Proceedings of the Twenty-Sixth International Conference on
Automated Planning and Scheduling (ICAPS 2016)

Optimal Scheduling of a Constellation of Earth-Imaging Satellites, for Maximal Data Throughput and Efficient Human Management

Sean Augenstein, Alejandra Estanislao, Emmanuel Guere, Sean Blaes
Google Inc.
1600 Amphitheatre Parkway
Mountain View, California 94043

Abstract

A mixed-integer linear program (MILP) approach to scheduling a large constellation of Earth-imaging satellites is presented. The algorithm optimizes the assignment of imagery collects, image data downlinks, and "health & safety" contacts, generating schedules for all satellites and ground stations in a network. Hardware-driven constraints (e.g., the limited agility of the satellites) and operations-driven constraints (e.g., guaranteeing a minimum contact frequency for each satellite) are both addressed. Of critical importance to the use of this algorithm in real-world operations, it runs fast enough to allow for human operator interaction and repeated rescheduling. This is achieved by a partitioning of the problem into sequential steps for downlink scheduling and image scheduling, with a novel dynamic programming (DP) heuristic providing a stand-in for imaging activity in the MILP when scheduling the downlinks.

1 Introduction

Terra Bella (formerly Skybox Imaging) provides access to timely, high-resolution satellite imagery and analytics of Earth, to improve understanding of daily patterns in the world. During the course of 2016, Terra Bella will launch 10+ satellites to add to the existing 2 satellites currently in orbit. These additional assets are critical to Terra Bella achieving its business goals of analytics applications built around timely geo-spatial information. Such a large coordinated satellite constellation is unique in the commercial high-resolution remote sensing business, and presents novel challenges to efficient operation. The job of the scheduling software is to automatically assign imaging and downlink tasks to each satellite to maximize throughput of imagery over a set scheduling horizon (typically 8-12 hours). This paper presents the automatic constellation scheduling algorithm in use at Terra Bella. The algorithm addresses all of the desired constraints and objectives (enumerated below), and allows for the seemless integration of even more satellites and ground stations beyond those immediately planned.

Orbits and Opportunities

Each satellite completes an orbit of the Earth once every 90-100 minutes, collecting imagery over desired targets and downlinking the data during ground station passes. Opportunities to image are dictated by the set of targets and when each satellite's orbit takes it in view of these targets. Similarly, opportunities to downlink are dictated by the fixed locations of the ground stations and when each satellite's orbit takes it in range of these locations. Figure 1 shows an example orbit over western Asia and eastern Africa, with a downlink when in range of a Terra Bella ground station in Norway, and an imaging pass when in view of a target in South Africa.

Figure 1: Examples of downlink opportunity (Norway) and imaging opportunity (South Africa)

Satellites must align to a specific orientation to point the camera or antenna at a location on the Earth (as indicated in Figure 1). Opportunities are defined by this required pointing orientation as well as the time at which they occur.

Opportunities over a ground station serve double duty for Terra Bella operations. During a ground station pass, imagery data stored onboard the satellite is downlinked to the ground, to be processed into images. In parallel, the operations team can 'contact' the satellite to receive engineering telemetry or issue new commands, for health and safety monitoring and upkeep of the constellation. So a downlink opportunity and a contact opportunity are one and the same.

As motivated further below, in the Terra Bella scheduling problem all opportunities are discrete entities; a pass over a target or a ground station will be broken down into a finite set of opportunities, each with a different fixed starting time, ending time, starting orientation, and ending orientation. Scheduling is then a selection problem of choosing which opportunities are in or out of the schedule. For imaging targets, the time duration of these opportunities is fixed by the exact geometry to be imaged and the camera's characteristics. For ground stations (where downlinks/contacts can occur for any continuous period as long as the satellite is in range), the time duration of these opportunities is fixed to be one minute. They are set to be adjacent to each other and abutting in time, so that a continous downlink of k minutes is achieved by selecting k adjacent downlink opportunities.

The Constellation Scheduling Problem

The hardware-driven constraints on the problem are:

- Satellites can only do one thing at a time, so same-satellite opportunities that overlap in time are mutually exclusive.
- Satellites have finite agility, where agility is the ability to change orientation over time. Thus, some pairs of non-overlapping same-satellite opportunities are also mutually exclusive (i.e., when the satellite is incapable of transitioning from the ending orientation of one opportunity to the starting orientation of another opportunity in the time in-between).
- Ground stations can only do one thing at a time, and need a non-zero time to reset themselves between different satellite passes. Same-ground station, different-satellite opportunities that are too close together or overlap in time are mutually exclusive.

Beyond hardware-dictated constraints, a variety of additional constraints and objectives arise when operating such a large imaging constellation. A human factors problem emerges for the Terra Bella operations team. Rather than rely on enough human bandwidth to monitor each satellite's scheduled ground station passes, the sheer volume of passes necessitates that the team be less involved on a per-satellite basis in real-time.

As a result, the operations team requires that the scheduling algorithm guarantee a minimum contact frequency with each satellite ("at least every nth orbit, talk for at least m minutes", where typically $n = 3$ and $m = 3$), to ensure enough time is reserved for health and safety monitoring. They also require the automatic balancing of imaging with downlinking, so that the satellites should *neither* pass up too many high value image targets *nor* build up any large backlogs of undownlinked image data.

In case of an unplanned occurrence requiring human intervention, the operations team needs the ability to force the scheduling system to lock a downlink/contact opportunity into a satellite's schedule (or lock it out of a satellite's schedule). Finally, to make the system as user friendly as possible, it's required that the algorithm run fast enough that an operator can interact with the scheduling system and regenerate a new schedule in a short period of time (where "short" in operations is considered to be on the order of 15 seconds or less).

Summarizing, the operations-driven constraints and objectives are:

- Satisfy the minimum contact frequency for each satellite
- Balance the competing objectives of image collecting and data downlinking
- Provide human operators the ability to manually "lock-in" or "lock-out" a given contact
- Constellation-wide optimization occurs in a short period of time (less than 15 seconds)

Finally, a business objective is to maximize the image collection of priority-weighted targets. As the algorithm is already required to run fast, it is straightforward to alter a target's priority weight and regenerate a schedule due to dynamic changes in target desirability (e.g., clouds).

The paper is organized as follows. The next section discusses previous research into satellite scheduling, and why it was inapplicable to operating the Terra Bella constellation. Section 3 outlines a new approach, a mixed-integer linear program (MILP) for constellation-wide scheduling. It specifies exactly how the various constraints enumerated above get mathematically encoded as part of a MILP, to yield a scheduler that simultaneously assigns both downlinks and image collects. It also explains why this 'global' MILP has issues with meeting the performance goals desired. This motivates a reformulation into a 'sequential' scheduler, with successive scheduling of first the downlinks and then the image collects. Section 4 describes this sequential approach. In particular, it details the novel dynamic programming (DP) heuristics that are used to allow downlinks to get scheduled in a MILP with awareness of how much and when imaging activity is taking place, *without* explicitly having to schedule image collects until a subsequent step. Section 5 provides comparison and analysis of the two algorithms and Section 6 concludes.

2 Background

A large body of research exists on optimal scheduling of Earth-imaging satellites. Maximizing imaging productivity has been a topic of heavy focus, with research into optimizing the imaging schedule of a single satellite (Bensana et al. 1996; Gabrel et al. 1997; Vasquez and Hao 2001; Gabrel and Vanderpooten 2002; Lemaitre et al. 2002; Gabrel and Murat 2003; Vasquez and Hao 2003; Lin et al. 2005; Tangpattanakul, Jozefowiez, and Lopez 2015) or of multiple satellites in a coordinated Earth-imaging constellation (Bianchessi et al. 2007; Yao et al. 2010; Shea and Nasgovitz 2011). Research has also been conducted on optimizing the downlink activity between a ground station network and a single satellite or multiple satellites (Burrowbridge 1999; Barbulescu et al. 2004; Zufferey, Amstutz, and Giaccari 2008; Marinelli et al. 2011; Karapetyan et al. 2015).

In contrast with the above work, in the Terra Bella use case both imaging and downlinking scheduling must be addressed in the same algorithm, because the satellites only undertake one activity at a time. Anytime both an imaging target and a ground station are accessible by a satellite, the scheduling algorithm must choose between these competing objectives.

In general the problem of co-optimizing the imaging and downlinking of a large satellite constellation based on high-fidelity data models has been less thoroughly studied. Some work has made assumptions that a downlink opportunity over a ground station is always selected as part of the schedule (Globus et al. 2004), regardless of how much data is onboard and/or whether there are conflicting image opportunities that would be more desirable to schedule instead. Other articles cover constellations where the satellites can simultaneously image and downlink (Bianchessi and Righini 2008; Wang et al. 2011), so don't address the additional aspect of having to choose between the two. Finally, something not previously studied in satellite constellation scheduling is optimizing imaging and downlinking while also selecting ground station opportunities to ensure a minimum contact constraint is met.

Recall that runtime is a chief concern. The algorithm presented in this paper continues in the spirit of previous mathematical programming solutions to satellite scheduling problems (Bensana et al. 1996; Gabrel and Murat 2003; Gabrel 2006; Bogosian 2008). The continuous span of time that a satellite is in view of a target is turned into a set of multiple feasible discrete opportunities. This results in a combinatorial optimization problem to select the opportunities to be included in the schedule. Discretization sets this algorithm and its antecedents apart from a number of the other approaches listed above. While it necessarily incurs accepting some theoretical sub-optimality, it allows the design of the algorithm to commence from a starting approach that runs in polynomial time rather than being NP-hard.

Discrete Graph-Based Approaches

Optimal opportunity selection for a single satellite (without any additional objectives or constraints) can be viewed as a directed acyclic graph (DAG) problem, where satellite opportunities are nodes in a graph and directed edges encode feasible transitions. An example is the graph-based approach taken at Terra Bella during the first two years of satellite operations (Augenstein 2014). DAGs are formed by ordering all opportunities chronologically and then using an agility model heuristic to determine which pairs of opportunities should have edges connecting them. Figures 2 and 3 give a pictorial view of this process. (Note that for simplicity only one opportunity is depicted for each imaging target, when in reality there are several opportunities per target per pass.)

The highest weighted path through the graph is the optimal schedule of the satellite over the time period of the DAG, and can be calculated via either dynamic programming (DP) (Leiserson et al. 2001) or integer/linear programming (IP/LP) (Nemhauser and Wolsey 1988). A brief review of highest weighted path DP is given here, as one of the schedulers in this paper makes use of it as a heuristic (explained in detail in Section 4). Let c_j be the priority of opportunity j and let C_j be the cumulative priority of the

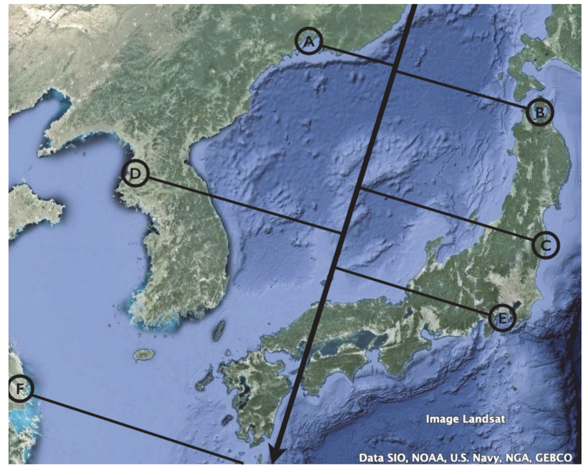

Figure 2: Example Opportunities Near Sea of Japan

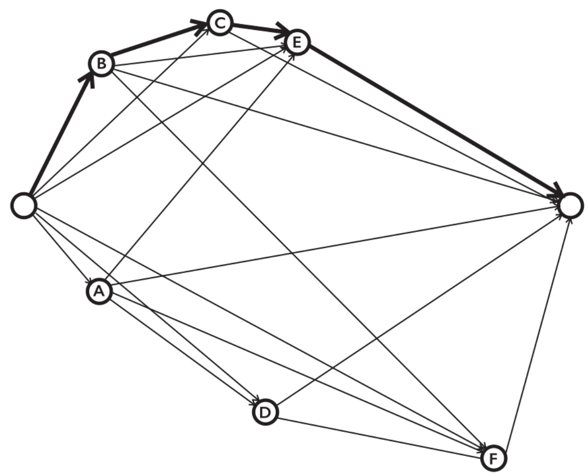

Figure 3: Sea of Japan Opps. arranged in DAG

highest weighted path between the start node and node j. Let p_j be a pointer to the node that preceeds node j along this highest weighted path. From a node-centric perspective, the calculation at each node j (to determine C_j and p_j) is:

$C_j = 0$
for all $\{i \mid edge(i,j) = true\}$ **do**
  if $C_i + c_j > C_j$ **then**
    $C_j \leftarrow C_i + C_j$
    $p_j \leftarrow i$
  end if
end for

This calculation is performed for each node in chronological order, from start node to end node. After completing, backtracking via the values of p_j (starting with p_{end}) yields the opportunities in the optimal satellite schedule. Such a DP process is repeated for multiple DAGs on each satellite, together spanning the timespan between the current time and the end of the scheduling horizon. The highest weighted path through each graph is calculated, and these output paths sequentially form the total schedule of a particular satellite.

Constellation-Scale Constraints

If DP is naively applied to the constellation, this process is repeated for all the satellites (see Figure 4).

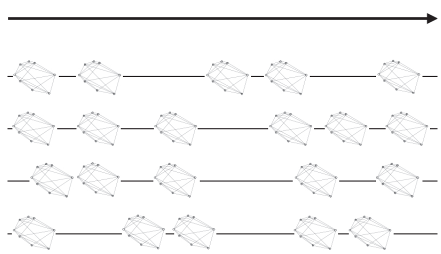

Figure 4: Multiple DAGs across many Satellites

However, the DP approach starts to reveal significant drawbacks as the transition to constellation-scale operations is made. A major aspect lacking is any awareness of multi-satellite constraints; each satellite is planned independently and without consideration of the others in the constellation. For example, given that only one satellite can talk to a given ground station at a time, anytime there are two satellites with conflicting downlinks requires an ad-hoc deconfliction. As Table 1 shows, with constellation growth the amount of downlink passes requiring deconfliction rises significantly, necessitating a lot of ad-hoc behavior.

# of satellites	# of ground stations	likelihood that station pass overlaps w/ another satellite's
2	2	0.199
13	2	0.888

Table 1: Likelihood of Pass Conflicts vs. Constellation Size

What is desired instead is an approach where decisions such as which satellites get to downlink to a ground station (and for how long) are based on objectives globally considered over all satellites and all ground stations over the entire set of opportunities in the scheduling horizon. For handling a constellation-scale optimization with the various features enumerated in Section 1, the need for richer constraints and objectives requires a change to an IP/LP approach. The combinatorial optimization can be encoded as a mixed-integer linear program (MILP) (Nemhauser and Wolsey 1988) with boolean variables for opportunity selection and continuous variables modeling amount of onboard data. The next section explains in detail how this is done.

3 Constellation Scheduling as a MILP

Let S_{img} be the set of image opportunities, let S_{dl} be the set of downlink opportunities, and let S_{all} be the set of all opportunities ($S_{img} \cup S_{dl}$).

Each opportunity i has a boolean variable (x_i) indicating inclusion in the schedule. Each opportunity i has a positive real (y_i) for amount of onboard data at end time of opportunity. Each imaging opportunity i has a priority weight ($c_i > 0$) indicating how important capturing that image is (for all downlink opportunities, $c_i = 0$). Opportunities can also have a penalizing cost of onboard data storage, d_i.

Objective

The objective function represents the twin goals of collecting the largest amount of prioritized imagery (minimizing the sum of $-c_i x_i$) and continually lowering the amount of undownlinked data onboard the satellite (minimizing the sum of $d_i y_i$):

$$
\begin{aligned}
\underset{x_i}{\text{minimize}} \quad & \sum_{i \in S_{all}} (-c_i x_i + \alpha d_i y_i) \\
\text{subject to:} \quad & x_i \in \{0, 1\} \\
& y_i \in \mathbb{R} \\
& y_i \geq 0 \\
& \vdots
\end{aligned}
$$

A good choice for how to set the storage penalty d_i is based on an objective of minimizing the "wait time" that image data spends onboard the satellite before it gets downlinked. Consider a scheme where for all imaging opportunities d_i is zero and for all downlink opportunities d_i is the time duration until the next downlink opportunity occurs. Then image data is penalized if it is still present onboard the satellite *after* the earliest opportunity at which it could be downlinked. The penalty is proportional to the amount of data and the duration of time it will be present onboard.

The coefficient α is a parameter that controls the relative importance of imaging versus downlinking. It is set empirically, and can be adjusted situationally by the Terra Bella operations team to indicate a stronger preference for one objective over the other.

Satellite and Ground Station Mutual Exclusion Constraints

Let there be a set, $S_{sat-mutex}$, containing all pairs of same-satellite opportunities that are mutually exclusive (e.g., due to lack of agility). Let there be another set, $S_{gs-mutex}$, containing all pairs of same-ground station opportunities that are mutually exclusive (e.g., due to overlapping time). Note that these sets are quickly precomputed before the optimization runs. Then:

$$
\begin{aligned}
x_k + x_l \leq 1 \ \ \forall \{k, l | \langle k, l \rangle \in S_{sat-mutex}\} \\
x_k + x_l \leq 1 \ \ \forall \{k, l | \langle k, l \rangle \in S_{gs-mutex}\}
\end{aligned}
$$

Onboard Data Model Continuity Constraints

Consider the opportunities to be ordered chronologically by end time, so i is the opportunity whose end time is soonest after the end time of $i-1$. Another set of constraint equations enforce that the amount of onboard data y increases (when imaging) and decreases (when downlinking). These equations are linear and a function of whether opportunity i is included ($x_i = 1$) or excluded ($x_i = 0$):

$$y_i = y_{i-1} + a_i x_i \;\; \forall\, i \in S_{img}$$
$$y_i \geq y_{i-1} - a_i x_i \;\; \forall\, i \in S_{dl}$$

For an imaging opportunity, a_i is the amount of data that would be collected if scheduled. For a downlink opportunity, a_i is the maximum amount of data that could be downlinked if scheduled. The inequality exists to handle the case where downlink capacity a_i is larger than the amount of data coming into the ground station pass y_{i-1} (the exiting amount of data y_i can never get less than zero).

Note that if opportunity i is not scheduled, y_i is simply the same as y_{i-1} (no change in amount of onboard data).

Minimum Contact Frequency Constraints

Additional indicator variables are used to formulate the minimum contact constraints. Recall from Section 1 that all downlink opportunities are one minute in duration and abut other downlink opportunities to the same ground station. Let a boolean variable w_i be an indicator of whether contiguous downlink opportunities i, $i+1$, and $i+2$ have all been selected for inclusion in the schedule, forming a 3-minute period of continuous contact between the satellite and the ground station. Let j index the orbits of a satellite, and let a boolean variable v_j be an indicator of whether at least 3 contiguous minutes of contact will occur in orbit j. Then the constraint to enforce that at least once every 3 orbits, a satellite contacts a ground station for at least 3 minutes is encoded as:

$$v_j \in \{0, 1\}$$
$$w_i \in \{0, 1\} \;\; \forall\, i \in S_{dl}$$
$$x_i + x_{i+1} + x_{i+2} \geq 3w_i \;\; \forall\, \{i | e_i = s_{i+1}, e_{i+1} = s_{i+2}\}$$
$$\sum_{i \in O_j} w_i \geq v_j$$
$$v_j + v_{j+1} + v_{j+2} \geq 1$$

In the above expression, the variables s_i and e_i are the starting and ending times (respectively) of opportunity i, so the condition of $e_i = s_{i+1}$ is a statement that opportunity i and $i+1$ are contiguous in time.

Also above, O_j represents the set of all downlink opportunities occuring during orbit j.

Opportunity Lock-In/Lock-Out

Lock-ins and lock-outs are simply the assignment of an indicator variable to be 0 or 1. (Note this assumes logic has been applied to ensure a user cannot lock-in two mutually exclusive opportunities).

$$x_i = 1 \;\; \forall\, i \in S_{lock-in}$$
$$x_i = 0 \;\; \forall\, i \in S_{lock-out}$$

Summary of Global MILP

Putting the objective and all of the above constraints together, the global MILP is:

$$\begin{aligned}
\underset{x_i}{\text{minimize}} \quad & \sum_{i \in S_{all}} (-c_i x_i + \alpha d_i y_i) \qquad (1) \\
\text{subject to:} \quad & x_i \in \{0, 1\} \\
& y_i \in \mathbb{R} \\
& y_i \geq 0 \\
& x_k + x_l \leq 1 \;\; \forall\, \{k, l | \langle k, l \rangle \in S_{sat-mutex}\} \\
& x_k + x_l \leq 1 \;\; \forall\, \{k, l | \langle k, l \rangle \in S_{gs-mutex}\} \\
& y_i = y_{i-1} + a_i x_i \;\; \forall\, i \in S_{img} \\
& y_i \geq y_{i-1} - a_i x_i \;\; \forall\, i \in S_{dl} \\
& v_j \in \{0, 1\} \\
& w_i \in \{0, 1\} \;\; \forall\, i \in S_{dl} \\
& x_i + x_{i+1} + x_{i+2} \geq 3w_i \;\; \forall\, \{i | e_i = s_{i+1}, e_{i+1} = s_{i+2}\} \\
& \sum_{i \in O_j} w_i \geq v_j \\
& v_j + v_{j+1} + v_{j+2} \geq 1 \\
& x_i = 1 \;\; \forall\, i \in S_{lock-in} \\
& x_i = 0 \;\; \forall\, i \in S_{lock-out}
\end{aligned}$$

Computational Considerations

The MILP formulated above assigns image and downlink opportunities simultaneously. Because of the onboard data model, there is a dense set of constraints linking all opportunities. In practice, this has negative runtime performance implications that preclude the use of this global MILP (Section 5 will present specific results). The next section presents an alternative to the global approach; it sequentially schedules downlinks and then image collects, with a novel DP heuristic providing the downlink scheduling MILP with 'awareness' of imaging activity.

4 A Sequential Algorithm

A path forward to acceptable runtime performance exists by separating the problems of downlink opportunity selection and imaging opportunity selection. This uncouples the variables that are present in the majority of the constraint equations (the downlink opportunities) from the variables that are in the majority (the imaging opportunities). Downlink opportunity selection is done first and imaging opportunity selection is done second. These two sequential steps are now presented.

Downlink Scheduling as a DP-aided MILP

While a reduced MILP that schedules only the downlink opportunities rather than all opportunities will clearly be faster, it is still essential to select the downlink opportunities in a way that matches the imaging activity occurring on each satellite. An acceptable algorithm preserves the preference for more downlink time when the satellite has just flown over a target rich area (e.g., western Europe, where it has likely collected a lot of images) and for less downlink time when the satellite has just flown over a target sparse area (e.g., the Pacific Ocean).

This can be achieved via a heuristic that approximates imaging activity without explicitly scheduling image opportunities. Specifically, two properties need to be approximated. First, the amount of priority-weighted imagery that a satellite would encounter during a particular time period is needed for making trade-off decisions between imaging and downlinking when opportunities would be mutually exclusive (e.g., image targets in view of a satellite while a ground station is also in range). Second, a profile of how much imagery a satellite would collect as a function of time is needed to adequately model how much data is onboard, and thus estimate the relative demand for downlink.

Approximated Priority-Weighted Imagery A modified priority weight $\tilde{c}_i$ for downlink opportunities is now introduced. Where in the previous section the downlink opportunities had zero priority weight ($c_i = 0$), now the downlink opportunities have associated to them a value $\tilde{c}_i$ representing the approximate summed priority weight of all imaging that could have feasibly been acquired if the downlink opportunity was *not* used (a "cost of missing out").

The objective function of the downlink scheduling problem is now expressed as a desired to both "not miss out on imagery" (minimizing the sum of $\tilde{c}_i x_i$) while also continually lowering the amount of undownlinked data onboard the satellites (minimizing the sum of $d_i y_i$):

$$\begin{aligned} \underset{x_i}{\text{minimize}} \quad & \sum_{i \in S_{dl}} (\tilde{c}_i x_i + \alpha d_i y_i) \\ \text{subject to:} \quad & x_i \in \{0,1\} \\ & y_i \in \mathbb{R} \\ & y_i \geq 0 \\ & \vdots \end{aligned}$$

Approximated Image Data Profile As earlier, a linear set of constraint equations enforce that the onboard data y is modeled to be consistent with imaging and downlink activity (increasing and decreasing the total data amount, respectively).

Recall that the opportunities can be considered ordered chronologically by end time. In contrast to the global MILP, now the constraint equations are written between only downlink opportunities ($i \in S_{dl}$). The constraint equation that links all consecutive downlink opportunitives is:

$$y_i \geq y_{i-1} - \tilde{a}_i x_i + \tilde{b}_i \;\; \forall \, i \in S_{dl}$$

The baseline imaging variable $\tilde{b}_i$ represents an estimate of how much imaging data would be acquired between the end of downlink opportunity $i-1$ and the end of downlink opportunity i, if the time was entirely devoted to imaging.

The $\tilde{a}_i$ term includes the amount of data that could be downlinked during opportunity i (same as above in the global MILP). It also includes an estimate of the amount of imaging data that is *not* acquired between the start and end of opportunity i, because the satellite is instead occupied with downlinking.

Putting them together, the value of $-\tilde{a}_i + \tilde{b}_i$ is equal to the net change in onboard data given the image data that is added via collecting between end of downlink $i-1$ and start of downlink i and given the image data that is subtracted via downlinking during downlink i.

DP as a Heuristic for Imaging Activity A good imaging heuristic that provides values for $\tilde{c}_i$, $\tilde{a}_i$, and $\tilde{b}_i$ can be found in the basic single satellite DP algorithm discussed in Section 2. Consider a DAG (like the one depicted in Figure 3) that contains a downlink opportunity i as one of the nodes in the graph.

Let C_{end} be the cumulative priority weight of the highest weighted path, and let $C_{end \mid i}$ be the cumulative priority weight of the highest weighted path that *must go through* opportunity i. Then:

$$\tilde{c}_i = C_{end} - C_{end \mid i}$$

Similarly, let Y_{end} be the cumulative amount of image data collected when scheduling the image opportunities that form the highest weighted path, and let $Y_{end \mid i}$ be the cumulative amount of image data collected on the highest weighted path that must go through i. Also, let Y_i be the cumulative amount of image data collected up until downlink opportunity i when the highest weighted path must go through i. Then:

$$\begin{aligned} \tilde{a}_i &= Y_{end} - Y_{end \mid i} + a_i \\ \tilde{b}_i &= (Y_{end \mid i-1} - Y_{i-1}) - (Y_{end \mid i} - Y_i) \end{aligned}$$

The first term reflects how much data is *not* collected due to inclusion of downlink opportunity i in the schedule ($Y_{end} - Y_{end \mid i}$), as well as the amount of data a_i that could be downlinked during i. The second term includes the amounts of imagery data that would be collected after $i-1$ and after i (respectively, $Y_{end \mid i-1} - Y_{i-1}$ and $Y_{end \mid i} - Y_i$), and considers the difference to be an estimate of how much imaging could be collected between $i-1$ and i.

As the DP approach yields an extremely fast calculation (the DAGs are separable and the highest weighted paths can be calculated in parallel), this approach provides a computationally efficient heuristic for determining $\tilde{c}$, $\tilde{a}$, and $\tilde{b}$.

Summary of DP-aided MILP Putting the objective and all of the above constraints together, the DP-aided MILP algorithm used in the sequential approach is:

$$\begin{aligned}
\underset{x_i}{\text{minimize}} \quad & \sum_{i \in S_{dl}} \tilde{c}_i x_i + \alpha d_i y_i \qquad (2) \\
\text{subject to:} \quad & x_i \in \{0,1\} \\
& y_i \in \mathbb{R} \\
& y_i \geq 0 \\
& x_k + x_l \leq 1 \;\; \forall \{k,l | \langle k,l \rangle \in S_{sat-mutex}\} \\
& x_k + x_l \leq 1 \;\; \forall \{k,l | \langle k,l \rangle \in S_{gs-mutex}\} \\
& y_i \geq y_{i-1} + \tilde{a}_i x_i + \tilde{b}_i \;\; \forall\, i \in S_{dl} \\
& v_j \in \{0,1\} \\
& w_i \in \{0,1\} \\
& x_i + x_{i+1} + x_{i+2} \geq 3 w_i \;\; \forall \{i | e_i = s_{i+1}, e_{i+1} = s_{i+2}\} \\
& \sum_{i \in O_j} w_i \geq v_j \\
& v_j + v_{j+1} + v_{j+2} \geq 1 \\
& x_i = 1 \;\; \forall\, i \in S_{lock-in} \\
& x_i = 0 \;\; \forall\, i \in S_{lock-out}
\end{aligned}$$

Image Scheduling

Image scheduling is trivial after the downlink MILP. With the assurance that the downlink scheduling has worked around imaging as best possible, image scheduling can take place via a variety of solution methods (DP, IP/LP, etc.). For the results in the next section, image scheduling was performed via DP as described in Section 2, with any imaging opportunities that are mutually exclusive with scheduled downlink opportunities removed from consideration.

5 Analysis

The rationale for the sequential scheduling algorithm formulated in the last section is that it is computationally fast enough to meet the runtime demands of Terra Bella constellation operations. In this section, the performance differences are quantified between the global MILP model (1) and the sequential model with DP-aided MILP (2). Three test cases were formulated, each representative of different steps in a realistic ramp-up in operating capacity during 2016. The first test case is with 3 satellites, the second test case is with 7 satellites, and the third test case is with 13 satellites. All cases involve a 10-hour scheduling horizon with a representative set of approximately 7000 imaging targets (which can be imaged by any satellite). Table 2 summarizes the cases.

test case	# of satellites	# of opportunities (imaging & downlink)
i	3	13659
ii	7	30516
iii	13	52695

Table 2: Test Cases

Tests were on a machine with 3.2GHz Intel Xeon E5 processor and 32GB of memory, using SCIP (Achterberg 2009), a fast open-source MILP software package in wide use.

Aside from algorithm runtime, other important metrics for algorithm validation are the sum priority of the scheduled imagery collects (C), and the average wait time that image data sits onboard the satellites ($\bar{D}$). It is desired that the sum priority of all scheduled collects be as high as possible, reflecting that the satellites' imaging activities are focused on the largest number of and highest priority targets. It is calculated via:

$$C = \sum_{i \in S_{img}} (c_i x_i)$$

The wait time is the time between earliest possible image downlink time and actual image downlink time. It is desired that wait time be as low as possible, reflecting that the satellite constellation and ground station network are utilized to promptly transmit image data to the ground. Using the wait time storage penalties d_i defined in Section 3, average wait time $\bar{D}$ is calculated via:

$$\bar{D} = \frac{\sum_{i \in S_{dl}} (d_i y_i)}{\sum_{i \in S_{dl}} (y_i)}$$

Table 3 displays the experimental results, with runtime, C, and $\bar{D}$ given for each test configuration.

case	algorithm	runtime (s)	C	$\bar{D}(s)$
i	global	*3600*	661×10^3	1585.29
i	sequential	2.86	783×10^3	1550.74
ii	global	*3600*	119×10^4	1720.33
ii	sequential	5.66	177×10^4	1649.56
iii	global	*3600*	183×10^4	1780.35
iii	sequential	12.41	312×10^4	1664.67

Table 3: Operational Metrics vs. Algorithm Used
(Italics indicate timeout before convergence)

All results in Table 3 indicate the superior performance of the sequential scheduler with DP-aided MILP over the global MILP. With regards to runtime, the global scheduler never converged to optimal for any of the test cases within one hour (which is already far too long for operational use at Terra Bella), while the sequential scheduler converged to an optimal solution quickly (e.g., in less than 13 seconds for the biggest test case). In addition, the sum priority collected C and average wait time of data $\bar{D}$ that the sequential scheduler produces is always more desirable than that from the global scheduler.

The superiority in $\bar{D}$ of sequential over global is consistent but minor; for all of the test cases the difference is never more than two minutes in wait time, which does not have a large impact on image production. However, the superiority in C of sequential over global is striking. For the 3-satellite test case, the sum priority imagery scheduled by the sequential algorithm in less than 3 seconds runtime is over 18% more than the sum priority imagery scheduled by the global algorithm given one hour of runtime. The percentage improvements in C are even greater for the larger constellation test cases.

6 Conclusion

This paper has presented a scheduling algorithm to operate an Earth-imaging constellation with a maximum of automation and a minimum of intensive human oversight. Solving this challenge was motivated by the unique size and rapid growth of Terra Bella's satellite constellation. Given a varied set of hardware- and operations-derived objectives and constraints, two MILP approaches were described. One schedules downlink and imaging opportunities simultaneously, while the other schedules downlink and imaging separately and sequentially. This latter algorithm leverages a DP approach to single-satellite scheduling to provide a heuristic on cumulative image priority and amounts of imagery data collected. Use of this heuristic allows a reduced MILP to be solved that just schedules the downlink opportunities, enabling a massive computational speedup to achieve the performance required for effective operations.

References

Achterberg, T. 2009. Scip: Solving constraint integer programs. *Mathematical Programming Computation* 1(1):1–41.

Augenstein, S. 2014. Optimal scheduling of earth-imaging satellites with human collaboration via directed acyclic graphs. In *AAAI Spring Symposium on the Intersection of Robust Intelligence and Trust in Autonomous Systems*.

Barbulescu, L.; Watson, J.-P.; Whitley, L. D.; and Howe, A. E. 2004. Scheduling space–ground communications for the air force satellite control network. *Journal of Scheduling* 7(1):7–34.

Bensana, E.; Verfaillie, G.; Agnese, J.; Bataille, N.; and Blumstein, D. 1996. Exact & inexact methods for the daily management of an earth observation satellite. In *Space Mission Operations and Ground Data Systems-SpaceOps' 96*, volume 394, 507.

Bianchessi, N., and Righini, G. 2008. Planning and scheduling algorithms for the cosmo-skymed constellation. *Aerospace Science and Technology* 12(7):535–544.

Bianchessi, N.; Cordeau, J.-F.; Desrosiers, J.; Laporte, G.; and Raymond, V. 2007. A heuristic for the multi-satellite, multi-orbit and multi-user management of earth observation satellites. *European Journal of Operational Research* 177(2):750–762.

Bogosian, J. 2008. Image collection optimization in the design and operation of lightweight, low areal-density space telescopes. Master's thesis, Massachusetts Institute of Technology.

Burrowbridge, S. E. 1999. Optimal allocation of satellite network resources.

Gabrel, V., and Murat, C. 2003. Mathematical programming for earth observation satellite mission planning. *Operations research in space and air* 103.

Gabrel, V., and Vanderpooten, D. 2002. Enumeration and interactive selection of efficient paths in a multiple criteria graph for scheduling an earth observing satellite. *European Journal of Operational Research* 139(3):533–542.

Gabrel, V.; Moulet, A.; Murat, C.; and Paschos, V. 1997. A new single model and derived algorithms for the satellite shot planning problem using graph theory concepts. *Annals of Operations Research* 69:115–134.

Gabrel, V. 2006. Strengthened 0-1 linear formulation for the daily satellite mission planning. *Journal of combinatorial optimization* 11(3):341–346.

Globus, A.; Crawford, J.; Lohn, J.; and Pryor, A. 2004. A comparison of techniques for scheduling earth observing satellites. In *AAAI*, 836–843.

Karapetyan, D.; Mitrovic-Minic, S.; Malladi, K. T.; and Punnen, A. P. 2015. The satellite downlink scheduling problem: A case study of radarsat-2. In *Case Studies in Operations Research*. Springer. 497–516.

Leiserson, C. E.; Rivest, R. L.; Stein, C.; and Cormen, T. H. 2001. *Introduction to algorithms*. The MIT press.

Lemaitre, M.; Verfaillie, G.; Jouhaud, F.; Lachiver, J.; and Bataille, N. 2002. Selecting and scheduling observations of agile satellites. *Aerospace Science and Technology* 6(5):367.

Lin, W.; Liao, D.; Liu, C.; and Lee, Y. 2005. Daily imaging scheduling of an earth observation satellite. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 35(2):213–223.

Marinelli, F.; Nocella, S.; Rossi, F.; and Smriglio, S. 2011. A lagrangian heuristic for satellite range scheduling with resource constraints. *Computers & Operations Research* 38(11):1572–1583.

Nemhauser, G. L., and Wolsey, L. A. 1988. Integer and combinatorial optimization.

Shea, P., and Nasgovitz, N. 2011. Scheduling of a constellation of imaging satellites with usage constraints. In *Proceedings of SPIE*, volume 8044, 804404.

Tangpattanakul, P.; Jozefowiez, N.; and Lopez, P. 2015. A multi-objective local search heuristic for scheduling earth observations taken by an agile satellite. *European Journal of Operational Research* 245(2):542–554.

Vasquez, M., and Hao, J.-K. 2001. A "logic-constrained" knapsack formulation and a tabu algorithm for the daily photograph scheduling of an earth observation satellite. *Computational Optimization and Applications* 20(2):137–157.

Vasquez, M., and Hao, J.-K. 2003. Upper bounds for the spot 5 daily photograph scheduling problem. *Journal of Combinatorial Optimization* 7(1):87–103.

Wang, P.; Reinelt, G.; Gao, P.; and Tan, Y. 2011. A model, a heuristic and a decision support system to solve the scheduling problem of an earth observing satellite constellation. *Computers & Industrial Engineering* 61(2):322–335.

Yao, F.; Li, J.; Bai, B.; and He, R. 2010. Earth observation satellites scheduling based on decomposition optimization algorithm. *International Journal of Image, Graphics and Signal Processing* 1:10–18.

Zufferey, N.; Amstutz, P.; and Giaccari, P. 2008. Graph colouring approaches for a satellite range scheduling problem. *Journal of Scheduling* 11(4):263–277.