collisions. In order to avoid collisions, a user u_i chooses to transmit its feedback message during its assigned time-slot with probability $p_i(k)$. The procedure for choosing collision avoidance probability values is discussed in the subsequent sections. The set $\hat{\mathcal{U}}_k$ is constructed as:

$$\hat{\mathcal{U}}_k = \{u_i : x_{ij}^k = 1 \cap \mathbf{1}(\mathcal{E}_i^k)\}, \quad (7)$$

where $\mathbf{1}(\cdot)$ is the indicator function, \mathcal{E}_i^k is the event that u_i transmits in period k , such that $E(\mathbf{1}(\mathcal{E}_i^k)) = p_i(k) \forall i$.

In the following section, we discuss the important properties of this policy formulation that enable efficient feedback suppression in the context of CSI collection.

B. Policy Properties

MSSP is a NP-hard problem, and we assume that *best-fit-decreasing* (BFD) approximation is implemented for the solution [5], not only because it is one of the fastest heuristics, but also it has desirable properties in achieving the goals of FIS policy.

Property 1: BFD algorithm starts the selection and allocation of the items first from the ones with the largest weights. The algorithm always selects the first M_k largest items. The remaining items are selected only if there is room.

Property 1 states that the first M_k users with the highest reported attenuation levels are always selected and assigned

to a time-slot in the current collection period. The remaining users face the possibility of suppression depending on their attenuation levels and the number of available time-slots. Assuming channel state varies slowly, this property effectively reduces the size of the user set that will be selected in the next period.

Property 2: Items with larger weights, if selected, are more likely to occupy a bin that has few other items since they fill most of the capacity of a bin.

Property 2 decreases the probability that a user with a high reported attenuation level will be involved in a collision in the current collection period, since it is more likely to be assigned to a time-slot with few other users.

Property 3: Items with smaller weights are only selected if there is available bin space after placement of items with larger weights.

Finally, Property 3 states that users with lower reported attenuation levels are *suppressed* by the *selection* process. The suppression becomes more severe as the number of available uplink time-slots decreases, since the MSSP algorithm runs with fewer bins. It may be argued that selecting the largest M_k items and allocating them into bins one-by-one for the collection period k would have been a better strategy, since that would have avoided collisions between users. However, note that the allocation decision at the NOC is based solely on the *last successfully reported* attenuation level of each user, rather than the current value measured by the user. Such a strategy would cause a user with a previously low attenuation level to remain suppressed for a long time, even if it currently measures a high attenuation level, and consequently reduce the robustness of FIS algorithm against changes in the channel conditions. This observation also is the basis for our collision avoidance strategy which we describe in the next section.

C. Collision Avoidance Strategy

At every collection period, an active user only knows its channel assignment for that period and the value of its current attenuation level measurement. We assume that users can also listen to their own transmissions and detect collisions. Therefore, all users keep a record of the value of the last attenuation sample they successfully transmitted to the NOC. This value is equal to the value of the state variable kept at the NOC for the user, which the NOC always uses in calculating the maximum as well as in determining the channel assignment vector for the FIS policy. The transmission probability $p_i(k)$ of the user u_i should be proportional to the discrepancy between its current measured level $s_i(kT)$ and the last reported level, $\hat{s}_i[k-1]$, since the error would not only affect the selection and suppression criteria in the subsequent collection periods, but more importantly would increase the error between the true maximum and the calculated maximum. Therefore, the user behaves more aggressively in transmitting its feedback message, if there is a significant change in its condition. However, in order to put a cap to the aggressiveness of a user, the functional relationship between the difference and probability of transmission should saturate as the difference

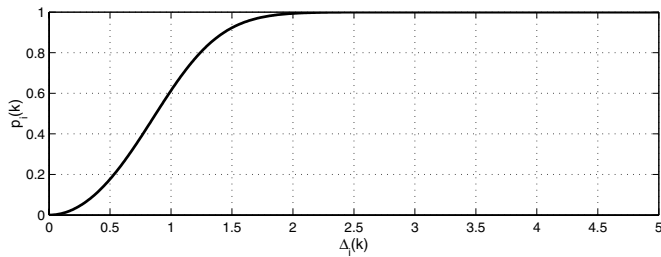


Fig. 2. Probability of transmission as a function of Δ (dB).

goes to infinity. In this paper, we choose:

$$p_i(k) = \frac{2}{(1 + e^{-\Delta_i(k)^2/b})} - 1.0, \quad (8)$$

for a given b , where $\Delta_i(k) = |\hat{s}_i[k-1] - s_i(kT)|$. In Fig. 2, we plot this curve for $b = 0.7$.

V. PERFORMANCE ANALYSIS

A. Channel model

The performance of our FIS policy depends on the statistical behavior of the underlying channel. In this paper, we use a Ka-band channel model that is based on the simulator developed at DLR (German Aerospace Center), Institute for Communications and Navigation [6]. The model is based on specific channel model parameters from the DLR measurement campaign carried out at Oberpfaffenhofen near Munich, Germany, in the years 1994 till 1997 with the 40 GHz beacon of the Italian satellite ITALSAT.

B. Analysis Results

In this section, we first look at the performance of the FIS algorithm for a set of $U = 200$ users over a simulated time of 1280 minutes. This test duration corresponds to a total of $K = 1200$ CSI collection periods with an interval of $T = 64$ seconds, which is the granularity of our channel simulator. The test is repeated 100 times, and average results are reported.

During the simulation, a separate instance of the channel simulator is run for each user, modeling its attenuation level. The attenuation time series generated for a user is independent from other users. In general, this assumption may not always be valid, since attenuation due to rain typically occurs in cells of certain diameter that may include more than one user. However, a study by Fukuchi [7] found that this correlation decreases rapidly with increasing distance, and that instantaneous rainfall rates at two locations more than about 100km apart can be regarded as independent. Also, we may argue that such a correlation would only improve the performance of our policy by decreasing the number of users with different attenuation measurements.

In Fig. 3, we plot a realization of the time series of maximum channel attenuation calculated from the user samples (dashed-line), versus the maximum channel attenuation level calculated from state variables at the NOC (solid-line), for $M = 2, 10, 20$ uplink time-slots. We observe that the two time-series are in close agreement for $M = 20$ and $M = 10$, but the error becomes more significant for $M = 2$. We calculate the

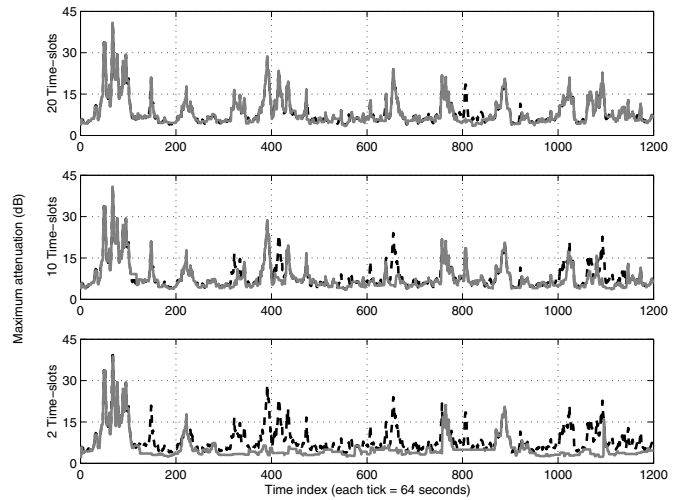


Fig. 3. Time-series plot of actual channel attenuation level (dashed-line) versus the level calculated at the NOC (solid-line) for a set of 200 users.

average absolute error between the two curves as a function of the number of available uplink time-slots:

$$error(M) = \frac{1}{K} \sum_{k=1}^K |s_{\max}^M[k] - \hat{s}_{\max}^M[k]|, \quad (9)$$

which we tabulate in Table I.

M	30	20	10	2	1
Error (dB)	0.3973	0.7117	1.4855	3.7733	4.5522

TABLE I
AVERAGE ABSOLUTE ERROR

In Fig. 4(a), we plot the fraction of the total users suppressed by the FIS algorithm as a function of the number of available uplink time-slots. A user's feedback may be suppressed in one of the two ways: (i) it is not selected by the MSSP solution, or (ii) its feedback experiences a collision and is not reported successfully to the NOC. Figure 4(a) shows the contribution of the two cases in each instance. We observe that, 95 – 99% of users are suppressed on the average by the FIS algorithm. More users are suppressed through the selection process than channel collisions as the number of available time-slots decreases. When there are available time-slots, MSSP formulation selects and assigns more users. This leads to more collisions on the average per user in the assigned time-slots, however, also gives the opportunity to sample a larger user space if some feedback successfully reaches to the NOC. When there are very few time-slots, MSSP formulation suppresses most of the users. This leads to fewer collisions per user, however, increases the error if the wrong subset of users are assigned to the time-slots (compare to Table I). In Fig. 4(b), we look at the channel-usage profile as a function of available uplink time-slots. We observe that a time-slot is utilized successfully 30 – 35% of the time. This percentage remains relatively constant independent of the ratio of the number of available time-slots to number of users.

In order to assess the performance of our policy under different user group sizes, we plot, in Fig. 5, the average

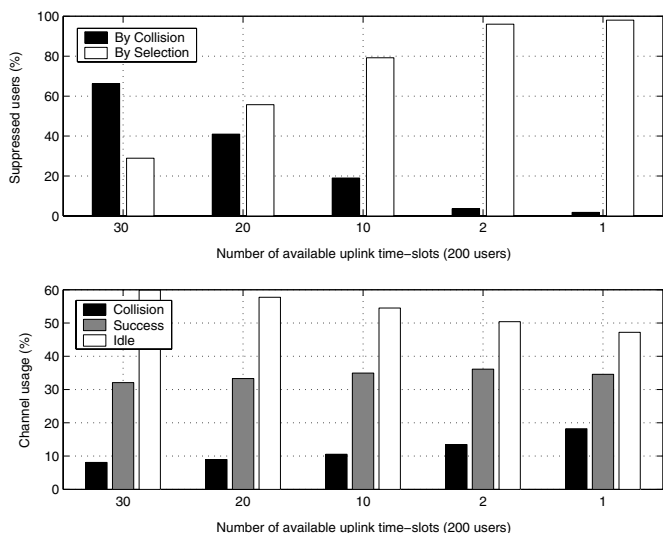


Fig. 4. (a) Fraction of the users suppressed by the FIS algorithm as a function of the number of available uplink time-slots (200 users). (b) Channel usage profile as a function of the number of available uplink time-slots (200 users).

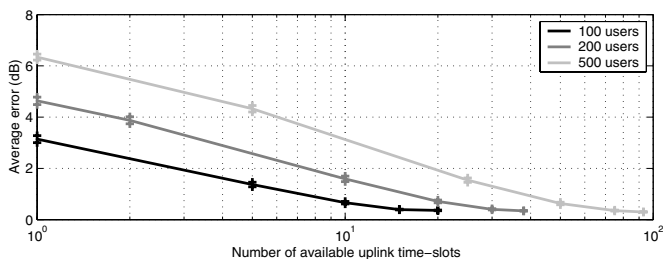


Fig. 5. Average error versus number of available uplink time-slots as a function of number of active users.

error in the calculated attenuation level versus the number of available uplink time-slots for a range of user groups. We observe that for a given number of time-slots, the average error of FIS algorithm increases as the number of users increase. This result is expected since the algorithm is forced to select and allocate users from a larger set to a fix number of time-slots. Therefore, more users are suppressed by the algorithm, resulting in less accurate calculation of the maximum attenuation level. However, more importantly, if the ratio of available time-slots to number of users is kept constant, then we observe that the algorithm performs equally well for all group sizes.

Although average error in the calculated attenuation level is an indicator of the performance of the FIS algorithm, from a system provider's or user's perspective, a more important performance metric would be the effect of this error as an input in fade compensation. In order to assess this performance, we look at the use of the attenuation level as an input to a simple downlink power control scheme. We assume that the system has a 10dB link-budget power margin that is used to compensate for the signal attenuation due to atmospheric and rain fading. The system allocates this power margin based on the calculated maximum attenuation in order to meet a target bit-error-rate (BER). With no FIS in place, the system will not be able to meet this BER target, only when the

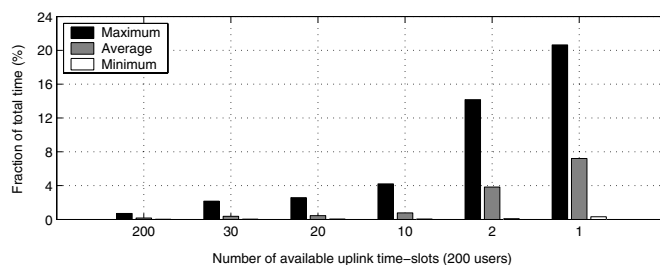


Fig. 6. Percentage of total time BER target is not met by a user as a function of number of available uplink time-slots (200 users).

maximum attenuation level is more than the 10dB power margin. However, with FIS in place, the error in calculating the maximum attenuation level may cause the system to underestimate the actual attenuation level. This will result in additional time instances when the BER target is not met by some users. In Figure 6, we look at the maximum, average and minimum percentage of time a user will not meet the BER target as a function of the number of available uplink time-slots for a user groups of 200.

We observe that, average performance degradation experienced by a user remains acceptable even if the number of available time-slots is as low as the 5% of the total user population.

VI. CONCLUSIONS

In this paper, we presented a feedback implosion suppression algorithm that uses a multiple subset sum (MSS) formulation in order to select and assign users to a limited number available uplink channels (time-slots). It complements a CSI collection protocol, which collects feedback from active users in an attempt to track down the user with highest channel attenuation level. FIS algorithm successfully achieves this task, even when the number of available uplink time-slots is as low as 5% of the user group size. Current research work is underway to integrate this algorithm into an adaptive power allocation scheme for fade compensation.

REFERENCES

- [1] J. V. Evans, "Communication satellite systems for high speed Internet access," *IEEE Antennas and Propagation Magazine*, vol. 43, no. 5, October 2001.
- [2] L. Castanet, A. Bolea-Almanac, and M. Bousquet, "Interference and fade mitigation techniques for Ka and Q/V band satellite communication systems," in *Proc. 2nd International Workshop of COST Action 280*, May 2003.
- [3] G. Akkor, M. Hadjitheodosiou, and J. S. Baras, "Transport protocols in multicast via satellite," *Int. J. Sat. Comm.*, vol. 22, no. 6, pp. 611-627, November/December 2004.
- [4] S. Cho and I. F. Akyildiz, "A poker-game-based feedback suppression algorithm for satellite reliable multicast," in *Proc. of IEEE GLOBECOM 2002*, November 2002.
- [5] M. R. Garey and D. S. Johnson, *Computers and intractability: a guide to the theory of NP completeness*. San Francisco: W. H. Freeman and Company, 1979.
- [6] U.-C. Fiebig, "A time-series generator modelling rain fading and its seasonal and diurnal variations," in *Proc. 1st International Workshop of COST-Action 280*, Malvern, UK, 2002.
- [7] H. Fukuchi, "Quantitative analysis of the effect of adaptive satellite power control as a rain attenuation countermeasure," in *Proc. IEEE International Symposium on Antennas and Propagation*, vol. 2, 1994, pp. 1332-1335.