Evaluation of Resource Allocation Methods for Integrated Satellite-Terrestrial Systems

Hyein Lee and Sooyoung Kim
IT Convergence Research Center, Div. of Electronics Eng.
Jeonbuk National University
Jeonju, Korea
{leehyein96,sookim}@jbnu.ac.kr

Daesub Oh
Radio & Satellite Research Division
ETRI
Deajeon, Korea
trap@etri.re.kr

***Abstract*—In this paper, we evaluate resource allocation methods for an integrated satellite-terrestrial (IST) system. We propose a computationally efficient resource allocation method for an interference limited IST network, by considering that the objective function and the constraints of the conventional resource allocation method had a steep slope which hindered the optimal solution search. A new objective function is proposed to solve the problems of the conventional schemes. The simulation results evaluated using sequential quadratic programming reveal that the proposed resource allocation method provides improved allocation performance as well as computational efficiency, compared to the existing methods.**

***Index Terms*—integrated satellite-terrestrial, optimization, resource allocation, SQP**

I. Introduction

The scarcity of the satellite resources requires efficient operation of multi-beam satellite communication systems with the aid of a frequency reuse technique. The frequency reuse technique is a method to share the same frequency band among geographically separated cells [1][2]. Several studies have been conducted on a resource allocation method for the integrated satellite and terrestrial (IST) systems utilizing frequency reuse technique over a multi-beam satellite system [3]–[6]. It was discussed that, even if the cells using the same frequency band were sufficiently far apart on the ground, inter-component interference between a terrestrial cell and satellite beam was unavoidable. Therefore, the main objective of previous studies was to minimize interference between components, by allocating the minimum power for quality of service (QoS) [5][6].

The Lagrangian method has been considered as the most representative approach to solve the optimal resource allocation problem [3]–[5]. Optimal joint power and bandwidth allocation (OPOB) methods based on duality theory were proposed to meet traffic demands from satellite beams [3][4]. These algorithms found a solution even in the case when the required traffic demand is greater than the system capacity, by reducing the traffic demand in a way to maximize the system capacity. However, it was found that the OPOB method invoked serious power inefficiency [6]. Furthermore, the computational complexity exponentially increased by the frequency reuse factor and the number of components.

On the other hand, a resource allocation method that minimizes the total power requirement for an IST system was proposed [5]. This method utilized a linear system which relates the power and bandwidth for an interference limited system, and provided an analytic solution of the Lagrangian function for the optimization problem. However, this method often resulted invalid solutions. As a solution for this problem, a new approach was proposed based on simple linear machine learning (ML) methods [6], and thus it required pre-trained perceptron and linear regression tool.

In this paper, we consider the same IST system as in the previous studies [5][6], and propose an efficient resource allocation problem which solves the existing problem in [5]. By considering that the objective function for an interference-limited IST system caused a steep slope with discontinuous points, we propose a new resource allocation problem by revising the objective function and constraints. The proposed scheme tackles the problem firstly by regulating the searching range of the optimal solution, and secondly by smoothing the slopes of the objective function. The implementation algorithm for the proposed scheme is presented by using sequential quadratic programming (SQP) algorithm, and its performance and computational complexity are investigated.

This paper is organized as follows. Section II describes an IST system configuration and various existing resource allocation problems. In Section III, we introduce the proposed resource allocation method and present its implementation with the SQP algorithm. Section IV presents comparison results of the proposed method with the existing methods in terms of various performance measures as well as computational complexity. Finally, Section V draws the conclusion.

II. Resource allocation for intergrated satellite and terrestrial system

The IST system has been regarded as an effective communication model that can safely serve a wide range of multimedia

This work was supported by the Institute for Information and Communications Technology Promotion Grant funded by the Korea Government (MSIT, Development of the spectrum sharing technology for Non-GSO satellite system) under Grant 2021-0-00719.

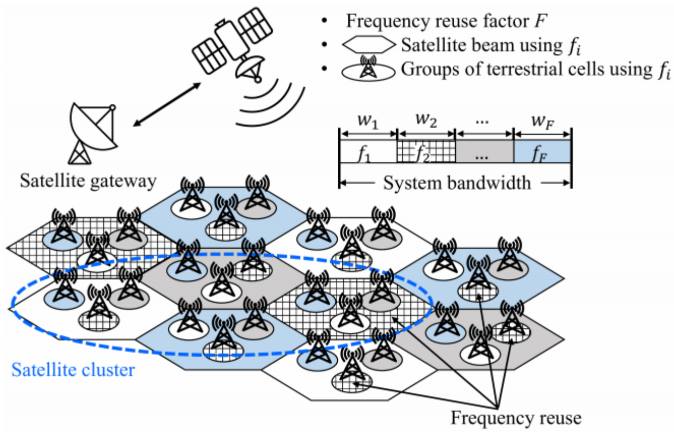

Fig. 1. A system configuration of the IST system using frequency reusing method [6].

services at low cost [5]. In this system, the terrestrial cells and the satellite beams can share the same bandwidth, extending satellite converage. In addition, both the terrestrial cells and the satellite beams are managed by a common resource control system, thus, it is possible to allocate resources that simultaneously satisfy the requirements from each of the terrestrial cells and the satellite beams.

Figure 1 depicts a configuration of the IST system, where frequency reuse factor is F [6]. The total available system bandwidth, W is divided by F, having subbands $\mathbf{w} = [w_1, \ldots, w_i, \ldots, w_F]^T$ with its center frequency, $f_1, \ldots, f_i, \ldots, f_F$, respectively. Within a satellite cluster, F satellite beams use a different bandwidth from each other, and within a satellite beam, $(F-1)$ groups of terrestrial cells use a bandwidth from another terrestrial cell and satellite beam so that all components can avoid interference caused by bandwidth overlap. We assume that M components (either beams or terrestrial cells) utilizing each subband w_i at an instance and there is no time slot re-use. Even though there is no inter-component interference by avoiding bandwidth overlap within a satellite beam, the system may suffer from inter-component interference from the adjacent beams or terrestrial cells which share the same frequency band in a satellite cluster.

For the IST network, the previous study formulated resource allocation problem in a way to minimize the total consumed transmission power as follows [5]:

$$\begin{aligned} &\underset{\mathbf{w}}{\arg\min} \quad \sum_{i}^{F}\sum_{j}^{M}(P_t)_i^j \\ &\text{s.t.} \quad \sum_{i}^{F} w_i \leq W, \quad w_i \geq 0, \quad (i = 1, \ldots, F) \end{aligned} \tag{1}$$

where the transmit power of the jth component using f_i in the unit of dBW, $(P_t)_i^j$ can be expressed as follows:

$$(P_t)_i^j = \frac{N_0 R_i^j \gamma_i^j}{G_i^{j,j}}. \tag{2}$$

In (2), N_0 is the noise spectral density, R_i^j is the traffic demand from the jth component using f_i, and $G_i^{k,j}$ is the channel gain from the kth component to jth component using f_i. In addition, γ_i^j denotes the bit energy to noise spectral density ratio (E_b/N_0) of the jth component using f_i [6].

Although the solution of the above problem could be found in some cases, a plain application of (1) to an optimization algorithm often results in invalid solutions. This is mainly caused by the following two reasons. The first case is when the traffic demand is beyond the system capacity. The second case is when the optimization algorithm misses the solution because of the steep slope of the objective function.

On the other hand, the OPOB methods could be able to find the optimal solutions in a way that the best satisfies the requirements of each component, i.e., they reduce the traffic demand if the demand is greater than the system capacity. The problem can be formulated for the IST system as follows [3][4]:

$$\begin{aligned} &\underset{\mathbf{P}_t, \mathbf{w}}{\arg\min} \quad \sum_{i}^{F}\sum_{j}^{M}\left(R_i^j - C_i^j\right)^2 \\ &\text{s.t.} \quad \sum_{i}^{F} w_i \leq W, \\ &\qquad \sum_{i}^{F}\sum_{j}^{M}(P_t)_i^j \leq P_{\text{max}}, \\ &\qquad w_i \geq 0, \quad (P_t)_i^j \geq 0, \end{aligned} \tag{3}$$

where $\mathbf{P}_t$ and P_{max} are a transmit power vector $\mathbf{P}_t = [(P_t)_1^1\,(P_t)_1^2 \ldots (P_t)_1^M \ldots (P_t)_i^j \ldots (P_t)_F^M]^T$ and the maximum available transmission power of system, respectively. In addition, C_i^j is a Shannon bounded capacity allocated to the jth component using f_i which can be expressed as follows [3][4]:

$$C_i^j = w_i \log_2\left(1 + \frac{(P_t)_j^j}{w_i N_0 + I_i^j}\right), \tag{4}$$

where I_i^j is a received interference power at the jth component using f_i from the other components sharing the same subband, w_i.

III. Computationally Efficient Resource Allocation Method

A. Formulation of the objective function

Here, we modify the optimization problem in (1) in order to find a solution more efficiently. First, the bandwidth-related constraint, $\sum_{i=1}^{F} w_i \leq W$ is changed to an equality equation, $\sum_{i=1}^{F} w_i = W$, by assuming the total bandwidth W was secured by the IST system. This is because the wider bandwidth indicates less interference as well as better spectral efficiency [7].

Meanwhile, we use the following linear equation which relates the required power with the traffic demand [6]:

$$\begin{bmatrix} \left((P_t)_i^1 G_i^{1,1}\right) / \left(R_i^1 N_0\right) \\ \left((P_t)_i^2 G_i^{2,2}\right) / \left(R_i^2 N_0\right) \\ \vdots \\ \left((P_t)_i^M G_i^{M,M}\right) / \left(R_i^M N_0\right) \end{bmatrix} = \mathbf{\Delta}_i^{-1} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{5}$$

where

$$\mathbf{\Delta}_i = \begin{bmatrix} 1/\rho_i^1 & -g_i^{2,1}\eta_i^2 & \dots & -g_i^{M,1}\eta_i^M \\ -g_i^{1,2}\eta_i^1 & 1/\rho_i^2 & \dots & -g_i^{M,2}\eta_i^M \\ \vdots & \vdots & \ddots & \vdots \\ -g_i^{1,M}\eta_i^1 & -g_i^{2,M}\eta_i^2 & \dots & 1/\rho_i^M \end{bmatrix}. \tag{6}$$

In (6), $g_i^{k,j} = G_i^{k,j}/G_i^{j,j}$, $\rho_i^j = (E_b)_i^j / \left((I_0)_i^j + N_0\right)$, where $(E_b)_i^j$ and $(I_0)_i^j$ represent the bit energy and the interference power spectral density at the jth component using f_i, respectively. In addition, $\eta_i^k = R_i^k/w_i$ denotes the spectral efficiency of the kth component using f_i. According to the Shannon capacity theorem, the relationship between ρ_i^j and η_i^j is as follows:

$$\rho_i^j = (2^{\eta_i^j} - 1)/\eta_i^j, \tag{7}$$

and thus the solution finding process of (5) requires iterative search. Furthermore, the inversion operation of $\mathbf{\Delta}_i$ associated with (7) results in a complex discontinuous function of power with respect to the bandwidth.

In order to investigate the relationship between the allocated power and the bandwidth, we define the total sum of the allocated power for the components using f_i, $P(i)$ as follows:

$$P(i) = \sum_j^M (P_t)_i^j. \tag{8}$$

Accordingly, we numerically investigate the relationship between $P(i)$ and w_i. Figure 2 shows a typical example of the graph showing $P(i)$ versus w_i. It is important to note that $P(i)$ has invalid negative values for $w_i < a_i < W$, where there is an asymptotic line at a_i which can be obtained by numerical search. This implies that the lower bound, $w_i \geq 0$ is neither a sufficient nor an efficient constraint. For this reason, we replace the constraint $w_i \geq 0$ with $10\log_{10}(P(i)) \geq 0$. Here, the logarithm operation not only regulates the valid positive value range of the objective function but also set the lower bound of $P(i)$, i.e., $P(i) > 1$.

Next, we consider that the steep slope of the graph in Figure 2 is caused by the exponential term in (7). This requires a small step size during the iterative solution finding process, which should be sufficiently small enough amount not to skip the optimum point, and eventually leads to an increase in the number of iterations of solution finding process. By taking logarithm operations on the power-related constraint and the objective function, their exponentially increasing gradients can be mitigated. Consequently, it helps to converge to the optimal solution with fewer iterations.

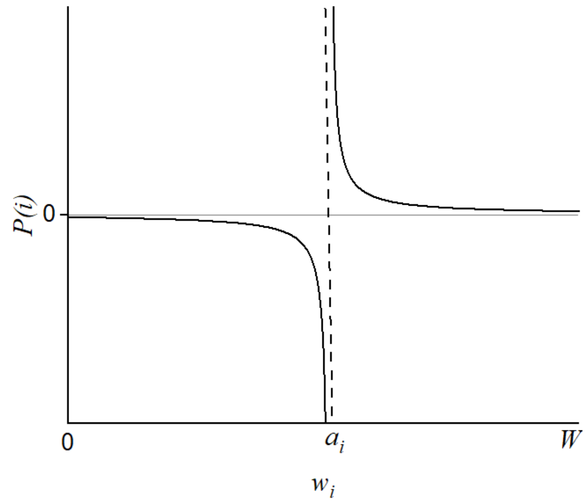

Fig. 2. A typical example of $P(i)$ versus w_i.

By integrating the above considerations, the modified optimization problem can be formulated as follows:

$$\begin{aligned} \underset{\mathbf{w}}{\arg\min} \quad & 10\log_{10}\left(\sum_i^F P(i)\right) \\ \text{s.t.} \quad & \sum_i^F w_i = W, \\ & 10\log_{10}(P(i)) \geq 0. \end{aligned} \tag{9}$$

B. Implementation with SQP

In this section, we construct an algorithm to implement the optimization problem using (9). First, we construct the standard optimization form of (9), as follows:

$$\begin{aligned} \text{Minimize} \quad & f(\mathbf{w}) = 10\log_{10}\left(\sum_i^F P(i)\right) \\ \text{s.t.} \quad & h(\mathbf{w}) : \sum_i^F w_i - W = 0, \\ & g_i(\mathbf{w}) : -10\log_{10}(P(i)) \leq 0, \quad (i = 1, \dots, F) \end{aligned} \tag{10}$$

where f and h represent the objective function and the equality constraint, respectively, and g_i denotes the inequality constraint for $P(i)$.

The quadratic programming (QP) subproblem is a simple nonlinear problem with linear constraints. At each iteration, the SQP algorithm defines an appropriate search direction $\Delta\mathbf{w}$, as a solution to the QP subproblem which can be expressed as [8]:

$$\begin{aligned} \text{Minimize}\ & \tilde{f}(\Delta\mathbf{w}) : \boldsymbol{\nabla} f(\mathbf{w}_{it})^T \Delta\mathbf{w} + \frac{1}{2}\Delta\mathbf{w}^T \boldsymbol{\nabla}^2 f(\mathbf{w}_{it})\Delta\mathbf{w} \\ \text{s.t.}\ & \tilde{h}(\Delta\mathbf{w}) : h(\mathbf{w}_{it}) + \boldsymbol{\nabla} h(\mathbf{w}_{it})^T \Delta\mathbf{w} = 0, \\ & \tilde{g}_i(\Delta\mathbf{w}) : g_i(\mathbf{w}_{it}) + \boldsymbol{\nabla} g_i(\mathbf{w}_{it})^T \Delta\mathbf{w} \leq 0, \\ & \qquad (i = 1, \ldots, F) \end{aligned} \tag{11}$$

where it is the iteration index and $\mathbf{w}_{it}$ is a subband vector at iteration it, and it is updated by using previous subband vector as well as search direction, i.e., $\mathbf{w}_{it+1} = \mathbf{w}_{it} + \Delta\mathbf{w}$.

Consequently, the generalized Lagrangian function of (11) can be formulated as follows:

$$\begin{aligned} \text{Minimize}\ & L(\boldsymbol{\Delta}\mathbf{w}, \lambda, \beta_1, \ldots, \beta_F) \\ & = \tilde{f}(\Delta\mathbf{w}) + \lambda\tilde{h} + \beta_1\tilde{g}_1 + \ldots + \beta_F\tilde{g}_F, \end{aligned} \tag{12}$$

where λ and β_i denote the Lagrangian multipliers of h and g_i, respectively, and $\lambda, \beta_i \geq 0$.

Referring to (12), there are $2F + 1$ unknown variables, and thus we need $2F + 1$ equations to find the solution. We use the Karush-Kuhn-Tucker (KKT) condition to set up $2F + 1$ equations. First, we set up F equations as follows:

$$\frac{\partial L}{\partial w_i} = \frac{\partial \tilde{f}}{\partial w_i} + \lambda \frac{\partial \tilde{h}}{\partial w_i} + \beta_1 \frac{\partial \tilde{g}_1}{\partial w_i} + \ldots + \beta_F \frac{\partial \tilde{g}_F}{\partial w_i} = 0, \quad (i = 1, \ldots, F). \tag{13}$$

Second, one equation can be obtained directly from the equality constraints as follows:

$$\tilde{h}(\Delta\mathbf{w}) = 0. \tag{14}$$

Finally, we set up the remaining F equations by considering that $\beta_i\tilde{g}_i$ should be 0 to satisfy (12), i.e.,

$$\beta_i\tilde{g}_i = 0, \quad (i = 1, \ldots, F). \tag{15}$$

To find the solutions, we note that there are two possible sets of solution to satisfy (15), that is

$$\begin{cases} \beta_i = 0, & \tilde{g}_i < 0, \\ \beta_i > 0, & \tilde{g}_i = 0. \end{cases} \tag{16}$$

The above leads to 2^F cases for $(\beta_1, \ldots, \beta_F, \tilde{g}_1, \ldots, \tilde{g}_F)$. Assuming that all the possible 2^F cases comprise a set $S = \{S_1, \ldots, S_q, \ldots, S_{2^F}\}$, e.g., $S_1 = (\beta_1 = 0, \ldots, \beta_F = 0, \tilde{g}_1 < 0, \ldots, \tilde{g}_F < 0)$, each case S_q should go through the KKT condition check with (13)-(16). This process would occupy the major computational burden.

The following Algorithm 1 details the implementation of the proposed resource allocation method with (9) in combination with SQP. The algorithm stops if the $\Delta\mathbf{w}_q$ converge within a tolerance ε or it reaches the maximum iteration, T_{max}. In addition, the initial searching point of the bandwidth $\mathbf{w}_{\text{init}}$ is set proportionally to the traffic demand to reduce the searching time, i.e., $\mathbf{w}_{\text{init}} = \left[\sum_j^M R_1^j, \ldots, \sum_j^M R_F^j\right]^T / \sum_i^F \sum_j^M R_i^j$.

Algorithm 1 The SQP algorithm to find optimal solution

Input: $N_0, R_i^j, G_i^{j,j}$
Output: $\mathbf{w}$

$\mathbf{w}_{\text{init}} \leftarrow R_i^j$
$\mathbf{w}_{it} \leftarrow \mathbf{w}_{\text{init}}$
for $it \leftarrow 1$ to T_{max} **do**
  Call the QP problem(11) $\leftarrow N_0, R_i^j, G_i^{j,j}, \mathbf{w}_{it}$
  for $q \leftarrow 1$ to 2^F **do**
    $\Delta\mathbf{w}_q \leftarrow$ the result of KKT condition check (13)-(16) with S_q
    if $h = 0$ and KKT conditions are satisfied **then**
      for $i \leftarrow 1$ to F **do**
        if $g_i \leq 0$ **then**
          if $i = F$ **then**
            $\mathbf{w}_{it+1} = \mathbf{w}_{it} + \Delta\mathbf{w}_q$
          end if
        end if
      end for
    end if
  end for
  if $\Delta\mathbf{w}_q{}^T \Delta\mathbf{w}_q < \varepsilon$ **then**
    $\mathbf{w} = \mathbf{w}_{it+1}$
  else if $it = T_{\text{max}}$ **then**
    $\mathbf{w} = \mathbf{0}$
  end if
end for
return $\mathbf{w}$

IV. Simulation results and complexity comparison

A. Computational complexity comparison

Computational complexity of the proposed method is compared with the existing methods. Table I compares the complexity of the algorithms in terms of the number of design variables, N_d, and the number of cases to be investigated to check whether KKT conditions are satisfied, N_c. For the conventional and proposed optimization problems, there are F design variables to find, i.e, $w_i, 1 \leq i \leq F$. On the other hand, the OPOB scheme has to find F and MF design variables for w_i and $(P_t)_i^j$, respectively. At every iteration, the optimization algorithm requires N_c investigations, this is the major factor to determine the computational complexity.

TABLE I
Comparison of computational complexity

	N_d	N_c
Conventional (1)	F	2^{F+1}
OPOB (3)	$F + MF$	2^{MF+F+2}
Proposed (9)	F	2^F

Figure 3 shows the comparison of N_c. The number of investigations for the OPOB scheme depends on F as well

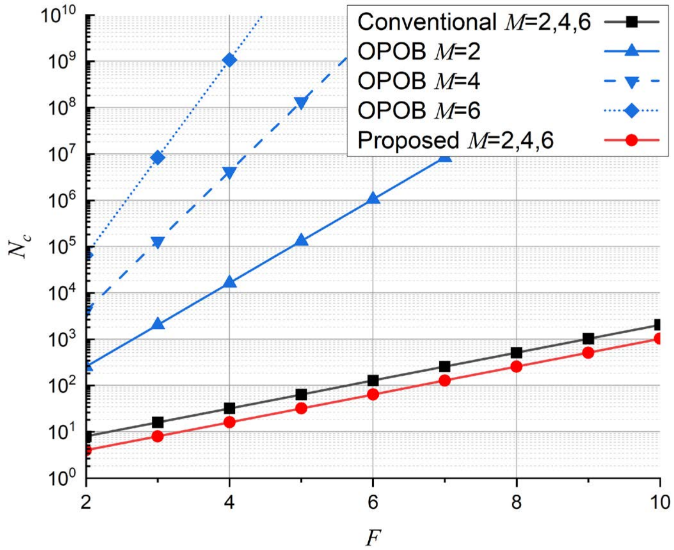

Fig. 3. Comparison of the number of investigations.

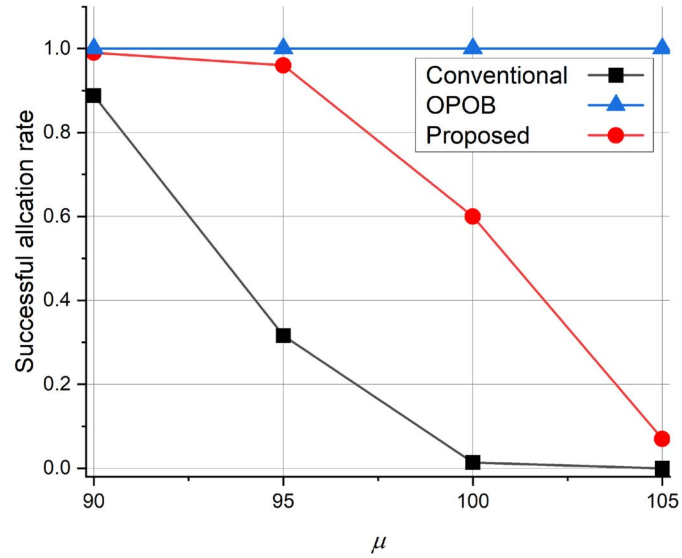

Fig. 4. Comparison of successful allocation rate.

as M, and thus it causes the highest complexity regardless of F and M. On the other hand, the proposed method has lower complexity than the conventional method. This is because the proposed method has less N_c by replacing the inequality constraint of the bandwidth in (1) with the equality constraint in (9).

B. Performance comparison

By applying the SQP algorithm to various resource allocation problems, we compare the performance in terms of the successful allocation rate and power efficiency. We assume that $F = M = 4, W = 140$ MHz and the system is composed of F satellite beams and $M - 1$ terrestrial cells in each beam, thus we have all 16 components where traffic demand from each component is dynamically changing. In order to estimate the performance, we generated traffic demands which follow Gaussian distribution $\mathcal{G}(\mu, \sigma^2)$ where μ is the mean of the traffic demand, i.e., $\mu = E[R_b]$, and $\sigma^2 = 10^2$ is the variance of the traffic demand. We set $T_{\text{max}} = 20, \varepsilon = 10^{-8}$, and we use P_{max} value of 29.54 (dBW) for the OPOB method. We investigated for 500 traffic demand cases for each $\mu = E[R_b]$.

Figure 4 compares the successful allocation rate. The OPOB method achieves rate 1 regardless of μ because it can flexibly adjust not only w_i but also $(P_t)_i^j$. As μ increases, each component requires more bandwidth, and eventually resource allocation becomes impossible. Therefore, the successful allocation rates for the conventional and proposed method gradually decrease as μ increases. Especially for the conventional scheme, the search range easily includes an invalid negative range $w_i < a_i$, and the solution can be easily missing. This leads to degraded successful allocation rate compared to the proposed method.

Figure 5 shows the comparison of the average required transmit power when the allocation is made, $E[P_t|\text{success}]$. For the OPOB method, the $E[P_t|\text{success}]$ converges to the average transmit power per component, i.e., $P_{\text{max}} - 10\log_{10}(FM) \approx 17.5$ (dBW), as μ increases, and it requires much higher power than the conventional and proposed method. On the other hand, the proposed method requires more power as μ increases, because the interference between component becomes more severe. Lastly, the conventional method seems to require the smallest power, but this is the case only for very small fraction of traffic demand. It should be noted that the proposed scheme requires much smaller power compared to the OPOB method when both methods achieve 100% successful allocation rate.

V. Conclusion

This paper evaluated resource allocation methods for integrated satellite-terrestrial systems. We first considered that the conventional resource allocation method often failed to find optimal solutions due to a steep slope of the objective function and the constraints including incorrect solution search ranges. The proposed method in this paper used a modified problem formulation with a logarithmic operation that can easy the steep slope, and also used a constraint setting strategy that can eliminate invalid ranges of the solution. The simulation results evaluated using the SQP show that the proposed method highly improves successful allocation rate, with lower computational complexity, compared to the conventional schemes.

References

[1] E. Yaacoub and Z. Dawy, "A survey on uplink resource allocation in ofdma wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 322–337, 2012.

[2] E. Castañeda, A. Silva, A. Gameiro, and M. Kountouris, "An overview on resource allocation techniques for multi-user mimo systems," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 239–284, 2017.

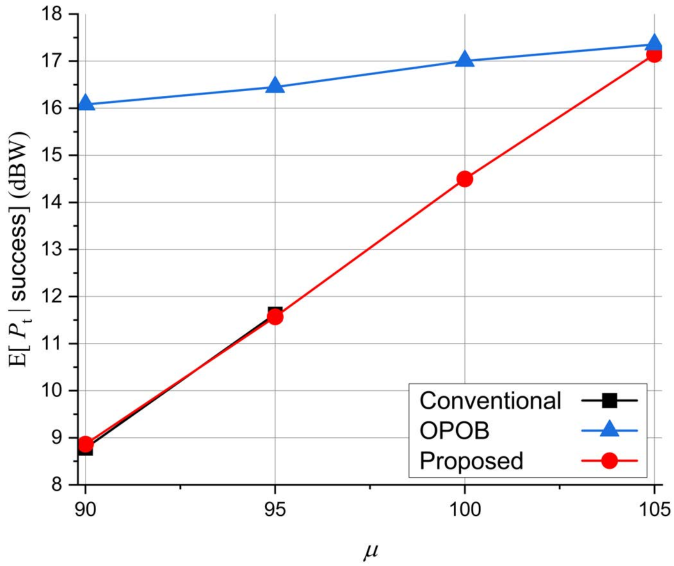

Fig. 5. Comparision of the average required P_t for successful allocation.

[3] M. Jia, X. Zhang, X. Gu, Q. Guo, Y. Li, and P. Lin, "Interbeam interference constrained resource allocation for shared spectrum multibeam satellite communication systems," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6052–6059, 2019.

[4] H. Wang, A. Liu, X. Pan, and J. Li, "Optimization of power allocation for a multibeam satellite communication system with interbeam interference," *J. Appl. Math.*, vol. 2014, pp. 1–8, 2014.

[5] U. Park, H. W. Kim, D. S. Oh, and B.-J. Ku, "Interference-limited dynamic resource management for an integrated satellite/terrestrial system," *ETRI J.*, vol. 36, no. 4, pp. 519–527, 2014. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.4218/etrij.14.0113.0079

[6] S. Chan, H. Lee, S. Kim, and D. Oh, "Intelligent low complexity resource allocation method for integrated satellite-terrestrial systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 5, pp. 1087–1091, 2022.

[7] K. Nakahira, K. Kobayashi, and M. Ueba, "Capacity and quality enhancement using an adaptive resource allocation for multi-beam mobile satellite communication systems," in *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006.*, vol. 1, 2006, pp. 153–158.

[8] P. Venkataraman, *Applied Optimization with MATLAB Programming*, 2nd ed. Wiley Publishing, 2009.