Multi-Resource Coordinate Scheduling for Earth Observation in Space Information Networks

Yu Wang, Min Sheng, *Senior Member, IEEE*, Weihua Zhuang, *Fellow, IEEE*, Shan Zhang, Ning Zhang, Runzi Liu, Jiandong Li, *Senior Member, IEEE*

Abstract

Space information network (SIN) is a promising networking architecture to significantly broaden the observation area and realize continuous information acquisition for Earth observation. Over the dynamic and complex SIN environment, it is a key issue to coordinate multi-dimensional heterogeneous network resources in the presence of multi-resource variations and severe conflicts, such that diverse Earth observation service requirements can be satisfied. To this end, this paper studies the multi-resource coordinate scheduling problem in SINs. Specifically, observation resource and transmission resource are jointly considered, and an optimization problem based on an event-driven time-expanded graph is formulated to maximize the sum priorities of successfully scheduled tasks. To solve the problem, an iterative optimization technique is employed to decompose the problem into separate observation scheduling and transmission scheduling sub-problems, which can be efficiently solved by an extended transmission time sharing graph and directed acyclic graph methods, respectively. Simulation results are provided to validate the effectiveness of the proposed algorithm and evaluate the performance impacts of different network parameters.

Index Terms

Earth observation, space information networks, satellite networks, multi-resource coordination, scheduling, time-expanded graph, optimization

I. INTRODUCTION

Earth observation serves as a fundamental function in environment monitoring, intelligence reconnaissance, and natural disaster surveillance. Due to the inherent large coverage area, potential overflight ability and high survivability, Earth observation satellites (EOSs) are widely employed for diverse Earth observation missions. The EOSs typically operate in the sun-synchronous low-Earth orbits (LEOs) and acquire high-resolution image data with onboard sensors. Currently, different types of EOSs separately collect information and lack interactions among one another. This inevitably leads to under-utilization of scarce network resources [1]. Moreover, with the unprecedented growth of Earth observation traffic (e.g., 27.9 TB/day traffic on average for NASA Earth observing system [2]), traditional standalone EOS systems are unable to accommodate such a huge amount of traffic and provide satisfactory service guarantee. The problem is exaggerated especially in emergency situations, e.g., earthquakes, where EOSs are expected to rapidly react to user requests and continuously provide useful Earth observation data of concerning targets (e.g., certain observation area on the ground). The violent 8.0 magnitude earthquake that occurred in Wenchuan has demonstrated the insufficient service capability of Earth observation system at that time [3].

To cater for the aforementioned issues, the concept of space information network (SIN) emerges [4]–[6]. A SIN is composed of heterogeneous EOS systems in different orbits that perform cooperative Earth observation, and geostationary-Earth orbit (GEO) satellites for timely observation data delivery between EOSs and destination ground stations. With the deployment of SINs, near-real-time data acquisition and transfer can be achieved [6]. However, the multi-dimensional heterogeneous resources (e.g., observation and transmission opportunities) are normally unbalanced in SINs. For example, a typical EOS can access a certain ground location in less than 10 minutes within the system period of approximately 100 minutes. This constraint makes EOSs have limited chances to observe the targets of interest and unable to send all of their collected data to the destination. Therefore, how to design appropriate cooperative scheduling mechanisms and achieve efficient network resource utilization to maximize the utility of the whole system become a key issue for SIN operation.

It is technically challenging to develop efficient multi-resource coordinate scheduling strategies for SINs, due to several reasons: 1) Because of the dynamic but predictable network connectivity, resource availability varies continuously and periodically. Moreover, there are various types of resources including observation resources, storage resources and transmission resources. It is difficult to precisely represent multi-dimensional resources in both time and space domains and dictate the correlation relationships among multiple heterogeneous resources; 2) There can be severe observation and transmission conflicts. On one hand, observation conflicts can occur frequently if the setup time between two successively observed targets in an EOS is not sufficient, and such conflicts vary in different EOSs. On the other hand, multi-resource limitations can lead to infeasibility of all potential observation and transmission opportunities, and thus induce observation and transmission conflicts [7]; and 3) As different Earth observation tasks have diverse quality-of-service (QoS) requirements in terms of observation time duration and end-to-end delay, multi-resource scheduling should satisfy the differentiated service requirements in the dynamic and complex SIN environment.

In this paper, to address the technical challenges, we study the multi-dimensional resource scheduling problem for SINs. Specifically, an event-driven time-expanded graph (EDTEG) is proposed to characterize multi-resource variations over the complex environment. Based on the EDTEG, a joint observation and transmission scheduling optimization framework is formulated, with the objective of maximizing the sum priorities of successfully scheduled tasks. Due to the NP-completeness of the optimization problem, an iterative optimization approach is proposed to decompose it into a separate transmission scheduling sub-problem and observation scheduling sub-problem. For the transmission sub-problem, we utilize an extended time sharing graph to properly allocate transmission time among multiple EOSs. For the observation scheduling sub-problem, we use an acyclic directed graph to model the observation conflicts and a column generation method to solve the observation sub-problem, wherein the underlying generation problem is solved by multi-constrained optimal path based solutions. The two sub-problems are then updated iteratively by redistribution of surplus transmission time. The convergence of the proposed algorithm is proved and its computational complexity is analyzed in detail. Extensive

simulation results are provided to demonstrate the performance gains of the proposed algorithm over existing benchmarks.

In a nutshell, the main contributions of this paper are summarized as follows:

1) By exploiting an event-driven time-expanded graph approach, a joint optimization framework of observation scheduling and transmission scheduling is formulated to maximize the sum priorities of successfully scheduled observation tasks;
2) An iterative optimization technique is utilized to decompose the problem into separate observation scheduling and transmission scheduling sub-problems, which are solved by acyclic directed graph and extended transmission time sharing graph, respectively;
3) Extensive simulation results are provided to validate the effectiveness of our proposed scheduling algorithm. The effects of different network parameters on the network performance are evaluated.

The remainder of this paper is organized as follows. Section II gives an overview of related works. Section III introduces the SIN system model under consideration and gives the detailed problem formulation. We propose an approximate multi-resource scheduling algorithm in Section IV. The performance evaluation by simulations is presented in Section V, followed by concluding remarks and future research in Section VI.

II. Related Works

Resource scheduling plays a critical role in efficiently utilizing the network resources, and gains significant attentions in the development of cooperative SINs. It can be classified into two main categories, namely single-resource scheduling and multi-resource coordinate scheduling.

A. Single-Resource Scheduling

Single-resource scheduling algorithms focus on scheduling only one type of network resources, i.e., observation resources or transmission resources. The observation resource scheduling is to allocate a subset of observation targets to multiple EOSs, and has received substantial interests in the literature. Existing studies include static scheduling algorithms for pre-planned targets and

dynamic scheduling algorithms for real-time targets. The static observation scheduling problem has been investigated for both single-EOS [8], [9] and multiple-EOS [3], [10]. Various algorithms including Lagrangian relaxation and graph techniques are proposed to obtain an approximate solution. On the other hand, dynamic scheduling handles aperiodic observation targets whose arrival times are not known a priori. Under such a circumstance, rescheduling principles, e.g., backward shift and rehabilitation strategics [11] and target merging strategies [12], are developed to cope with random arrivals of new observation targets. In [13], an agent-based dynamic scheduling is presented to improve the global optimization and load balancing of observation resources. A comprehensive survey on observation scheduling solutions is given in [12].

The transmission scheduling problem focus on data exchange from EOSs to multiple destinations. In [14], the problem is investigated to resolve potential transmission conflicts, taking account of both the fairness requirement and overall transmission capacity. As an extension to [14], the data delivery time and resource usage are improved through making use of traffic information [15]. By considering the time-varying downlink channel quality, delay-optimal and throughput-optimal data downloading strategies are studied in [16] and [17], respectively. A throughput-optimal collaborative data downloading scheme is proposed to exploit data offloading opportunities between EOSs via inter-satellite links [18].

The existing single-resource scheduling algorithms can result in low efficiency, if the observation and transmission resources/opportunities are not properly matched. For instance, when an EOS scheduled to collect a large amount of observation data has a very short transmission time with a destination, the EOS is unable to download all its observation data to the destination. Therefore, each EOS should adjust the amount of data it collects during the observation scheduling process to match the data transmission time allocated in the transmission scheduling process, such that the network performance can be improved.

B. Multi-Resource Coordinate Scheduling

Multi-resource coordinate scheduling algorithms schedule multiple types of resources simultaneously, to provision satisfactory service in SINs. The joint observation and transmission

scheduling problem for the COSMO-SkyMed constellation is first introduced in [19]. A heuristic scheme with look-ahead and back-tracking capabilities is then devised to produce feasible scheduling solutions. In [20], a constraint satisfaction optimization model is used to describe EOS observation tasks and data transmission jobs in an integrated way. Accordingly, a genetic algorithm based meta-heuristic is proposed. However, the transmission time sharing among multiple EOSs is neglected in the transmission scheduling procedure, which to some extent restricts the efficient utilization of transmission resources. We exploit an extended time-expanded graph method and propose an analytical framework to characterize multi-resource evolution over the complex SIN environment in [21]. The framework allows us to investigate the impact of different factors on the network performance, e.g., delay-limited throughput. Furthermore, a primal decomposition approach is presented in [22] to solve the optimization problem by taking advantage of its special structure. Although the proposed optimization framework and solution technique in [21], [22] are useful in deriving upper bounds of network throughput performance, they cannot capture the specific characteristics in terms of observation conflicts caused by insufficient setup time between multiple targets as well as non-preemptive observation time duration requirements. For a realistic SIN with a large number of targets, the computational complexity of the proposed algorithms is unacceptable.

To sum up, existing works either fail to model multi-resource evolution over dynamic SIN or incur high computational complexity. Most of them do not consider the diverse service requirements of observation tasks, e.g., observation time duration and end-to-end delay. Different from these works, here we study joint scheduling of observation resource and transmission resource to improve network performance, while taking into account SIN's diverse requirements of multiple observation tasks.

III. System Model and Problem Formulation

In this section, we first model the system by exploiting an EDTEG approach. Based on EDTEG, we then give the detailed problem formulation.

A. System Model

1) Network Model: Consider a SIN with three types of components: 1) a set $\mathcal{I} = \{1, \ldots, I\}$ of I point targets (e.g., observation area) on the ground that need to be observed; 2) a set $\mathcal{K} = \{1, \ldots, K\}$ of K EOSs moving in the LEOs to acquire observation data (e.g., high-resolution images) for the targets of interest and finally transmit those data to destinations; and 3) a set $\mathcal{N} = \{1, \ldots, N\}$ of N destinations (i.e., relay satellites or ground stations) which serve as the sinks for all the observation data. The considered time horizon $\mathcal{T} = \{1, ..., T\}$ is discretized into T time slots, each with a constant time duration τ. An example SIN with 2 targets, 2 EOSs and 1 destination relay satellite is depicted in Fig. 1(a).

There is a task, i, associated with each observation target, i. We use the terms target and task interchangeably in this paper when no ambiguity is caused[1]. All tasks are independent, non-preemptive and aperiodic [11]. New tasks normally arrive in a batch. Denote $\mathcal{I}_t \subset \mathcal{I}$ as the subset of tasks that arrive at the network at time slot t. A task $i \in \mathcal{I}$ is described by a tuple comprised of four elements, i.e., $\{\mu_i, a_i, nd_i, g_i\}$, where μ_i, a_i, nd_i and g_i denote task i's priority, arrival time slot, required continuous observation time duration, and expected finish time slot (i.e., the deadline when its observation data should be transmitted to destinations), respectively. As in [9], accurate information of all tasks is known a priori in the time horizon $\mathcal{T}$.

There are two phases, namely observation phase and transmission phase, to complete a task. Firstly, in the observation phase, an EOS is scheduled to collect observation data using its onboard imaging sensor (e.g., optical camera) when it is in the line-of-sight of the associated target. The time interval during which a target can be observed by an EOS is referred to as observation window (OW). Secondly, in the transmission phase, an EOS stores the observation data onboard, and downloads them to a destination after it enters the coverage of the destination. The time interval when an EOS moves into the transmission range of a destination is termed as transmission window (TW). As shown in Fig. 1(a), target 1 is first observed by EOS 1 in the observation phase, and then its observation data are delivered to the destination in the

[1]A target corresponds to an observation area on the ground. A task refers to both the observation and transmission phases of a target.

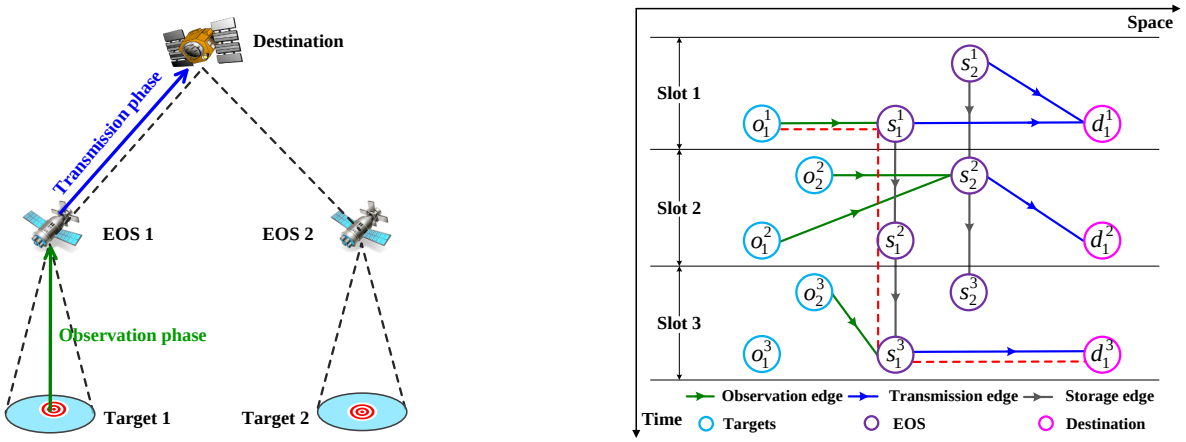


Figure 1. (a) An example SIN with 2 targets, 2 EOSs and a destination relay satellite. Targets 1 and 2 arrive at the first and second time slot, respectively. (b) Corresponding EDTEG spanned over 3 time slots for the example SIN, where $\{o_1^t, o_2^t\}$, $\{s_1^t, s_2^t\}$ and $\{d_1^t\}$ represent the set of targets, EOSs and destination at time slot t ($t \in \{1, 2, 3\}$), respectively.

transmission phase.

2) Event-Driven Time-Expanded Graph: We herein use EDTEG[2] $\mathcal{G}^{\text{TEG}} = (\mathcal{V}^{\text{TEG}}, \mathcal{E}^{\text{TEG}})$ to model all the available OWs and TWs, where $\mathcal{V}^{\text{TEG}}$ and $\mathcal{E}^{\text{TEG}}$ represent the set of vertices and edges, respectively. The EDTEG is based on the predictable mobility trace [21] for the concerned SIN. We use the following procedures to construct the EDTEG. Firstly, we build a unified two-dimensional time-space basis, wherein vertices correspond to targets, EOSs and destinations at different time slots. Secondly, edges denote the availability of different resources, i.e., OWs and TWs. If an edge exists, its corresponding resource is available and vice versa. Finally, we utilize a path formed by a set of contiguous vertices in the directed graph to capture the chronological relationship between observation phase and transmission phase.

To be specific, there are T layers in the constructed EDTEG, with each layer indicating network status at a single time slot. Within a time slot, the network is static, i.e., the status of an OW or a TW does not change. However, the network status may instantaneously change during time slot transitions. At time slot t $(1 \leq t \leq T)$, target $i \in \mathcal{I}$, EOS $k \in \mathcal{K}$, and destination

[2]The reason beneath using EDTEG modeling is twofold. First, it facilitates multi-resource coordinate scheduling optimization problem formulation [23], [24]. Second, for a small-scale SIN, the optimal solution can be obtained to the optimization framework based on EDTEG [1]. It thus serves as a useful benchmark to evaluate the efficiency of proposed scheduling algorithms.

$n \in \mathcal{N}$ are represented by vertices o_i^t, s_k^t, and d_n^t in the EDTEG, respectively.

There are three different types of directed edges in the EDTEG, namely observation edge, transmission edge and storage edge. Within a time slot t, an observation edge (o_i^t, s_k^t) exists, if an observation opportunity[3] is present between target i and EOS k during the time slot. Let $\mathcal{E}_t^{\text{ob}}$ denote the set of all observation edges in the graph at time slot t. Each observation edge, $(o_i^t, s_k^t) \in \mathcal{E}_t^{\text{ob}}$, is associated with a weight, $w(o_i^t, s_k^t)$, representing the amount of observation data volume acquired during time slot t. Similarly, a transmission edge (s_k^t, d_n^t) exists, if a transmission opportunity between EOS k and destination n is available during the time slot. Let $\mathcal{E}_t^{\text{ob}}$ denote the set of transmission edges in the graph at time slot t. Such type of transmission edge $(s_k^t, d_n^t) \in \mathcal{E}_t^{\text{tr}}$ is associated with a weight, $w(s_k^t, d_n^t)$, equal to the amount of data volume that can be delivered within the time slot. Let $w(s_k^t, d_n^t, i)$ denote the data volume delivered for target i on (s_k^t, d_n^t), where $\sum_{i=1}^{I} w(s_k^t, d_n^t, i) = w(s_k^t, d_n^t)$ holds. On the other hand, during time slot transitions, a storage edge (s_k^t, s_k^{t+1}) is drawn to model that an EOS k can physically carry its data forward from time slot t to time slot $t+1$. A storage edge is assumed to have a weight of infinity, i.e., $w(s_k^t, s_k^{t+1}) = \infty$, indicating that EOS k's onboard buffer size is infinite.

Next, a path is a sequence of distinct vertices in the directed graph. It originates from a vertex representing the target and ends at a vertex representing the destination. The set of edges that it traverses can thus capture the sequentially chained multi-dimensional resources (i.e., OWs and TWs). An example of its derivation is given in Fig. 1(b). The time horizon is set to 3 time slots. There are 3 vertices created for an EOS or a destination, but only two vertices, i.e.,$\{o_2^2, o_2^3\}$, for target 2. This is because vertices corresponding to target 2 are created only upon its arrival. As indicated by the observation edge (o_2^2, s_2^2), target 2 can be observed by EOS 2 at time slot 2. Similarly, with transmission edge (s_1^3, d_1^3), a transmission opportunity exists between EOS 1 and the destination at time slot 3. The path, $\{o_1^1, s_1^1, s_1^2, s_1^3, d_1^3\}$, represented by the red dotted line in the graph, indicates that EOS 1 can first observe target 1 at time slot 1, carry those observation data at time slot 2, and then transmit them to the destination at time slot 3.

[3]Note that an OW $[ow^{\text{bt}}, ow^{\text{ft}}]$ is generally represented by $(ow^{\text{ft}} - ow^{\text{bt}} + 1)$ continuous observation opportunities (i.e., observation edges) in the graph, where ow^{bt} and ow^{ft} denote the start time and ending time of the observation window, respectively. Likewise, a TW is captured by several continuous transmission opportunities.

B. Problem Formulation

1) Basic Constraints: Herein, we formulate some basic constraints on the constructed EDTEG.

Observation constraints. Define an observation scheduling vector $\mathbf{X} = \{x(o_i^t, s_k^t) | t \in \mathcal{T}, (o_i^t, s_k^t) \in \mathcal{E}_t^{\text{ob}}\}$ to reflect the mapping of I targets to K EOSs, where the binary element $x(o_i^t, s_k^t) = 1$ if target i is observed by EOS k at time slot t, and $x(o_i^t, s_k^t) = 0$ otherwise. Considering that an EOS can process at most one target at a time slot, we have

$$\sum_{o_i^t:(o_i^t,s_k^t)\in\mathcal{E}_t^{\text{ob}}} x(o_i^t, s_k^t) \leq 1, \ \ \forall k, t. \tag{1}$$

Meanwhile, if target i is observed, its non-preemptive observation duration requirement should be satisfied, which yields

$$\sum_{t=bt_{i,k}}^{ft_{i,k}} \sum_{s_k^t:(o_i^t,s_k^t)\in\mathcal{E}_t^{\text{ob}}} x(o_i^t, s_k^t) = nd_i, \ \ \forall i \tag{2}$$

where $bt_{i,k}$ and $ft_{i,k}$ denote the observation beginning time and ending time of target i in EOS k, respectively, and $ft_{i,k} = bt_{i,k} + nd_i$ holds. According to [19], [25], both $bt_{i,k}$ and $ft_{i,k}$ can be predetermined[4], because they depend on the flight parameters of EOS k. Note that (2) actually implies that, from time slot $bt_{i,k}$, contiguous nd_i time slots within available OWs in EOS k are allocated together to complete observing target i.

When a target is observed, an EOS requires a setup time to maneuver its position or sensor orientation [9]. To capture this sequence-dependent feature, we further define a vector $\mathbf{Z} = \{z_{i,j}\}$, where the binary variable $z_{i,j} = 1$ if target j is observed after target i by the same EOS, and $z_{i,j} = 0$ otherwise. Since there should be sufficient setup time to observe successive targets in an EOS, we have

$$ft_{i,k} + \delta_{i,j,k} \leq bt_{j,k} + (1 - z_{i,j})H, \ \ \forall i, j, k \tag{3}$$

where H is a sufficiently large positive constant, and $\delta_{i,j,k}$ is the setup time from executing target i to target j by EOS k.

[4]The proposed optimization framework can be extended to the case that both $bt_{i,k}$ and $ft_{i,k}$ are decision variables.

Transmission constraints. Due to limited resources (e.g., transponders) in EOSs, only a subset of potential transmission opportunities within TWs can be utilized for data transmission [26]. To this end, we define a transmission scheduling vector $\mathbf{Y} = \{y(s_k^t, d_n^t) | t \in \mathcal{T}, (s_k^t, d_n^t) \in \mathcal{E}_t^{\text{tr}}\}$, where $y(s_k^t, d_n^t) = 1$ if EOS k is scheduled to transmit to destination n at time slot t, and $y(s_k^t, d_n^t) = 0$ otherwise. Following [22], we assume that a destination can support one transmission at a time slot. In addition, an EOS can transmit to a destination at a time slot. Thus, the following constraints should be satisfied:

$$\sum_{s_k^t:(s_k^t,d_n^t)\in\mathcal{E}_t^{\text{tr}}} y(s_k^t, d_n^t) \leq 1, \ \ \forall n, t \tag{4}$$

$$\sum_{d_n^t:(s_k^t,d_n^t)\in\mathcal{E}_t^{\text{tr}}} y(s_k^t, d_n^t) \leq 1, \ \ \forall k, t. \tag{5}$$

Flow conservation constraints. The total amount of data transmitted by EOS k to destinations must not exceed the amount of data that it acquired, by the end of time slot $t\,(t \in \{1, \ldots, T-1\})$. That is,

$$\sum_{t=1}^{\theta} \sum_{o_i^t:(o_i^t,s_k^t)\in\mathcal{E}_t^{\text{ob}}} w(o_i^t, s_k^t) - \sum_{t=1}^{\theta} \sum_{d_n^t:(s_k^t,d_n^t)\in\mathcal{E}_t^{\text{tr}}} w(s_k^t, d_n^t) \geq 0, \ \ \forall \theta \in \{1, \ldots, T-1\}, k \tag{6}$$

where the inequality permits the EOS to hold data and carry them forward into future time slots. In (6), $w(o_i^t, s_k^t)$ and $w(s_k^t, d_n^t)$ satisfy:

$$w(o_i^t, s_k^t) = x(o_i^t, s_k^t) \cdot r_k^{\text{ob}} \cdot \tau, \ \ \forall t, (o_i^t, s_k^t) \in \mathcal{E}_t^{\text{ob}} \tag{7}$$

$$w(s_k^t, d_n^t) = y(s_k^t, d_n^t) \cdot r_k^{\text{tr}} \cdot \tau, \ \ \forall t, (s_k^t, d_n^t) \in \mathcal{E}_t^{\text{tr}} \tag{8}$$

where r_k^{ob} and r_k^{tr} represent the data collection rate and data transmission capacity for EOS k, respectively.

Finally, we impose that the total observation data volume for a scheduled task should equal to that delivered to destinations before its expected finish time, which is

$$\sum_{t=a_i}^{g_i} \sum_{s_k^t:(o_i^t,s_k^t)\in\mathcal{E}_t^{\text{ob}}} w(o_i^t, s_k^t) = \sum_{t=a_i}^{g_i} \sum_{(s_k^t,d_n^t)\in\mathcal{E}_t^{\text{tr}}} w(s_k^t, d_n^t, i), \ \ \forall i. \tag{9}$$

2) Optimization Problem Formulation: Given the constructed EDTEG, the problem under consideration is to select and schedule a subset of targets to different OWs, while considering the data transmission scheduling of multiple TWs. Our objective is to maximize the sum priorities of successfully scheduled tasks. A task is successfully scheduled if and only if its associated target is observed and the required observation data are transmitted to destinations before the expected finish time. We formulate it as the following optimization problem (P1):

$$(\mathbf{P1}) \quad \max_{\mathbf{X},\mathbf{Y},\mathbf{Z}} \sum_{t\in\mathcal{T}} \sum_{(o_i^t,s_k^t)\in\mathcal{E}_t^{\text{ob}}} \frac{\mu_i}{nd_i} x(o_i^t, s_k^t)$$
$$\text{s.t.} \quad (1)-(9). \tag{10}$$

In problem (P1), the objective is to maximize the total successfully scheduled tasks weighted by their priorities, subject to observation constraints (1)-(3), transmission constraints (4)-(5), and flow conservation constraints (6)-(9). Notice that the coefficient $\frac{\mu_i}{nd_i}$ in the objective function stands for the average priority received from one time slot observation.

Lemma 1. *Problem (P1) is NP-complete to solve.*

Proof: Observe that problem (P1) is an integer linear programming problem. Consider a generalized case where all the transmission constraints (4)-(5) and flow conservation constraints (6)-(9) are relaxed. In this case, problem (P1) is reduced to the satellite range scheduling problem with non-identical machines [12], which is already an NP-complete problem. Accordingly, the original problem (P1) is NP-complete to solve as well. ■

In addition to its NP-completeness, there are some other aspects that increase the computational complexity of solving the multi-resource scheduling problem (P1). The observation and transmission constraints are tightly coupled by the flow conservation constraints. It is challenging to directly decompose such couplings. On one hand, as indicated by (6), the coupling relationship is time dependent and spans nearly over the entire time horizon; on the other hand, because both the observation and transmission scheduling problems involve integer decision variables, it will inevitably lead to non-negligible optimality gap if traditional decomposing techniques, e.g.,

Lagrangian decomposition technique, are used [27].

For a small-scale SIN, existing optimization toolkits can provide an optimal solution to the above optimization problem [1]. However, for a large-scale SIN, it is computationally prohibitive to directly employ existing algorithms without taking the specific structure of problem (P1) into consideration. This is because the scheduling problem is generally an oversubscribed problem and involves a large number of decision variables [28]. In view of these, we aim to devise an approximate scheduling algorithm with acceptable complexity to solve the optimization problem.

IV. Approximate Multi-Resource Schedule

In this section, we propose an approximate multi-resource scheduling (AMRS) algorithm to solve problem (P1). In AMRS algorithm, the original problem (P1) is decomposed into a transmission scheduling sub-problem and an observation scheduling sub-problem. Then, an iterative optimization method is utilized to asymptotically reach the desired solution. The transmission scheduling sub-problem is solved by exploiting an extended transmission time sharing graph method. The observation scheduling sub-problem is tackled by the column generation approach, wherein the underlying generation problem is determined by finding multi-constrained optimal paths on a constructed acyclic directed graph (ADG). The above two sub-problems are then updated through a procedure of redistribution surplus transmission time.

A. Transmission Scheduling Sub-problem

We employ an extended transmission time sharing graph method [18] to deal with the transmission scheduling between multiple EOSs and destinations (i.e., multiple TWs). The extended transmission time sharing graph is constructed as follows. As shown in Fig. 2, the time horizon is divided by the start-times and end-times of all potential TWs. Two adjacent time points on the time horizon form a segment, denoted by $TS_m\,(m = 1, 2, \ldots, M)$. Let bt$(TS_m)$ and ft(TS_m) be the beginning time and ending time of segment TS_m. The time duration, ψ_m, of segment TS_m can be expressed as $\psi_m = \text{ft}(TS_m) - \text{bt}(TS_m)$. The total time duration ψ_m is shared among EOSs at the same destination over segment TS_m. As depicted in Fig. 2, segment TS_3 for destination

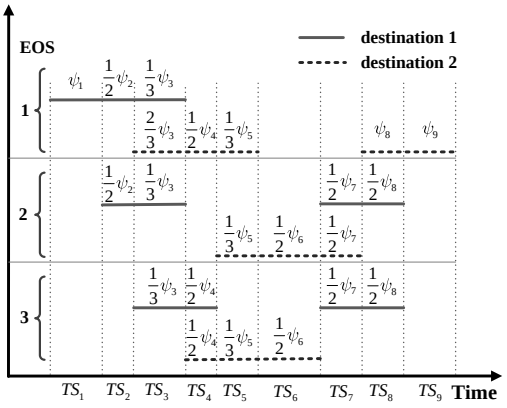


Figure 2. An example of extended transmission time sharing graph with 3 EOSs and 2 destinations. The whole time horizon includes 9 segments. An initial allocation of transmission time is labeled on the segments.

1 is shared by EOS 1, EOS 3 and EOS 4. For an EOS, if TWs for different destinations overlap at a segment, transmission time over the segment should also be properly shared. For instance, transmission time of EOS 1 over segment TS_3 is distributed between destinations 1 and 2.

Define capacity region $\mathcal{C} = \{(dt^1_{1,1}, dt^1_{1,2}, \ldots, dt^m_{k,n}, \ldots, dt^M_{K,N})\}$ as a multidimensional region of all the transmission time combinations that multiple TWs can support for data transmission, where $dt^m_{k,n}$ denotes the amount of allocated transmission time in segment TS_m between EOS k and destination n. We use the following lemma to characterize $\mathcal{C}$.

Lemma 2. *Capacity region $\mathcal{C}$ is constituted by the set of transmission time $(dt^1_{1,1}, dt^1_{1,2}, \ldots, dt^M_{K,N})$ such that*

$$\sum_{k \in \mathcal{K}} dt^m_{k,n} \leq \psi_m, \ \ \forall m, n \tag{11}$$

$$\sum_{n \in \mathcal{N}} dt^m_{k,n} \leq \psi_m, \ \ \forall m, k. \tag{12}$$

Proof: Capacity region $\mathcal{C}$ should be bounded by the following two conditions: 1) for destination n, the allocated transmission time to all EOSs sharing segment TS_m should not exceed the maximum available transmission time, i.e., ψ_m, represented by (11); and 2) the sum transmission time from EOS k to all destinations over segment TS_m is less than ψ_m, represented by (12). Putting (11) and (12) together, we can obtain capacity region $\mathcal{C}$. ■

As the initial allocation of transmission scheduling between EOSs and destinations, we equally

distribute the transmission time among the EOSs that share the segment with the same destination. It can be observed that segment TS_2 is shared by EOS 1 and EOS 2 at destination 1, so the transmission time allocated to each EOS is $\frac{1}{2}\psi_2$, i.e., $dt^2_{1,1} = dt^2_{2,1} = \frac{1}{2}\psi_2$. For an EOS, the total available transmission time over a segment for different destinations should not exceed the length of the segment. As can be seen from Fig. 2, the allocated transmission time over segment TS_3 for EOS 1 at destination 2 is $\frac{2}{3}\psi_3$, i.e., $dt^3_{1,2} = \frac{2}{3}\psi_3$, because the rest time is already given to destination 1 with $dt^3_{1,1} = \frac{1}{3}\psi_3$. Let $D^{\text{tr}}_{k,m}$ denote the actual transmission time allocated to EOS k in segment TS_m. D^{tr}_k is the sum of transmission time allocated to EOS k in all segments during the whole time horizon, i.e., $D^{\text{tr}}_k = \sum_m \sum_n dt^m_{k,n} = \sum_m \psi_m$. For example, the total TWs for EOS 2 comprise six segments, TS_2, TS_3, TS_5, TS_6, TS_7 and TS_8. Segments TS_2, TS_3, TS_5, TS_6 and TS_8 are shared by 2, 3, 3, 2 and 2 EOSs, respectively, while segment TS_7 is shared by transmissions at both the destinations. Consequently, we have $D^{\text{tr}}_2 = \frac{1}{2}\psi_2 + \frac{1}{3}\psi_3 + \frac{1}{3}\psi_5 + \frac{1}{2}\psi_6 + \psi_7 + \frac{1}{2}\psi_8$.

Note that based on the results from the observation scheduling sub-problem described later, the pre-allocated transmission time should be readjusted. The transmission time redistribution procedure is detailed in Section IV-C.

B. Observation Scheduling Sub-problem

Given an allocated transmission time vector $(D^{\text{tr}}_{1,1}, D^{\text{tr}}_{1,2}, \ldots, D^{\text{tr}}_{K,M})$, we need to deal with the observation scheduling sub-problem, i.e., schedule a subset of observation targets for each EOS. The original problem (P1) reduces to the following optimization problem (P2):

$$
\begin{aligned}
(\mathbf{P2}) \quad & \max_{\mathbf{X},\mathbf{Z}} \sum_{t\in\mathcal{T}} \sum_{(o_i^t, s_k^t)\in\mathcal{E}_t^{\text{ob}}} \frac{\mu_i}{nd_i} x(o_i^t, s_k^t) \\
& \text{s.t.} \quad \text{(1)-(3),(7),(9)} \\
& \qquad \sum_{t=1}^{\theta} \sum_{o_j^t:(o_j^t, s_k^t)\in\mathcal{E}_t^{\text{ob}}} w(o_j^t, s_k^t) \geq D^{\text{tr}}_{k,t}, \ \ \forall \theta \in \{1, \ldots, T-1\}, k \qquad (13)
\end{aligned}
$$

where $D^{\text{tr}}_{k,t}$ is the allocated transmission time to EOS k before time slot t. Noticeably, given $(D^{\text{tr}}_{1,1}, D^{\text{tr}}_{1,2}, \ldots, D^{\text{tr}}_{K,M})$, $D^{\text{tr}}_{k,t}$ can be obtained by

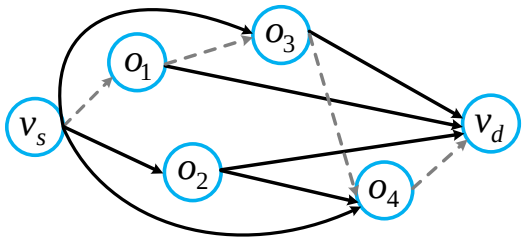


Figure 3. An example of directed acyclic graph to model possible observation sequences for an EOS.

$$D^{\text{tr}}_{k,t} = \sum_{m:\text{bt}(TS_m)\le t<\text{bt}(TS_{m+1})} \psi_m + t - \text{bt}(TS_m), \ \ \forall k, t. \tag{14}$$

Compared with problem (P1), transmission constraints (4) and (5) are removed in problem (P2). Also, (6) and (8) are represented by (13) in problem (P2). Problem (P2) can be reformulated and solved based on a useful definition of observation sequence.

Definition 1. An observation sequence $\ell = \{\ldots, o_i, o_j, \ldots\}$ is a set of ordered targets that can be sequentially observed by an EOS, with any two adjacent elements, (o_i, o_j), satisfying:

$$bt_{i,k} + nd_i + \delta_{i,j,k} \le bt_{j,k}, \ \ \forall (o_i, o_j). \tag{15}$$

It can be found that (15) replaces (2). Let $\mathcal{L}_k$ be all the possible options for observation sequences on EOS k. The set of all possible observation sequences for EOS k is obtained using an ADG $\mathcal{OG}_k = (\mathcal{V}_k^{\text{ADG}}, \mathcal{E}_k^{\text{ADG}})$ [25], where $\mathcal{V}_k^{\text{ADG}}$ and $\mathcal{E}_k^{\text{ADG}}$ are the set of vertices and edges, respectively. If a target i can be observed by EOS k, a vertex $o_i \in \mathcal{V}_k^{\text{ADG}}$ is created in $\mathcal{OG}_k$. If two targets i and j can be observed successively on EOS k (i.e., the setup time constraint (3) holds), a directed edge $(o_i, o_j) \in \mathcal{E}_k^{\text{ADG}}$ exists. We also add two virtual vertices v_s and v_d to represent the common source and destination in $\mathcal{OG}_k$, respectively. A path from v_s to v_d in the graph thus corresponds to an observation sequence. An example ADG is given in Fig. 3. There is an observation conflict between targets 1 and 2, since edge (o_1, o_2) does not exist in the graph. Meanwhile, the example path $\{v_s, o_1, o_3, o_4, v_d\}$ corresponding to $\ell = \{o_1, o_3, o_4\}$ means that, the EOS can observe targets 1, 3, and 4 successively.

Based on the preceding notations, we can reformulate the optimization problem (P2). The sum priorities of sequence ℓ for EOS k becomes $f_\ell^k = \sum_{i\in\mathcal{I}} \mu_i \rho_{i,\ell}^k$, where $\rho_{i,\ell}^k = 1$ if sequence ℓ of EOS k contains target i, and $\rho_{i,\ell}^k = 0$ otherwise. Thus, the following equation holds:

$$\sum_{t\in\mathcal{T}} \sum_{(o_i^t, s_k^t)\in\mathcal{E}_t^{\text{ob}}} \frac{\mu_i}{nd_i} x(o_i^t, s_k^t) = \sum_{k\in\mathcal{K}} \sum_{\ell\in\mathcal{L}_k} a_\ell^k f_\ell^k \tag{16}$$

where $a_\ell^k = 1$ if observation sequence ℓ is assigned to EOS k, and $a_\ell^k = 0$ otherwise. Define $\mathbf{A} = \{a_\ell^k\}$ as the assignment vector. Problem (P2) is equivalent to problem (P3), given by

$$\begin{aligned}
(\mathbf{P3}) \quad & \max_{\mathbf{A}} \sum_{k\in\mathcal{K}} \sum_{\ell\in\mathcal{L}_k} a_\ell^k f_\ell^k \\
& \text{s.t.} \quad \sum_{k\in\mathcal{K}} \sum_{\ell\in\mathcal{L}_k} a_\ell^k \rho_{i,\ell}^k \le 1, \ \ \forall i && (17)\\
& \qquad \sum_{\ell\in\mathcal{L}_k} a_\ell^k \le 1, \ \ \forall k && (18)\\
& \qquad a_\ell^k \in \{0,1\}, \ \ \forall \ell && (19)\\
& \qquad \sum_{i:g_i>\text{ft}(TS_m)} nd_i \cdot \rho_{i,\ell}^k \cdot r_k^{\text{ob}} \le \sum_{\vartheta=1}^{m} D_{k,m}^{\text{tr}} \cdot r_k^{\text{tr}}, \ \ \forall \vartheta \in \{1,\dots,M\}, k && (20)
\end{aligned}$$

where (17) allows each task to be processed by an EOS at most once, (18) ensures that each EOS should have only one assignment appeared at the final optimal solution, and (20) states that obtained data volume for scheduled tasks with finish time more than $\text{ft}(TS_m)$ in EOS k should be transmitted to destinations timely. This replaces (9) and (10) in problem (P2).

As the number of targets increases, listing all observation sequences in constructed ADGs is not scalable because the number of observation sequences grow exponentially with the number of targets. To circumvent this difficulty, the column generation method [29], [30] is employed. For problem (P3), the algorithm decomposes (P3) into a master problem (P3-M) and a generation problem (P3-G). Master problem (P3-M) solves the linear programming with a selected subset of observation sequences. Since the number of observation sequences in master problem (P3-M) is much smaller than that of the original problem (P3), the complexity in solving master problem (P3-M) is significantly reduced. In generation problem (P3-G), we use duality theory to verify the optimality of master problem (P3-M). Then, a new observation sequence is selected and

added to master problem (P3-M) to improve the results.

The master problem, (P3-M), is formulated as follows:

$$
\begin{aligned}
(\mathbf{P3-M}) \quad \max_{\mathbf{A}} \ & \sum_{k\in\mathcal{K}}\sum_{\ell\in\mathcal{L}_k} a_\ell^k f_\ell^k \\
\text{s.t.} \ & \sum_{k\in\mathcal{K}}\sum_{\ell\in\mathcal{L}_k} a_\ell^k \rho_{i,\ell}^k \le 1, \ \forall i && (21)\\
& \sum_{\ell\in\mathcal{L}_k} a_\ell^k \le 1, \ \forall k && (22)\\
& a_\ell^k \in \{0,1\}, \ \forall \ell. && (23)
\end{aligned}
$$

A lemma is introduced below to show the property of problem (P3-M).

Lemma 3. *Problem (P3-M) is a weighted set packing problem.*

Proof: First, we show that (21) can be removed in problem P3-M without affecting its optimality by contradiction. Assume that there exists an EOS, k, satisfying $\sum_{\ell\in\mathcal{L}_k} a_\ell^k > 1$. Accordingly, at least two observation sequences, e.g., ℓ_1 and ℓ_2, are selected for the EOS, i.e., $a_{\ell_1}^k = a_{\ell_2}^k = 1$. If both ℓ_1 and ℓ_2 contain the same target i, we derive $\sum_{\ell\in\mathcal{L}_k} a_\ell^k \rho_{i,\ell}^k > 1$. Clearly, (21) is violated. Otherwise, we can equivalently substitute a new observation sequence $\ell = \{\ell_1 \cup \ell_2\}$. Thus, problem (P3) can be reformulated by removing (22), wherein an observation sequence ℓ corresponds to a set, and f_ℓ^k is its weight. To this end, problem (P3-M) turns into a weighted set packing problem [31]. ■

Problem (P3-M) can be approximately solved by local search algorithms [31], [32] or simply by the continuous relaxation technique. The solution to problem (P3-M) can be fractional if relaxation technique is applied. In this case, it is possible to round up the fractional solution to get a feasible solution by setting $a_\ell^k = \lfloor a_\ell^k \rfloor$. Denote λ_i as the optimal dual variable for constraint (11) in the master problem. By solving problem (P3-M), we can obtain λ_i. Subsequently, the optimality condition of obtained results are verified by the following inequality

$$
\sum_{i\in\mathcal{I}} (\lambda_i - \mu_i)\rho_{i,\ell}^k \ge 0, \ \forall k, \ell. \tag{24}
$$

A generation problem should be triggered if (24) does not hold. For EOS k, the generation

problem, (P3-G), is expressed as

$$(\mathbf{P3-G}) \quad \min \sum_{i\in\mathcal{I}} \sum_{\ell\in\mathcal{L}_k} (\lambda_i - \mu_i)\rho_{i,\ell}^k \quad \text{s.t.} \quad (20). \tag{25}$$

To solve problem (P3-G), we first give the following definition with respect to multi-constrained optimal path (MCOP) problem [33].

Definition 2. Consider an edge-weighted directed graph $\mathcal{OG}_k = (\mathcal{V}_k^{\text{ADG}}, \mathcal{E}_k^{\text{ADG}})$, with a primary cost parameter $c(e)$, and Q additional non-negative real-valued weights $\omega_q(e)$, $1 \le q \le Q$, associated with each edge $e \in \mathcal{E}_k^{\text{ADG}}$; a constraint vector $W = (W_1, \ldots, W_Q)$ where each W_q is a positive constant; and a source-destination node pair (v_s, v_d). A MCOP problem is to find a path, π, from v_s to v_d such that $c(\pi) = \sum_{e\in\pi} c(e)$ is minimized, subject to the constraints $\omega_q(\pi) = \sum_{e\in\pi} \omega_q(e) \le W_q$, $1 \le q \le Q$.

Lemma 4. *Problem (P3-G) can be equivalently reformulated as a MCOP problem.*

Proof: By properly associating each edge, $e \in \mathcal{E}_k^{\text{ADG}}$, with a cost $c(e)$ and $Q = M$ weights $\omega_q(e)$, $1 \le q \le M$, problem P3-G can be transformed into a MCOP problem. As for the cost, if $e = (o_i, o_j)$, we set $c(e) = (\lambda_j - \mu_j)\rho_{j,\ell}^k$. We also let $c(e) = (\lambda_i - \mu_i)\rho_{i,\ell}^k$ for $e = (v_s, o_i)$ and $c(e) = 0$ for $e = (o_i, v_d)$. The objective function in problem (P3-G) thus becomes finding a path, π, such that the total cost of the path $c(\pi) = \sum_{e\in\pi} c(e)$ is minimized. Besides, each edge, $e = (o_i, o_j)$, is associated with M weights, $\omega_q(e)$, $1 \le q \le M$. In terms of $e = (o_i, o_j)$, set $\omega_q(e) = nd_j \cdot \rho_{j,\ell}^k \cdot r_k^{\text{ob}}$ and $W_q = \sum_{\vartheta=1}^{q} D_{k,\vartheta}^{\text{tr}} \cdot r_k^{\text{tr}}$ for $q < g_j$, while we set $\omega_q(e) = 0$ and $W_q = 0$ for $q \ge g_j$. To this end, it can be observed that constraints $\omega_q(\pi) = \sum_{e\in\pi} \omega_q(e) \le W_q$ for $1 \le q \le M$ are equivalent to (20). This completes the proof. ■

According to Lemma 4, we can use existing pseudo-polynomial time approximation algorithms [34], [35] to solve problem (P3-G). After solving the overall observation scheduling sub-problem, we can obtain a set of observation sequences $\mathcal{L} = \{\ell_1, \ell_2, \ldots, \ell_K\}$, with ℓ_k corresponding to the observation sequence for EOS k. Denote $D_{k,m}^{\text{ob}}$ as the required transmission time in segment

TS_m for delivering all the observation data if ℓ_k is selected for EOS k, which is given by

$$D_{k,m}^{\text{ob}} = \frac{r_k^{\text{ob}}}{r_k^{\text{tr}}} \sum_{i \in \mathcal{I}_m} nd_i, \tag{26}$$

where $\mathcal{I}_m$ represents the subset of tasks transmitted in segment TS_m. Accordingly, the total required transmission time D_k^{ob} for EOS k is expressed as $D_k^{\text{ob}} = \sum_m D_{k,m}^{\text{ob}}$.

C. Redistribution of Surplus Transmission Time

In each iteration, a set of observation sequences $\mathcal{L} = \{\ell_1, \ell_2, \ldots, \ell_K\}$ is generated. Based on the result, we need to readjust the pre-allocated transmission time to match the required transmission time. Specifically, the total remaining transmission time for EOS k is $\max\{0, D_k^{\text{tr}} - D_k^{\text{ob}}\}$. Denote $\mathcal{K}_m^{\text{tr}}$ as the set of EOSs that posses surplus transmission time over segment TS_m, i.e., $\max\{0, D_{k,m}^{\text{tr}} - D_{k,m}^{\text{ob}}\} > 0,\ k \in \mathcal{K}_m^{\text{tr}}$. Similarly, the total required transmission time for EOS k is $\max\{0, D_k^{\text{ob}} - D_k^{\text{tr}}\}$. Denote $\mathcal{K}_m^{\text{ob}}$ as the set of EOSs that need more transmission time over segment TS_m, i.e., $\max\{0, D_{k,m}^{\text{ob}} - D_{k,m}^{\text{tr}}\} > 0,\ k \in \mathcal{K}_m^{\text{ob}}$. For each EOS, $k \in \mathcal{K}_m^{\text{tr}}$, the surplus transmission time $D_{k,m}^{\text{tr}} - D_{k,m}^{\text{ob}}$ over segment TS_m is equally distributed to a subset of EOSs $\mathcal{K}_m^{\text{ob}}$ that requires transmission time in the segment. We set $D_{k,m}^{\text{tr}} - D_{k,m}^{\text{ob}} = 0$ for EOS $k \in \mathcal{K}_m^{\text{tr}}$ if all its surplus transmission time in TS_m is distributed, or there are no EOSs need additional transmission time. The available transmission time for EOS $k \in \mathcal{K}_m^{\text{ob}}$ over segment TS_m is updated via

$$D_{k,m}^{\text{tr}} \leftarrow D_{k,m}^{\text{tr}} + \sum_{l \in \mathcal{K}_m^{\text{tr}}} \frac{D_{l,m}^{\text{tr}} - D_{l,m}^{\text{ob}}}{|\mathcal{K}_m^{\text{ob}}|},\ \forall k, m. \tag{27}$$

Then, the next iteration starts. The iteration process terminates until $\max\{0, D_k^{\text{tr}} - D_k^{\text{ob}}\} = 0$ holds for all $k \in \mathcal{K}$.

D. Algorithm Description and Analysis

By summarizing the preceding descriptions, the detailed procedure of AMRS algorithm is given in Algorithm 1. First, the algorithm initializes the iteration counter h, successfully scheduled task set $\mathcal{I}^*$, observation sequence set $\mathcal{L}^*$, and transmission time vector D^*. An initial

Algorithm 1 Approximate Multi-Resource Schedule (AMRS)

// Initialization: Set $h = 1$, $\mathcal{I}^* = \emptyset$, $\mathcal{L}^* = \emptyset$, and $D^* = \mathbf{0}$. Initiate a small set of observation sequences $\mathcal{L}_0$ and set $\mathcal{L} = \mathcal{L}_0$.
repeat
  // Transmission scheduling:
  if $h = 1$ **then**
    Construct an initial extended time sharing graph and compute $(D^{\text{tr}}_{1,1}, D^{\text{tr}}_{1,2}, \ldots, D^{\text{tr}}_{K,M})$.
  else
    // Redistribution of surplus transmission time
    Equally distribute the surplus transmission time $D^{\text{tr}}_{k^*,m} - D^{\text{ob}}_{k^*,m}$ from EOS $k^* \in \mathcal{K}^{\text{tr}}_m$ to corresponding subset of EOSs, i.e., $\mathcal{K}^{\text{ob}}_m$;
    Add $D^{\text{ob}}_{k^*}$ into D^*, and set $D^{\text{tr}}_{k^*,m} - D^{\text{ob}}_{k^*,m} = 0$;
    For EOS $k \in \mathcal{K}^{\text{ob}}_m$, update $D^{\text{tr}}_{k,m}$ using (27);
  end if
  // Observation scheduling:
  Find optimal observation sequence ℓ^* in the constructed ADG by employing MCOP solutions;
  $\mathcal{L} \leftarrow (\mathcal{L} \cup \ell^*) \setminus \mathcal{L}_{k^*}$;
  $\mathcal{L}^* \leftarrow \mathcal{L}^* \cup \{\ell_{k^*}\}$, $\mathcal{I}^* \leftarrow \mathcal{I}^* \cup \{i | i \in \ell^*\}$;
  $h = h + 1$;
until $\max\{0, D^{\text{tr}}_k - D^{\text{ob}}_k\} = 0$ holds for $\forall k \in \mathcal{K}$;
Return coordinate scheduling results $\mathcal{L}^*$ and D^*.

extended time sharing graph is built as stated in Section IV-A. A small feasible set of observation sequences $\mathcal{L}_0$ is chosen for problem (P3). For instance, $\mathcal{L}_0$ can be constructed by simply assigning a target to the EOS that has the earliest OW for it without violating constraint (20). Second, the transmission scheduling process starts when needed. Surplus transmission time is distributed from EOS $k^* \in \mathcal{K}^{\text{tr}}_m$ to EOS $k \in \mathcal{K}^{\text{ob}}_m$. Third, given an allocated transmission time vector, observation scheduling is done for the set of unscheduled EOSs utilizing state-of-the-art MCOP routing solutions, e.g., [35]. Finally, the iteration continues until all EOSs have no more remaining transmission time.

Herein, we analyze the proposed AMRS algorithm in terms of its convergence and computational complexity by the following theorems.

Theorem 1. *The AMRS algorithm will eventually terminate.*

Proof: In an iteration, if every EOS, $k \in \mathcal{K}$, satisfies $\max\{0, D^{\text{tr}}_k - D^{\text{ob}}_k\} > 0$, we naturally

set $\max\{0, D_k^{\text{tr}} - D_k^{\text{ob}}\} = 0$ in the end of the iteration, since no EOS needs additional transmission time. Therefore, the algorithm terminates. Otherwise, there must exist a subset of EOSs, e.g., $k^* \in \mathcal{K}_m^{\text{tr}}$, whose $\max\{0, D_{k^*}^{\text{tr}} - D_{k^*}^{\text{ob}}\} > 0$. In this case, remaining transmission time in EOS k^* over segment TS_m will be reallocated to the subset of EOSs, i.e., $\mathcal{K}_m^{\text{ob}}$, that need it. Thus, after distributing the surplus transmission time from EOS k^* to others, we have $\max\{0, D_{k^*}^{\text{tr}} - D_{k^*}^{\text{ob}}\} = 0$. As a result, in each iteration, there is at least one EOS k^* whose $D_{k^*}^{\text{tr}}$ becoming zero. By extension, it takes at most K iterations for all EOSs to reach $\max\{0, D_{k^*}^{\text{tr}} - D_{k^*}^{\text{ob}}\} = 0$, where the algorithm terminates. Hence, theorem 1 is proved. ■

Theorem 2. *The AMRS algorithm has a computational complexity of* $O(K^3I^3 + K^3M)$.

Proof: The complexity of the AMRS algorithm consists of three main parts: 1) transmission scheduling subproblem with the complexity of $O(KN)$—This complexity comes from building an extended transmission time sharing graph. It takes $O(KN)$ to construct the graph for a SIN with K EOSs and N destinations; 2) observation scheduling subproblem with the complexity of $O(K^2I^3)$—It corresponds to complexity of master problem $O(K^2)$ and complexity of generation problem $O(I^3)$. For the master problem, an algorithm with $O(K^2)$ complexity can be employed to solve the relaxed linear programming [10]. As for the generation problem, it first takes complexity of $O(I^2)$ to construct an ADG. Then, existing approximation algorithms with complexity of $O(I^3)$ can be devised to solve the equivalent MCOP problem on the ADG [34]. Therefore, the overall computational complexity of generation problem is $O(I^2) + O(I^3) = O(I^3)$; and 3) redistribution of surplus transmission time with complexity of $O(K^2M)$—In the worst case, for each EOS, its surplus transmission time over M segments are distributed to all other $K - 1$ EOSs. This incurs a complexity of $O(K^2M)$.

The algorithm runs iteratively to obtain the desired solution. It can be seen from theorem 1 that at most K iterations can take place. Within each iteration, the complexity is computed as $O(KN)+O(K^2I^3)+O(K^2M) \approx O(K^2I^3+K^2M)$. We neglect the term $O(KN)$ as the number of targets I is generally far more than that of destinations N. As a result, the total complexity of the proposed AMRS algorithm is approximately $K \cdot O(K^2I^3+K^2M)$, namely $O(K^3I^3+K^3M)$,

which is polynomial time complexity much less than that of the optimal solution. ■

V. Performance Evaluation

In this section, extensive simulation results are presented to evaluate the performance of our proposed AMRS algorithm. We quantitatively compare it with two baseline algorithms, namely separate scheduling (SS) algorithm and heuristic coordinate scheduling (HCS) algorithm. In SS algorithm, the observation scheduling problem is solved via the tabu search meta-heuristic, while the overlapping segment of transmission resources is equally allocated to corresponding EOSs. While in HCS algorithm, the overlapping part of conflicting TWs is first shared equally among its components. This constraint is then imposed to observation scheduling, which is further solved via the MCOP algorithm embedded in AMRS algorithm. We evaluate the system performance by the following two widely-used metrics in the literature [11]–[13]. The first is sum priorities of all successfully scheduled tasks. The other is guarantee ratio defined as (Total number of successfully scheduled tasks / Total number of tasks) ×100%.

We conduct experiments in a realistic SIN environment where the targets are randomly spread on the Earth's surface with latitude between $0°$N and $60°$N and longitude between $30°$E and $90°$E. A number of targets are properly chosen, and the priority of a target is a random integer distributed in the interval $[1, 10]$. The observation time duration of a task randomly takes a value from the set $\{1, 2, 3\}$ time slots. The length of a time slot is set to 1min. A number of EOSs are uniformly distributed over 2 sun-synchronous orbits at a height of 619.6km and with inclination $97.86°$. We set two relay satellites d_1 and d_2 which lie at nominal longitudes of $176.76°$E and $16.65°$E as the destinations. The data collection rate and the transmission rate for all EOSs are set to 300Mb/s. The set of available OWs and TWs are obtained by using the Satellite Tool Kit (STK). All the simulations are coded in Matlab simulator.

A. Optimality Gap Evaluation in a Small-Scale SIN

To demonstrate the feasibility and correctness of the developed AMRS algorithm, we first use a small-scale SIN where optimal solution to the optimization problem can be obtained using

Table I
SETTINGS OF THE SIMPLE TEST CASE.

Task No.	Priority	Observation Duration (Slots)
1	5	1
2	8	3
3	4	3
4	6	2
5	3	2

Table II
SCHEDULING RESULTS OF DIFFERENT ALGORITHMS.

Algorithm	Scheduled Task No.	Guarantee Ratio	Sum Priority
SS	1,4,5	60%	14
HCS	1,2,3	60%	17
AMRS	1,2,4,5	80%	22
Optimal	1,2,4,5	80%	22

an off-the-shelf solver [1]. There are 2 EOSs and a destination d_1 in the small-scale SIN. The considered time horizon is 10mins, i.e., 10 time slots. The number of targets is 5-10 with a step of 1. All targets arrive at the beginning of the time horizon. We do not impose delay requirements, i.e., expected finish times, for all tasks in this case.

An instance for a SIN with 5 targets is shown. Targets 1, 2, 3 and 4 can be observed by EOS 1, while targets 1, 3, 4 and 5 can be observed by EOS 2. In addition, target 3 and target 4 are conflicting with each other in both EOSs. The priority and required observation time duration are given in Table I. The maximum available transmission time for EOS 1 and EOS 2 to the destination are 5mins and 6mins, respectively, with 3mins that can be shared by both EOSs. Table II shows the scheduling results for four different algorithms. In AMRS algorithm, 4 out of 5 tasks can be successfully scheduled. It can be seen that our proposed AMRS algorithm produces the optimal results in the tested small-scale SIN. In SS and HCS algorithms, only 3 tasks can be completed, which is less than that in AMRS algorithm. Thus, the guarantee ratio and sum priority performance of SS and HCS algorithms is lower than that of AMRS algorithm.

Fig. 4(a) and 4(b) depict the sum priority and guarantee ratio of different algorithms with

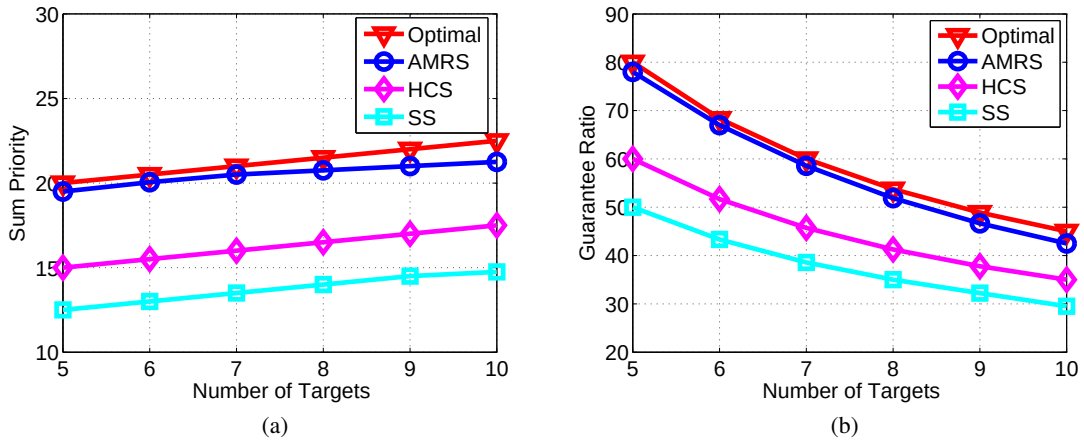


Figure 4. Evaluation of optimality gap in a small-scale SIN.

varying number of targets in small-scale SINs. It can be found that AMRS algorithm can achieve a network performance close to that of optimal. We attribute the result to the balanced matching between observation and transmission resources in AMRS algorithm, which boosts the resource utilization. In parallel, SS algorithm performs worst in terms of both the sum priority and guarantee ratio. This is accounted by the imbalance between observation and transmission resources in SS algorithm. Specifically, an EOS scheduled to observe a large number of targets does not receive sufficient transmission time. Consequently, as verified in Fig. 4(b), only a small number of tasks can be successfully scheduled. In HCS algorithm, the observation decisions are made based on the pre-allocated transmission resources. Since the transmission distribution scheme is sub-optimal, HCS algorithm yields lower network performance compared with that of AMRS algorithm.

B. Network Performance in a Large-Scale SIN

Herein, we use a large-scale SIN scenario to investigate the effects of different parameters on network performance. Six EOSs and two destinations d_1 and d_2 are deployed in the tested SIN. The whole time horizon is two hours from 10 Jun 2017 04:00:00 to 10 Jun 2017 06:00:00.

1) Performance impact of task number: In this experiment, we investigate the performance impact of the number of tasks that increases from 100 to 200 with an increment of 20. The

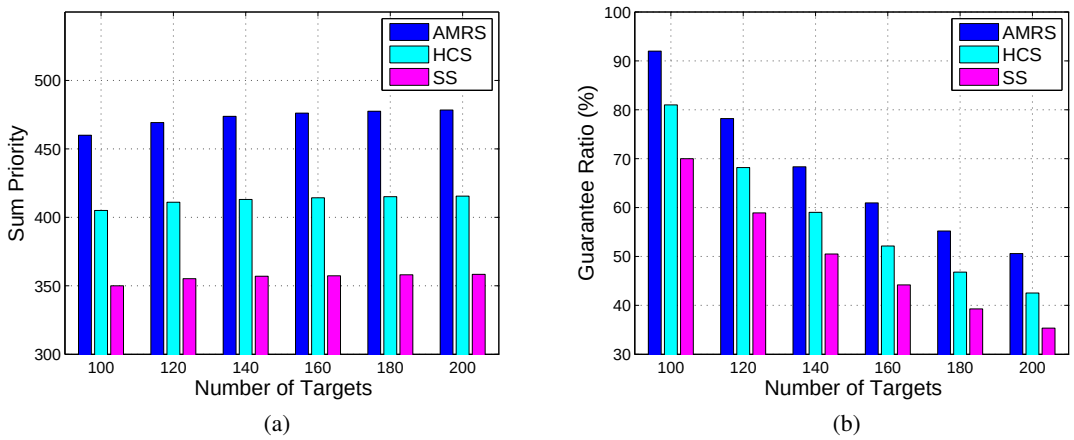


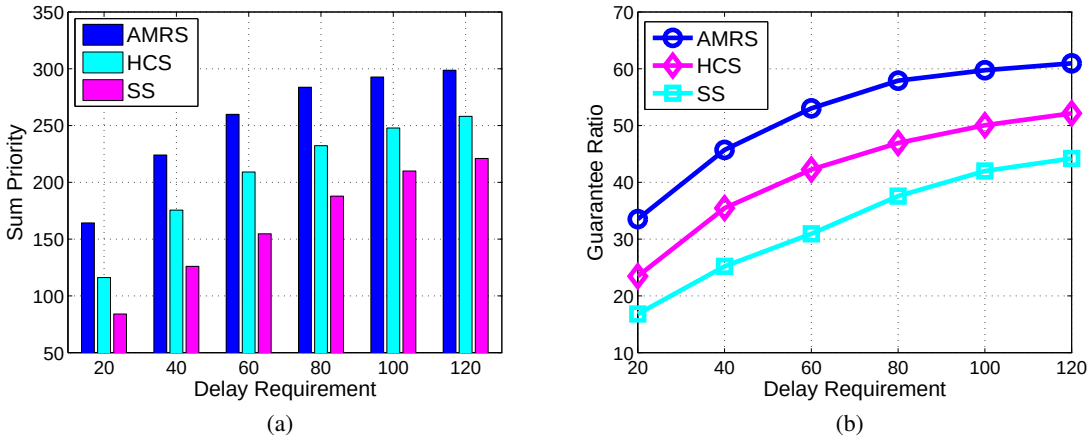
Figure 5. Network performance versus the number of targets.

delay requirements of all tasks are set to 120mins.

Fig. 5 demonstrates that the sum priority of the three algorithms gradually goes up with the increasing number of targets, because more high priority targets are accepted by exploiting the priority diversity. However, when the number of targets grows large enough, the sum priority does not improve any more. The reason lies in that transmission resources become the bottleneck in this case and additional observation data cannot be downloaded. Besides, Fig. 5(b) shows that with the increase of the number of targets, the guarantee ratios of all the three algorithms decrease. This trend is expected, because given limited resources, increasing the number of targets makes the tested system more heavily loaded, which in turn reduces the guarantee ratios.

2) Performance impact of delay requirement: Now we investigate the impact of the tasks' delay requirements on the performance of the three algorithms. A batch of 160 tasks arrive at the beginning of the time horizon. The delay requirement of a task varies from 20mins to 120mins with an increment of 20mins. The simulation results are plotted in Fig. 6.

Fig. 6(a) demonstrates that the sum priority of the three algorithms show ascending trends with the increase of delay requirement. The reason is that with the delay requirements of tasks becoming loose, the available observation and transmission opportunities for tasks increases. Meanwhile, more observation data can be stored onboard to reduce the transmission conflicts among multiple EOSs. Consequently, as can be seen in Fig. 6(b), more tasks can be accommo-



(a) (b)

Figure 6. Network performance versus delay requirements.

dated, and the guarantee ratios are improved. This further leads to an increase in the obtained sum priorities. Another observation is that AMRS algorithm achieves superior performance with respect to sum priority and guarantee ratio compared with that of SS and HCS algorithms. The underlying reason is that the proposed AMRS algorithm considers diverse service requirements of different tasks when making scheduling decisions, while both SS and HCS algorithms neglect the delay requirements. Subsequently, some tasks being scheduled can not meet their deadlines, and the allocated observation and transmission resources for the tasks are wasted.

VI. Conclusions and Future Work

In this paper, we investigate the problem of multi-resource coordinate scheduling for SINs. Based on an iterative optimization method, a low complexity joint scheduling algorithm, i.e., the AMRS algorithm, is proposed to properly balance the observation and transmission resources. As a result, each EOS can effectively observe a subset of targets according to its available transmission resources. Extensive simulations have been conducted to verify that the newly proposed AMRS algorithm performs close to the optimal solution in the tested small-scale SINs. In large-scale SINs, AMRS algorithm is shown to outperform two benchmark algorithms in terms of both the sum priority and guarantee ratio.

There are a few open issues to be addressed for future studies. First, we tend to take time-

varying capacity of both the observation and transmission resources into consideration. Second, we will extend the SIN scenario considered in this paper to include high altitude platforms (HAPs) as essential counterparts. Finally, we plan to develop new scheduling algorithms for dynamic observation tasks whose arrival times are not known in advance.

REFERENCES

[1] M. Sheng, Y. Wang, J. Li, R. Liu, D. Zhou, and L. He, "Toward a flexible and reconfigurable broadband satellite network: resource management architecture and strategies," *IEEE Wireless Commu.*, DOI: 10.1109/MWC.2017.1600173, 2017.

[2] H. Ramapriyan, "The role and evolution of NASA's earth science data systems," http://ntrs.nasa.gov, 2015.

[3] G. Wu, W. Pedrycz, H. Li, M. Ma, and J. Liu, "Coordinated planning of heterogeneous earth observation resources," *IEEE Trans. Trans. Syst., Man, Cybern: Syst.*, vol. 46, no. 1, pp. 109–124, Jan. 2016.

[4] C. Jiang, X. Wang, J. Wang, H.-H. Chen, and Y. Ren, "Security in space information networks," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 82–88, Aug. 2015.

[5] Q. Yu, W. Meng, M. Yang, L. Zheng, and Z. Zhang, "Virtual multi-beamforming for distributed satellite clusters in space information networks," *IEEE Wireless Commu.*, vol. 23, no. 1, pp. 95–101, Feb. 2016.

[6] J. Du, C. Jiang, Y. Qian, Z. Han, and Y. Ren, "Resource allocation with video traffic prediction in cloud-based space systems," *IEEE Trans. Multimedia*, vol. 18, no. 5, pp. 820–830, May 2016.

[7] J. A. Fraire and J. M. Finochietto, "Design challenges in contact plans for disruption-tolerant satellite networks," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 163–169, May 2015.

[8] W. J. Wolfe and S. E. Sorensen, "Three scheduling algorithms applied to the earth observing systems domain," *Management Science*, vol. 46, no. 1, pp. 148–168, Jan. 2000.

[9] D. Y. Liao and Y. T. Yang, "Imaging order scheduling of an earth observation satellite," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 37, no. 5, pp. 794–802, Sep. 2007.

[10] Z. Zhang, C. Jiang, S. Guo, Z. Ni, and Y. Ren, "Optimal satellite scheduling with critical node analysis," in *Proc. IEEE WCNC*, San Francisco, CA, USA, 2017, pp. 1–6.

[11] J. Wang, X. Zhu, D. Qiu, and L. T. Yang, "Dynamic scheduling for emergency tasks on distributed imaging satellites with task merging," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 9, pp. 2275–2285, Jun. 2014.

[12] X. Zhu, J. Wang, X. Qin, J. Wang, Z. Liu, and E. Demeulemeester, "Fault-tolerant scheduling for real-time tasks on multiple earth-observation satellites," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 11, pp. 3012–3026, Nov. 2015.

[13] X. Zhu, K. M. Sim, J. Jiang, J. Wang, C. Chen, and Z. Liu, "Agent-based dynamic scheduling for earth-observing tasks on multiple airships in emergency," *IEEE Syst. Journal*, vol. 10, no. 2, pp. 661–672, Jun. 2016.

[14] J. A. Fraire, P. G. Madoery, and J. M. Finochietto, "On the design and analysis of fair contact plans in predictable delay-tolerant networks," *IEEE Sensors Journal*, vol. 14, no. 11, pp. 3874–3882, Nov. 2014.

[15] J. Fraire, P. Madoery, and J. Finochietto, "Routing aware fair contact plan design for predictable delay tolerant networks," *Elsevier Ad-Hoc Networks*, vol. 25, pp. 303–313, Feb. 2015.

[16] K. Kaneko, Y. Kawamoto, H. Nishiyama, N. Kato, and M. Toyoshima, "An efficient utilization of intermittent surface-satellite optical links by using mass storage device embedded in satellites," *Performance Evaluation*, vol. 87, pp. 37–46, May 2015.

[17] Y. Wang, M. Sheng, J. Li, X. Wang, R. Liu, and D. Zhou, "Dynamic contact plan design in broadband satellite networks with varying contact capacity," *IEEE Commu. Letters*, vol. 20, no. 16, pp. 2410–2413, Dec. 2016.

[18] X. Jia, T. Lv, F. He, and H. Huang, "Collaborative data downloading by using inter-satellite links in LEO satellite networks," *IEEE Trans. Wireless Commu.*, vol. 16, no. 3, pp. 1523–1532, Mar. 2017.

[19] N. Bianchessi and G. Righini, "Planning and scheduling algorithms for the COSMO-SkyMed constellation," *Aerospace Science and Technology*, vol. 12, pp. 535–544, 2008.

[20] H. Chen, J. Wu, W. Shi, J. Li, and Z. Zhong, "Coordinate scheduling approach for EDS observation tasks and data transmission jobs," *Journal of Systems Engineering and Electronics*, vol. 27, no. 4, pp. 822–835, Aug. 2016.

[21] R. Liu, M. Sheng, K.-S. Lui, X. Wang, Y. Wang, and D. Zhou, "An analytical framework for resource-limited small satellite networks," *IEEE Commu. Letters*, vol. 20, no. 2, pp. 388–391, Feb. 2016.

[22] D. Zhou, M. Sheng, X. Wang, C. Xu, R. Liu, and J. Li, "Mission aware contact plan design in resource-limited small satellite networks," *IEEE Trans. Commu.*, 2017, DOI: 10.1109/TCOMM.2017.2685383.

[23] Y. Li, C. Song, D. Jin, and S. Chen, "A dynamic graph optimization framework for multihop device-to-device communication underlaying cellular networks," *IEEE Wireless Commu.*, vol. 21, no. 5, pp. 52–61, Oct. 2014.

[24] F. Malandrino, C. Casetti, C.-F. Chiasserini, and M. Fiore, "Optimal content downloading in vehicular networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1377–1391, Jul. 2013.

[25] V. Gabrel, A. Moulet, C. Murat, and V. T. Paschos, "A new single model and derived algorithms for the satellite shot planning problem using graph theory concepts," *Annals of Operations Research*, vol. 69, pp. 115–134, 1997.

[26] J. A. Fraire and J. M. Finochietto, "Design challenges in contact plans for disruption-tolerant satellite networks," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 163–169, May 2015.

[27] V. Gabrel, "Strengthened 0-1 linear formulation for the daily satellite mission planning," *J. Combinat. Optim.*, vol. 11, no. 3, pp. 341–346, 2006.

[28] N. Bianchessi and *et al.*, "A heuristic for the multi-satellite, multi-orbit and multiuser management of earth observation satellites," *Eur. J. Oper. Res.*, vol. 177, no. 2, pp. 750–762, 2007.

[29] H. Vance, C. Barnhart, E. Johnson, and G. L. Nemhauser, "Solving binary cutting stock problems by column generation and branch-and-bound," *Computational Optimization and Applications*, vol. 3, pp. 111–130, May 1994.

[30] C. Mancel and P. Lopez, "Complex optimization problems in space systems," in *Proc. 13th International Conf. Automated Planning & Scheduling (ICAPS'03)*, Trento, Italy, Jun. 2003, pp. 1–5.

[31] B. Chandra and M. M. Halldorsson, "Greedy local improvement and weighted set packing approximation," *Journal of Algorithms*, vol. 39, no. 2, pp. 223–240, May 2001.

[32] P. Berman, "A d/2 approximation for maximum weight independent set in d-claw free graphs," *Algorithm Theory*, pp. 214–219, Berlin, Germany: Springer, Jul. 2000.

[33] T. Korkmaz and M. Krunz, "Multi-constrained optimal path selection," in *Proc. IEEE INFOCOM*, 2001, pp. 834–843.

[34] M. Song and S. Sahni, “Approximation algorithms for multiconstrained quality-of-service routing,” *IEEE Trans. Computers*, vol. 55, no. 5, pp. 603–617, May 2006.

[35] G. Xue, W. Zhang, J. Tang, and K. Thulasiraman, “Polynomial time approximation algorithms for multi-constrained QoS routing,” *IEEE/ACM Trans. Netw.*, vol. 16, no. 3, pp. 656–669, Jun. 2008.