Cross-Layer Optimization for Satellite-Terrestrial Heterogeneous Networks

(Invited Paper)

Jihwan P. Choi* and Changhee Joo†

*Dept. of Information and Communication Eng.
Daegu Gyeongbuk Institute of Science & Technology (DGIST), Korea

†School of Electrical and Computer Engineering,
Ulsan National Institute of Science and Technology (UNIST), Korea

Email: jhchoi@dgist.ac.kr, cjoo@unist.ac.kr

Abstract*—The advanced multibeam satellite equipped with phased array antenna can effectively serve a large number of users over its coverage area by beamforming narrow spotbeams and managing interbeam interference. In this paper, we investigate a cross-layer scheme of resource allocation and user scheduling when the multibeam satellite has a choice between routing signals to a ground gateway/feeder antenna/gap filler or sending directly to end user terminals. The satellite scheduler faces a problem of choosing a better signal path by considering transmission diversity to improve reliability and throughput. We solve a general utility problem based on channel conditions and potential interbeam interference that is primarily determined by geographical distribution of users and gateway locations. The optimum solution chooses users to be served at each transmit opportunity and the optimum transmission strategy that gives the highest marginal return of the utility in terms of transmit power allocation and signal path selection. Our cross-layer optimized scheme for satellite-terrestrial heterogeneous networks takes into joint account the PHY layer beamforming, the MAC layer scheduling, and the Network layer routing path selection. From the optimum scheduling policy, a computationally efficient algorithm is proposed. Simulation results show that the use of gateway/feeder antenna can increase the sum capacity of the multibeam satellite regardless of the interference level.

***Keywords*- multibeam satellite, cross-layer, satellite-terrestrial heterogeneous network, resource allocation, user scheduling.**

I. Introduction

The importance of efficient satellite communications is ever-growing, especially in case of disaster recovery and rescue missions. When tragic disasters such as earthquakes, tsunamis, and hurricanes may disrupt terrestrial network services, satellite networks can still provide robust regional coverage for victims and rescuers. The capability of satellites can be enhanced when they are interlinked with ground components, which can fill some coverage holes and improve system throughput. This "heterogeneous" network can perform better when the different types of components over the sky and on the ground are interconnected seamlessly by protocols and algorithms that consider the heterogeneity of the system.

One applicable scenario is when a disaster recovery team is deployed in the area where terrestrial infrastructure is heavily damaged [1]. The recovery team sets up mountable antennas that can play a role of gateways or relay stations between the satellite and end user terminals. Each end user terminal for a recovery team member can also receive the satellite signal directly from the satellite. The multibeam satellite is given path diversity for each beam and the scheduler faces a problem of choosing a better signal path, or even considering transmission diversity for higher reliability and packet segmentation for better throughput. The quality of service (QoS) requirements and priority of the packets should be taken into consideration as well. Path diversity can make the system performance less dependent of channel conditions, especially when some users are under a bad channel condition for a long time.

Not only in emergencies but in daily uses, do satellite networks carry more and more IP-based data packets in addition to traditional telephony, data trunking by circuit switching, and broadcasting of video/audio. However, conventional satellite network architecture for long duration circuit traffic is not sufficient for supporting bursty unscheduled computer data. For data satellite networks to be cost-competitive with respect to terrestrial systems, packet (datagram) service over satellites requires efficient utilization of precious resources, such as power, transmitter, receiver and amplifier.

In our previous work [2], [3], we showed that the advanced multibeam satellite equipped with phased array antenna can effectively serve a large number of users over its coverage area by scheduling narrowly beamformed spotbeams and managing interbeam interference. Multibeam satellites can provide high throughput by increasing the received power density and thus, the power efficiency for data rates. In this paper, we extend the previous results into the satellite-terrestrial heterogeneous system by incorporating ground stations to receive satellite signals, and investigate a scheme of resource allocation and user scheduling when the multibeam satellite has a choice of routing signals to gateways/feeder antennas/gap fillers and/or sending directly to end user terminals. We formulate a general utility problem and derive a resource allocation and user scheduling solution based on channel conditions and potential interbeam interference that is primarily determined by geographical distribution of users and gateway locations. By comparing direct transmission to users and routing via gateways, the scheduler chooses the users to be served at each transmit opportunity and the optimum transmission strategy that gives the highest marginal return of the utility in terms of transmit power allocation and signal path selection.

In Section II, we formulate a resource allocation problem for a satellite aided by ground gateway/feeder antenna. In Section III, we present optimum antenna gain patterning results for very close-in users from our previous work [2], [3]. The cross-layer problem is analyzed and a computationally efficient algorithm is suggested in Section IV. In Section V, the performance of the satellite-terrestrial heterogeneous system is compared with that of the satellite-only network. We conclude the paper and suggest future work in Section VI.

II. Formulation

Traditionally, gateway links have used different frequencies from user links for communication satellites such as Globalstar. There are only a few conceptual patents that propose the frequency reuse for gateway and user beams [4], and the use of repeaters for combining satellite and repeater signals at user terminals [5]. Here, we consider small ground stations and assume that both types of links may suffer from significant interbeam interferences by sharing the same frequency. For ease of exposition, we start from a simple model where there is only one gateway/feeder antenna/gap filler. Later, we will generalize our solution taking into account more practical conditions.

We consider a scenario where a satellite with K active spotbeams has a gateway/feeder antenna/gap filler in addition to M users on the Earth (Fig. 1). Each user is assumed to receive its downlink signal either directly from the satellite or via a gateway/feeder antenna/gap filler. Then, the satellite scheduler faces a problem of choosing a better signal path between the two. Let P_i denote the power allocated to the direct signal to user i and Q_i the power allocated to the gateway/feeder antenna/gap filler that will forward the received data to user i eventually. Given the total power constraint P_{total} of the satelllite, we maximize a general utility U that is a jointly concave function of P_i and Q_i. Roughly speaking, the concave property of U holds not only for each P_i and Q_i, but for a pair of (P_i, Q_i) jointly. We can formulate the scheduling problem as follows:

$$\text{maximize} \quad U(P_i, Q_i) \tag{1}$$

$$\text{subject to} \quad \sum_{i=1}^{M} (P_i + Q_i) \leq P_{total} \tag{2}$$

$$\text{and} \quad P_i \geq 0, \quad Q_i \geq 0 \quad \text{for every } i, \tag{3}$$

where the total power P_{total} of the satellite constrains the sum of power allocated to multiple beams. All the power amounts of P_i, Q_i, and P_{total} have the values scaled by free space loss to the ground from far-field transmission.

We assume the use of phased array antenna (Fig. 2(a)) that uses solid state power amplifiers (SSPA) and can linearly superimpose signals at array elements by controlling an antenna-patterning matrix. SSPAs are frequently used to feed a large number of array antenna elements. Signal power can be divided among multiple channels up to the total power of the array. While the multiple beam antenna (Fig. 2(b)) with travelling wave tube amplifiers (TWTAs) has a fixed beam size due to the fixed size of feedhorn for each signal, the phased array antenna can have any size and/or shape of beam by feeding many array elements with the same signal. Moreover, the phased array antenna together with transmission scheduling can cycle much more rapidly ($\ll$ msec) than the multiple beam antenna and is advantageous in meeting time deadlines via fast

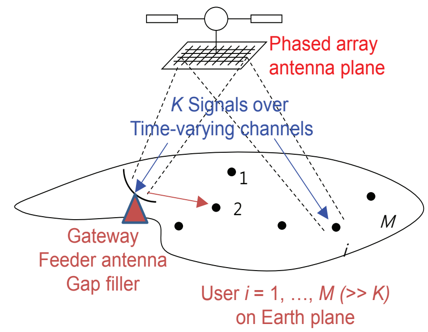

Fig. 1. A multibeam satellite with K active spotbeams and a gateway/feeder antenna/gap filler on the Earth, to serve M users

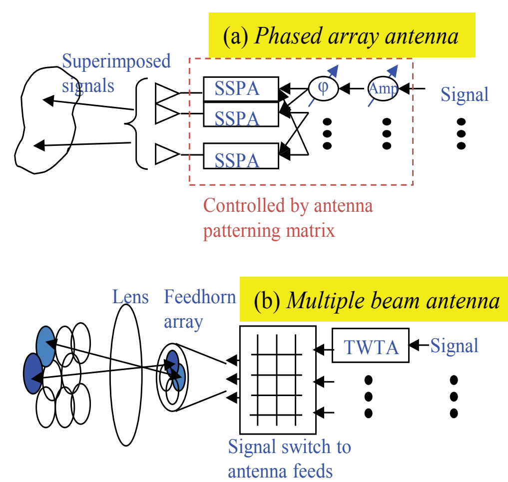

Fig. 2. Schematics of (a) phased array antenna and (b) multiple beam antenna

switching of resources. Flexible antenna gain patterning allows for simultaneous service of the users in a populated area by suppressing possibly significant interbeam interference.

III. Antenna Gain Patterning

One advantage of using the phased array antenna is to provide flexible antenna gain patterning. We here compare three schemes of managing interference: 1) complete cancellation in the form of transmission zero-forcing, 2) optimum suppression that maximizes throughput and 3) sequential service that serves users one by one, not simultaneously. For a simple example, suppose that two active users are separated by distance $l < \frac{\lambda L}{D}$, which is the mainlobe width of the narrowest sportbeam achievable by using the carrier of wavelength λ, the antenna of diameter D and the satellite located at the altitude of L. We define the carrier-to-interference ratio, CIR $= -20 \log_{10} \left[\text{sinc}\left(\frac{Dl}{\lambda L}\right)\right]$ (dB), where the sinc function is modeling far-field transmission from uniform field distribution over the whole antenna aperture with unit power, as we assume that the worst-case interference results from the other active user at the same signal-noise ratio (SNR) but with no interference suppression. In Fig. 3 we compare capacities of the schemes as changing the CIR. The scheme without interference suppression is also shown as a benchmark. At the SNR of 10.2 dB the gap between optimum suppression

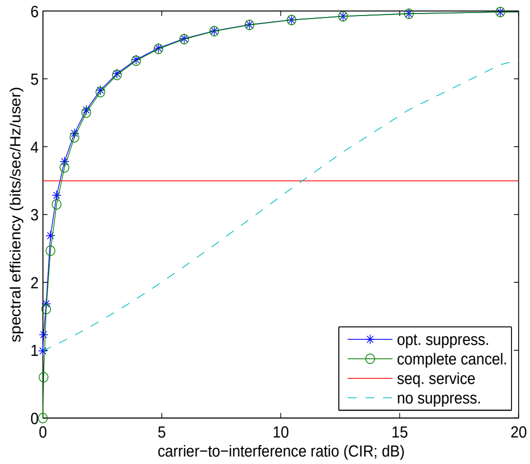

Fig. 3. Capacity of one user as a function of CIR at the SNR of 10.2 dB

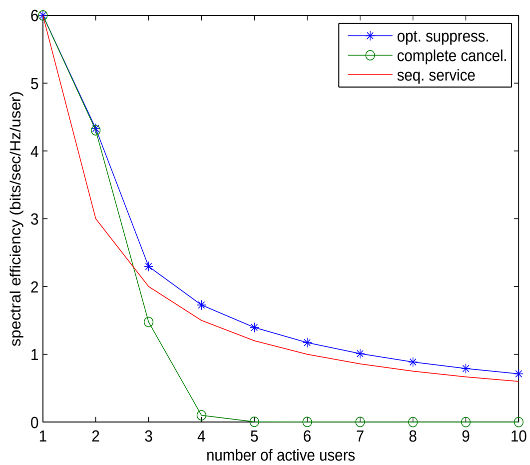

Fig. 4. Capacity of one user as a function of the number of active users within one beamwidth at the SNR of 10.2 dB

and complete cancellation is very small. If the CIR is smaller than some threshold of CIR* that is decided by comparing interference suppression and sequential service (CIR$^* = 0.58$ dB in this example), signal degradation is too severe due to co-channel interference and it is better to provide sequential service. If CIR $>$ CIR* and $l < \frac{\lambda L}{D}$, active users share the bandwidth and timeslots, and appropriate optimum antenna gain patterning is deployed with interference suppressed.

Fig. 4 shows performance comparison of the different schemes in terms of the number of uniformly located active users within $0 < l \leq \frac{\lambda L}{D}$. We observe the advantage of multiple signals over a single beam of sequential service for two or three active users, in spite of power loss from interference suppression. The complete cancellation scheme is very vulnerable to more than three active users. As the number of active users increases, the gap between optimum suppression and sequential service decreases and is negligible. Eventually, sequential service outperforms optimum suppression for a large number of active users (not shown in the

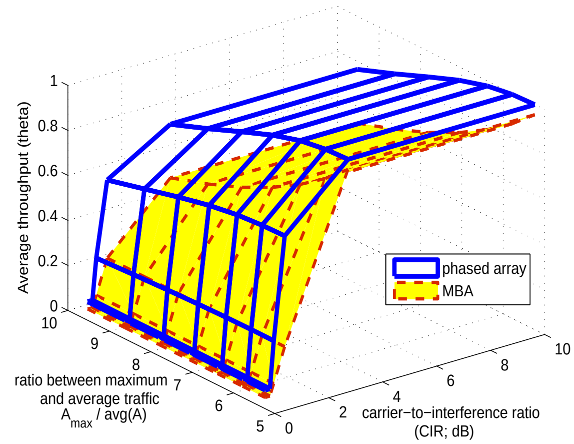

Fig. 5. Steady-state throughput performance comparison of the phased array antenna and the multiple beam antenna as a function of traffic distribution and CIR

figures). This is consistent with what is shown in the Gaussian interference channel problem and our two-user example in Fig. 3: sequential service is better than SDM (with or without interference suppression) under severe interference. Thus, the use of simple sequential service can be recommended when an area is crowded with many users.

Fig. 5 compares the performances of the phased array antenna and the multiple beam antenna with respect to traffic distribution and the CIR. To mitigate interference as the CIR decreases, the multiple beam antenna reduces the number of active beams since interference suppression cannot be supported. On the other hand, the phased array antenna considers possible interference between close-in users, and then selects the better of interference suppression and sequential service. The result shows that the phased array antenna always outperforms the multiple beam antenna. The advantage of the phased array antenna over the multiple beam antenna can be observed especially when there is a moderate level of interference between active users, so that interference suppression can be used (0.58 dB $<$ CIR $<$ 2.42 dB).

IV. Beam Scheduling

A. Analysis

To see which signal path should be selected in the optimization problem (1)-(3), we apply the Kuhn-Tucker condition [6], and observe which P_i and/or Q_i have a non-zero value. The corresponding Lagrangian function is $J(P_i, Q_i) = U - \Lambda \cdot (\sum_i P_i + \sum_i Q_i - P_{total}) - \sum_i \mu_i \cdot (-P_i) - \sum_i \nu_i \cdot (-Q_i)$, where Lagrangian variable Λ (≥ 0) is for the total power constraint, μ_i (≥ 0) and ν_i (≥ 0) are for non-negative constraints of P_i and Q_i respectiely, and added to see which users should be served with non-zero P_i and/or Q_i. The Kuhn-Tucker condition yields $\mu_i > 0$ if $P_i = 0$ and $\mu_j = 0$ if $P_j > 0$, which applies to ν_i and Q_i in the same manner. Differentiating $J(P_i, Q_i)$ with respect to P_i and Q_i respectively gives

$$\frac{\partial J}{\partial P_i} = \frac{\partial U}{\partial P_i} - \Lambda + \mu_i, \tag{4}$$

and

$$\frac{\partial J}{\partial Q_i} = \frac{\partial U}{\partial Q_i} - \Lambda + \nu_i. \tag{5}$$

We first consider the case where only a single path is allowed for each user, in which there are three possibilities of optimum power allocation: user i has $P_i^* > 0$ and $Q_i^* = 0$, user j has $P_j^* = 0$ and $Q_j^* > 0$, and user k has $P_k^* = 0$ and $Q_k^* = 0$. In other words, user i receives a packet directly and user j's packet is sent to the gateway while user k is not served at this opportunity. At the optimum power $P_i^* > 0$ and $Q_i^* = 0$, user i has $\mu_i = 0$ and $\nu_i \geq 0$. This gives

$$\left.\frac{\partial U}{\partial P_i}\right|_{(P_i=0,Q_i=0)} > \left.\frac{\partial U}{\partial P_i}\right|_{(P_i=P_i^*,Q_i=0)} = \Lambda, \tag{6}$$

where the inequality holds due to the concavity of U in terms of P_i and the equality comes from setting Eq. (4) equal to zero. Similarly at the optimum power $P_j^* = 0$ and $Q_j^* > 0$, user j has $\mu_j \geq 0$ and $\nu_j = 0$. Again from Eq. (5), we obtain

$$\left.\frac{\partial U}{\partial Q_j}\right|_{(P_j=0,Q_j=0)} > \left.\frac{\partial U}{\partial Q_j}\right|_{(P_j=0,Q_j=Q_j^*)} = \Lambda. \tag{7}$$

If both P_k^* and Q_k^* have zero values, user k has $\mu_k \geq 0$ and $\nu_k \geq 0$, and thus,

$$\Lambda \geq \left.\frac{\partial U}{\partial P_k}\right|_{(P_k=0,Q_k=0)}, \tag{8}$$

and

$$\Lambda \geq \left.\frac{\partial U}{\partial Q_k}\right|_{(P_k=0,Q_k=0)}. \tag{9}$$

If there are only $K(< M)$ active beams available in the practical multibeam satellite, we should select K options out of $2M$ candidate pairs of user and path. The above results imply that the optimal user and path selection can be obtained based on the marginal returns of the utility function at zero power of each path. That is, through (6) and (9), Λ can be interpreted as the threshold level of the partial derivative, which should be less than the marginal return of the utility function at zero power, if the path has a non-zero power value in the optimal power allocation vector. Hence, if only a single path is allowed for each user, we choose K highest values of $\left.\frac{\partial U}{\partial P_i}\right|_{(P_i=0,Q_i=0)}$ and $\left.\frac{\partial U}{\partial Q_j}\right|_{(P_j=0,Q_j=0)}$ such that $i \neq j$.

We now relax the single path assumption for each user, in which case we should consider the case that a user m has non-zero values for both P_m and Q_m. With both $P_m^* > 0$ and $Q_m^* > 0$, user m has $\mu_m = \nu_m = 0$, and thus,

$$\left.\frac{\partial U}{\partial P_m}\right|_{(P_m=P_m^*,Q_m=Q_m^*)} = \left.\frac{\partial U}{\partial Q_m}\right|_{(P_m=P_m^*,Q_m=Q_m^*)} = \Lambda, \tag{10}$$

where the marginal returns of U at P_m^* and Q_m^* respectively are equal. From the assumption of a jointly concave function U with respect to P_m and Q_m, user m should have

$$\left.\frac{\partial U}{\partial P_m}\right|_{(P_m=0,Q_m=0)} > \left.\frac{\partial U}{\partial P_m}\right|_{(P_m=P_m^*,Q_m=Q_m^*)} = \Lambda, \tag{11}$$

and

$$\left.\frac{\partial U}{\partial Q_m}\right|_{(P_m=0,Q_m=0)} > \left.\frac{\partial U}{\partial Q_m}\right|_{(P_m=P_m^*,Q_m=Q_m^*)} = \Lambda. \tag{12}$$

Again, we choose the K highest values of $\left.\frac{\partial U}{\partial P_i}\right|_{(P_i=0,Q_i=0)}$ and $\left.\frac{\partial U}{\partial Q_j}\right|_{(P_j=0,Q_j=0)}$, but without any restriction to i and j. To summarize, we see that the marginal return at zero power can be used for a metric to allocate power and decide the path for data to be sent over.

The next question is how to model the utility function of power P_i and Q_i. We assume the data rate achieved by P_i as Shannon capacity:

$$C_i = W \log\left(1 + \frac{\alpha_i^2 H_i P_i}{W N_0}\right), \tag{13}$$

where α_i^2 (≤ 1) is signal power attenuation due to the weather effects [7], [8] that only change quasi-statically over the link to the i^{th} user, W is the bandwidth used, and N_0 is the additive white Gaussian noise power density. The value of H_i represents the power loss incurred by deploying zero-forcing transmission strategy. For active users located within one spotbeam width, we model

$$H_i = 1 - \mathrm{sinc}^2\left(\frac{l}{\lambda L/D}\right), \tag{14}$$

where we do not provide antenna gain patterning with interference suppression for more than three active users in a crowded area as shown in Fig 4. Advantages of this model are the decoupling of different signals and their capacities with respect to a set of allocated power, and the derivation of a convex optimization problem.

Modeling of data rates to gateways/feeder antennas with respect to Q_i is not so straightforward. There are a few factors to be considered. The first is whether the modulator can be shared between two signal paths for one user. In other words, the same signal can be sent via both paths to the user directly and the gateway/feeder antenna. The second is whether signals for multiple users can be aggregated to one big packet and how the packet is transmitted if aggregation is possible. Since gateways/feeder antennas usually have more antennas with bigger size than user terminals, they can take advantage of better link budget and transmission strategies, such as larger signal constellation size and bigger scale of multi-input multi-output (MIMO) antennas. Thus, it is reasonable to assume that the same modulated signal for one user cannot be shared between two paths. If one user receives her/his signal via both paths, it is assumed that a different modulator should be used for each signal path. Further we assume that multiple signals, if any, will be sent via MIMO channels, which is modelled as parallel pipelines by the singular value decomposition (SVD) method [9], for example. We represent the data rate of aggregated signals to the gateway/feeder antenna as a sum of Shannon capacities:

$$C_g = \sum_i^M G_i = \sum_i^M W \log\left(1 + \frac{\alpha_g^2 H_g Q_i}{W N_0}\right), \tag{15}$$

where $G_i = W \log\left(1 + \frac{\alpha_g^2 H_g Q_i}{W N_0}\right)$ is the gateway/feeder antenna capacity allocated to user i, α_g^2 (≤ 1) is signal power attenuation over the link to the gateway/feeder antenna, and H_g represents the power loss incurred for the gateway/feeder antenna by deploying zero-forcing transmission strategy.

With the above modeling of C_i with respect to P_i and C_g with respect to Q_i, the marginal returns at $P_i = 0$ and $Q_i = 0$ are as follow respectively:

$$\left.\frac{\partial U}{\partial P_i}\right|_{(P_i=0,Q_i=0)} = \frac{\alpha_i^2 H_i}{N_0} \cdot \left.\frac{\partial U}{\partial C_i}\right|_{(C_i=0,G_i=0)}, \tag{16}$$

and

$$\left.\frac{\partial U}{\partial Q_i}\right|_{(P_i=0,Q_i=0)} = \frac{\alpha_g^2 H_g}{N_0} \cdot \left.\frac{\partial U}{\partial G_i}\right|_{(C_i=0,G_i=0)}, \tag{17}$$

where the chain rule is applied with $\left.\frac{\partial C_i}{\partial P_i}\right|_{P_i=0} = \frac{\alpha_i^2 H_i}{N_0}$ and $\left.\frac{\partial G_i}{\partial Q_i}\right|_{Q_i=0} = \frac{\alpha_g^2 H_g}{N_0}$.

B. Algorithm

The results derived so far imply that we have to select the pairs of user and path as follow:

- better channel conditions with higher α_i^2 or α_g^2,
- less interference with higher H_i or H_g, and
- higher marginal returns of the utility at zero capacity with higher $\left.\frac{\partial U}{\partial C_i}\right|_{C_i=0}$ or $\left.\frac{\partial U}{\partial C_g}\right|_{C_g=0}$.

The selection of signal paths changes the values H_i or H_g of other paths, and then the changes of the interference values can impact the selection of signal paths. Implementing the results as derived is quite complicated. Instead, we develop a sub-optimum but computationally efficient algorithm, which will be compared to the one with no gateway/feeder antenna support [2], [3].

Among $2M$ pairs of user and path combinations, the scheduler selects the best K pairs by the following rule:

1) With every $H_i = H_g = 1$ for initial values, select a pair of user and path with the biggest value of $\alpha_i^2 \cdot \left.\frac{\partial U}{\partial C_i}\right|_{C_i=0}$ or $\alpha_g^2 \cdot \left.\frac{\partial U}{\partial C_g}\right|_{C_g=0}$ and update H_i and H_g of the remaining pairs not to interfere with the selected pair.
2) For the remaining pairs with the values of H_i and H_g updated, select the biggest value of $\alpha_i^2 H_i \cdot \left.\frac{\partial U}{\partial C_i}\right|_{C_i=0}$ or $\alpha_g^2 H_g \cdot \left.\frac{\partial U}{\partial C_g}\right|_{C_g=0}$ if the user/gateway satisfies $l > l^*$, where l is the minimum distance between the user/gateway of the biggest value of $\alpha_i^2 H_i \cdot \left.\frac{\partial U}{\partial C_i}\right|_{C_i=0}$ or $\alpha_g^2 H_g \cdot \left.\frac{\partial U}{\partial C_g}\right|_{C_g=0}$ and other active users/gateway already selected, and l^* is the distance threshold below which interference suppression cannot prevent significant power loss due to proximity between active users/gateway. Update interference level H_i and H_g for the remaining pairs. If $l \leq l^*$, reject the pair and repeat Step 2.
3) Repeat Step 2 until either
   i) selecting K pairs or
   ii) scanning all $2M$ pairs (even if less than K pairs are only served).

We allocate the optimum power $P_i^\star$ such that every selected user has an identical marginal return as in Eq. (6), (7) and (10). The required inputs to the algorithm are user locations, link conditions, and QoS requirements. In reality, for rapidly time-varying compared to long feedback latency, parameter mismatch can result in some performance loss. Note that the number of paths to a gateway/feeder antenna, denoted as N_a is limited to the order of MIMO antennas deployed. We can generalize this to the case of multiple gateways, by considering $(N_g+1)M$ pairs of paths and users, where N_g is the number of gateways deployed and each signal is assumed to be able to be routed to any gateway.

V. Simulation Results

This section provides simulation results and compares the scheme with a gateway/feeder antenna to the one without it. We consider $M = 25$ users/gateway on the 5 by 5 planar grid and the maximum number of active beams is $K = 20$. The gateway/feeder antenna is located at one point of the grid. All the 25 users/gateway are assumed in the moderately rainy channel condition of the Ka band. Rain attenuation of each user/gateway is modeled as an independent and identically distributed (i.i.d.) lognormally Gaussian random process with mean of 7.1205 dB and standard deviation of 1.8629 dB [7]. The number of the MIMO path to the gateway is assumed to be $N_a = 8$. The total power to noise ratio of $\frac{P_{total}}{W N_0}$ is chosen to be 15 dB on the ground. We change the distance between adjacent users/gateway from $0.1\frac{\lambda L}{D}$ to $1\frac{\lambda L}{D}$, i.e., the area of the total planar grid from $0.16(\frac{\lambda L}{D})^2$ to $16(\frac{\lambda L}{D})^2$. As the distance decreases, we can simulate the situation where more and more users are crowded in a small area and the impact of interbeam interference can be severe. Two values of threshold l^* are tested: $0.4\frac{\lambda L}{D}$ and $0.7\frac{\lambda L}{D}$ respectively. Simulations are iterated 10,000 times and averaged per distance and l^*. The result is compared with that of the case when there is no gateway/feeder antenna but another user in the same location. The utility U into consideration is a sum capacity of

$$U = \sum_{i=1}^{M} C_i + C_g = \sum_{i=1}^{M} (C_i + G_i). \tag{18}$$

As seen in Fig. 6, the use of gateway/feeder antenna gives extra sum capacity, compared with the case without it. The gain is observed all over the interference level that is the function of the distance between users/gateway. When the scheduler selects the gateway/feeder antenna, the total of $N_a \cdot C_g$ replaces the sum of C_i's of the last N_a users among selected if there is no gateway/feeder antenna. Testing different values of l^* sheds light on how to schedule users/gateway that are densely located within a beamwidth. A small value of l^* implies that a large number of active users are served with antenna gain patterning of interference suppression applied. On the other hand, a large value of l^* prefers a small number of active users being served and does not rely on interference

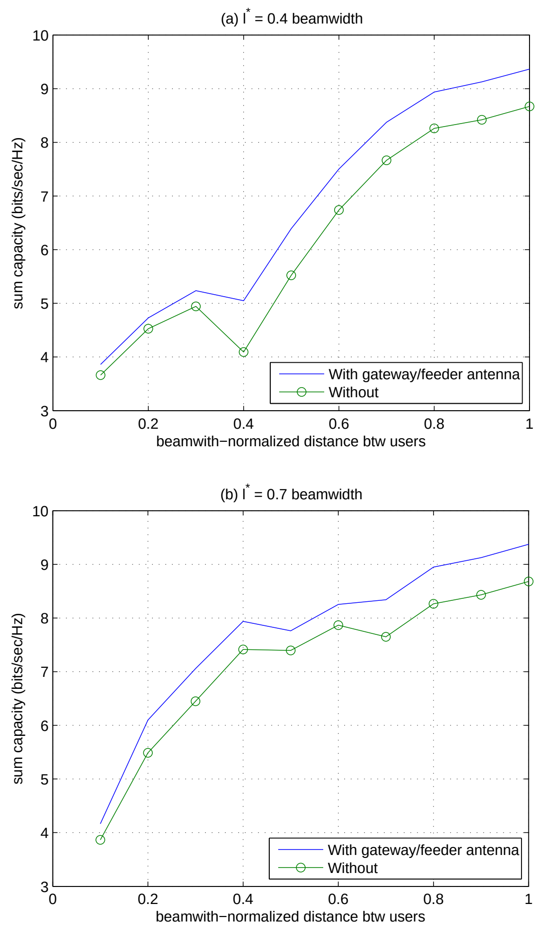

Fig. 6. Comparison of the sum capacities with and without the use of gateway/feeder antenna: (a) $l^* = 0.4\frac{\lambda L}{D}$ (b) $l^* = 0.7\frac{\lambda L}{D}$

suppression too much. Our results show that the large value of $l^* = 0.7\frac{\lambda L}{D}$ gives higher sum capacity, especially in the middle range of the distance between 0.2 and 0.7 beamwidth. With $l^* = 0.4\frac{\lambda L}{D}$ the sum capacity decreases sharply at the adjacent user/gateway distance equal to l^*, implying that the aggressive deployment of interference suppression in the crowded area results in poor performance for sum capacity. However, if additional QoS constraints are considered such as latency and provisioning the minimum capacity for every user has higher priority than maximizing the sum capacity, more aggressive scheme of antenna gain patterning for interference suppression should still be a viable solution.

VI. Conclusion and Future Work

The multibeam satellite can provide robust global and regional communication networking service when disasters strike terrestrial infrastructure. Its effectiveness can be enhanced when ground stations such as gateways, gap fillers or feeder antennas are also utilized. Resource allocation of the space and ground components should be jointly coordinated and user scheduling should be considered together with the decisions at other layers such as power/rate allocation and data routing path selection.In this paper, we modeled the simple scenario where user packets can be sent either directly to the user or via the gateway/feeder antenna/gap filler. We formulated a utility maximization problem and derived simple criteria to choose the optimum pairs of user and path for multiple beams. A computationally efficient algorithm was suggested and simulations results were compared with those of the case of no gateway/feeder antenna/gap filler. The use of MIMO ground station increases the sum capacity of the multibeam satellite.

Future work may include taking QoS requirements, such as average latency, and traffic characteristics, such as voice, data, broadcasting, unicasting, etc., into account for modeling of the utility. The problems of additional latency due to the introduction of ground station and complicated signal combining between direct and relayed signals in the form of hybrid automatic retransmission request (HARQ) can be investigated. Various receiver processing techniques should be utilized as well [10].

Further cross-layer optimization can be performed by adding the issue of congestion control in the Transport layer as well. So far, the use of proxies at the boundary of satellite and terrestrial links has been popular even if the end-to-end control principle of the TCP (Transmission Control Protocol) is violated. We may have to revisit another type of Transport layer protocols for satellite-terrestrial heterogeneous networks: e.g., rate-control based and network-assisted, so that each node is as responsible as the two end nodes for flow control. This new approach can overcome the main disadvantage of having long latency by implementing end-to-end control over long-distance satellite channels.

Acknowledgments

This work was supported by the DGIST R&D Program of the Ministry of Science, ICT and Future Planning (13-BD-0404).

References

[1] M. Berioli, A. Molinaro, S. Morosi and S. Scalise, "Aerospace Communications for Emergency Applications," *Proceedings of the IEEE*, vol. 99, no. 11, pp. 1922-1938, Nov. 2011.

[2] J. P. Choi, "Resource Allocation and Scheduling for Communication Satellites with Advanced Transmission Antennas," PhD Thesis, EECS, MIT, September 2006.

[3] J. P. Choi and V. W. S. Chan, "Resource Management for Advanced Transmission Antenna Satellites," *IEEE Trans. on Wireless Communications*, vol. 8, no. 3, pp. 1308-1321, March 2009.

[4] Mark D. Dankberg, Mark J. Miller, Kristi Jaska, and Robert Wilson, "Frequency re-use for service and gateway beams," US Patene Application No. 2009/0291633 A1, Nov. 26, 2009.

[5] S. J. Campanella, "Digital broadcast system using satellite direct broadcast and terrestrial repeater," US Patent No. 6,944,139 B1, Sept. 13, 2005.

[6] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1995.

[7] J. P. Choi, "Channel Prediction and Adaptation over Satellite Channels with Weather-Induced Impairments," Master's Thesis, EECS, MIT, May 2000.

[8] J. P. Choi and V. W. S. Chan, "Predicting and Adapting Satellite Channels with Weather-Induced Impairments," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 779-790, July 2002.

[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005

[10] S. Kim, "Evaluation of Cooperative Techniques for Hybrid/Integrated Satellite Systems," *IEEE International Conference on Communications (ICC)*, June 2011.