Communications Network Architecture for Space-Based Sensor Networks[1]

Loren P. Clare, Jay L. Gao, Esther H. Jennings, Clayton Okino
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
{Loren.P.Clare, Jay.L.Gao, Esther.H.Jennings, Clayton.M.Okino}@jpl.nasa.gov

Abstract—Numerous planned and proposed future space exploration missions will employ multiple spacecraft that perform multipoint sensing. Distributed space-based sensing missions can significantly benefit from incorporation of cross-link communications capabilities, thereby forming space-based networks, by enabling continuous access to any/all spacecraft via single ground contact, real-time coordinated observations, and autonomous processing among in situ spacecraft. In this paper, we present a communications architecture for space-based sensor networks. Because of the large inter-spacecraft distances, directional antennas are used, with a single half-duplex transceiver per spacecraft to achieve low cost. Orbital motion induces a dynamic albeit predictable geometry (and topology) among the spacecraft. Primary offered traffic is sensor telemetry destined to the Earth ground station, although other traffic patterns are also treated. We present a technique that derives the link activation schedule (transmit/receive mode and communications neighbor selection) and routes used for efficient traffic relay through the network. An illustrative example is presented consisting of a 16-satellite sensor network. Throughput and latency performance is evaluated. An extension to the networking method is described that is traffic adaptive.

TABLE OF CONTENTS

1. INTRODUCTION	1
2.DISTRIBUTED SPACECRAFT MISSIONS	2
3. MOTIVATION FOR NETWORKING	2
4. SPACE-BASED NETWORK MODEL	3
5. SPECIAL CASES	7
6. SCHEDULING/ROUTING METHOD	8
7. TRAFFIC ADAPTIVE EXTENSION	14
8. CONCLUSIONS	16
REFERENCES	16

1. INTRODUCTION

An explosion of sensor network research and development has occurred in recent years, owing to the recognition of the power they can achieve. Technological advancements in microsensors, low-power electronics, cooperative processing among distributed nodes and wireless communications protocols have contributed to the rapid progress made. Sensor networks have been applied across a variety of applications, including battlefield surveillance, condition-based maintenance in factory automation, and habitat and environmental monitoring. In this paper, we focus on sensor networks in the space exploration domain, and provide a communications architecture that is well suited for this application. One key characteristic of distributed spacecraft missions is the large distances between spacecraft, which (coupled with the need for low energy consumption) drives the use of directional transmit and receive antennas.

In the following sections, we first describe the need for distributed sensing in space. We next identify the rationale for networking the distributed spacecraft. In Section 4 we provide the formal definition of the system characteristics, the general assumptions and approach for developing the network architecture. Section 5 provides results for simple special cases, which are useful for building intuition and bounding performance for more general situations. Section 6 presents the algorithm used to generate schedules that provision network bandwidth across the spacecraft. A 16-satellite example is used to illustrate the execution of the procedure, and compared with alternative methods from sensor networking and mobile ad hoc network (MANET) researchers. In Section 7 we present an extension of the method to operate in a demand-driven fashion, which will provide substantial improvements in efficiency for bursty offered traffic. Conclusions and future work are presented in the closing section of the paper.

[1] 0-7803-8155-6/04/$17.00 © 2004 IEEE
IEEEAC Paper Number: 1471, Updated October 8, 2003

2. DISTRIBUTED SPACECRAFT MISSIONS

The need for multi-spacecraft sensing has long been recognized. In 1974, the Space Science Board of the National Academy of Sciences recommended multi-spacecraft missions for "space weather" investigations. Towards this aim, the Interplanetary Monitoring Platform (IMP-7 and IMP-8) spacecraft were launched in the early 70s, and the International Sun-Earth Explorer (ISEE) mission consisting of three spacecraft were launched in the late 70s [1]. These missions yielded "simultaneous coordinated measurements to permit separation of spatial from temporal irregularities in the near-Earth solar wind, the bow shock, and inside the magnetosphere" [2] and were "able to break the space-time ambiguity inevitably associated with measurements by a single spacecraft on thin boundaries which may be in motion, such as the bow shock and the magnetopause." [3] The Dynamics Explorer (DE) mission, consisting of two spacecraft, were launched 1981, and many subsequent missions followed (e.g., GEOTAIL, WIND, INTERBALL, SOHO, POLAR). In the Space Studies Board (NRC) 2003 decadal strategy, 7 of 9 of the recommended moderate-class programs are multi-spacecraft [4]. The NASA 2003 Space Sciences Enterprise Strategy [5] states "Constellation technology is needed to enable efficient, simultaneous data collection at dispersed locations" and "Spacecraft systems for affordable clusters and constellations of small, ultra-low-power satellites providing multi-point measurements." Many additional distributed spacecraft missions have been identified within the context of NASA's Earth Science Enterprise, with the "sensor web" concept being a key strategic goal.

In general, a sensor network consists of multiple nodes, where each node contains one or more sensors, a radio, some degree of processing capability, and an energy supply. The sensor nodes are spatially distributed and perform multipoint sensing of the environment, i.e., simultaneously sense from their individual perspectives. There are a number of reasons for performing multipoint sensing (see e.g. [6,7]). A broad classification of multipoint sensing types has been presented [7]; the three basic classes are (1) pixellation/voxellation, in which each sensor perceives its local area "pixel" so that collectively the overall "picture" of the phenomenon of interest is derived, (2) beamformation, where an emission by an object of interest is received by multiple sensors (each signal having taken a different path) and the signals are combined to achieve higher SNR or target localization, and (3) tomography/rendering, where a target object is viewed from different perspectives corresponding to projections, and these are combined to form a higher dimensional representation. All three basic classes have application in the space-based sensor network context.

Some of the current missions involving a pixilation/voxellation class of multipoint sensing are the Magnetospheric Constellation (MagCon; also known as DRACO) [8] where the objective is to determine how the magnetotail stores, transports, and releases matter and energy, Magnetospheric Multiscale Mission (MMS) [9], Geospace Electrodynamic Connections (GEC) [10], Living With a Star (LWS) Geospace Mission [11] incorporating Radiation Belt Mappers, Ionospheric Mappers and Solar Sentinels; Global Precipitation Measurement (GPM) [12], and ESA missions Cluster [13], ACE+ [14], and SWARM [15]. Examples of the beamformation class multipoint sensing based missions are the group of interferometry missions such as Starlight/Terrestrial Planet Finder (TPF) [16] (leading to Planet Imager), Stellar Imager (SI) [17], Laser Interferometer Space Antenna (LISA) [18], MAXIM [19], ALFA [20], and Constellation X [21]. Tomography class missions include Leonardo (combining Bidirectional Reflectance Distribution Function signals) [22], Solar Terrestrial Relations Observatory (STEREO) [23], and COnstellation for Aerosols and Cloud Heights (COACH). Another is the Radio Tomography Imaging mission concept [24], in which radio signals are transmitted between spatially distributed spacecraft to form the sensor waveforms modulated by the magnetopause.

3. MOTIVATION FOR NETWORKING

Often, a sensor network is designed to collect information at a single base station. In space-based networks, spacecraft data will be relayed to one or more ground terminals and routed via terrestrial network to a Ground Data System. Thus conceptually the Earth is the base station.

The question arises whether there is a need for inter-satellite links (ISLs) and space-based networks. For example, in one design concept for the Magnetospheric Constellation, each satellite stores its data until it is near perigee, at which time it establishes a link with a ground terminal and dumps its storage contents. Thus the cost of ISLs is avoided. However, considerable functionality is sacrificed, as will be discussed next.

There are three basic capabilities enabled by space-based networking. First, mission operations users are provided continuous access to any and all spacecraft via a single ground contact with any single member of the constellation. This greatly increases ground operations efficiency and permits operation of the whole fleet as a single mission. The ability for the overall constellation to "act as a single mission spacecraft for coordinated observations" is a critical goal.

Second, space-based networking enables *real-time* coordinated observations by virtue of rapid dissemination of sensor observations. For example, ground-based sensors may be alerted while the phenomena of interest are still active. Instances of this have already been demonstrated, such as the gamma burst detected March 29, 2003 by the High-Energy Transient Explorer (HETE) that cued the European Southern Observatory's Very Large Telescope,

which confirmed a correlated supernova explosion [25]. Real-time coordinated sensing may also occur within the sensor constellation itself, in which interaction among the space assets is generated by sensed events (e.g., alert neighboring satellite to raise its sensing sensitivity/resolution, or corroborate/fuse data prior to transport to Earth). For example, early detection of a magnetic storm (arising from a Coronal Mass Ejection event) may be relayed to spacecraft further distant from the Sun, so that appropriate procedures may be taken (e.g., safing from the impending onset of radiation; or temporary downgrading of the presumed the accuracy of magnetic-based navigation sensors).

A third important capability enabled by space-based networking is autonomous local processing among the distributed spacecraft to achieve system functions beyond the sensor information exchange already mentioned. For example, many distributed spacecraft missions require precise, virtually rigid positioning in a formation flying configuration. Inter-spacecraft communications are essential to achieve formation control and reconfiguration functions. Additional capabilities include precise network time (for time-stamping sensor data) and precision navigation.

It is also noted that use of space-based networking allows the spacecraft to quickly move the data out of onboard buffers, clearing the contents to make room for new sensor information. Therefore significantly smaller sized buffers are required.

4. Space-Based Network Model

Space-based sensor networks differ from conventional sensor networks in that (1) long ranges exist between adjacent nodes, which impacts the physical layer design, and (2) the nodes may move along known trajectories. In studies where nodes are mobile, the typical assumption is that they do so in an unpredictable manner, and therefore require ad hoc network methods to communicate. In space-based sensor networks, the motion and resulting topology is dynamic but predictable; this fact can be used to design the network to operate with greater efficiency. It is also noted that because of the orbital motion of spacecraft, there is a natural energy load-balancing that arises, particularly for co-orbital spacecraft. (A number of sophisticated methods (e.g., routing) for balancing energy consumption across sensor network nodes have been presented in the literature.)

In this section we present the model for the space-based sensor network. Key components of the model consist of the transceivers, antennas, spatial motion, network timing, and offered traffic. We generally assume homogeneity across all spacecraft. If there is more than one ground terminal, all ground terminals have similar capabilities.

Half Duplex Communications Transceiver

Because we wish to minimize spacecraft cost, especially when they form large constellations, we assume that each spacecraft utilizes a single half duplex transceiver. Extension to multiple transceivers per spacecraft is left for future work. It is noted that other authors have considered multiple transceivers (not necessarily half duplex) per node for satellite and general ad hoc networks [26-31].

Antenna Models

Possible directional antenna models for spacecraft cross-links are (1) patch antennas on n sides (e.g., 4 sides of a nanosat cube parallel to a common axis), where one patch is selected for use at a time; (2) a single fixed mount (e.g., patch) antenna in which the inertial spin of the spacecraft body yields coverage; (3) an antenna array with electronic beam steering; and (4) a mechanically gimbaled antenna. These choices are depicted in Figure 1 below.

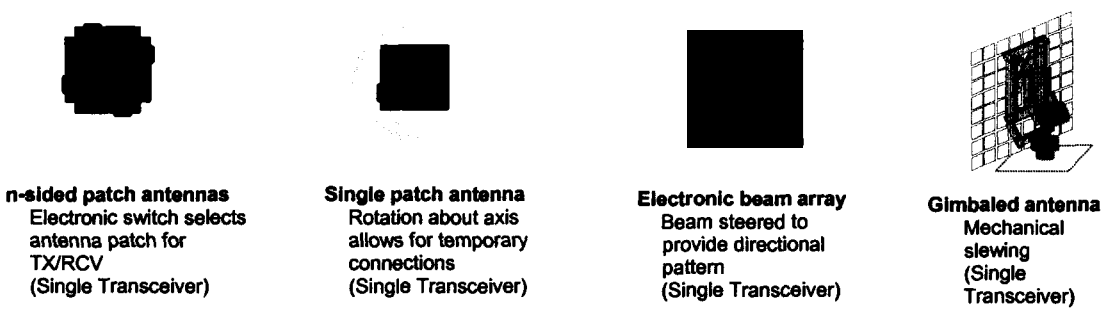


Figure 1. Directional Antenna Models

The least expensive alternative is the spinning single patch case. However, intercommunications is only possible when both link endpoints are pointed within antenna beamwidths, which may require careful control of spin rates and phasing across the network. Gimbaled antennas provide greater pointing flexibility but size/mass demands may conflict with the low-cost requirements of many space missions. Electronic beam arrays offer even greater pointing agility with smaller size and ability to conform to spacecraft shape, however, current electronics implementation costs likely preclude this choice for most missions. The n-sided patch antenna case appears to offer

the most viable alternative, although each mission will require careful trade analysis to determine the optimal design.

In the following, we assume that each spacecraft is always able to point toward any of its neighbors, and the delay to repoint to a different neighbor is insignificant. In addition, we assume that the beamwidths and spacecraft spacing are such that multi-access interference may be ignored.

Potential Connectivity Based on Maximum Range

It is assumed that there is a maximum range between spacecraft for which ISL communications may be achieved. Although two spacecraft are in range of each other (and not occluded by an object such as the Earth), they must both point their antennas at one another, as well as agree which endpoint is transmitting (and the other receiving), in order to achieve communications. Because of these additional constraints, we refer to spacecraft within range of each other as having a *potential* link between them.

The maximum range at which communications may take place from any spacecraft to Earth may be different than the maximum inter-satellite communications range. This is because the Earth terminal may have greater capabilities in antenna aperture, pointing, and receiver sensitivity.

We assume that a single data rate is used, independent of inter-spacecraft distance. It is noted that shorter range links may be closed with lower transmit power levels. We do not further consider such energy-conserving methods in this paper.

Spatial Dimensionality of Different Configurations

Distributed spacecraft missions may be generally distinguished as being either formation flying or loosely coupled. Formation flying missions (e.g., TPF, MAXIM or COACH) are typically configured with the spacecraft in precise fixed positions relative to one another. Loosely coupled missions (e.g., Magnetospheric Constellation) may require knowledge of their positions but need not maintain relative fixed positions.

Distributed spacecraft missions may also be distinguished by their spatial dimensionality. A "string-of-pearls" formation flying mission is essentially a linear, one-dimensional array. On such a mission, simple body-mounted antennas might be used. The Magnetospheric Constellation is an example where the spacecraft are all contained within orbits that are coplanar, and therefore a two-dimensional arrangement occurs. In this case, use of antenna patches along the sides parallel to a common axis perpendicular to the orbit plane may be used. An example of a three-dimensional configuration is MMS.

Examples of 1D, 2D, and 3D scenarios of distributed spacecraft configuration as depicted below in Figure 2.

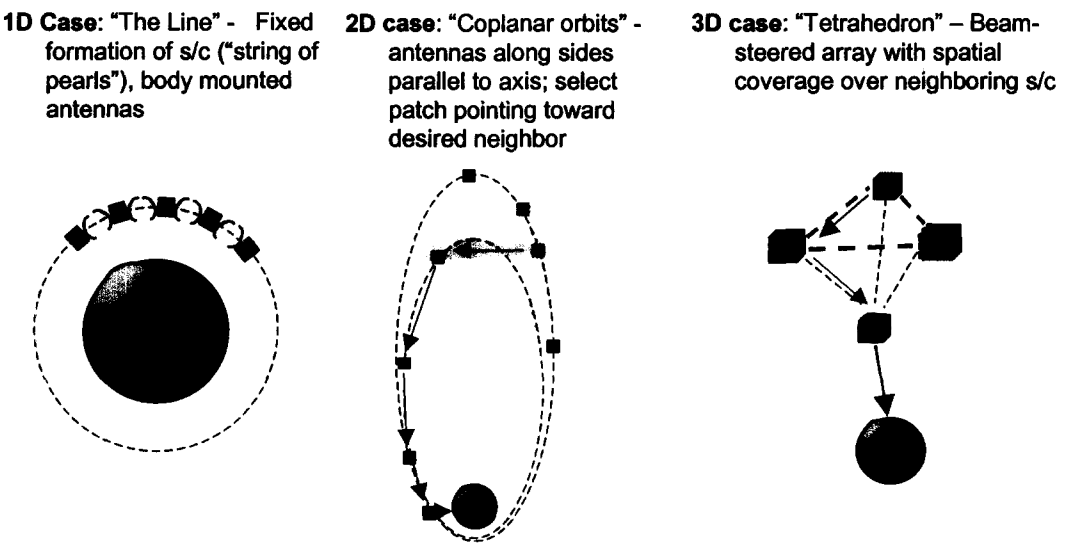


Figure 2. 1D, 2D, and 3D Mission Scenarios

Slotted Communications Assumption

It is assumed that all spacecraft know the common network time, and in particular are globally synchronized to periodic "slot" time boundaries. All transmissions are assumed to be initiated at the beginning of slot boundaries. The slot width is equal to the transmission of one fixed-length packet plus a time guard band. The time guard band is at least the propagation delay at maximum communications range. Receptions occur anytime within the window having the time guard band width (using a burst preamble to acquire the packet). These timing assumptions are depicted in Figure 3 below.

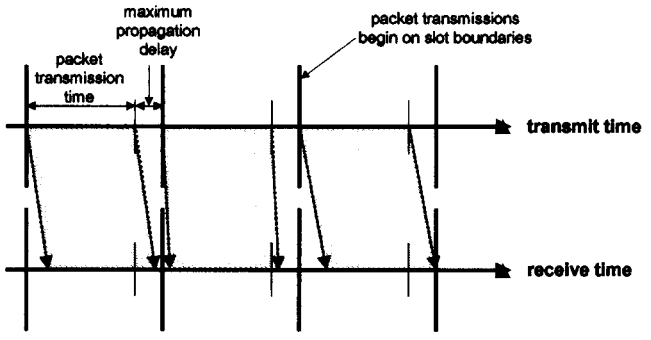


Figure 3. Slotted operation assumption

Model of Orbital Mechanics

The sensor spacecraft move in space generally in a nonlinear manner that is coupled with the communications system, causing dynamic effects in resource (connectivity) capabilities. However, the motion is deterministic and can be pre-computed using well-known Keplerian laws. Figure 4 below illustrates the variable satellite speed viewed in terms of the eccentric anomaly E, which is the ellipse-centered angle between line of apsides (the major axis) and the vertical projection of the satellite onto the circumscribed reference circle.

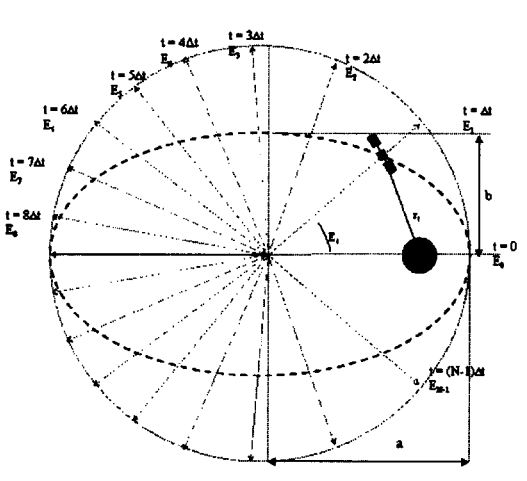


Figure 4. Spacecraft in Elliptical Orbit

A basic requirement for a sensor network could be continuous-time coverage over a continuous spatial region. A constellation of "evenly placed" co-orbital satellites may provide the solution. Co-orbital deployments can be achieved cost effectively using a single launch vehicle (position changes within an orbit require much less energy than orbit plane charges).

An even deployment of M satellites along a common orbit is defined as follows: For any fixed point on the orbit, the satellites pass the point at equally spaced time instants. This implies that the satellite rate past a fixed point is P/M where P = orbital period. The orbit period is determined by the parameters

a = semi-major axis
b = semi-minor axis
e = eccentricity
μ = gravitational parameter of Earth

The time interval between satellites moving to an adjacent satellite's location is denoted
$\Delta t = P/M$
and we increment over time in unit steps of this amount, denoting the time step index by i. The different time-phase positions along the orbit are $i\,\Delta t$, $i = 0,\ldots,M\text{-}1$.

We solve for the eccentric anomaly, E_i for each i.

$$i\Delta t = \frac{a^{3/2}}{\mu^{1/2}}\left(E_i - e\sin E_i\right)$$

Using E_i, we can solve for the radius and angle at each i.

$$r_i = a(1 - e\cos E_i)$$

and

$$\theta_i = 2\tan^{-1}\left(\left(\frac{1+e}{1-e}\right)^{1/2}\tan\frac{E_i}{2}\right)$$

The positions of the M spacecraft are given by the set of vectors over the index i.

Greater spatial coverage may be achieved by building the constellation out of multiple sets of evenly distributed co-orbital spacecraft. The number of satellites per orbit is generally different. An example is illustrated in Figure 5.

To complete the definition of the network geometry, the position(s) of the Earth terminal(s) most be given.

Topology

A potential link occurs between nodes within range of each other. As indicated earlier, this constraint may be

different for ISL and Earth-satellite links. Let D_{IS} and D_{ES} be the two maximum ranges that correspond to inter-satellite links and Earth-satellite links respectively. Generally we expect $D_{IS} <= D_{ES}$

Suppose that there is a single Earth terminal and M satellites. The potential connectivity matrix G_t, consisting of binary indicators for each possible node pair at time t, may be represented as

$$G_t = \begin{bmatrix} 1 & c_t(0,1) & & \cdots & & c_t(0,M) \\ c_t(1,0) & 1 & c_t(1,2) & & & c_t(1,M) \\ \bullet & & \bullet & & & \\ \bullet & & & \bullet & & \\ \bullet & & & & \bullet & \\ & & & & 1 & c_t(M-1,M) \\ c_t(M,0) & & \cdots & & c_t(M,M-1) & 1 \end{bmatrix}$$

where the Earth terminal corresponds to index = 0, and $c_t(i,j)$ is the symmetric indicator function of being within range (and not occluded), where constraint D_{ES} is used if the index = 0 and D_{IS} otherwise. The binary matrix G_t will remain constant except at occasional discrete time instants when a link will form or die.

Figure 5 below is an example constellation consisting of two sets of co-orbital satellites. This configuration is similar to that proposed for the Radio Tomography Imaging mission [24]. In this example, there are 16 satellites in two coplanar orbits, both having inclination 90 degrees. The inner orbit, containing 6 satellites, has semi-major axis = 62,187 km and eccentricity = 0.89, while the outer orbit of 10 satellites has semi-major axis = 87,700 km and eccentricity = 0.93. There are three Earth terminals which are chosen to be at the Deep Space Network sites. The maximum inter-satellite communications range D_{IS} = 31,891 km while the maximum Earth-satellite communications range D_{ES} = 63,781 km. The figure was generated using the Satellite Orbit Analysis Program (SOAP) tool developed by the Aerospace Corp.

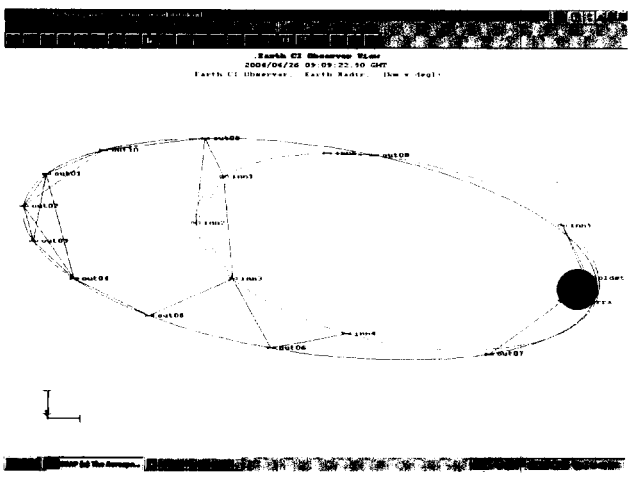


Figure 5. Example Multiple Orbit Constellation

In the following we assume a single terminal fixed on the surface of the Earth. It is noted that if there were multiple Earth terminals, a reasonable model would treat a packet as "delivered" if it is transported to any of the Earth terminals (anycast model).

Traffic Model

We assume that the primary traffic corresponds to a sensor network, in which each sensor node (spacecraft) originates sensor telemetry destined to a single ground station. Thus, the "traffic topology" is a directed star, topology with all traffic flows directed toward the hub (ground station).

Traffic arrivals occur as packets, where one packet may be transmitted in a single time slot. Generally, we provision bandwidth "pipes" to serve source-destination traffic, based on prescribed packet arrival rates. The demand-adaptive scheduling extension of Section 7 is analyzed under the assumption that arrivals occur as a general i.i.d. time process for each spacecraft, are also independent over the set of spacecraft.

A realizable system will also require support for commanding each spacecraft by the ground system. We therefore also consider traffic that is of a dual nature to the model above, in that the directionality of the traffic is reversed. This traffic is also assumed to be i.i.d. over time and over the set of spacecraft, as well as independent of the sensor telemetry traffic.

We also consider the possibility of broadcast command traffic, in which the ground station issues a single broadcast message that is received by all spacecraft. Furthermore, we consider the "dual" of this traffic flow type, which corresponds to the use of "in-network aggregation" (e.g., see [32]). "Aggregated" traffic occurs when multiple messages are received by a forwarding node and combined into a single message that is relayed toward the ground (base) station. This is generally associated with processing, such as summarization or fusion of the sensor information (e.g., maximum value).

In Section 6 we present the combined routing/scheduling that may be applied to these different traffic models. Once one has derived such a schedule for a given traffic type, the "dual" case (same load distribution but in the reverse direction) is trivially derived by simply running the schedule in reverse. Thus there are actually two scheduling algorithms presented.

It is noted that transport of a command to an individual node may be accomplished by issuing the command as a broadcast message in which the message payload indicates the intended recipient, so that all other nodes receiving the message simply discard the message (i.e., filter at a higher

layer). This introduces some degree of throughput inefficiency but is operationally preferable if command messages are relatively small (as is typical). Also, if each node has its own message to transfer to the base station, but the message is small relative to the packet overhead, a reasonable implementation would form a single "aggregate" message consisting of a vector of fields in which each individual node inserts its own data as it is transferred toward the base station. We will utilize this latter method in the signaling portion of the demand-driven case described in Section 7.

Finally, we also consider local traffic which occurs between satellites that are nearby one another. For simplicity, we assume that bidirectional traffic is generated between node pairs whenever they are in range of one another, based on the same range as used for communications (D_{IS}). Thus in this case the "traffic topology" exactly matches that of the physical potential topology.

Error Control

Forward error correction (FEC) as well as automatic repeat request (ARQ) error control techniques will influence the required allocation of bandwidth, as well as impact transport latency performance. We assume that the appropriate factors have been incorporated for link and end-to-end error control overhead.

Performance Metrics

The primary metrics of interest are throughput and latency. Throughput is to be considered in terms of successfully transporting a general end-to-end load vector over all source-destination offered traffic. The maximum capacity of the network occurs at the boundary of total aggregate load over which the system is no longer stable (i.e., queue sizes would grow without bound). Latency is time interval from the instant a new packet arrives at the network node until it completes transport to its final destination.

5. Special Cases

We consider a few special scenarios for which scheduling are simple. These will provide bounds on performance, and offer some intuition in generating heuristics applied to NP-hard general cases. Numerical examples are provided for illustration.

1. No Cross-links Case

Suppose there are no inter-satellite links. Data is stored onboard and each satellite dumps its data to Earth when it is near perigee. Data delivery latency is therefore approximately equal to the orbital period of the spacecraft. For example, for the Magnetospheric Constellation mission, the worst case is

$$T = \frac{2\pi a^{3/2}}{\mu^{1/2}} = \frac{2\pi\left(40\times 6378.137\times 10^3 m\right)^{3/2}}{\left(3.986004\times 10^{14} m^3/s^2\right)^{1/2}} = 1282453.89795s = 14.84 \text{ days}$$

$\left(\mu = GM = 3.986004\times 10^{14}\ \text{m}^3/\text{s}^2, a = 40R_E \text{ for largest orbit in MC mission}\right)$

Note that the storage requirements are substantial, in addition to age of data.

2. Line Topology (without and with parallelism)

Consider the case where each satellite has a single packet to transport to the Earth terminal over the multi-hop network. Assume that the network possesses a line topology as shown below in Figure 5. In this case the average path length that a packet takes is $(M+1)/2$ hops to arrive to its destination, which can be seen to be a worst case topology.

Figure 5. Linear Topology Case

A simple scheduling method proceeds without providing "spatial reuse", i.e. assumes an operation where only one satellite (directional) link is active at a time. For example, may execute a schedule from the outmost node Earthward one link at a time. Suppose that there are M spacecraft. Including the Earth terminal implies that there are $M+1$ nodes in the network. The network delay δ, also referred to as schedule length, is defined as the latency in delivering all M packets to the Earth terminal. Letting R = Data Rate, Λ = packet length (bits); D = maximum link range (assumed to be equal for intersatellite and satellite-Earth links); and c = speed of light, then for this simple schedule

$$\delta = \frac{(M+1)}{2}(M)\left(\frac{\Lambda}{R}+\frac{D}{c}\right)$$

Note that the term

$$\left(\frac{\Lambda}{R}+\frac{D}{c}\right)$$

represents the slot duration.

For illustration, we choose the following numerical values: M=50 (a large distributed spacecraft mission); RF maximum link range $D \sim 2\ R_E$ (where R_E = the radius of the Earth, or about 6371 km), data rate R = 100 kb/s; and packet length Λ = 16 kb, we obtain

$$\begin{aligned}\delta &= \frac{(M+1)}{2}(M)\times\left(\frac{\Lambda}{R}+\frac{D}{c}\right)\\ &= \frac{51}{2}(50)\left(\frac{16\times 10^3}{10^5}+\frac{2\times 6378.137\times 10^3 m}{2.99792458\times 10^8 m/s}\right)\\ &= 1275\times(.1600+.04255) = 258.25\,\text{sec} = 4\,\text{min}\,18.25\,\text{sec}\end{aligned}$$

Therefore, we deliver 50 x 16kb = 800 kb in 258 sec, or .1936 pakcets/sec or 3.098 kb/s.

Note that for this example the overhead for the time guard band is .04255/.20255 = 21%

Next, consider the same linear topology case but now allow simultaneous active links in the schedule. Again assuming each satellite has one telemetry packet, the optimal schedule is straightforward: one activates all "odd" links (counting outward from Earth), then all "even" links, and alternates thereafter until all packets are relayed to the Earth terminal. In this case the schedule length is

$$\delta = (2M - 1) \times \left(\frac{\Lambda}{R} + \frac{D}{c} \right)$$

For the numerical values used in the previous examples,

$$\delta = (99) \times (.2055) = 20.344 \sec$$

for a throughput of about 40kb/s, i.e., an order of magnitude improvement over the non-parallel schedule.

It is noted that use of full duplex transceivers would not improve the schedule length, unless it were possible to transmit and receive with different neighbors (which would be a difficult capability to implement).

3. Star Topology

Now consider a constellation in which the network is formed by a star topology as shown in Figure 6 below.

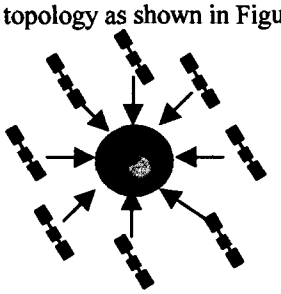


Figure 6. Star Topology Case

This case provides the optimal topology, and the optimal schedule is obvious, where each satellite takes its turn in round robin to send its packet to Earth. In this case, for M satellites, the delay is $\delta = M$ time slots. Using the same values as above, we obtain

$$\delta = M \times \left(\frac{\Lambda}{R} + \frac{D}{c} \right) = 50 \times .20255 = 10.1275 \sec$$

so that the throughput is 50 pkts/ 10.13 sec = 4.94 pkts/sec which is equal to 78 kb/s.

3. Double Line Topology

Suppose that the network topology is as shown in Figure 7 below, where there are M_1 spacecraft forming one chain and the remaining $M_2 = M - M_1$ spacecraft on the other.

Figure 7. Double Line Topology Case

Then the Earth terminal can be kept active for M_{min} = min(M_1,M_2) consecutive slots by alternating the Earth link between the chains, and remaining packets can be delivered in an additional 2 M_{max} - 1 slots, where M_{max} = M - M_{min}. Thus the total schedule length is determined simply by the longer chain length:

$$\delta = 2 \times M_{\max} \times \left(\frac{\Lambda}{R} + \frac{D}{c} \right)$$

In particular, when the linear segments have the same length (M_1=M_2=M/2) then the same optimal performance as the star topology case can be achieved (although more link activations and hence energy is needed than the star case).

6. Scheduling/Routing Method

The network's ability to transport the offered traffic will be realized by defining the combined link activation schedule and routing procedure. The link activation schedule determines for each node for each time slot whether the node should be idle or whether it should communicate with a neighboring node. If the node is to communicate, the schedule also dictates which neighboring node it should communicate with and therefore point its antenna toward, as well as whether it should be in transmit or receive mode. The routing rules determine which packets are forwarded during each transmission slot. The method presented in this section derives the integrated link activation and routing schedule.

For a fixed formation flying mission, the potential topology will be fixed. The potential topology of loosely coupled distributed spacecraft configurations will be dynamic, changing at discrete deterministic instants as the ranges between nodes cross the maximum range threshold. In this latter case, we compute the potential topology and the time interval during which it holds, for each potential topology configuration that arises over the overall period of interest.

We use the approach that the schedule/routing is calculated and used throughout the duration when the potential topology (or "visibility matrix") G_t is constant. This approach has been used in other related works, e.g., [29].

It is noted that we assume that the operational procedure is constant while the potential topology is constant, including adherence to transmissions on slot boundaries. Other works [33-35] incorporate the stronger assumption that inter-node distances are continuously taken into

account to adjust for propagation delay effects. For channels with large time-bandwidth products, the additional complexity in managing transmission instants may be justified in providing additional communications performance.

Our objective is to derive an efficient schedule that provisions bandwidth to each node equal to its prescribed traffic load L_i, $i = 1,\ldots,M$. The types of flows assumed are those described in the "Traffic Model" of Section 4, namely, unicast traffic from each spacecraft to Earth, broadcast traffic from Earth to all spacecraft, and their duals (reverse direction traffic). Initially, we consider the unicast case, which is expected to be characteristic of the primary flows in a sensor network.

The basic problem is to find a schedule that maps the required "traffic topology" onto the network's physical topology. The traffic topology, in all cases considered here, has a star topology nature, with all traffic either terminating or originating at the Earth node. Therefore, when this is optimally overlaid onto the physical topology (having any general form), one can expect that the resulting subgraph generated within the physical topology will be dominated by tree-like structure. Our approach therefore is to begin by pruning the potential topology to derive a tree subgraph, using heuristics to obtain a good tree. We can then determine the best schedule for the given offered load and tree based on existing algorithms that are known to be optimal for tree structures.

Unicast Scheduling

We derive the schedule for sensor traffic that originates at each spacecraft and is destined for Earth. Assume that the amount of bandwidth allocated to spacecraft i is denoted L_i, $i = 1,\ldots,M$. (Even distribution of bandwidth occurs as a special case where L_i is constant with respect to i.)

<u>*Step 0.*</u> Assume that the potential topology has been derived (see Section 3); for a dynamic topology case this will be valid over a specific time interval during which the potential topology is constant.

<u>*Step 1.*</u> Using the connectivity matrix G, we determine all nodes that are directly connected to Earth. Suppose there are B such nodes, indexed in some fashion over $b = 1,\ldots,B$. We will partition the potential topology, excluding Earth, into min-hop tree subgraphs, one for each of the B nodes that connects directly to Earth; we refer to the B subgraphs as branches. We work outward from Earth, adding a node at a time, after which the node is marked as processed for this step. Every node has an associated attribute vector indicating to branch it has been assigned and its hop distance from Earth; initially the attributes are set to 0 to indicate that the node hasn't yet been assigned to a branch. Let h be the min-hop distance for a node to Earth. Then B is the number of nodes at distance $h = 1$; attribute the branch ID and hop distance accordingly for these nodes. The procedure begins with $h = 1$.

For each branch $b = 1,\ldots, B$,
For each node in branch b at distance h,
Using the connectivity matrix G, find all unassigned nodes that are directly connected to it and assign them to that branch together with hop distance h+1.
Increment h and loop.
The process terminates when all nodes are assigned. The maximum hop distance for each branch $h_{max}(b)$ and the overall graph h_{max} (graph diameter) are also determined in this process. It is noted that this step does not depend on the offered load, but only on the potential topology. Subsequent steps require the load distribution $\{L_i, i = 1,\ldots M\}$.

<u>*Step 2.*</u> Now we work inward toward Earth to determine an additional attribute $L_{subtree}$ for each node that represents the total load associated with the subtree formed by the node and all its children. The procedure begins with $h = h_{max}$.

For each branch $b = 1,\ldots, B$,
For each node at distance h, initially set its $L_{subtree}$ value to its own load L_i. Then, add the $L_{subtree}$ value for each of its immediate children, which are determined by being directly connected (from G), at distance h+1, and assigned to the same branch b.
The procedure terminates when all nodes (except Earth) are processed. In particular, it will result in determining the load associated with each branch, namely the $L_{subtree}$ value for each of the B nodes with $h = 1$. We denote these values by $L_{branch}(b)$, $b = 1,\ldots, B$.

<u>*Step 3.*</u> The total load offered to the network is denoted L_{total}. If $B = 1$ (only one branch) or if $L_{branch}(b) <= L_{total}/2$ for all $b = 1,\ldots, B$, then skip to Step 4. Otherwise, this step will attempt to improve the balancing of the loads across the branches by moving nodes (subtrees) between them.

a.) We begin by finding the branch b_{max} having the maximum load $L_{branch}(b_{max})$. If this branch's hop-distance $h_{max}(b_{max}) = 1$ then this step terminates (go to Step 4). Otherwise
for $h = 2,\ldots, h_{max}(b)$
For all nodes in branch b_{max} that are h hops from Earth, process them in the order of decreasing $L_{subtree}$ value. Denote the node n^*. and denote its subtree value as $L_{subtree}(n^*)$.
For this node n^*, determine from G whether it has any potential links connecting to a node in any other branch. If not, go to the next node.
If there is at least one link connecting to another branch, choose the link that connects to the branch having the smallest branch weight. Denote this branch as b_{conn}, and let the node in b_{conn} that shares the link with n^* be denoted

n_{conn}. The weight of the connecting branch is denoted $L_{branch}(b_{conn})$.

We decide to move the subtree rooted at n^* from branch b_{max} to branch b_{conn} if

$$L_{branch}(b_{max}) - L_{branch}(b_{conn}) > |L_{branch}(b_{max}) - L_{branch}(b_{conn}) - 2L_{subtree}(n^*)|$$

which implies that the "out-of-balance" excess of the maximum weight branch will improve (and may achieve the minimum). (Remark: the algorithm may be improved by allowing equality at this step, provided measures are taken to prevent unending loops. Another possible refinement would consider the branch hop-length h_{max}, as this can have a secondary impact on schedule length) If the inequality above is not satisfied, then step through the loop to consider the next candidate node for n^*.

If the above inequality holds, then we move the subtree rooted at n^* from branch b_{max} to branch b_{conn}, connecting to node n_{conn}. This requires that the attributes associated with all affected nodes are updated The affected nodes are (1) the ancestors of n^* in branch b_{max} [update $L_{subtree}$], (2) the subtree rooted at n^* (now in branch b_{conn}) [update h], (3) node n_{conn} [update $L_{subtree}$], and (4) all ancestors of node n_{conn} [update $L_{subtree}$]. In each case, the L We also update at the branch level: updating the branch weights by adding/subtracting $L_{subtree}(n^*)$ to/from branches b_{conn}/b_{max}, and updating h_{max} for both as required. Once a subtree move (including a single leaf) has been made, the loop breaks out and returns to the test at Step 3a).

If we loop through all nodes in branch b_{max} without finding a subtree to move, this Step terminates and we go to Step 4.

<u>Step 4.</u> We now have found a subgraph of the original potential topology that is a tree rooted at Earth. We may now find the schedule that will enable the given load to be relayed to Earth using the algorithm of Florens et. al [36]. This algorithm is known to be optimal in providing the minimum-length schedule for any load distribution over any tree. The schedule is derived from the root (Earth) outward as though the traffic flow occurs as a one-to-each dissemination from the root; since we desire the "dual" case where flows are actually inbound, the final schedule we desire is simply found by reversing its temporal order of execution.

<u>Operation and Performance:</u> Let F be the schedule length. The schedule is executed repeatedly. Effectively we have a network TDMA frame of length F that accommodates each node transmission or reception and between whom. This results in a constant bandwidth that will carry L_i packets from node i to Earth every F time slots. Random traffic may be offered by each node and arrival bursts locally queued, with stable performance resulting for node i provided its arrival rate is less than L_i/F packets per slot (and each node adheres to commutating relay traffic in its allocated slots). Generalization to include other traffic types is achieved by interleaving into a superframe structure.

Related Work

The network architecture and scheduling/routing algorithms above may be compared with those that have been developed for general wireless sensor networks and ad hoc networks. Wireless sensor networks typically have traffic characterized by flows from each sensor node to a central base station, which corresponds to the traffic assumption for the unicast scheduling presented above. Traffic for ad hoc networks is usually assumed to have general any-to-any characteristics.

In ad hoc networks, nodes are assumed to move unpredictably; and therefore procedures must be introduced to handle the random dynamics of the topology. Nodes in sensor networks are typically assumed to be non-mobile. While space-based sensor networks may be relatively non-mobile (i.e., formation flyers), generally the nodes move, albeit along predictable orbit(s) generated from predetermined parameters.

A key impact of the assumptions on mobility are that a schedule-driven networking approach is amenable when the nodes are non-mobile (hence portray a fixed topology) or follow predictable motions that permit a priori computation of the topology dynamics. The unpredictability of topology dynamics in ad hoc networks justifies the use of random access approaches, as in 802.11 wireless LANs operating in DCF mode.

As indicated earlier, the large inter-satellite distances give rise to use of directional antennas. Although wireless network problems have been studied extensively with the assumption of omni-directional antennas, directional antennas in ad-hoc and sensor networks have recently gained interest. Examples of research developments in directional ad hoc networks are [30-31, 37-43]. For example, in [40] and [41], the 802.11 MAC handshaking mechanisms (Request-to-Send and Clear-to-Send) are modified to accommodate the use of directional antennas, resulting in performance improvements.

A schedule-driven solution has been presented in [44] (see also [45]) of an ad hoc network using directional antennas. The topology is assumed known, and node mobility is not considered. While we focus traffic patterns arising in the context of sensor networks, they consider a general traffic flow model, with flow rates parameterized for every possible source-destination pair. We follow a similar approach to theirs, in that initially path routes are found, and then efficient schedules are generated. They use a shortest cost routing algorithm to generate a graph, where different cost metrics (such as energy use) may be used. The resulting graph (analogous to our tree) is then scheduled using maximum weighted

matching. Our heuristic for deriving the routing structure utilizes the nature of the sensor network traffic assumption by crafting a tree structure within the procedure. Our investigations of several example cases have found that our algorithm provides a noticeable improvement in performance (for sensor network traffic). However, further improvements are anticipated and being pursued in ongoing research.

In [44], the potential communication graph is pruned down to a shortest cost routing graph where the cost is defined as a function of the energy consumption of the specific paths; optimal routes are chosen by directing traffic along paths which are shortest with respect to link lengths that also depend on the flow carried by the links. Since our goal is not to minimize energy consumption but delay, we prune the potential communication graph into a shortest path (min-hop) graph. Our traffic topology is a star, so it is natural to use a tree as the routing structure where in [44] the routing structure is a directed acyclic graph (DAG). The similarity between a tree and a DAG is that there is only one route between any source and destination pair, i.e., neither algorithm uses multiple paths to route packets between the same source and destination pair (traffic balancing). We may consider a tree to be a special case of a DAG, where a DAG may contain more links than a tree, in general.

In terms of scheduling, [44] computed the amount of time each individual link need to stay up and then applied maximum weighted matching to obtain the schedule. More specifically, the links are partitioned into disjoint sets such that each set is a matching. When the set of matching is produced that covered all the links, a schedule is obtained by giving x time slots to each matching where x is the longest duration a link need to stay up in that particular matching. Note that all the links in the same matching can be scheduled to transmit during the same x time slots. In our algorithm, we applied the optimal tree scheduling algorithm of [36] to obtain the schedule. One key difference between the scheduling algorithm of [44] and [36] is that the first algorithm would produce a schedule that keeps an antenna pointing in the same direction as long as possible to minimize the energy needed to re-point the antenna. The implication is that large quantities of packets must be stored at intermediate nodes during the multi-hop relay. However, the scheduling algorithm we use will re-point as often as possible to minimize the buffering of packets. In essence, this mechanism works as a pipeline to achieve minimum delay for the packets.

16-Satellite Example of Unicast Scheduling

We present here a numerical example to illustrate the scheduling method described above. Assume that there are M = 16 satellites that are evenly distributed geometrically along a single orbit. There is a single Earth terminal, so that the total number of nodes in the network is $M+1$ = 17.

The orbit has zero inclination, has eccentricity = 0.5, and the semi-major axis is 21378 km. The orbit period is found to be P = 8 hr 38 min 56 sec.

The Earth terminal is located on the equator. It is noted that the network topology dynamics are largely periodic with period equal to $P/16$ = 31136/16 = 1946 sec = 32 min 26 sec., corresponding to the duration when satellites recur in similar positions but re-indexed, however, perturbations arise due to the rotation of the terminal on Earth. Periodicity could be enforced by adjusting the orbit period to be harmonic with the rotation period of the Earth.

The maximum link range for both inter-satellite and Earth-satellite links is assumed to be D = 20,000 km.

For a single orbit period and particular Earth terminal phasing, we find that a total of eight distinct potential topologies arise for this case. The initial potential topology is illustrated in Figure 8 below.

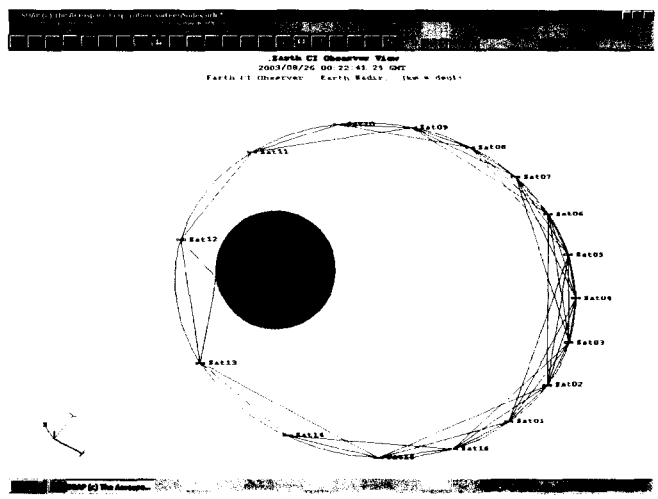


Figure 8. Potential Topology for Example Case

The topology of Figure 8 is equivalent to that depicted in the first (top left) topology of Figure 9. The remaining graphs shown correspond to the rest of the potential topologies that arise in one orbit period. On each figure, the corresponding relative time interval is also provided, where time 0 corresponds to a topology change instant. Generally exactly one link is added or deleted between adjacent topologies. The shortest length intervals occur for topologies 3 and 5 (with durations of 12 sec and 17 sec respectively).

Our goal is provision bandwidth for the sensor telemetry across all satellites. We assume that even distribution of bandwidth is desired, so that L_i = constant over all i.

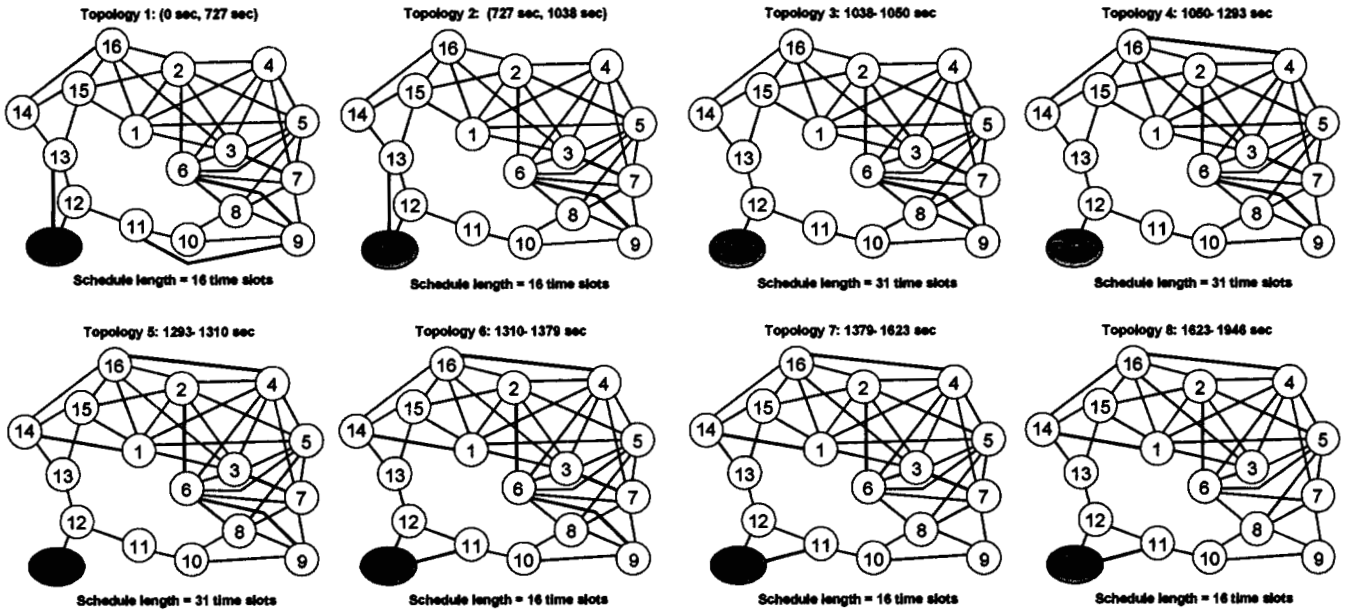


Figure 9. Potential Topologies During a Single Orbit

We derive the schedule for the first topology shown. Similar schedules can be found for the other eight time intervals during which the topology is constant. We expect that the time to go through one execution of a schedule (perhaps 10s of slots) is small compared to the duration during which the schedule remains active (i.e., the interval during which the topology is unchanging). Also, since we expect some "softness" in the maximum link ranges, the issue of synchronizing schedules across topology changes is assumed to be manageable and is not further considered here.

The topology of interest is shown in Figure 10 below. Links that do not lie on a shortest-hop path from some node to Earth are shown as dotted lines, and are pruned away.

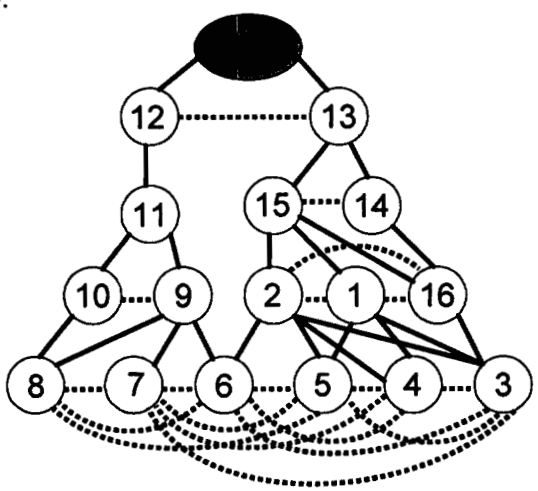


Figure 10. Potential Topology for 16-Satellite Example

Because we assume even bandwidth provisioning over all satellites, the algorithmic step that prunes the graph to a tree results in this case in that shown in Figure 11.

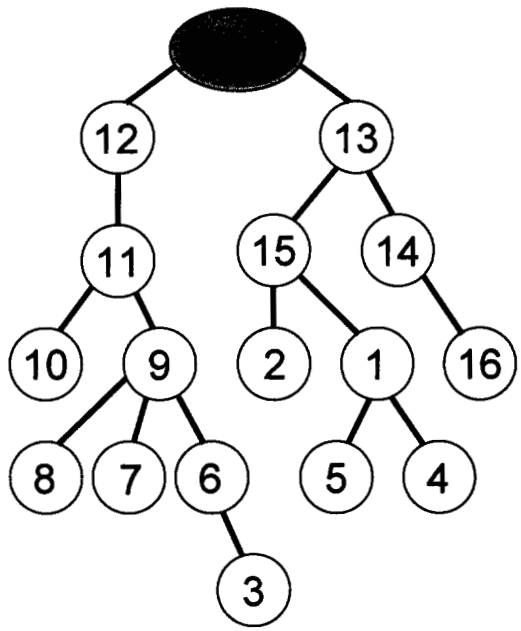


Figure 11. Tree Subgraph Derived from Algorithm

Application of the algorithm of Florens and McEliece results in the schedule having length 17 time slots given by Table 1. An entry in the table indicates the satellite (or Earth if the entry is "E") that would receive a transmission from the satellite associated with the entry's column.

Time\Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1												E				
2											12		E			
3									11			E			13	
4	15					9					12		E			
5				1	6				11			E		13		
6						9					12		E			14
7			16							11		E			13	
8		15						10			12		E			
9									11			E		13		
10							9				12		E			14
11	15									11		E			13	
12											12		E			
13									11			E			13	
14											12		E			
15												E		13		
16													E			

Table 1. Unicast Schedule for Example Topology

In comparison, we applied the algorithm of [44] and found the schedule to have length 19. Thus both our algorithm and that of [44] are efficient.

We similarly derive the schedules for each of the remaining topologies that arise during the course of one orbit period. The schedule lengths for all eight topologies are provided in Table 2, together with the time-averaged schedule length (i.e., weight each summation term by its proportion of the orbit period), which is found to be 18.95 time slots.

Topology	Duration	Schedule length (even provisions)
1	12 min 7 sec	16 slots
2	5 min 12 sec	16 slots
3	0 min 12 sec	31 slots
4	4 min 0 sec	31 slots
5	0 min 17 sec	31 slots
6	1 min 9 sec	16 slots
7	4 min 3 sec	16 slots
8	5 min 23 sec	16 slots
Overall Mean Schedule Length		18.08 slots

Table 2. Schedule Lengths over Entire Orbit Period

The time-averaged throughput capacity of the system is therefore one packet originating at each of the 16 satellites and all destined to Earth every 18.08 time slots, or 16/18.08 = .885 packets per slot. (Recall that there is additional overhead associated with the time guard band within each slot.) For example, if we use the communications parameters of Section 5, where

- data rate R = 100 kb/s and
- packet length = 16 kb,

together with the assumed maximum link range of 20,000 km, we find that the duration of a time slot is .227 sec so that the network throughput of .885 packet per slot corresponds to 62.5 kb/s (source-to-destination carried load).

Scheduling for Broadcast/Aggregate Traffic

We now present a straightforward modification of the scheduling method presented earlier in this section to accommodate in-network aggregation (aggregate-to-one) traffic. This may be applied to derive the dual case of broadcast (one-to-all) traffic by simply reversing the order of execution of the schedule for the aggregate traffic case.

Two changes are made to earlier scheduling procedure. First, a modification is made to "Step 3b" regarding the metric associated with a path from a node to Earth. For the traffic aggregation case, no matter what the number of incoming links is, the total outgoing load is simply one, because the traffic will be aggregated. Otherwise, the pruning to derive a tree proceeds as before.

Once we have the desired tree, we replace "Step 4" by the following simple algorithm to derive the schedule: We determine which nodes will transmit on the nth time slot of the schedule, beginning with $n = 1$. Consider all leaves of the tree. For all leaf nodes having a common parent, choose one arbitrarily to transmit during this n^{th} slot. (Any node with no siblings will be selected to transmit.) Now remove all nodes from the tree that have been assigned to transmit, increment n, and repeat this process until no nodes remain.

Recall that we assume that inter-node distances are large and directional antenna beamwidths sufficiently narrow that multi-access/broadcast physical layer characteristics do not arise. Thus methods such as those presented in [46] cannot be used to improve the schedule efficiency.

16-Satellite Example of Broadcast Scheduling

We present here a numerical example to illustrate the scheduling method for broadcast traffic, using the same 16-satellite mission scenario as earlier. Again we concentrate on the topology of Figure 8; the schedules for broadcast/aggregate traffic for the other topologies are computed similarly.

Node / Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1									8		12		1	2		
2	3									8		16			2	
3		4						7								3
4			5	6												
5					E	6										
6						E										

Table 3. Schedule for Aggregated Traffic, 16-Satellite Example

Scheduling for Local Traffic

The schedule that is formed for the local inter-satellite traffic is generated using a maximum matching algorithm. For the 16-satellite case having the same topology of

Figures 9a and 10, we find that all local traffic may be serviced in 7 slots. Figure 12 below illustrates the schedule by coloring the links, with all the links of the same color being active during a single slot (any ordering of the colors may be used).

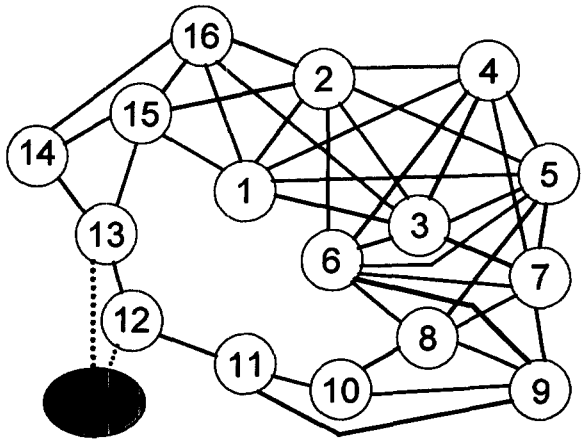


Figure 12. Link Activation Coloring for Local Traffic

7. Traffic-Adaptive Extension

The above fixed-provisioning type of scheduling is expected to provide excellent performance traffic sources that generate packets at relatively even rates. However, when traffic is spatio-temporally bursty, one may pursue greater system performance by attempting to allocate resources dynamically to satisfy short-term demand conditions.

Such is our goal in this section. For simplicity, assume that the topology is constant (formation flyers). Our approach is to generalize the operation from periodic repetition of a fixed schedule corresponding to a TDMA frame to operating with dynamic frame lengths. This will be accomplished by extending the frame structure into two subframes: one for "signaling" (our "reservations") and one for actual data transfer. The general form of system operation is depicted by Figure 13.

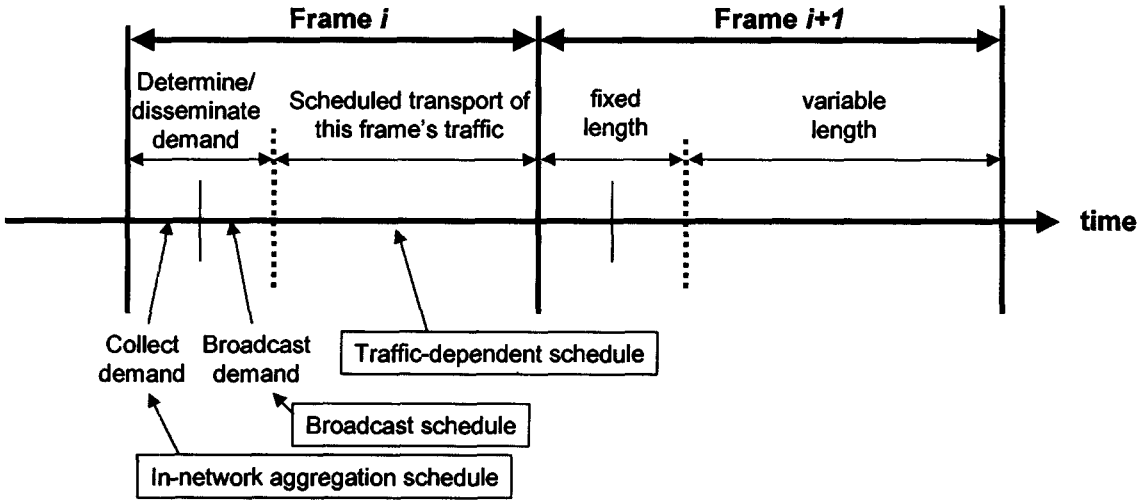


Figure 13. Demand-Driven Network Operation

The "signaling" subframe consists of two parts. The first part essentially polls all the satellites to determine their current load. The process proceeds by first collecting the total load vector at the Earth terminal. This could be accomplished by each satellite sending a single packet with its current traffic load state. However, one expects that the entire load vector can be represented in a small number of bits relative to the overhead associated with a single packet (e.g., the time guard band). Therefore we propose that such a representation be used, with each node updating the packet with its own state information. The appropriate schedule is then for aggregated traffic, which is more compact.

The second part of the signaling subframe is used to broadcast the global load vector to all nodes. Each node would then determine the global schedule based on the given global load vector. This schedule would be customized to carry exactly the given load vector, with minimum schedule length. (Remark: Alternatively, the ground station could compute the schedule and broadcast that instead of the load vector, thereby relieving the satellites of the processing task to generate the schedule.)

The second subframe of the frame would be used to transport all traffic that was pending at the onset of the current frame. The next frame will begin with a load vector created from all new packet arrivals that occurred during the frame preceding that time instant.

The performance benefits of the traffic adaptive scheme will depend on the burstiness of the offered traffic both in space and time. To obtain an indication of the potential magnitude of the improvement, we consider the following traffic model:

$$L_i = L_{even} \quad \forall i \neq m^*$$
$$L_{m^*} = L_{even} + L_{burst}$$

Therefore, all nodes have an "even" amount of traffic (L_{even} each), and one particular node, designated m^*, has an additional amount of traffic L_{burst}. The total amount of traffic is therefore

$$L_{Total} = M \times L_{even} + L_{burst}$$

We assume that the node having the additional burst of traffic changes slowly and unpredictably, moving uniformly among all the satellites, and therefore a predefined schedule cannot be used. The static schedule that provisions a single slot to each satellite is designated S_{unit}. Let $|S_{unit}|$ denote the length of schedule S_{unit} then we know that for any physical topology $|S_{unit}| \leq M$ (equality occurring, e.g., for a star topology). We may clear all the traffic by using schedule S_{unit} repeatedly $L_{even} + L_{burst}$ times (corresponding to the load L_{m^*}). Therefore a lower bound on the time to clear all traffic using fixed, even provisioning is $|S_{fixed}| \geq M(L_{even} + L_{burst})$.

If on the other hand we use traffic adaptive scheduling, the schedule can be tuned to accommodate the additional load at node m^*. One method (which is not necessarily optimal) would use schedule S_{unit} repeatedly L_{even} times to clear all the "even" traffic, and then use schedule S_{burst} that is designed to clear the remaining traffic of node m^*. The worst case (longest length) schedule S_{burst} to clear traffic from a single node occurs with a line topology, for which the length is $(2M-1)L_{even} + h^* + 2(L_{burst}-1)$, where h^* is the hop distance from node m^* to Earth. Since $h^* \leq M$, we have that an upper bound on the traffic adaptive schedule length is $|S_{adapt}| \leq M\, L_{even} + (2M-1)\, L_{even} + M + 2(L_{burst}-1)$. However, the traffic-adaptive technique also requires the the additional overhead of the signaling subframe needed to collect and disseminate the current demand. The signaling messages may be substantially smaller than telemetry packets, so that smaller slot sizes could be used for greater efficiency. Also, the aggregation and broadcast schedules are generally considerably shorter than the unicast schedule, although the length is roughly dependent on the number of satellites M. In the following, we approximate the length of the signaling subframe by M, in order to obtain a general indicator of performance impact.

Bounds on throughput are easily obtained from the total carried load L_{Total} divided by the schedule length bounds ($|S_{fixed}|$ or $M+|S_{adapt}|$). The relative gain through the use of the traffic adaptation is the ratio of the demand-driven case over the fixed scheduling method. The relative gain throughput by using the demand adaptive technique is depicted in Figure 14 below in terms of the relative spatial "burstiness" in the traffic

$$Burstiness = \frac{L_{burst}}{L_{Total}}$$

The figure depicts performance for the number of satellites M = 10, 20 and 40, and for total offered load of L_{Total} = 10 or 40. Performance gain occurs whenever the throughput ratio exceeds 1.

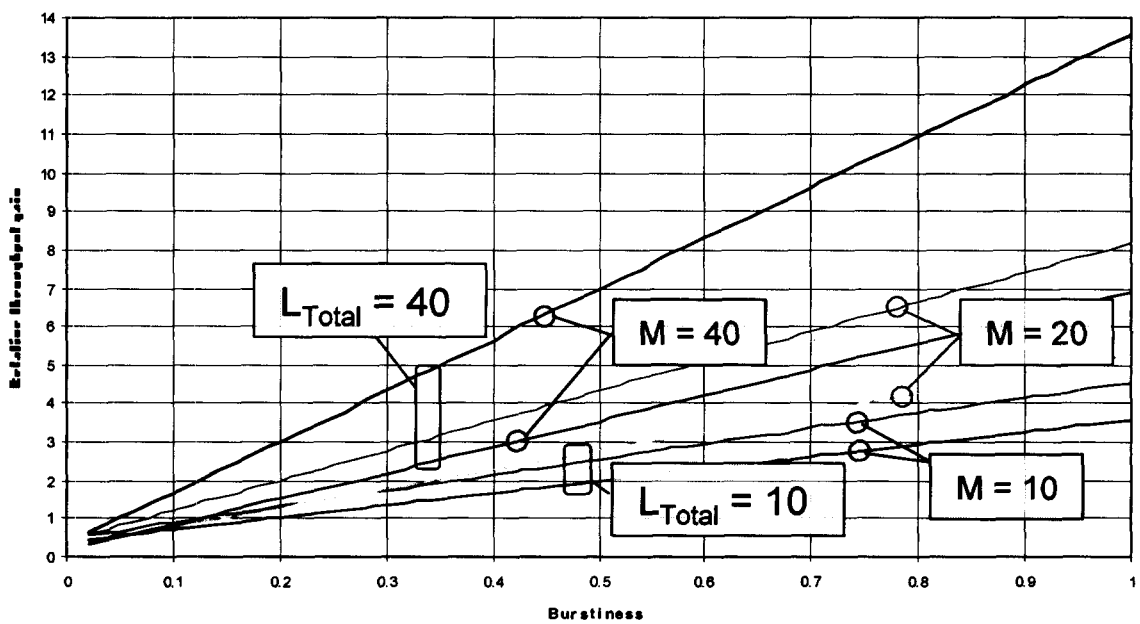


Figure 14. Relative Throughput Gain Using Adaptation

Further capacity analysis of the network under the demand-driven operation is ongoing at the time of this writing. Under the assumptions of the arrival process possessing independent increments, the sequence of frame lengths forms a Markov chain. Stable operation will occur if this Markov chain is positive recurrent and possesses a stationary probability distribution.

8. Conclusions and Future Work

In this paper, we presented a network architecture for space-based sensor networks. Components of the system model include node motions and topology dynamics, use of inexpensive half-duplex radios with directional antennas for crosslinks, network timekeeping, and a protocol that coordinates each node's operation for each time slot (point antenna to neighbor x and transmit, point antenna to neighbor y and receive, or be inactive).

An algorithm was presented for efficiently scheduling the communications resources to satisfy the sensor network traffic with minimum latency. The approach derives a tree with load balanced over its main branches, and leverages the Florens and McEliece algorithm for scheduling tree networks. A numerical example was illustrated. The algorithm is not limited to sensor networks that are space-based.

The network operation was generalized to be traffic-adaptive, by incorporating a signaling capability that allowed the network to autonomously allocate resources dynamically to meet current demand.

There are a great number of opportunities for extensions of this work. For example, the assumption on slot synchronization will clearly need to be addressed for higher data rates. In relaxing assumptions, other applications may be considered, such as sensor networks deployed on the ground (surface). Ongoing work is investigating refinements to better accommodate specific traffic needs of precision formation flying mission, including real-time control traffic and multi-way exchange of navigation data types.

Contractual Acknowledgement

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology and was sponsored by the Army Research Laboratory Advanced Sensors Collaborative Technology Alliance and the National Aeronautics and Space Administration.

References

[1] D. N. Baker and R. Carovillano, "IASTP and Solar-Terrestrial Physics," *Section 1. IASTP Missions and Programs* of RESULTS OF THE IASTP PROGRAM Volume 20, Issues 4-5, *Proc. of the Symposium D0.1 of COSPAR Scientific Commission D, Thirty-First COSPAR Scientific Assembly, Birmingham, United Kingdom, 14-21 July, 1996,* http://www-ssc.igpp.ucla.edu/IASTP/01/

[2] http://nssdc.gsfc.nasa.gov/database /MasterCatalog?sc=1977-102A

[3] http://www-ssc.igpp.ucla.edu/ssc/isee.html

[4] The Sun to the Earth – and Beyond, A Decadal Research Strategy in Solar and Space Physics, 2003, National Academies Press., Washington D.C.

[5] 2003 Space Science Enterprise Strategy, available at http://spacescience.nasa.gov/admin/pubs/strategy/2003 /index.html

[6] D. Estrin, L. Girod, G. Pottie, M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2001),* Salt Lake City, Utah, May 2001.

[7] J. R. Agre and L. P. Clare, "Distributed Sensor Networks," IEEE Computer, vol. 33. No. 5, pp. 106-108 (Integrated Engineering column), May 2000.

[8] http://stp.gsfc.nasa.gov/missions/mc/mc.htm

[9] http://stp.gsfc.nasa.gov/missions/mms/mms.htm

[10] http://stp.gsfc.nasa.gov/missions/gec/gec.htm

[11] http://lws.gsfc.nasa.gov/geospace.htm

[12] http://gpm.gsfc.nasa.gov/

[13] http://sci.esa.int/science-e/www/area /index.cfm?fareaid=8

[14] http://www.esa.int/export/esaLP/ace.html

[15] http://www.esa.int/export/esaLP/swarm.html

[16] http://planetquest.jpl.nasa.gov/TPF/tpf_index.html

[17] http://hires.gsfc.nasa.gov./%7Esi/

[18] http://lisa.jpl.nasa.gov/

[19] http://maxim.gsfc.nasa.gov/

[20] http://sgra.jpl.nasa.gov/html_dj/ALFA.html

[21] http://constellation.gsfc.nasa.gov/

[22] S. P. Hughes and L. M. Mailhe, "A preliminary formation flying orbit dynamics analysis for Leonardo-BRDF," in Flight Mechanics Symposium, (NASA Goddard Space Flight Center, Greenbelt, MD), 2001.

[23] http://stp.gsfc.nasa.gov/missions/stereo/stereo.htm

[24] R. E. Ergun et al., "Radio Tomography Mission Concept," Science Closure and Enabling Technologies for Cosntellation Class Missions, ed. by V. Angelopoulos and P. V. Panetta, UC Berkeley, CA, 1998, http://sprg.ssl.berkeley.edu /ConstellationClassMissions/ergun.pdf

[25] http://www.gsfc.nasa.gov/topstory/2003 /0618rosettaburst.html

[26] M. D. Noakes, J. B. Cain, J. W. Nieto and E. L. Althouse, "An adaptive link assignment algorithm for dynamically changing topologies", IEEE Trans. On Commun., Vol. 41, No. 5, pages 694-706, May 1993.

[27] K. Harathi, P. Krishna ,R. E. Newman-Wolfe, and R. Y. C. Chow, "A fast link assignment algorithm for satellite communication networks", Proc. 12th Ann. Int'l. Phoenix Conf. on Computers and Commun., 23-26 March 1993, pp. 401-408.

[28] J. Huang et al., "Crosslink Slot Assignment in LEO Networks," IEEE Trans. on Aerospace and Electronic Systems, Jan. 1997.

[29] Hong Seong Chang, Byoung Wan Kim, Chang Gun Lee, Sang Lyul Min,Yanhee Choi, Hyun Suk Yang, Doug Nyun Kim and Chong Sang Kim, "FSA-Based Link Assignment and Routing in Low-Earth Orbit Satellite Networks", IEEE Transactions on Vehicular Technology, vol. 47, no. 3, pages 1037-1048, Aug, 1998.

[30] Lichun Bao and J. J. Garcia-Luna-Aceves, "Transmission scheduling in ad hoc networks with directional antennas", Proceedings of ACM MOBICOM 2002, Sept. 23-26, 2002, Atlanta, GA.

[31] C. Peraki and S. Servetto, "On the Maximum Stable Throughput Problem in Random Networks with Directional Antennas," MobiHoc'03, June 1-3, 2003, Annapolis, MD.

[32] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann, "Impact of network density on data aggregation in wireless sensor networks," International Conference on Distributed Computing Systems (ICDCS-22). November 2001.

[33] R. Binder et al., "Crosslink architectures for a multiple satellite system," Proc. IEEE, Jan 1987.

[34] R. P. Kosowsky, I. M. Jacobs and K. S. Gilhousen, "ARNS: A New Link Layer Protocol," IEEE MILCOM'88, October 1988, pp. 515-519.

[35] M. Carr and B. Hajek, "Scheduling with asynchronous service opportunities with applications to multiple satellite systems," IEEE Trans. Automatic Control, Dec. 1993.

[36] Cedric Florens and Robert McEliece, "Scheduling algorithms for wireless ad-hoc sensor networks," Proceedings of IEEE GLOBECOM 2002, Dec. 1-5, 2002.

[37] J. Zander, "Slotted ALOHA Multihop Packet Radio Networks with Directional Antennas," Electronics Letters, vol. 26, no. 25, December 1990, pp. 2098-2099.

[38] Marvin Sanchez, Jens Zander and Tim Giles, "Combined routing and scheduling for spatial TDMA in multihop ad hoc networks," 2nd Swedish Workshop on Wireless Ad-Hoc Network Proceedings, March 5-6, 2002.

[39] T.-Y. Liu and R. A. Scholtz, "Link Activation Protocols for Mobile Communication Network with Directive/Adaptive Antennas," *IEEE Trans. on Commun.*, vol. 48, no. 1, January 2000.

[40] A. Nasipuri, S. Ye, and R. E. Hiromoto, "A MAC Protocol for Mobile Ad Hoc Networks Using Directional Antennas," Proc. IEEE Wireless Communications and Networking Conference (WCNC), Chicago, IL, Sep. 23-28 2000.

[41] Young-Bae Ko, V. Shankarkumar and N. H. Vaidya, "Medium access control protocols using directional antennas in ad hoc networks", Proceeding of INFOCOM 2000. Nineteenth Annual Joint Conference, pages 13-21, March 26-30, 2000.

[42] Kang Wang, Carla F. Chiasserini, Ramesh R. Rao and John G. Proakis, "A distributed joint scheduling and power control algorithm for multicasting in wireless ad hoc networks," Proceedings of IEEE ICC 2002, May 11—15, 2003.

[43] S. Yi, Y. Pei, S. Kalyanaraman, "On the Capacity Improvement of Ad Hoc Wireless Networks Using Directional Antennas," MobiHoc'03, June 1-3, 2003, Annapolis, MD, pp. 108-116.

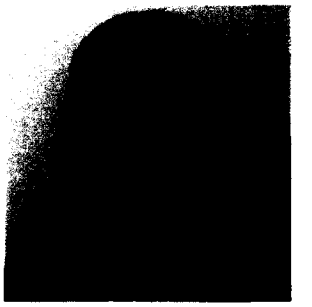
[44] Akis Spyropoulos and Cauligi S. Raghavendra, "Energy efficient communications in ad hoc networks using directional antennas," Proc. IEEE Infocom 2002.

[45] Akis Spyropoulos and Cauligi S. Raghavendra, "Asymptotic capacity bounds for ad-hoc networks revisited: the directional and smart antenna cases," to be published in Proceedings of IEEE GLOBECOM 2003, Dec 2003.

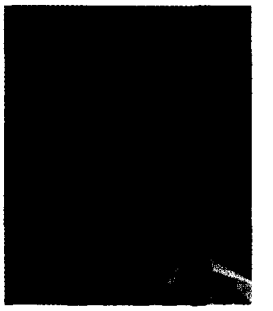
[46] J. E. Wieselthier et al., "Energy-Aware Wireless Networking with Directional Antennas: The Case of Session-Based Broadcasting and Multicasting," *IEEE Trans. on Mobile Computing*, vol. 1, no. 3, July-Sept. 2003, pp. 176-191.

Loren Clare *is the Technical Group Supervisor for the Communications Networks Group at the Jet Propulsion Laboratory. He obtained the Ph.D. in System Science from the University of California, Los Angeles in 1983. His research interests include wireless communications protocols, self-organizing systems, network systems design, modeling and analysis, and realtime distributed control systems. Prior to joining JPL in May 2000, Dr. Clare was a senior research scientist at the Rockwell Science Center. Recent work focused on distributed sensor networks, including low-power protocol design as well as development of a layered architecture for cooperative signal processing. He has extensive experience in tactical network design and analysis, as well as satellite networking. He has also developed communications protocols for realtime networks supporting industrial automation.*

Jay L. Gao *is a research staff with the Communication Systems and Research Section of the Jet Propulsion Laboratory. He is currently working on protocol development and performance evaluation for in-situ surface-to-surface and surface-to-orbit communications in planetary missions, and sensor network research and modeling. He received his B.S., M.S., and Ph.D. in Electrical Engineering from UCLA in 1993, 1995, and 2000, respectively.*



Esther Jennings *is a research staff with the Communication Systems and Research Section of the Jet Propulsion Laboratory. She received a Ph.D. in Computer Science from Luleå University of Technology, Sweden in 1997. From 1997-1999, she was a postdoctoral fellow at the Industrial Engineering Department at Technion, Israel Institute of Technology. From 1999-2001, she has been an assistant professor at the Computer Science Department of California State Polytechnic University, Pomona. Her research interests are in distributed graph algorithms, reliable multicast protocols, energy-efficient algorithms for wireless network and algorithm simulations.*



Clayton Okino *received a BS in Electrical Engineering at Oregon State University in 1989, a MS in Electrical Engineering at Santa Clara University in 1993, and a Ph.D. in Electrical and Computer Engineering from the University of California, San Diego in 1998. After receiving his Ph.D., Dr. Okino accepted a position as an assistant professor in Thayer School of Engineering at Dartmouth College, where he pursued research in communication and wireless networks, emphasizing on performance and security. In 2001, Dr. Okino accepted a position as a Senior Member of the Technical Staff in the Digital Signal Processing group at Jet Propulsion Laboratory where his current research is in wireless link layer, network, and access algorithms, reconfigurable sensors, wireless QoS and location based processing techniques.*