Power, Data Latency, and Radio Frequency Interference Issues in Mars Relay Network Scheduling[1][2][3]

Charles H. Lee[4] and Kar-Ming Cheung
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
818-393-0662
HungLee@fullerton.edu
Kar-Ming.Cheung@jpl.nasa.gov

Abstract— Mars will be continuously crowded this decade and beyond with many missions. Along with the current orbiters of Mars Global Surveyor and Odyssey 2001, future Mars missions within the next few years include Mars Exploration Rovers, Mars Express, and Mars Beagle in 2003, Mars Reconnaissance Orbiter in 2005, G. Marconi, Mars Netlanders, Mars Scouts in 2007, and Mars Smart Lander in 2009. At different time periods in the future, these missions are overlapped. Previous studies indicate that during such periods, existing deep space communication infrastructure cannot handle all Mars communication needs. There has been much coordination between various Mars projects and the Deep Space Network to ensure communication resources are effectively utilized so that valuable science and engineering data from Mars orbiters and landers can be accommodated. A plausible solution is to take into account the end-to-end communication performances of network along with operational constraints, and optimize the resource usage by scheduling communication at highest possible data throughputs. As a result, shorter communication time is required and more missions can be accommodated. This principle is demonstrated in this paper for a Mars relay communication network; a network consisting of multiple surface units and orbiters on Mars and the Deep Space Stations. A relay communication network around Mars can increase the overall science data return of the surface elements, reduce surface elements' direct-to-earth communication demands, and enable communication even when the surface elements are not facing Earth. It is the objective of this paper to take advantage of the relay operation to efficiently plan and schedule the network communications. Our previous results in relay network planning and scheduling include (i) modeling and simulating the overall end-to-end network link capabilities as time-varying resources by incorporating spacecraft dynamics, telecom configurations and other limiting factors such as planet occultation, weather, etc.; (ii) developing mathematical formulations for operational constraints such as daylight operations, one-to-one communication, time for acquisition and calibration, science data volume return requirement, onboard storage capacity, etc; (iii) formulating and solving the Mars relay network planning and scheduling as linear and nonlinear constrained optimization problem. In this paper, we address several issues that arise in Mars relay network operations. Major operation issues that we investigate in this paper include radio frequency interference, surface unit's battery limitations, and data return latency. Particularly, we (a) develop mathematical conditions, based on the geometry of the orbiters and surface units, to identify links with potential radio frequency interference and impose constraints on the links so that the optimal network scheduling is free from interference; (b) impose both the Sun angle constraint and the transmission duration constraint on the surface unit's battery; and (c) associate each orbiter with a latency function that allows the surface unit to judiciously select its orbiter to minimize the data latency. Numerical studies for a sample Mars relay network will also be presented.

Table of Contents

---

[2] Aerospace Conference 2003 Paper No. 215; Abstract submitted July 1, 2002. Version: Tuesday, October 22, 2002

[3] The work in this article was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

[4] Charles H. Lee is a JPL faculty-part-time researcher from the Department of Mathematics at California State University Fullerton

1. Introduction

The deployment of orbiting and landing assets infrastructure on Mars would broaden the opportunity to explore Mars by many folds. An example of such infrastructure is displayed in Figure 1. By taking advantage of the short-range UHF communications between the landing elements and the orbiting spacecrafts, end-to-end communications efficiency between Mars and Earth can be improved substantially. Particular advantages include more scientific activities, faster data return at larger throughputs and, more importantly, the capabilities to support future in-situ navigations and to enhance the success for Mars sample-return missions. Starting in late 2003 timeframe, although they are equipped with direct-to-Earth communication capability, the twin Mars Exploration Rovers (MER-A and MER-B) and the Beagle 2 will begin testing and sharing the UHF relay links offered by the orbital Mars Odyssey and Mars Global Surveyor (MGS) spacecrafts. The infrastructure at Mars will continue with a number of missions supporting relay link capability such as 2003 Mars Express, 2005 Mars Reconnaissance Orbiter, 2007 G. Marconi Orbiter and 2007 CNES Orbiter. Also we expect in the future, Mars relay communications will be more mature and reliable so that direct-to-Earth communications can be an option. As a result, landing assets can be miniaturized and more affordable. Mars relay network can also enable a new class of network science that involves multiple landing elements that coordinate and perform simultaneous measurements at several landing sites. Examples of such missions for this decade include the 2007 Mars Premier NetLanders and the 2007 Mars Scouts. Mars Premier will consist of four landers, which are expected to be in operation for one Martian year after their landing. The Mars Scout missions will include a series of landers, aerial gliders, and aerial balloons. Throughout their missions at Mars, communications for the landing assets will solely rely on the relay capability of the Mars orbital infrastructure.

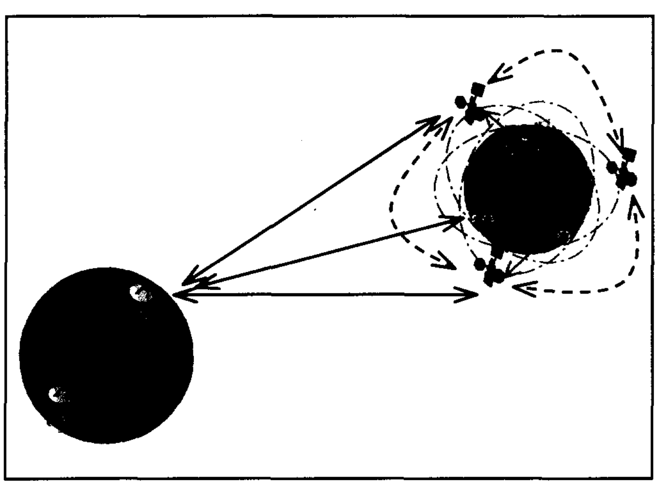

Figure 1 – A Relay Communication Link Network of Multiple Landing Assets, Orbiters and Ground Stations

There are many technical challenges associated with the Mars relay network. One area of interest is the effective planning and scheduling of the communication links between the orbiters, landing assets, and Earth communication stations. Unlike a point-to-point communication, a relay link consists of a concatenation of links, in which the data transfer mechanism can be either real-time or store-and-forward. In addition, the Mars infrastructure connectivity is highly dynamic, as the landing assets could be mobile or stationary and the orbiters could complete up to twelve orbits in a day. In our study, we have developed a high-fidelity modeling and simulation environment that predicts, for any time period of interest, the link capabilities between all possible links within the Mars network. Such information is much needed for the optimal planning and scheduling process. In addition to relay challenges are the mission requirements, link priorities, spacecraft activities, as well as deep space communications, navigation and system constraints. The final difficulty for a relay network is the optimal planning so that objectives such as minimal the total transmitting time and delaying time maximal network throughputs are met while the resulting schedule is conflict-free.

Our previous works in this area include modeling, simulating, and optimal scheduling for a network of multiple Low-Earth-Orbit satellites and the Earth communication stations (Cheung et. al. [Aerospace Conference 2002]). We also extended our results to a sample Mars relay communications network where several objectives and constraints are considered (Cheung et. al. [SpaceOps 2002 Conference]). In this paper, we address several additional issues that arise in Mars relay network operations. This includes landing asset's battery power limitations, radio frequency interference, and network latency. In the next two subsections we will address the radio frequency interference and data latency issues. In Section 3 we will describe the framework for a Mars relay communication network. Landing assets onboard power limitations will be discussed in Section 4. We will end the paper with simulation and optimization of a sample Mars relay communication network.

Radio Frequency Interference in a Mars Relay Network

Although the chance for radio frequency interference for the near-future Mars communication network is rare, its occurrence can pose great threats or jeopardize a mission in the real-time operation of a critical event. Problems associated with radio frequency interference range from losing lock to false locking, which in turn can cause severe signal degradation or communication blockage.

Let us first describe the scenarios where radio frequency inference is possible in a Mars communication network. For demonstrating purpose, we suppose that there are two pairs of communicating transmitters and receivers and each receiver is anticipating the signals from its corresponding transmitter. The following four conditions, when occurred concurrently, shall constitute a radio frequency interference phenomenon:

(i) One receiver is inside the communicating cone of another pair of transmitter-receiver,
(ii) Both pairs of receiver-transmitter are operating in overlapping frequency bandwidths (in-band or out-of-band),
(iii) The interfering power surpasses certain threshold
(iv) The interfering duration exceeds the requirement.

For a forward link, signals from an orbiter to a designated landing asset could turn into unwanted signals to another landing asset. As displayed in Figure 2, the bottom lander could suffer severe received SNR degradation or locking onto the wrong orbiter. Such events happen more frequently when the landing assets are nearby, the interfering orbiters are in high altitudes, and the orbiters' beamwidths are broad.

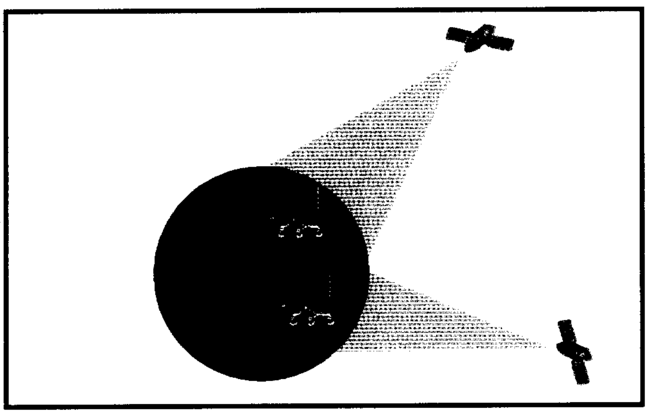

Figure 2 – Forward-Link Radio Frequency Interference

Similarly, the return relay telemetry from a landing asset to an intended orbiter could induce interference to another. As sketched in Figure 3, the interfered orbiter (right) drifted in two communicating cones. This is often the case, especially when the landing assets are equipped with an omni-directional relay antenna. Other contributing factors include nearby landing assets and high-altitude orbiters.

Our approach in eliminating RFI consists of two parts. We first identify, within the time period of interest, all possible RFI occurrences for the network. We then take into account the RFI constraints in the Mars network planning and scheduling process to resolve all potential radio frequency interferences.

Since communication between the orbiters and landing assets are at UHF bandwidths, RFI occurrence, based on our four conditions, will solely depend on the geometry of the landing assets and orbiters, landing assets' terrain mask angles, and the antenna beamwidths. In our study,

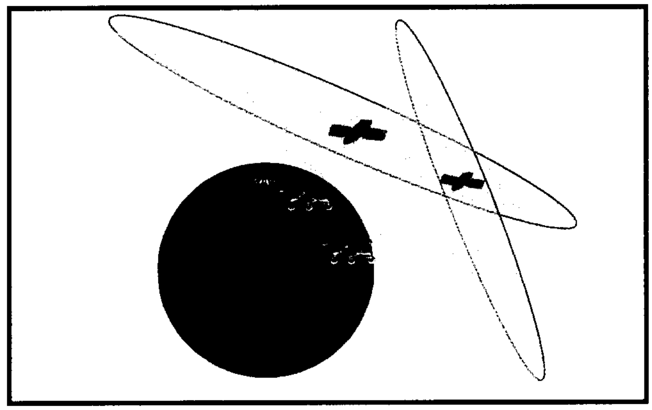

Figure 3 – Return-Link Radio Frequency Interference

we will follow the mathematical criteria derived in Lee et.al. [2001] to locate the RFI starting time, ending time, and the involving pairs of orbiters and landing assets.

It should also be noted that there exist techniques that are capable of suppressing RFI at the signal processing level (Tsou [1988]). However, due to their limiting size and weight, we will assume that the landing assets are not equipped with the RFI mitigation capability, and RFI has to be resolved at the planning level. Details of the planning process, which eradicate RFI in the Mars relay network, will be addressed in Section 3.

Data Latency in a Mars Relay Communication Network

Command files to a spacecraft are often small unless software uploading is needed; whereas telemetry data to Earth are large and are of higher demand. In-situ operations, and real-time observations occasions such as Mars seismic events, require immediate and prudent attentions. Direct to Earth transmission, if exists, of a large data file could, depending on the link's pipeline, be time consuming. With the virtual benefit of the network structure, such large data files at the source can broken up into small packages, which will be sent over the network's links using some optimal routing schemes. At the receiving end, the file is reconstructed from the packages. Such procedure is a common practice, and has been extensively used for the Internet and the File Transfer Protocol (FTP) applications. Our greatest concern in this subsection is to minimize time its takes to deliver a data file from one node of the network to another.

Two fundamental obstacles to minimizing the overall network data latency are connectivity of the network and the routing-and-packaging process. As discussed earlier, in a Mars relay communication network, link capability varies wildly with time and connectivity could at times be sporadic due to the complicated interaction between the spacecraft and the celestial bodies. This poses a challenge to network planning and scheduling. The following scenario is demonstrated to give the readers the latency

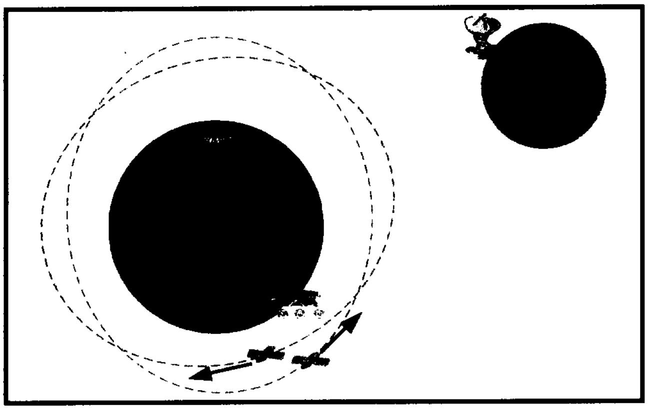

Figure 4 – The left orbiter is about to be occulted by Mars and thus relaying data via the right orbiter will reduce overall network latency

perspectives for a Mars relay communication network. Let us suppose a landing asset needs to transmit a large telemetry file to Earth in an urgent manner and it is about to lose its connection with Earth in 10 minutes. Its direct-to-Earth communication could be non-existed or its bandwidth and onboard power are so limited that the file cannot be transmitted completely in time. If the landing assets has the relay capability and it can locate several orbiters that are currently in-view. Note that although its contact time with the orbiters is relatively short compared to Earth, its data rates to the orbiters at UHF bandwidths are much higher than the direct-to-Earth links at S or X bands. In addition, when a landing asset is in view with multiple orbiters, it can choose to route telemetry data to orbiters whose contact with Earth will continue beyond the required relaying time. Routing the data to an orbiter that is about to go behind Mars can result in a 40 minutes or more latency (see Figure 4). In this paper we propose an approach to incorporate data latency requirement into our optimal planning and scheduling based on the connectivity and pipelines of the network. The resulting schedule is optimized in a sense that it meets system/navigation constraints, and mission requirements, communicates at highest possible data rates, and has abundant network connectivity between entities of the network. Furthermore it provides the blueprint layout that can be fed into existing routing schemes to achieve the overall network latency. The packaging-and-routing process could be found in computer literature and will not be addressed in this paper. The details of the proposed approach will be discussed in Section 3.

2. FRAMEWORK FOR A MARS RELAY NETWORK

In [6], we formulate the Mars relay network planning process into a constrained optimization problem. This paper follows a similar approach, and incorporates additional constraints based on RFI, data latency, and onboard power. To make this paper more self-contained, we restate some of the mathematical formulations and derivations.

The Mars relay communication network describes in this paper assumes a set of N orbiters $\{Orbiter_1, Orbiter_2, \ldots, Orbiter_N\}$, L landing assets $\{Lander_1, Lander_2, \ldots, Lander_L\}$, and M monitoring Earth ground stations $\{GS_1, GS_2, \ldots, GS_M\}$ (Figure 1). In this paper we addresses the return link only, since the forward link can be modeled similarly, and the vast volume of science data demands more effective usage of the network's return link capability. Though possible, landing-asset-to-landing-asset and orbiter-to-orbiter communications are not implemented in a foreseeable future. Thus communications between a landing asset and Earth can either be direct or relayed via an orbiter. It follows that our relay network scenario considered in this paper is at most one hop.

The Mars relay network we consider here is different from the wireless mobile communications in many significant ways. Our network is simpler with a much small number of communication nodes. Also communications is done based on a priori arrangements and the mobile units, in our case spacecraft and ground stations, have known positions and communications configurations. Many operational constraints are either known in advance, or can be derived from simulation of spacecraft and celestial dynamics. As a result, communications link performance (e.g. signal-to-noise ratios) and operation scenario (in-view and out-of-view periods between spacecraft and ground stations) for our network can be accurately modeled. Link performance between a transmitter and receiver pair is determined by the link capability, which is usually expressed in terms of supportable data rates and radiometric tracking performance. Link performance changes constantly, and the link itself may at time become unavailable due to spacecraft planned activities or celestial dynamic events. When the link is possible, link performance is affected by communication configurations and requirements such as antenna pattern, operating frequency, weather attenuation, masking angle, elevation angle, range, modulation indices, etc. A sample time-dependent link performance between a spacecraft and a ground station is displayed in Figure 5. Notice that when the transmitter and receiver are out of view, the supportable data rate is identically zero. An entire duration of a continuous link is considered a pass. Thus for any given time period of interest, for example a one-week timeframe, there are K possible number of passes $\{P_k \mid k = 1, 2, \ldots K\}$ between all communicating entity pairs within the network; Each pass P_k represents the communicating window between a pair of transmitter and receiver, which could be from a landing asset to an orbiter, or from a landing asset to a ground station, or from an orbiter to a ground station. Associated with the pass P_k are its starting time T_0^k, end time T_f^k, and the supportable data rate $R^k_{XMT,RCV}(t)$, which is valid only on

$[T_0^k \;\; T_f^k]$. If communication is scheduled for the pass P_k, the actual transmission starting (on) time and end (off) time are denoted by t_0^k and t_f^k respectively. Our goal is to determine the optimal pair of start and end time $\{[t_0^k, \; t_f^k] \mid T_0^k \le t_0^k \le t_f^k \le T_f^k\}$ for each pass P_k so that the total transmitting and delaying time are minimized while the network throughputs are maximized and other navigation constraints and mission requirements including priority are not violated.

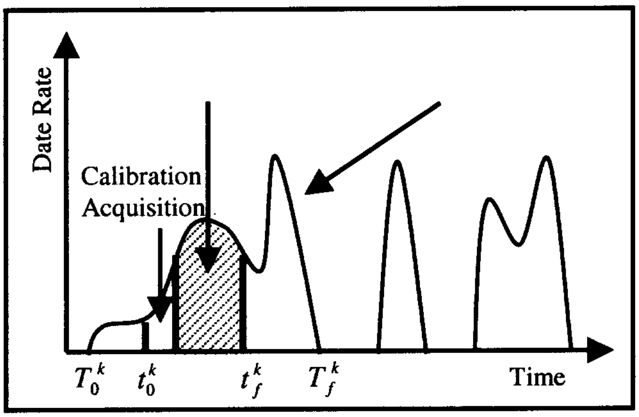

Figure 5 – Supportable Data Rate Sample During a Pass

Our optimal planning and scheduling will be formulated as a nonlinear constrained optimization problem. Let us start with the discussion of the mission objectives, operational constraints and mission requirements for the Mars Relay network assumed in this paper. We remark here that the size of our constrained optimization problem depends on the number of passes, which is dictated by the number of landers, orbiters, and as well as the time span. Our primary efforts are to eliminate mediocre passes and refine the qualified ones. To constitute a pass, the following constraints must be met:

(a) **Sun Angle Requirement** Due to the limited onboard power constraint, landing asset's solar panels might be required to be in direct Sun's illumination in order to transmit or receive any data. Thus the passes that do not entirely satisfy the Sun elevation angle constraint

$$\theta_{Lander\,l}^{Sun}(t) \le \Theta_{Lander\,l}^{Sun}, \quad t \in [T_0^k, \;\; T_f^k], 1 \le l \le L, \qquad (1)$$

are shortened, if not removed. Such requirement will be imposed throughout this decade for almost all Mars missions, except the Mars Smart Lander mission. For such a nuclear-powered landing asset, the Sun's elevation angles constraint in (1) can still be applied by setting $\Theta_{Lander\,l}^{Sun}$ to $-\pi/2$.

(b) **Exceeding Performance Passes** The data rate capability for each pass must exceed a certain performance threshold:

$$R_{XMT,RCV}^{k}(t) \ge R_{threshold}, \quad t \in [T_0^k, \;\; T_f^k]. \qquad (2)$$

In the event that the entire pass may not meet this requirement, only the portion that the supportable data rate is above the threshold is considered. This constraint applies only to S or X band links and not the UHF proximity links. In addition, this constraint can be replaced by the elevation angle requirement; That is, the constraint (2) can be imposed to require communications to occur only at sufficiently high elevations angles. Thus a pass might be eliminated, shortened, or broken into shorter passes. For high-frequency bandwidth such as the Ka-band, communication performance could be volatile due to rapid weather changing. In such a scenario, larger number of short passes will result. The next condition will further eliminate all the brisk passes.

(c) **Sufficiently Long Passes** A pass is valid if its entire duration from start to finish must last longer than some minimal required time T_{min}; otherwise it is not worth considering,

$$T_f^k - T_0^k \ge T_{min} \quad for \quad k = 1,2,\ldots K. \qquad (3)$$

(d) **Minimizing Network Latency** Next, we address the issues in network scheduling to minimize data latency due to relaying connectivity. Since we assume that there is at most one-hop in the relay framework, return link relaying communications with urgency can suffer significant latency if the landing asset is not in view with any orbiter or is in view with an orbiter that is about to lose contact with Earth. For such scenarios, network connectivity is the problem and long delay is inevitable. However, if the landing asset can be in view with a few orbiters, which remain in contact with Earth well beyond the completion of the passes, then data latency can be improved significantly. The constraint to minimize relay latency can be described as follows. Suppose P_k is a pass of interest, which relays data from the landing asset $Lander_l$ and the orbiter $Orbiter_n$. If such pass is scheduled, then network latency depends solely on the connectivity of the orbiter $Orbiter_n$ with Earth. Several scenarios for the connectivity of the orbiter to Earth can occur. First, consider the case when the orbiter is in direct contact with Earth. If the remain of the orbiter's connectivity with Earth is sufficiently long to relay the data, then no latency due to connectivity is resulted. However, if the orbiter is going to lose contact with Earth soon or is not connecting with Earth, then latency definitely results and the delay time is at least the time it takes for the orbiter $Orbiter_n$ to reconnect communications with

Earth. Thus each relay pass P_k from a landing asset to an orbiter is assigned with a look-at-head latency value Lat^k. If a landing asset has multiple competing relay passes $\{P_k\}$ that are overlapped, then lower priority will be given will be given to the passes with higher Lat^k values:

$$Priority\_Latency^k = \exp(-Lat^k). \qquad (4)$$

The above latency priority guarantees that communications in the network will suffer the least latency due to network connectivity.

Our previous conditions are imposed to screen out unqualified passes; an essential reduction for the dimensions of the search space. Our next efforts are to impose constraints on the scheduling time:

(e) **Scheduled Starting and Ending Time Constraints** Communication cannot start ahead of its pass,

$$t_0^k \geq T_0^k \quad for \quad k = 1,2,\ldots K. \qquad (5)$$

Communication end time must be smaller than the time allowed within a pass,

$$T_f^k \geq t_f^k \quad for \quad k = 1,2,\ldots K. \qquad (6)$$

Communication start time should not only exceed its end time,

$$t_f^k \geq t_0^k \quad for \quad k = 1,2,\ldots K. \qquad (7)$$

(f) **Lander's Power Constraint** For small landers such as those of the 2007 Mars Premier NetLanders, communications are restricted solely to the orbiters and power consumption is also limited. If a pass is utilized $(t_f^k > t_0^k)$, the total communication time should not exceed the maximal allowable communicating time,

$$t_f^k - t_0^k \leq T_{\max}^k \quad for \quad k = 1,2,\ldots K. \qquad (8)$$

(g) **Calibration/Acquisition/Tracking Time** If a pass is utilized (i.e. $t_f^k > t_0^k$), the total communication time should last not only beyond the required overhead time τ_{acq}^k for calibration/acquisition, but should also last longer than the minimal tracking time τ_{track}^k. Thus if the pass is used it must be long enough to worth the efforts. That is,

$$t_f^k - t_0^k \geq T_{\min}^k \quad for \quad k = 1,2,\ldots K, \qquad (9)$$

where $T_{\min}^k$ is the sum of τ_{acq}^k and τ_{track}^k.

(h) **One-to-one Communications** Payload and power are the major constraints which landing assets must struggle with. Thus in our relay infrastructure we will assume that the landing assets can only communicate with either Earth or an orbiter at any one time. There are occasions when a landing asset can link to Earth as well as multiple orbiters. Several overlapping passes will result and are resolved by requiring that the communicating time for overlapping passes be disjoint. That is, if the passes P_{k_1} and P_{k_2} originate from a same landing asset and

$$[T_0^{k_1}, T_f^{k_1}] \cap [T_0^{k_2}, T_f^{k_2}] \neq 0, \text{ then} \qquad (10a)$$

$$[t_0^{k_1}, t_f^{k_1}] \cap [t_0^{k_2}, t_f^{k_2}] = 0. \qquad (10b)$$

(i) **Radio Frequency Interference Avoidance** The following constraint is added to ensure that the network communication scheduling is RFI-free. Suppose P_{k_1} and P_{k_2} are the passes corresponding to two pairs of transmitters and receivers. Potential RFI can be identified using the geometric information of the transmitters and receivers (Lee [MILCOM 2000]). Particular, we verify geometrically whether the receiver is in the communicating cone of the transmitter. If RFI exists, let us denote by $[t_0^{RFI}, t_f^{RFI}]$ the corresponding time duration. In this case, the communicating time scheduled for the passes must be disjoint. That is,

$$[t_0^{k_1}, t_f^{k_1}] \cap [t_0^{k_2}, t_f^{k_2}] \cap [t_0^{RFI}, t_f^{RFI}] = 0 \qquad (11)$$

(j) **Data Volume Requirement** If communication is scheduled during the pass P_k, the data volume transmitted during such period is

$$DV_k = \int_{t_0^k + \tau_{acq}^k}^{t_f^k} R_{XMT,RCV}^K(t)dt, \qquad (12a)$$

where DV_k is the delivered data volume during the pass P_k and τ_{acq}^k is the overhead time required for calibration and acquisition. Note that due to constraint (7), the delivered data volume DV_k is always nonnegative. Also, multiple Earth ground stations may be able to receive data from a spacecraft at Mars; however, the delivered data volume can only be accounted by one station. Our next constraint can be viewed as an objective. That is, for any period of interest, each spacecraft must transmit at least the required amount of data volume. Let RDV_i be the minimally required data volume for the spacecraft SC_i (a landing asset or an orbiter) during the considered planning period. The constraint can be expressed as,

$$\sum_{\substack{k=1 \\ P_k \text{ involves } SC_i}}^{K} DV_k \geq RDV_i. \tag{12b}$$

(k) **Orbiter's Storage Capacity Constraint** The next requirement is taken into consideration so that every landing asset in our network receives the open line of communication necessary to relay its collected data. This constraint mirrors the technical limitation of orbiter's storage devices and is described by

$$S_{Orbiter_n}([\,t_o^1 \quad t_f^1 \quad \cdots \quad t_o^K \quad t_f^K\,], t) \leq C_{Orbiter_n}, \tag{13}$$

where $C_{Orbiter_n}$ is the storage capacity onboard the orbiter $Orbiter_n$ and $S_{Orbiter_n}([\,t_o^1 \quad t_f^1 \quad \cdots \quad t_o^K \quad t_f^K\,], t)$ is the data storage for $Orbiter_n$ at time t according the schedule $[\,t_o^1 \quad t_f^1 \quad \cdots \quad t_o^K \quad t_f^K\,]$. Note that $S_{Orbiter_n}([\,t_o^1 \quad t_f^1 \quad \cdots \quad t_o^K \quad t_f^K\,], t)$ is calculated by back-tracking the amount of data being stored (including its own mapping data and relaying data the from landing assets) and forwarded (downlink to Earth).

(l) **Mission Priority** To address the priority issues, we assign to each pass P_k with a scoring number $\sigma_k \geq 1$, which helps to rank the importance of the passes. In our study, when priority is not an issue, we set $\sigma_k = 1$. If the mission for a spacecraft is more critical, the score for such pass is higher. Note also that we allow priority scores to vary from pass to pass. These priority scores will guide the optimal solution toward meeting our objectives while paying greater attention to higher priority passes. The priority scores are incorporated into the next two objective functions.

(m) **Maximizing Data Return** If telemetry data volume from a landing asset or an orbiter is of greater importance, then the objective function which we minimize is of the form

$$J_{DV}(t_o^1, t_f^1, \ldots, t_o^K, t_f^K) = -\sum_{k=1}^{K} \sigma_k^{-1} \cdot DV_k, \tag{14}$$

where σ_k's are the priority scores.

(n) **Minimizing Total Transmitting Time** As discussed earlier, landing assets may have limited power supply and thus communication should be scheduled efficiently so that the total communicating time to be as small as possible and more missions can be supported with the same resources. The corresponding minimizing criterion is

$$J_{TIME}(t_o^1, t_f^1, \ldots, t_o^K, t_f^K) = \sum_{k=1}^{K} \sigma_k^{-1} \cdot (t_f^k - t_0^k). \tag{15}$$

In summary, the above constraints and objectives lay the foundation for our scheduling optimization problem. In the next section, we will cast our problem into a multi-objective nonlinear optimization problem.

4. Mathematical Modeling & Formulation for Operational and Communication Constraints

Let us next discuss all of the constraints and their corresponding mathematical formulations. First notice that the potential passes from our Mars relay network are screened a priori so that the considered passes must meet the Sun angles requirement, ensure network connectivity to avoid major latency and must be long and strong enough, i.e. the finalized passes $\{P_k \mid k = 1,2,\ldots K\}$ must satisfy the conditions in (1)-(4). Then the task of optimizing the relay communication network would require the decision of the starting and ending time pairs so that the objectives and constraints in (5)-(15) are satisfied. Equivalently, we seek for an optimal solution

$$\vec{X} = \begin{bmatrix} t_o^1 \\ t_f^1 \\ \vdots \\ t_o^K \\ t_f^K \end{bmatrix}, \tag{16}$$

that optimizes the objective functions and fulfills the constraints (5)-(15).

Consequently, conditions (5)-(8) can be translated into the following linear constraints

$$A\vec{X} \leq \vec{B}, \tag{17}$$

$$\vec{L}_B \leq \vec{X} \leq \vec{U}_B, \tag{18}$$

where

$$A = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}_{K \times 2K}; \quad \vec{B} = \begin{bmatrix} T_{\max}^1 \\ \vdots \\ T_{\max}^K \end{bmatrix}; \tag{19}$$

$$\vec{L}_B = \begin{bmatrix} T_o^1 \\ T_o^1 \\ \vdots \\ T_o^K \\ T_o^K \end{bmatrix}; \quad \vec{U}_B = \begin{bmatrix} T_f^1 \\ T_f^1 \\ \vdots \\ T_f^K \\ T_f^K \end{bmatrix}. \tag{20}$$

Let us start with the first nonlinear constraint by considering the scenario where two passes are overlapped. Let there be J incidents of overlapping passes among the K possible passes that we consider. We want our communicating time to be disjoined, namely satisfying constraint (10b). Particularly we assume that at any time if a landing asset possesses two overlapping passes, the communication time, if scheduled, should be done in the vicinity of the peak of the supportable data rate within the pass. That is, let us assume that the passes P_{k_1} and P_{k_2} are overlapped. Then any overlapping scheduling will register a violation to constraint (10b). The severity of the violation is determined by the measure of $[t_0^{k_1}, t_f^{k_1}] \cap [t_0^{k_2}, t_f^{k_2}]$ which we seek to eliminate by adopting the following nonlinear constraint

$$\vec{G}_1(\vec{X}) \leq \vec{0}, \quad (21a)$$

where

$$G_1^j(\vec{X}) = (t_f^{k_1} > t_0^{k_2})(t_f^{k_2} > t_0^{k_1})[\min(t_f^{k_1}, t_f^{k_2}) - \max(t_0^{k_1}, t_0^{k_2})], \quad (21b)$$

for $j = 1, 2, \ldots, J$. Notice that the expressions $(t_f^{k_1} > t_0^{k_2})$ and $(t_f^{k_2} > t_0^{k_1})$ return the Boolean values.

The minimal calibration, acquisition and tracking time constraint (9) can be expressed as follow,

$$\vec{G}_2(\vec{X}) \leq \vec{0}, \quad (22a)$$

where

$$G_2^k(\vec{X}) = (t_f^k > t_0^k)[T_{\min}^k - (t_f^k - t_0^k)], \quad (22b)$$

for $k = 1, 2, \ldots, K$.

The nonlinear constraint for the radio frequency interference condition (11) can also be handled similarly. That is, suppose there are I number of occasions where the passes are affected by RFI, then the constraint can be formulated by

$$\vec{G}_3(\vec{X}) \leq \vec{0}, \quad (23a)$$

where its component $G_3^i(\vec{X})$ is defined as

$$(t_f^{k_1} > t_0^{k_2})(t_f^{k_2} > t_0^{k_1})[\min(t_f^{k_1}, t_f^{k_2}, RFI_0^i) - \max(t_0^{k_1}, t_0^{k_2}, RFI_0^i)], \quad (23b)$$

for $i = 1, 2, \ldots, I$.

The data volume requirement constraint (12b) for each orbiter or landing asset can also be employed as a nonlinear constraint

$$\vec{G}_4(\vec{X}) \leq \vec{0}, \quad (24a)$$

where its components are

$$G_4^m(\vec{X}) = \sum_{\substack{k=1 \\ P_k \text{ involves } SC_m}}^{K} DV_k - RDV_m \quad m = 1, \ldots, L+N. \quad (24b)$$

Finally, the store-and-forward capacity constraint (13) is imposed by the nonlinear constraint

$$\vec{G}_5(\vec{X}) \leq \vec{0}, \quad (25a)$$

where

$$G_5^n(\vec{X}) = \int_{\substack{\text{scheduling} \\ \text{duration}}} \max\left\{ S_{Orbiter_n}(\vec{X}, t) - C_{Orbiter_n}, 0 \right\} dt, \quad (25b)$$

for $n = 1, \ldots, N$.

The problem (16)-(25) we are trying to solve belongs to a class of nonlinear constrained optimization [1]-[2], which can be set up and solved differently. Two particular constrained optimization approaches are considered. In the first approach, the optimization is performed by minimizing the cost functional

$$J(\vec{X}) = w\, J_{DV}(\vec{X}) + (1-w)\, J_{TIME}(\vec{X}), \quad (26)$$

for some weighing factor $\omega \in [0,1]$ subject to the linear constraint $A\vec{X} \leq \vec{B}$, the upper and lower bounds for the solution space $\vec{L}_B \leq \vec{X} \leq \vec{U}_B$, and a combined nonlinear constraint

$$\vec{C}(\vec{X}) = [\vec{G}_1(\vec{X}); \vec{G}_2(\vec{X}); \vec{G}_3(\vec{X}); \vec{G}_4(\vec{X}); \vec{G}_5(\vec{X})]^T \leq \vec{0}. \quad (27)$$

Our problem can also be transformed into the goal attainment optimization problem, where the scheduling is performed to satisfy the goals (or constraints). In which case, the problem can be expressed as nonlinear programming problem:

$$\min_{\vec{X}, \lambda} \{\lambda\} \quad \text{subject to} \quad (28)$$

$$\vec{G}(\vec{X}) - \lambda \vec{\omega} \leq \vec{0} \quad (29)$$

where,

$$\vec{G}(\vec{X}) = \begin{bmatrix} A\vec{X} - \vec{B} \\ \vec{L}_B - \vec{X} \\ \vec{X} - \vec{U}_B \\ \vec{G}_1(\vec{X}) \\ \vec{G}_2(\vec{X}) \\ \vec{G}_3(\vec{X}) \\ \vec{G}_4(\vec{X}) \\ \vec{G}_5(\vec{X}) \\ J_{DV}(\vec{X}) - DV_{goal} \\ J_{TIME}(\vec{X}) - XT_{goal} \end{bmatrix}; \quad (30)$$

is the goal function with the data volume goal DV_{goal} and the minimal total transmitting time XT_{goal} and $\vec{\omega} = [\omega_1 \ \cdots \ \omega_{10}]^T$ is some assigned weight vector.

Many commercial off-the-shelf software tools are capable of solving both problems. We particularly employ the MATLAB's FMINCON and FGOALATTAIN subroutines to solve the nonlinear constrained optimization (26)-(27) and the nonlinear programming problem (28)-(30), respectively. Numerical implementation for optimal scheduling a sample Mars relay communication network is presented in the next section.

4. MARS RELAY NETWORK SIMULATION & OPTIMAL SCHEDULING

For proof-of-concept purposes, we consider a relay communication network consisting of four Mars landing assets, five Mars orbiting spacecrafts and three monitoring **DSN stations - at Canberra, Goldstone, and Madrid.** Problems of different sizes can be investigated similarly. We assume that the mask angles for the ground stations on Earth are 15 degrees and the ground stations have the *multiple spacecraft per aperture* capability or have sufficient antennas to accommodate all spacecrafts of the network. We assume that the landing assets remain stationary throughout the planning and scheduling period although mobile landing assets can be used if their time-dependent locations on Mars are known. The locations for the considered landing assets on Mars at J2000 are described in Table 1.

Lander	Longitude	Latitude	Altitude	Mask Sun
[redacted]				

Table 1: Location of the four Mars landing assets at J2000

The five Mars orbiting spacecraft are chosen to be of low altitude and some of them are in sun-sync orbits. Their six orbital elements are given in Table 2.

Semi-major axis	Eccentricity	Inclin. Angle	Asc. Node	Arg. of Perigee	Time at Perigee
4384 km	0.010814	10	119	153	0
3972 km	0.011474	-11	97	144	0
4552 km	0.009869	-89	292	317	0
4408 km	0.012156	-53	-336	251	0
3890 km	0.056136	26	46	110	0

Table 2: Six orbital elements the five Mars orbiter at J2000

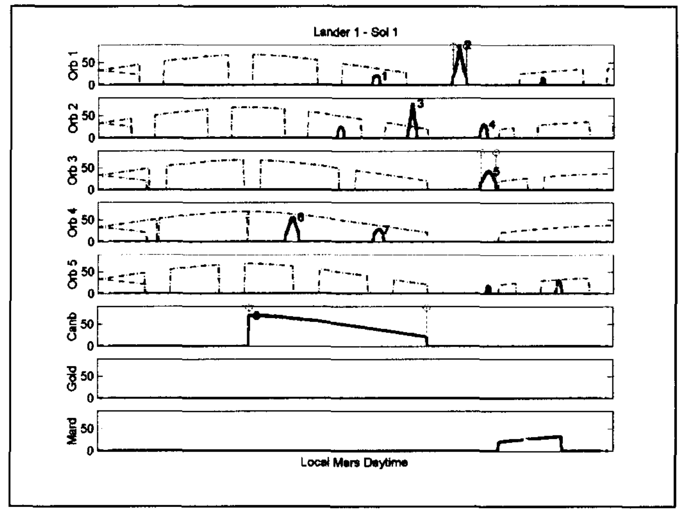

Figure 6. Considered passes from Lander 1 to the five orbiters and the DSN stations are highlighted and numbered.

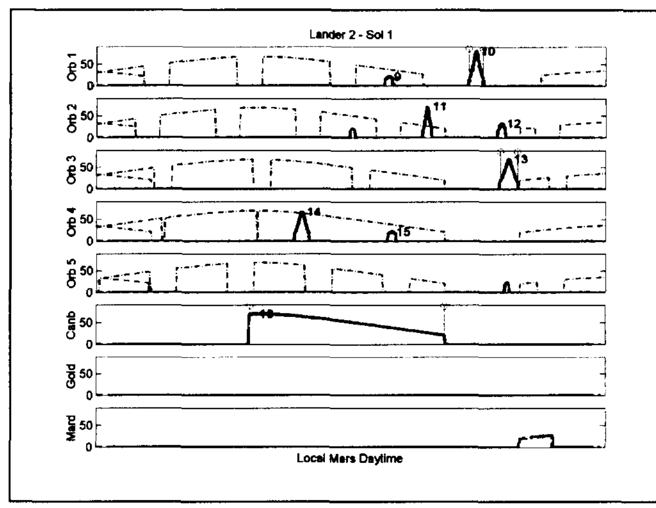

Figure 7. Considered passes from Lander 2 to the five orbiters and the DSN stations are highlighted and numbered.

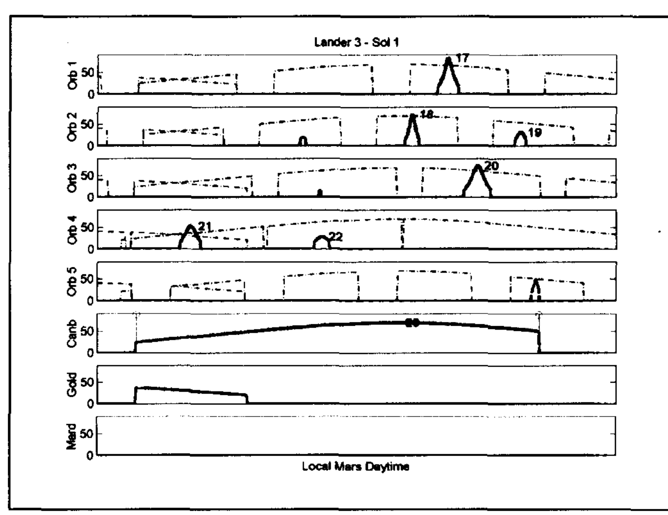

Figure 8. Considered passes from Lander 3 to the five orbiters and the DSN stations are highlighted and numbered.

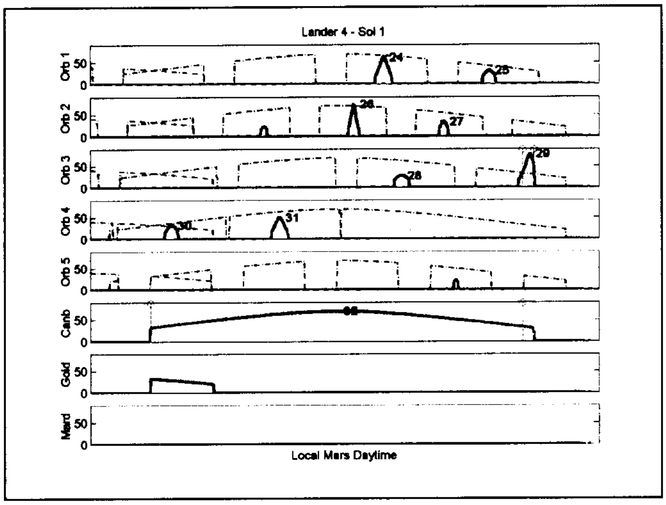

Figure 9. Considered passes from Lander 4 to the five orbiters and the DSN stations are highlighted and numbered.

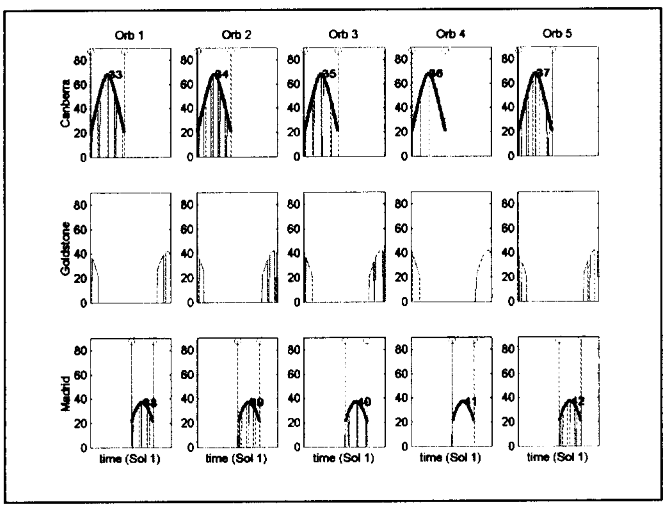

Figure 10. Considered passes from the five orbiters to the DSN stations.

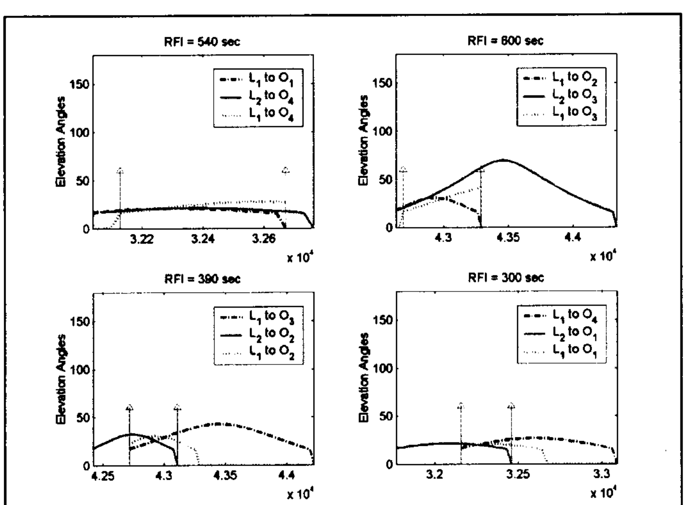

Figure 11. For the considered Mars communication network, four RFI occurrences were identified.

The orbiters move in three-dimensional elliptical orbits. Their states are simulated using the second-order Keplerian framework [3]-[4]. The landing assets are situated on Mars and their positions are determined with respect to the tilted rotating body of Mars. The time period of consideration is one Earth day after J2000. Based on the ephemeris of the landing assets and orbiters, and the geometry of the Sun, Mars, and the DSN stations, the time-dependent connectivity of the network is established and based on which the network pipeline measured in terms of supportable data rate between the entities of the network. To demonstrate the relay planning and scheduling process without going into great details of the framework with the link analysis models, we represent, to a first order assumption, the supportable data rate with the corresponding elevation angles of the pass. Displayed in Figures 6-10 are all possible passes for the relay network, which are then screened using conditions (1)-(3). Particularly, the local mask angles and the minimal Sun angle for the landing assets are assumed to be at 15 and 20 degrees, respectively. The minimal times for the passes are required as 10 minutes for a landing asset to an orbiter, four hours for landing asset to Earth, and five hours for orbiter to Earth. The numeric times and angles are chosen for flexibility demonstration purposes, these values can be modified according to mission requirements.

Figures 6-9 exhibit and number all the qualified passes from each landing asset during its local Mars daylight. The landing assets' Sun elevation angles are displayed in the background. To effectively identify the network latency due to connectivity, we also plot the corresponding orbiter's contact with Earth. For instance, in Figure 9, for Pass No. 24 from Landing Asset 4 to Orbiter 1, we can see the in view plot between the orbiter and Earth as it goes around Mars. Of the possible 42 qualified passes during one Earth day, eight have potential latency due to network connectivity. These passes were chosen because the corresponding relay orbiter is not in view with Earth or cannot completely transmit the relay data during its contact with Earth. These passes are displayed in Table 3.

Passes with Relay Latency (Land. to Orb.)	**Corresponding Relay Passes (Orb. To Earth)**	**Orbiter's Next Reconnection Time (sec)**
2	38	5310
3	39	8220
4	39	1230
5	40	300
10	38	5520
11	39	8400
12	39	1410
13	40	180

Table 3: Eight identified passes with potential latency.

Our approach to minimize such latency is demonstrated as follows. Since Pass 2 does not coincide with any other passes (see Figure 6), and if utilized, its latency is definite and unavoidable; whereas, Pass 3 and Pass 8 are overlapped and Pass 8 is direct to Earth and is thus has priority over Pass 3. Other passes with latency are treated in a similar manner.

For our considered Mars relay network, there are four occasions in which backward relay radio frequency interference occurs. The scenarios are captured in Figure 11. In one occasion, while Lander 1 is communicating with Orbiter 2 during Pass No. 4 and while Lander 2 is communicating with Orbiter 3 during Pass No. 13, Lander 1 is also in contact with Orbiter 3 for 600 seconds. We result the RFI issues by imposing constraint (23), i.e., Pass No. 4 and Pass No. 13 should not be scheduled simultaneously.

Let us next discuss the dimensions for the solution space and the constraints. Since there are 42 qualified passes, our solution space is of 84 dimensions. The number of linear constraints $Ax \leq B$ in (19) is fixed at 42. There are 26 constraints for overlapping passes from the landing assets; 42 calibration/acquisition requirements, 4 RFI avoidance constraints, 9 data volume requirements, and 5 onboard memory constraints. The resulting number of nonlinear constraints is 86.

To achieve the optimal planning and scheduling for the considered Mars relay communication network, we maximize the total data return while satisfy the temporal and power constraints (17)-(20), one-to-one communication constraint (21), the calibration/acquisition constraints (22), the radio frequency interference constraint (23), and the onboard data storage constraint (24). The optimal results are displayed in Figures 6-10 with the starting and ending tick marks, when a pass is utilized. Remarkable achievements can be summarized as:

1. The optimal solution yields the largest total transmitting data volume from the landing assets. For example, the optimization process chooses to end the direct to Earth Pass No. 32 prematurely to accommodate Pass No. 29 to the Orbiter 3, which, in return, results in more transmitting data volume.

2. The optimal solution falls within the allowable time during the pass. It can be verified graphically in Figures 6-10.

3. Scheduled passes from the landing assets never coincide.

4. Calibration/acquisition constraints are met.

5. Network latency due to connectivity is minimal. For instance Pass No. 11 can yield better data return, but was not utilized due to the latency minimization requirement.

6. The onboard data storage requirements for the orbiters are satisfied.

5. CONCLUSIONS

In this paper, we address major issues that a Mars relay communication network faces. Particularly, we address, propose solution, and resolve the power, latency, and radio frequency interference issues along with many others such as the communication constraints, mission requirements, temporal constraints. Our approach takes into consideration the end-to-end aspects of the space operations, it is capable of improving network productivity, efficiency and reducing operational costs. Important factors varying from the dynamics of the spacecrafts, planet occultation, locations of ground stations, to actual telecom configurations, including antenna pattern, operating frequency, weather forecast, etc. to navigation requirements such as spacecraft's planned activities and priorities are incorporated in our model. The result is an evolving network with dynamic communication links connecting between transmitters and receivers. We put great efforts to translate various communication-specific physical and operational requirements into mathematical constraints. We also convert our network objectives into cost functions. Such problem formulation approach is of great importance because the optimal results are obtained deterministically and mathematically. Above all, the resulting optimal scheduling allows the network to communicate its maximally possible performance. A sample Mars communications relay network system consisting of four Mars landing assets, five Mars orbiters, and the DSN stations at Canberra, Goldstone, and Madrid was simulated and optimized. Our result indicates significant promises as it satisfies the operational constraints while achieving remarkable communication efficiency.

ACKNOWLEDGEMENTS

This research was carried out jointly by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, and by the California State University Fullerton (CSUF), with funding from the CSUF Faculty Development Center Grant and Funds.

REFERENCES

[1] Wayne L. Winston, *Operations Research – Applications and Algorithm*, Belmont: Duxbury Press, 1994.

[2] Kaisa Miettinen, *Nonlinear Multiple Objective Optimization*, Boston: Kluwer Academic, 1999.

[3] Pedro Ramon Escobal, *Methods of Orbit Determination*, Florida: Krieger, 1976.

[4] Tom Logsdon, *Orbital Mechanics – Theory and Applications*, New York: John Wiley & Sons, 1998.

[5] K.M. Cheung, C.H. Lee, W.B. Gearhart, T. Vo, and, S. Sindi, "*Link-Capability Driven Network Planning and Operation*" IEEE Aerospace Conference – Advanced Concepts for Mission Operations, Big Sky, Montana (March, 2002)

[6] K.M. Cheung and C.H. Lee, "*Design and Architecture of the Mars Relay Network Planning and Analysis Framework*" World Space Congress – 2002 Space Ops, Houston, Texas (October, 2002)

BIOGRAPHIES

Dr. Charles H. Lee *is a professor of mathematics at the California State University Fullerton (CSUF) and a faculty part time staff in the Communications Systems Research Section (331) at the Jet Propulsion Laboratory. Before becoming a faculty member, he spent three years as a Post-Doctorate fellow at the Center for Research in Scientific Computation, Raleigh, North Carolina, where he was the recipient of the 1997-1999 National Science Foundation Industrial Post-Doctorate Fellowship. His research has been Computational Applied Mathematics with emphases in Control, Fluid Dynamics, Smart Material Structures and Telecommunications. He received his Bachelor of Science in 1990, Master of Science in 1992, and Doctor of Philosophy in Applied Mathematics in 1996 all from the University of California at Irvine.*

Dr. Kar-Ming Cheung *is a Technical Group Supervisor in the Communications Systems Research Section (331) at JPL. His group provides telecom analysis support for JPL missions, and develops the operational telecom analysis and predict generation tools for current and future JPL missions and the DSN. He received NASA's Exceptional Service Medal for his work on Galileo's onboard image compression scheme. He was the Manager of the Network Signal Processing Work Area of NASA's Deep Space Network (DSN) Technology Program. He has authored or co-authored 6 journal papers and over 20 conference papers in the areas of error-correction coding, data compression, image processing, and telecom system operations. Since 1987 he has been with JPL where he is involved in research, development, production, operation, and management of advanced channel coding, source coding, synchronization, image restoration, and link analysis schemes. He got his B.S.E.E. degree from the University of Michigan, Ann Arbor in 1984, his M.S. degree and Ph.D. degree from California Institute of Technology in 1985 and 1987 respectively.*