Energy Efficient Transmission Scheduling over Mars Proximity Links

Alessandro Tarello*, Jay Gao† and Eytan Modiano‡
* *Telecommunication Networks Group*
Electronics Engineeering Department
Politecnico di Torino
c.so Duca degli Abruzzi, 24
10129, Torino, Italy
+39-011-564-4194
tarello@mail.tlc.polito.it
† *Communication Network Group*
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91109-8099, USA
818-354-9528
jay.l.gao@jpl.nasa.gov
‡ *Laboratory for Information and Decision Systems*
Massachusetts Institute of Technology
77 Massachusetts ave.
Cambridge, MA 02139, USA
617-452-3414
modiano@mit.edu

Abstract— **We consider the scheduling of transmissions from an energy limited Mars lander to a Mars orbiter. Typically, the channel quality of the Mars proximity links is time-varying due to orbital dynamics, multi-path effects and the antenna positioning on the lander. Since the channel state determines the throughput obtained per unit of energy transmitted, it is desirable to select when, and at what data rate, to transmit based on channel conditions. In this paper we consider the dual problem of maximizing the throughput of a lander that has a limited amount of energy to be used for transmission; and minimizing the energy consumption used for transmission of data subject to delay constraints. In [1], [3] it was shown, using techniques from Dynamic Programming, that energy efficiency can be significantly improved through an adaptive rate and power allocation scheme. We apply the policies developed in [1] and [3] to the Mars proximity link and compare them with scheduling algorithms presently in use for proximity links. Our simulations, using channel measurement data obtained from NASA's 2001 Mars Odyssey orbiter and the Mars Exploration Rover (MER) project, show over an order of magnitude increase in throughput and decrease in energy consumption through the use of the Dynamic Programming based rate and power adaptation scheme.**

[1]0-7803-9546-8/06/$20.00(c) 2006 IEEE
[2]IEEEAC paper n.1171, Version 4, Updated December, 5 2005
[3]This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration, and at Massachusetts Institute of Technology under NASA grant number NAG3-2835. Alessandro Tarello was also supported by Politecnico di Torino, Torino, Italy.

CONTENTS

I. Introduction

A key aspect of the Mars Network relay operation is the telecommunications planning process. Detail models of antenna gain patterns and spacecraft ephemeredes are used by mission operators to predict the proximity link quality and the proper data rate for each pass between a Mars lander (or rover) and an orbiter. The selected data rate profile is then uploaded to the remote spacecraft days or sometimes weeks prior to the beginning of each pass. Such "earth-in-the-loop" operation is inherently labor intensive and incapable of adapting to real-time stochastic fluctuations in the channel and subtle geometric complexity in the RF signal propagation path. Performance optimization will be particularly challenging for communications between an energy-starved lander and science satellite whose orbit and capabilities were optimized for science objective instead of relay communications. Therefore, it is crucial to develop alternative approaches for transmission scheduling over the Mars proximity link.

Furthermore, for communication-on-the-move operations and reduced manual labor, it is desired to remove the operational dependency on a priori detailed telecommunications prediction, but instead use statistical data on the channel, (e.g., mean, variance, or other higher order statistics, or even known distribution) and combine it with real-time, in-situ channel measurement to achieve optimization. In this work we assume that some level of telecommunication prediction and characterization of the statistics of the channel will be available as part of design and mission planning; this alone is needed. It is desirable to eliminate the day-to-day detailed pass prediction and planning exercise where entire SNR profile is generated for each pass and possible orbiter-lander orientations.

In this paper we consider the dual problem of maximizing the throughput of a lander that has a limited amount of energy to be used for transmission; and minimizing the energy consumption used for transmission of data subject to delay constraints. First we consider the problem of maximizing the data throughput between the lander and a Mars orbiter during a pass. We consider a lander with limited amount of energy in the battery to be used for data transmission. Over the orbital duration, the lander will only be in view of the orbiter for a short period of time during which transmission can take place. Since the lander may not have sufficient energy to continuously transmit through this pass at full data rate, it is possible to significantly increase the data throughput of the lander by choosing when to transmit, and at what data rates, based on measured channel conditions.

Using techniques from Dynamic Programming, we develop an optimal algorithm for choosing when to transmit based on channel conditions. Our algorithm has a simple threshold form; so that when the channel quality is above the threshold, it is optimal to transmit at full power and when the channel is below the threshold it is best not transmit at all. Moreover, the thresholds are easy to compute, off-line, based on channel statistics. We show, using channel measurements from the Mars Odyssey, that total data throughput over a pass can be increased by nearly an order of magnitude through the use of our optimal scheduling algorithm.

We then consider the problem of minimizing the total energy consumption for transmitting a given amount of data by a deadline. Such a deadline may be imposed by real-time mission requirements; for example, the Mars Exploration Rover uses the morning pass, when the solar panel is operating, to send time-critical operational data such as terrain image and health/status information for daily science activity and mission planning purposes. Again, we use Dynamic Programming techniques to obtain the optimal transmission policy and our policy has a simple threshold form.

II. Scenario

Currently two Mars Exploration Rovers (MER), known as Spirit and Opportunity, are deployed on the planet surface. The mission of a Mars rover is to explore the planet surface and gather raw science data and analysis results (e.g., chemical and geological analysis of rocks and soil samples) to send back to Earth. The rover generates the power to cope with these basic activities using solar panels, and storing energy in internal batteries.

Typically a lander/rover will carry two communication systems. The first one is an X-band Direct-to-Earth (DTE), Direct-from-Earth (DFE) system designed to operate at lower data rate with higher reliability. It can be used to carry critical command/telemetry data or as a backup system in the case the relay infrastructure is not available. The other system is a higher-data rate UHF radio with omni-directional antenna designed to communicate with three satellites orbiting around Mars, the 2003 Mars Express, the 2001 Mars Odyssey and the 1997 Mars Global Surveyor. When passing overhead a rover, the orbiter signals the lander via a hailing message defined by the CCSDS Proximity-1 Space Link Protocol [4] and establishes radio communication for the duration of the pass, which can last as long as 20 minutes in the case of MER. The rover can send data during that time window, transmitting at 128 kilobits per second (typical data rate for MER) using automatic repeat request (ARQ) and forward error correction technique for error control. Once data is received onboard the satellite, it can be forwarded to Earth when the orbiter comes into view with one of the ground antenna sites of NASA's Deep Space Network (DSN).

The operational experience of MER clearly demonstrates the benefit of using the proximity links for relay to improve energy efficiency and ultimately extending the operating life span as well as the science return of the rover. Based on this experience, many future scout class missions will likely be "UHF only" missions in the sense that they rely primarily on the orbiters for communications with Earth instead of using X-Band DTE radio. Besides energy efficiency, there are also mass/weight issues relating to the power subsystems such as solar panel and batteries. Energy efficient operation will turn the mass reduction into savings on propellant cost or more resources for improving science instruments and mobility. The implementation of scheduling policies that maximize throughput and minimize energy consumption subject to deadline constraints is therefore an important aspect of the design for future Mars landers.

A new Orbiter, the Mars Reconnaissance Orbiter (MRO), has been launched on August 2005, and its relay service will start in 2007 to support the Phoenix lander that is scheduled to arrive. MRO carries a new UHF communication system called Electra, and the Phoenix lander is planning to conduct nominal operation via relay communication services provided by MRO and Odyssey. Due to the energy constraint of the lander, it will not transmit data for the entire duration of the pass, but rather will be more selective depending on the volume of data and energy budget. Also, due to MRO's science orbit, many passes will be short and at low elevation angle, which amplifies the multi-path effects from terrain features and near-field effects from science instruments co-located with the antenna. In this case, the need to optimize transmission over the Mars proximity link is even more critical.

Fig. 1 shows a sample channel realization with minimum variance among the set of realizations considered in this paper. It is taken from a database of recorded proximity link measurement for the MER-to-Odyssey link. Even though the stochastic variance of the channel condition in this particular sample is small compared to the rest of the sample set, it is clear that stochastic fluctuations still play a central role. A dynamic scheduling algorithm able to adapt the decision based on real-time observation of these fluctuations is necessary for energy efficient communications.

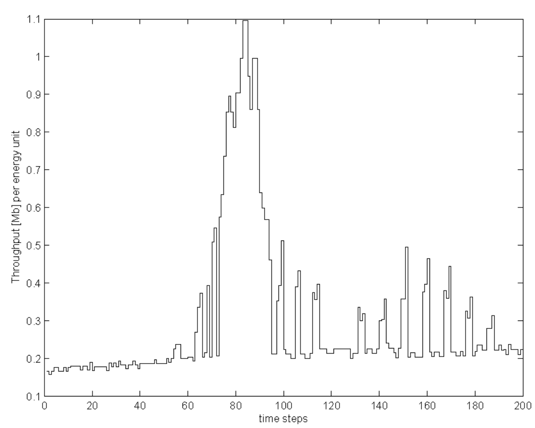

Fig. 1. Sample of the Mars proximity link as observed from the Mars Odyssey orbiter. The channel quality is given in terms of Mb achievable per unit of energy consumed, where one unit of energy corresponds to a trasmission at full power over an entire time slot (4-second long step).

The path shown in Fig. 1 is just one example of channel realization associated with a particular pass for an orbiter-rover pair. A sample of the channel realization represents the throughput achievable consuming one unit of energy over an entire time slot: the sample roughly equal to 1.1, for instance, indicates that consuming one unit of energy over that time slot an amount of data equal to 1.1Mb can be transmitted. In general, all the experiments reported in the rest of the paper are the result of testing the scheduling algorithms presented over *real* Mars proximity link realizations. In total we have the use of almost 400 channel realizations available, sampled during the communication between Mars Exploration Rovers and Odyssey. The data consists of samples taken at 4 second intervals and over a variable period of time depending on the elevation angle of the orbiter, the antenna patterns, and the stochastic fluctuations of that channel that affect the received signal strength. The duration of a pass window, namely the time during which the rover has a realistic chance of closing the proximity link with the orbiter, varies from 752 to 1176 seconds. In this paper we break up time into 4-second slots; this does not imply that a TDMA structure was assumed on the link layer, but simply refers to the time resolution of the available channel measurements. Since the channel is slowly varying, the 4-second time resolution does not impact the performance of the algorithm. However, if the channel was changing faster, it would be straightforward to use the algorithm with shorter time slots. For simplicity, from the entire set of realizations available, we considered a subset composed of at least 200 time slots, taking into account only the first 200 samples of each realization. Each sample associated with a time-slot and originally expressed as a signal-to-noise ratio has been converted into achievable throughput using full power during the entire time-slot, for clarity of presentation.

III. Maximum Throughput Problem

We consider a lander sending information to a satellite orbiting around Mars. Time is discretized; in each time slot the channel state changes according to a probabilistic model that can be realistically predicted. The channel state determines the throughput per unit of energy expended by the transmitter onboard the lander. The lander is extremely energy limited, therefore the transmitter has a limited amount of energy units to dedicate to transmission. The objective is to find a transmission schedule that maximizes the expected throughput, subject to a constraint on the total energy that can be expended, and a deadline by which the energy must be consumed. The time constraint is the result of the limited period of time in which the satellite orbiting around Mars can establish radio lock with the lander and the amount of solar energy and battery reserve allocated for the communication pass. A problem formulation matching the scenario just described was first presented by Fu, Modiano and Tsitsiklis in [1]. Let a_k be the available energy in the battery at time slot k. The battery starts with a_1 units of energy and the transmission must be completed by time slot n. The energy consumed at time slot k is c_k. It follows that, the available energy evolves according to $a_{k+1} = a_k - c_k$, with $a_k \geq c_k$ for all the time slots.

The throughput obtained, given a certain amount of consumed energy, depends on the channel fade state. Let q_k be the channel quality at time slot k, and let $f(c_k, q_k)$ be the throughput achieved at time slot k. The function f is in general assumed concave and non-decreasing with respect to c_k.

The objective is to maximize the expected data throughput given n time slots to transmit and a_1 units of initial energy. The problem is

$$\max \boldsymbol{E}\left[\sum_{k=1}^{n} f(c_k, q_k)\right] \quad (1)$$

subject to the constraints that $c_k \geq 0$ for all k and

$$\sum_{k=1}^{n} c_k \leq a_1 \quad (2)$$

In the following subsection the problem is studied under the condition that the channel quality q_k evolves according to a known distribution function $p(q)$; the channel quality sample per time slot is not revealed until just before the transmission at time k. In [1] the authors develop a dynamic programming algorithm that provides an optimal policy for the case where f is concave, and obtain an optimal closed-form policy for the special case where f is piecewise linear, namely linear with a power limitation. Recall that we are not assuming a complete knowledge of the channel in advance, an assumption that would reduce the problem to a waterfilling solution; we are assuming a rough prediction of the channel distribution and the knowledge of the channel quality sample at the beginning of each time slot.

A. Dynamic rate adaptation scheme

Let us assume the channel quality q_k is unknown until just before the transmission at time k. The sample q_k evolves according to a stationary distribution $p(q)$ that depends on the orbital dynamics of the particular pass (e.g., due to the elevation angle of the orbiter or the position of the lander). Hence, the channel distribution may change from one pass to another. Furthermore, we assume that the channel statistics can be reliably predicted (although accurate channel prediction is not required by our algorithm). We want to maximize the throughput achieved during a "pass" of the orbiter given an initial amount of energy on the lander. This problem is nearly identical to that considered in [1] where the objective was to maximize throughput subject to a deadline constraint. Here the deadline corresponds to the end of a "pass" during which the lander can communicate with the orbiter.

Under these assumptions, the dynamic programming technique of [1] can be used to find an optimal policy. Specifically, the value function $\mathcal{J}_k(a_k, q_k)$ provides a measure of the desirability of the transmitter having energy level a_k at time k, given that the current channel quality is q_k. The recursion of the DP formulation that links the value functions at each stage k is

$$\mathcal{J}_n(a_n, q_n) = f(a_n, q_n)$$

$$\mathcal{J}_k(a_k, q_k) = \max_{0 \leq c_k \leq a_k} \left[f(c_k, q_k) + \overline{\mathcal{J}}_{k+1}(a_k - c_k)\right] \quad (3)$$

where $\overline{\mathcal{J}}_{k+1}(a_k - c_k) = \boldsymbol{E}[\mathcal{J}_k(a_k - c_k, q_k)]$. The first term in the right hand side of equation (3), $f(c_k, q_k)$, represents the data throughput that can be achieved in the current stage consuming c_k units of energy, and given the channel quality sample q_k. The remaining energy at the next stage is then $a_k - c_k$, and $\overline{\mathcal{J}}_{k+1}(a_k - c_k)$ represents the expected throughput that can be obtained in the future if the amount of energy left is $a_k - c_k$.

Let us now assume the function $f(c_k, q_k)$ as a piecewise linear function of the expended energy, of the form $f(c_k, q_k) = q_k \min(c_k, P)$, where P represents a limit on the energy that can be expended per time slot, or power limit. Such a model is appropriate for space links that typically operate in the low SNR regime where the relationship between rate and power is linear.

Substituting into (3), the recursion becomes

$$\mathcal{J}_k(a_k, q_k) = \max_{0 \leq c_k \leq a_k} \left[q_k \min(c_k, P) + \overline{\mathcal{J}}_{k+1}(a_k - c_k)\right] \quad (4)$$

$$\mathcal{J}_n(a_n, q_n) = q_n \min(a_n, P).$$

At this point, an optimal policy and a closed-form formula implementing it can be derived.

Theorem 1: The expected value function $\overline{\mathcal{J}}_k(a_k)$, for $1 \leq k \leq n$, is piecewise linear with the form

$$\begin{aligned}\overline{\mathcal{J}}_k(a_k) &= \gamma_k^k \min(a_k, P)\\ &+ \gamma_k^{k+1}[\min(a_k, 2P) - \min(a_k, P)]\\ &+ \gamma_k^{k+2}[\min(a_k, 3P) - \min(a_k, 2P)]\\ &\quad\vdots\\ &+ \gamma_k^{n}[\min(a_k, (n-k+1)P) - \min(a_k, (n-k)P)] \quad (5)\end{aligned}$$

where the number of linear segments is equal to $(n-k+1)$ and where $\gamma_k^k, ..., \gamma_k^n$ are constants that give the slopes of each segment and are recursively determined. Starting from

$$\gamma_n^n = \boldsymbol{E}[q_n]$$

and moving backward up to $k = 1$, the slopes $\gamma_k^k, \gamma_k^{k+1}, ..., \gamma_k^n$ are calculated from $\gamma_{k+1}^{k+1}, ..., \gamma_{k+1}^n$ for $k < n$. The constants γ_k^k and γ_k^n are given by

$$\gamma_k^k = \boldsymbol{E}[\max(q_k, \gamma_{k+1}^{k+1})]$$

and

$$\gamma_k^n = \boldsymbol{E}[\min(q_k, \gamma_{k+1}^n)]$$

while $\gamma_k^{k+1}, ..., \gamma_k^{n-1}$ are given by

$$\gamma_k^I = \boldsymbol{E}[\min(q_k, \gamma_{k+1}^i) - \min(q_k, \gamma_{k+1}^{i+1})] + \gamma_{k+1}^{i+1}$$

The above form for the value function leads to a simple threshold based scheme for setting the energy consumption values, as follows:

Corollary 1: An optimal policy for $1 \leq k < n$ is to set the consumption c_k to:

$$\begin{cases} \min(P, a_k) & for\ \gamma_{k+1}^{k+1} < q_k \\ \min(P, \max(a_k - P, 0)) & for\ \gamma_{k+1}^{k+2} < q_k \leq \gamma_{k+1}^{k+1} \\ . \\ . \\ . \\ \min(P, \max(a_k - (n-k)P, 0)) & for\ q_k \leq \gamma_{k+1}^{n} \end{cases}$$

and $c_n = min(a_n, P)$.

The proof is given in [1]. This policy can be explained as follows. Assume a_k units of energy available at time k. At each time slot at most P units of energy may be consumed. If the channel quality q_k was known for all k, maximizing throughput would mean selecting the $\lceil \frac{a_k}{P} \rceil$ time slots with the best channels and scheduling the service during these time slots. Assuming that there are enough time slots available, this would entail consuming P units of energy for $\lfloor \frac{a_k}{P} \rfloor$ time slots and $a_k - \lfloor \frac{a_k}{P} \rfloor P$ for another time slot.

In our problem of interest, future channel qualities are of course not known. However, the constants γ_k^i are representative of expected channel qualities for future time slots as perceived just before the time slot k. The values γ_k^i are ordered, namely γ_k^k is the expectation of the best channel, and γ_k^n is the expectation of the worst one. If we associate, at time k, the list $\gamma_{k+1}^{k+1}, ..., \gamma_{k+1}^{n}$ to the actual expectations for future channel fade states, sorted in order of quality, we may derive an optimal policy based on the earlier case of known channel qualities. The policy is as follows: take the current channel fade state q_k and insert it into the ordered list. If q_k is among the best $\lfloor \frac{a_k}{P} \rfloor$ channel qualities, consume P units of energy; if q_k is exactly the $\lceil \frac{a_k}{P} \rceil$-th into the list, consume $a_k - \lfloor \frac{a_k}{P} \rfloor P$ units; otherwise, do not consume anything.

To conclude, notice that the threshold evaluation depends on the channel statistics. Since a particular channel distribution is associated with each pass of the orbiter, the threshold has to be computed for each pass. This evaluation can be accomplished prior to the start and then uploaded onboard the lander because the geometry of each pass, which is the most critical factor influencing the channel distribution, is known. This reduces the complexity on the spacecraft.

B. Numerical Results

In this subsection we present the performance evaluation of the *Dynamic rate adaptation* scheme, based on the Dynamic Programming formulation, compared to two other algorithms. Performances are evaluated over 50 channel realizations measured from the Mars Odyssey orbiter.

The two algorithms with which the DP algorithm is compared are based on the present scheme that uses a fixed transmission rate for the duration of the pass and a proposed scheme that adapts the transmission rate to the channel quality but does not attempt to optimize energy consumption by avoiding transmission during particularly bad slots. In particular the two schemes are as follows:

- *Full Power Fixed Rate* policy: this policy transmits at full power using a fixed ($BPSK$ in our examples) modulation scheme. When the initial amount of energy is not enough to transmit over the entire time window at full power, the policy transmits symmetrically around the center (peak) of the time window.
- *Full Power Multiple Rates* policy: this policy transmits at full power using different modulation schemes (i.e., different rates). When the initial amount of energy is not enough to transmit over the entire time window at full power, the policy transmits symmetrically around the center of the time window.

It is clear that the *Full power multiple rates* scheme is more efficient than the *Full power fixed rate* because it chooses the highest possible transmission rate for the given received SNR. These two policies were compared to the algorithm based on the Dynamic Programming formulation, in which not only the rate, but also the power consumption is adapted to the channel quality. Two versions of this *Dynamic rate adaptation* scheme have been adopted:

- *with empirical discrete distribution of the channel*: this version exploits the realizations from the Mars orbiter to obtain the channel distribution.
- *with Gaussian approximation of the channel*: this version uses a Gaussian approximation of the channel, with mean and standard deviation derived from the realizations from the Mars orbiter.

Recall that the distribution of the channel is needed to calculate the thresholds. While in the first case the distribution is empirically evaluated using real channel realizations, in the second one those realizations are only used to get means and standard deviations to use in the Gaussian approximation of the channel. In both cases, the algorithms' performances are then evaluated over the realizations from the Mars orbiter.

Let us consider a single realization of the channel from the Mars orbiter. Given a power limitation P of 1 unit per time slot (it may be equal, for instance, to 1 or 10 Watts), and a time window of 200 slots, the amount required to transmit at full power over the entire window is 200 units. Let us now assume, for example, that the initial amount of energy is 50 units. That would correspond to the amount of energy required for full power transmission for $\frac{1}{4}$ of the pass. In Fig. 2 the scheduling solution of the policy based on the DP formulation is shown for this example. Shown in the figure is the channel quality, the thresholds, and the energy consumption. Since the channel quality is expressed in Mbits per unit of consumed energy, we consider a continuously variable range of data rates. Notice, that the threshold values decrease as the deadline approaches. Also notice that the algorithm attempts to transmit during the best channel conditions.

In Fig. 3 we compare the performances of all the algorithms for the same initial amount of energy, $\frac{1}{4}$ of what is needed to transmit at full power during the entire pass, over a set of 50 realizations from the Mars Odyssey orbiter. In particular, for each channel realization the figure shows the throughput

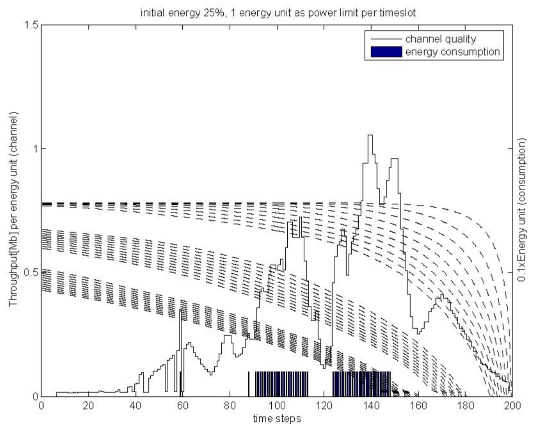

Fig. 2. Threshold values (represented by dashed lines), channel condition and energy consumption for the dynamic programming based rate adaptive scheme. Any time the channel condition is above the threshold, a transmission occurs; the amount of energy consumed depends on the quality of the channel compared to the set of thresholds, according to the algorithm described in Corollary 1.

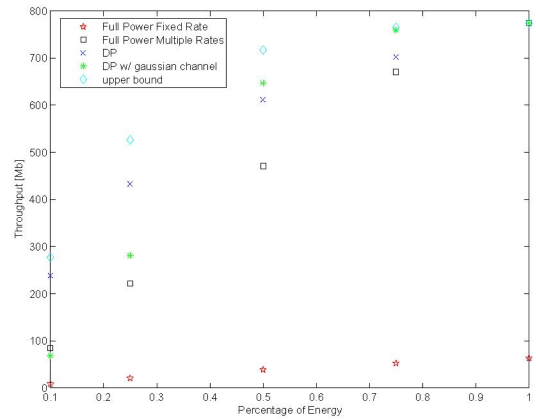

Fig. 4. Average Throughput achieved by different policies for different initial amounts of energy, expressed as a fraction of the amount of energy needed for full power transmission during the entire pass. The average performance is taken by applying the algorithm over 50 channel realizations measured from the Mars Odyssey orbiter.

achieved by the lander for each of the scheduling policies.

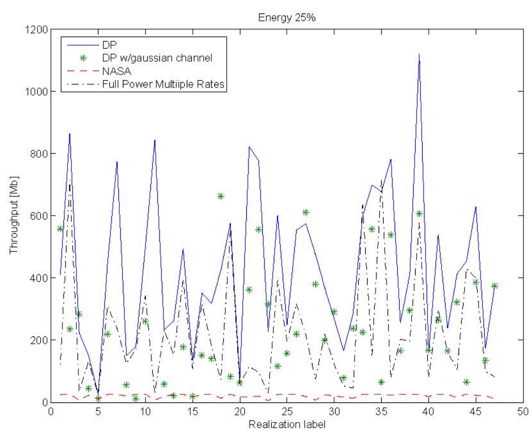

Fig. 3. The Throughput achieved by Mars lander sending data to the orbiter for each of 50 channel realizations.

To better appreciate the performance of the rate adaptation algorithm, in Fig. 4 we show the average performances over the entire set of realizations, for the same initial amount of energy. It is clear that the smaller is the initial amount of available energy, the larger is the benefit of using the dynamic programming rate adaptation scheme. This is because when the amount of energy is large (e.g., sufficient to transmit at full power for the duration of the pass), the benefit of being able to opportunistically select the time-slots in which to transmit diminishes. Finally, in Fig. 4 we also show the upper bound on throughput obtained by selecting the best fraction of time-slots during which to transmit if the channel conditions were fully known in advance for the entire duration of the pass.

IV. Minimum Energy Problem

We move now from a problem in which we have a given amount of energy, and we wish to maximize the expected throughput within a fixed period, to another problem in which we have an initial amount of data to transmit by a deadline, and we want to minimize the energy consumption. Let d_k be the amount of data remaining to be served at time k, and s_k the amount of data served during time slot k. Thus, $d_{k+1} = d_k - s_k$. The channel quality at time k is q_k. Transmitting s_k units of data requires $g(s_k, q_k)$ units of energy, and the rate-power curve is assumed convex and differentiable in s_k. Since the transmission must be completed by time n, the objective is to find a transmission policy that minimizes the expected energy

$$\boldsymbol{E}\left[\sum_{k=1}^{n} g(s_k, q_k)\right] \tag{6}$$

subject to the constraints that $s_k \geq 0$ for all k and

$$\sum_{k=1}^{n} s_k \geq d_1 \tag{7}$$

Also in this case a Dynamic Programming approach can be used to obtain an optimal policy [1]. Finally, an optimal policy in closed-form can be achieve when g is linear and subject to a power limit, and q_k only takes values which are integer multiples of a minimum channel quality q_{min}.

A. Dynamic rate adaptation scheme

Let us assume a special case where $g(s,q)$ is linear in $\frac{s}{q}$, so that q is proportional to the amount of data transmitted per unit of energy consumed. In addition, let us assume a power limit P on g. This power limit results in a limit of Pq_k on the throughput achievable at time k. If d_k is the amount of data remaining to be sent, the recursion of the

Dynamic Programming formulation is the following:

$$\mathcal{J}_k(d_k, q_k) = \min_{0 \le s_k \le \min(d_k, Pq_k)} \left\{ \frac{s_k}{q_k} + \overline{\mathcal{J}}_{k+1}(d_k - s_k) \right\} \tag{8}$$

where $\overline{\mathcal{J}}_{k+1}(d_k) = \boldsymbol{E}[\mathcal{J}_k(d_k, q_k)]$. We impose an infinite cost for not sending all the data by the deadline; the *terminal cost function* is

$$\mathcal{J}_{n+1}(d_{n+1}, q_{n+1}) = \begin{cases} 0 & for\ d_{n+1} \le 0 \\ \infty & for\ d_{n+1} > 0 \end{cases}$$

For any possible value of q_k, let $\phi_k(q_k)$ be a value of u_k that minimizes the expression

$$\frac{d_k - u_k}{q_k} + \overline{\mathcal{J}}_{k+1}(u_k)$$

over all $u_k \ge 0$. Thus

$$\phi_k(q_k) = \arg\min_{u_k \ge 0} \left[\overline{\mathcal{J}}_{k+1}(u_k) - \frac{u_k}{q_k} \right] \tag{9}$$

A value of s_k minimizing the right-hand side of equation (8) can be expressed in terms of $\phi_k(q_k)$, leading to the optimal policy as given in Theorem 2.

Theorem 2: There exists an optimal policy in which the amount of data to serve at time slot k has the following form:

$$s_k = \begin{cases} 0 & if\ d_k \le \phi_k(q_k) \\ \min(d_k - \phi_k(q_k), Pq_k) & if\ \phi_k(q_k) < d_k \end{cases}$$

The proof is given in [1]. In effect, $\phi_k(q_k)$ is a threshold under which the energy cost of sending data immediately exceeds the cost of saving data for later transmission. It does not depend on the remaining data d_k, so it is easy to compute. This property allows the development of numerical methods that considerably speed the process of calculating the value function, detailed in [2]. When q_k is discrete and is restricted in value to integer multiples of a constant q_{min}, it is possible to obtain a closed-form expression of the optimal solution [1].

Theorem 3: Suppose the channel quality q_k is restricted to integers multiple of q_{min}. The expeced value function is given by

$$\begin{aligned} \overline{\mathcal{J}}_k(d_k) &= \frac{1}{\eta_k^1} \min(d_k, Pq_{min}) \\ &+ \frac{1}{\eta_k^2}[\min(d_k, 2Pq_{min}) - \min(d_k, Pq_{min})] \\ &\quad \vdots \\ &+ \frac{1}{\eta_k^{n-k+1}}[\min(d_k, (n-k+1)Pq_{min}) - \min(d_k, (n-k)Pq_{min})] \end{aligned} \tag{10}$$

[1]This assumption is particularly reasonable for the Mars proximity links where the channel values over a pass can be quantized and represented as a integer multiple of some minimum value.

where the constants η_k^i, for $1 \le k \le n$ and $1 \le i \le n-k+1$, are defined in the following.

The proof is given in [1]. It turns out that the expected value function $\mathcal{J}_k(d_k)$ is a piecewise linear function with $n-k+1$ segments, each with slope $\frac{1}{\eta_k^i}$. Regarding the slopes, the base case $k = n$ (and $i = 1$) is given by

$$\frac{1}{\eta_n^1} = \boldsymbol{E}\left[\frac{1}{q_n}\right]$$

while $\eta_{k-1}^1, ..., \eta_{k-1}^{n-k+2}$ are recursively obtained from $\eta_k^1, ..., \eta_k^{n-k+1}$ resulting in

$$\left(\frac{1}{\eta_{k-1}^1}, ..., \frac{1}{\eta_{k-1}^{n-k+2}}\right) = \boldsymbol{E}\left[\theta\left(\frac{q_k}{q_{min}}, \frac{1}{q_k}, \frac{1}{\eta_k^1}, ..., \frac{1}{\eta_k^{n-k+1}}\right)\right] \tag{11}$$

where θ is defined in Definition 1.

Definition 1: Given an m-dimensional list $(\alpha_1, ..., \alpha_m)$ sorted in ascending order, and an i-dimensional list consisting of i repetitions of the same number x, let $\theta(i, x, \alpha_1, ..., \alpha_m)$ be the $(m+1)$-dimensional sorted list obtained by

1) merging and sorting the two list
2) keeping the smallest $m+1$ elements

The slopes $\frac{1}{\eta_k^1}, ..., \frac{1}{\eta_k^{n-k+1}}$ reflect the expected marginal cost of sending a data unit. At time k, data may be sent immediately for a cost of $\frac{1}{q_k}$ units of energy per unit of data. A maximum of Pq_k units of data can be sent per time slot. Alternatively, data may be sent in future stages for an expected cost given by $\overline{\mathcal{J}}_{k+1}$: $\frac{1}{\eta_{k+1}^1}$ for the first Pq_{min} units of data, and $\frac{1}{\eta_{k+1}^i}$ for each ith additional Pq_{min} units of data. The resulting value function $\mathcal{J}_k(d,q)$ is then a piecewise linear function with slopes

$$\theta\left(\frac{q_k}{q_{min}}, \frac{1}{q_k}, \frac{1}{\eta_{k+1}^1}, ..., \frac{1}{\eta_{k+1}^{n-k+1}}\right)$$

for $0 \le d \le (n-k+1)Pq_{min}$. In particular, the slopes for $\overline{\mathcal{J}}_{k+1}$ are calculated according to (11). Given this form of the value function, a closed-form algorithm implementing the optimal solution can be obtained as follows.

Corollary 2: An optimal policy at time k, for $1 \le k \le n-1$, is to set the amount of data to serve s_k equal to:

$$\begin{cases} \min(d_k, Pq_k) & if\ q_k > \eta_{k+1}^1 \\ \min(\max(d_k - Pq_{min}, 0), Pq_k) & if\ \eta_{k+1}^2 < q_k < \eta_{k+1}^1 \\ \min(\max(d_k - 2Pq_{min}, 0), Pq_k) & if\ \eta_{k+1}^3 < q_k < \eta_{k+1}^2 \end{cases}$$

and so on until $q_k < \eta_{k+1}^{n-k}$ where $s_k = \min(\max(d_k - (n-k)Pq_{min}, 0), Pq_k)$.

B. Numerical Results

In this subsection we present the performance evaluation of the *Dynamic rate adaptive scheme* with empirical discrete distributions of the channel, the *Full power fixed rate* scheme and the *Full power rate adaptive* scheme. Performances are evaluated on the same set of 50 realizations of the channel from the Mars Odyssey orbiter (see paragraph III-B); here the objective is to minimize energy consumption, given some initial amount of data to deliver. Let us consider a single realization of the channel from the Mars orbiter. We apply a channel quantization where q_{min} is equal to 10 kilobits and all the other values are integer multiple of q_{min} with 10 kilobits as minimum granularity. Given a power limitation P of 1 unit per time slot, a time window of 200 slots and an initial amount of data equal to 4Mb, in Fig. 5 the scheduling solution for the scheme based on the DP formulation is presented. Notice, again, that the

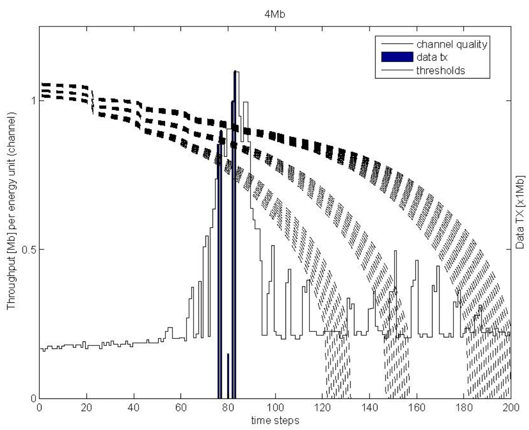

Fig. 5. Thresholds (dashed lines), channel conditions, and data transmission for the energy minimization problem with 4Mb of data to send.

threshold values decrease as the deadline approaches. Also, notice that the algorithm chooses to transmit during the best samples of the channel quality (although channel conditions are not known to the algorithm in advance).

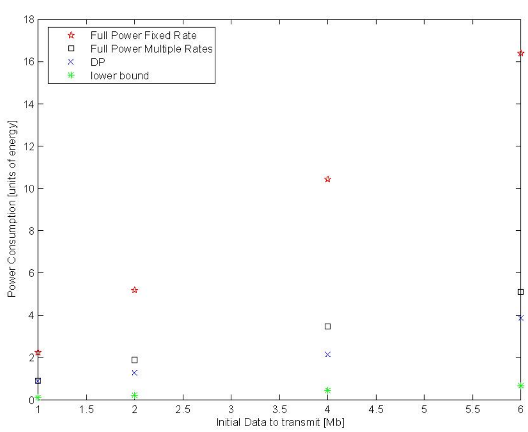

Fig. 6. Average Energy Consumption of the different policies for different initial amounts of data. The average performance is taken by applying the algorithm over 50 channel realizations measured from the Mars Odyssey orbiter.

In Fig. 6 we plot the average performances of the *Dynamic rate adaptation* scheme and the others schemes for different initial amounts of data. It can be noticed that the approach based on the DP formulation improves the gap from the other algorithms for larger and larger amounts of data to serve. Also in this case a lower bound is shown by assuming complete knowledge of the channel realization over the entire pass and transmitting during the best time slots. Again, the *Dynamic rate adaptation* scheme is reasonably close to the bound. The gap can be interpreted as the cost due to the lack of knowledge of the channel. Note that the ability to respond to the channel variation more rapidly would improve the absolute performance of all the algorithms. However, the gap between the lower bound and the DP algorithm is independent of the duration of the time slot. In this case, this gap only depends on the uncertainty about the future of the channel quality, even if we could measure the channel perfectly time slot after time slot.

V. THE MULTI-USER CASE OF THE MINIMUM ENERGY PROBLEM

In this section we extend the energy minimization problem to the multi-user (landers) case, using the approach described in [3]. We assume that the Mars orbiter has to communicate with more than one Mars lander, scheduling their transmission over the same time window. This assumption is particularly reasonable when multiple landers are near each other on the Mars surface.

We consider a system with a single receiver, the Mars orbiter, collecting data from N users, through independent time-varying channels. The channel state of each user is a random process and is assumed to be known at the beginning of each time slot; i.e., q_{jk} is assumed known at the beginning of time slot k for each user j, but is unknown for future time slots.

Each user has an amount of data that must be transmitted by a deadline. The transmitter controls the consumed energy by adjusting the rate allocation subject to the constraints that only one user can transmit in each slot and that all of the data must be transmitted by the deadline. The rate per time slot assigned to the user j at time slot k is μ_{jk}. For any given state q_{jk} of the channel j at time slot k, there is a rate-power curve $g(\mu_{jk}, q_{jk})$ representing the amount of energy required to transmit at rate μ_{jk} when the channel is in state q_{jk}.

The goal is to find a transmission schedule that minimizes the expected consumed energy, subject to a constraint on the minimum amount of data to serve for each user and a deadline by which it must be transmitted. For clarity of presentation we now let K represent the deadline constraint (i.e., the end of the pass) where $K \geq N$. The optimization problem becomes

$$\min \boldsymbol{E}\left[\sum_{j=1}^{N}\sum_{k=1}^{K} g(\mu_{jk}, q_{jk}) \cdot \tau_{jk}\right] \tag{12}$$

subject to the constraint that at least the initial amount of data d_j for each user is served within a finite time window:

$$\sum_{k=1}^{K} \mu_{jk}\tau_{jk} \geq d_j \ \forall j = 1, ..., N \tag{13}$$

$$\sum_{j=1}^{N} \tau_{jk} \leq 1 \; \forall k = 1, ..., K \tag{14}$$

where τ_{jk} is equal to 1 if the user j has been served during the time slot k, 0 otherwise. Inequality (13) expresses that the service of all the users has to be completed within the frame of K time slots, while (14) that at most one user per slot can be served.

The power expenditure is assumed to be a piece-wise linear function of the transmission rate, i.e. $g(\mu_{jk}, q_{jk}) = \frac{\mu_{jk}}{q_{jk}}$, with a power limitation. Again, this optimization problem can be formulated and solved using discrete-time Dynamic Programming techniques. However, in the case of multiple users, we are no longer able to find a simple solution to the Dynamic Program and even a numerical solution become prohibitively complex. Hence, we consider heuristic approaches based on our earlier work in [3]. The algorithm is briefly described below; more details can be found in [3].

A. Multi-user heuristic algorithm

For the moment, let us ignore the deadline constraint and focus on a *single user* with an average rate requirement. Our heuristic is based on solving the following optimization problem. Specifically, we want to minimize the expected consumed energy subject to an average rate guarantee to the user. Therefore, during each time slot, the transmitter need to choose a data rate based on the channel state during that time slot. The formulation of this problem is as follows:

$$\begin{aligned} &\min \sum_{q} \boldsymbol{P}(q) g(\mu_q, q) \\ &s.t. \sum_{q} \boldsymbol{P}(q) \mu_q = LTRG \end{aligned} \tag{15}$$

where the optimization is taken over the values of μ_q, the data rate chosen when the channel is in state q. The function to be minimized represents the expected energy cost averaged over all possible discrete channel values, and the constraint represents the average long-term rate guarantee ($LTRG$); where, $\boldsymbol{P}(q)$ is the probability mass function of the channel state. The sum in the objective function is a simple expectation of the consumed energy. The solution of this optimization problem yields, for any channel state q, the value μ_q of the data rate, in order to meet the long term rate guarantee, subject to minimizing the expected consumed energy.

Now, to solve the multi-users problem, the transmitter has to make the following two decisions at the beginning of each time slot: 1) to whom the time slot should be allocated; and 2) the transmission rate to use during the time slot. Our heuristic algorithm selects the user with the best channel state during each time slot; and uses the data rate obtained for the single user problem with appropriately chosen LTRG parameter and the channel condition for the chosen user. The LTRG parameter is updated at each time slot on the basis of the amount of data already served and the remaining time before the deadline.

B. Numerical results

In this subsection we compare the performance of the heuristic algorithm with the multi-user version of the *Full Power Fixed Rate* and the *Full Power Multiple Rate* approaches that we presented for the single-user scenario. The *Full Power* algorithms are adapted to the multi-user scenario using time-division-multiple-access, but are otherwise no different than the single user versions described earlier. We assume 2 users, an initial amount of data to transmit from 5 to 25 Megabits, and a time window constraint of 800 seconds. As can be seen from Fig. 7 the heuristic algorithm outperforms the *Full Power* schemes; however, the energy savings due to the heuristic algorithm are not nearly as large as in the single user scenario. This is most likely due to the sub-optimality of the heuristic algorithm. Hence, developing more efficient algorithms for the multi-user scenario remains an important area for future work.

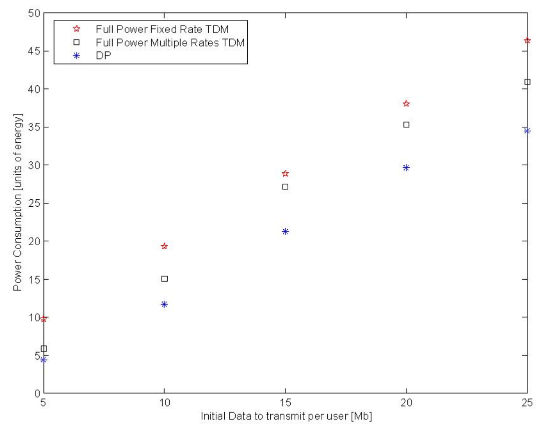

Fig. 7. Energy consumption per user with different initial amount of data to transmit from the landers to the Mars orbiter

VI. Conclusions

Efficient use of energy is critical for future robotic missions to Mars. Since communications is a major source of the energy drain for many such missions (e.g., scout missions); in this paper we focused on energy efficient transmission scheduling schemes. By adopting Dynamic Programming based algorithms first developed in [1] to the Mars scenario, we were able to demonstrate an order of magnitude increase in data throughput and similar decrease in energy consumption. Our approach takes advantage of the time-varying channel conditions to "opportunistically" transmit at time when the channel is relatively good. The benefits of our Dynamic Programming based approach is not only in the increased throughput and energy efficiency, but also in the simplicity of operation. Indeed, implementing our algorithms only requires statistical knowledge of the channel (e.g., distribution, or first and second moments), but does not require sophisticated channel prediction. This is in contrast with proposed approaches that attempt to develop accurate channel prediction models based on antenna patterns, orbital dynamics, and lander positions.

References

[1] A.Fu, E.Modiano, J.Tsitsiklis, "Optimal Transmission Scheduling over a fading channel with Energy and Deadline Constraints", *IEEE Transactions on Wireless Communications*,to appear.

[2] A.Fu, "Optimal Energy Allocationfor for Space and Wireless Communications", Ph.D. Thesis, MIT, 2002.

[3] A.Tarello, J.Sun, M.Zafer, E.Modiano, "Minimum Energy Transmission Scheduling subject to Deadline Constraints", *Proceedings of the 3rd Intl. Symposium in Mobile, Ad-Hoc and Wireless Networks, WiOPT 2005*, Riva del Garda (TN), Italy, April 2005.

[4] "Data Link Layer", CCSDS Recommendation for Space Data System Standards, *http://public.ccsds.org/publications/archive/211x0b3.pdf*

Alessandro Tarello Alessandro Tarello (S'03) received his M.S. degree in Communication Engineering from Politecnico di Torino, Torino, Italy, in 2002. He is currently enrolled in a Ph.D. program at Politecnico di Torino, under the supervision of proff. Emilio Leonardi and Marco Ajmone Marsan. He visited the Laboratory for Information and Decision Systems at MIT, Cambridge, MA, USA, from January to December 2004 and from July to October 2005. During Summer 2005 he also visited the Jet Propulsion Laboratory, Pasadena, CA, USA. He received the best student paper award at the Third International Symposium on Modeling and Optimization in Mobile, Ad-Hoc and Wireless Networks (WiOPT'05).

His research interests are in the fields of models for performance evaluation of the TCP protocol and scheduling algorithms and optimization techniques for wireless and ad-hoc networks.

Jay L. Gao Jay Gao joined the Jet Propulsion Laboratory in 2001 and is currently a member of the technical staff in the Communications Networks Group in the Telecommunication Research and Architecture section. His research is primarily focused on space-based wireless communications and networking, with emphasis on applications for the Mars Network. He is currently conducting research for developing quality-of-service (QoS)-aware protocols for the envisioned Interplanetary Network (IPN) and study optimization and protocols for deep space Ka-band communications. He also supports requirements definition and interface design activities for the Department of Defense's Transformational Communications MilSatcom project and system engineering effort for NASA's Exploration System and Mission Directorate (ESMD). Other research interests include optical-based sensor-web, discrete event simulation of distributed communication/sensor systems, energy efficient routing and self-organization algorithm for cooperative signal processing and sensor networks. He received his B.S., M.S., and Ph.D. degree in Electrical Engineering from UCLA in 1993, 1995, and 2000, respectively.

Eytan H. Modiano Eytan Modiano received his B.S. degree in Electrical Engineering and Computer Science from the University of Connecticut at Storrs in 1986 and his M.S. and Ph.D. degrees, both in Electrical Engineering, from the University of Maryland, College Park, MD, in 1989 and 1992 respectively. He was a Naval Research Laboratory Fellow between 1987 and 1992 and a National Research Council Post Doctoral Fellow during 1992-1993. Between 1993 and 1999 he was with MIT Lincoln Laboratory where he was the project leader for MIT Lincoln Laboratory's Next Generation Internet (NGI) project. Since 1999 he has been an Associate Professor in the Department of Aeronautics and Astronautics and the Laboratory for Information and Decision Systems (LIDS) at MIT. His research is on communication networks and protocols with emphasis on satellite, wireless, and optical networks.

He is currently an Associate Editor for Communication Networks for IEEE Transactions on Information Theory and for The International Journal of Satellite Communications. He had served as a guest editor for IEEE JSAC special issue on WDM network architectures; the Computer Networks Journal special issue on Broadband Internet Access; the Journal of Communications and Networks special issue on Wireless Ad-Hoc Networks; and for IEEE Journal of Lightwave Technology special issue on Optical Networks. He is the Technical Program co-chair for Wiopt 2006 and Infocom 2007.