Wireless Networks 11, 507–521, 2005

A Routing Protocol for Hierarchical LEO/MEO Satellite IP Networks

CHAO CHEN
Broadband & Wireless Networking Laboratory, School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332

EYLEM EKICI*
Department of Electrical & Computer Engineering, The Ohio State University, Columbus, OH 43210

Abstract. The rapid growth of Internet-based applications pushes broadband satellite networks to carry on IP traffic. In previously proposed connectionless routing schemes in satellite networks, the metrics used to calculate the paths do not reflect the total delay a packet may experience. In this paper, a new Satellite Grouping and Routing Protocol (SGRP) is developed. In each snapshot period, SGRP divides Low Earth Orbit (LEO) satellites into groups according to the footprint area of the Medium Earth Orbit (MEO) satellites. Based on the delay reports sent by LEO satellites, MEO satellite managers compute the minimum-delay paths for their LEO members. Since the signaling traffic is physically separated from the data traffic, link congestion does not affect the responsiveness of delay reporting and routing table calculation. The snapshot and group formation methods as well as fast reacting mechanisms to address link congestion and satellite failures are described in detail. The performance of SGRP is evaluated through simulations and analysis.

Keywords: satellite network, connectionless routing, performance, Low Earth Orbit (LEO), Medium Earth Orbit (MEO)

1. Introduction

Satellite systems have the advantage of global coverage and inherent broadcast capability, and offer a solution for providing broadband access to end-users. Compared to Geostationary (GEO) satellites, Low Earth Orbit (LEO) and Medium Earth Orbit (MEO) satellite networks have shorter round trip delays and lower transmission power requirements. In many constellations, direct inter-satellite links (ISLs) provide communication paths among satellites. They can be used to carry signaling and network management traffic as well as data packets [15].

Since satellites are constantly moving, LEO and MEO satellite networks have dynamic topologies. The ISL connectivity changes based on the distance and azimuth angle between the two end satellites. Hence, routing in this environment is a challenging problem. Most of the routing schemes developed for LEO satellite networks assume a connection-oriented network structure. In [3] and [14], the dynamic routing problem is tackled by a discrete time network model. In each equal-length interval, the satellite network is regarded as having a fixed topology so that optimal link assignments can be performed. Call statistics are exploited in [11] to maintain the initial paths and reduce the re-routing frequency so as to minimize the signaling overhead. In [7], a satellite over satellite (SOS) network architecture is proposed, which is composed of LEO and MEO satellite layers. Long distance-dependent traffic is carried in the MEO layer to reduce satellite hops and resource consumption. A LEO/MEO two-tier satellite network and the corresponding routing strategies are described in [6]. However, it is assumed that there is no direct ISL between any two LEO satellites; all network routing functions involve MEO satellites.

With the rapid growth of Internet-based applications, proposed broadband satellite networks will be required to transport IP traffic [16]. Routing protocols for IP-based LEO satellite networks have also been introduced. The datagram routing algorithm (DRA) [4] aims to forward the packets on minimum propagation delay paths. The satellite network is regarded as a mesh topology consisting of logical locations. Data packets are routed distributedly on this fixed topology. DRA causes no overhead since the satellites do not exchange any topology information. In [5], a link state packet is flooded only as far as the routing radius for a given satellite. Shortest path routing is used in the near vicinity of the destination, while data packets are routed based on the destination satellite's position when they are far away. The basic shortcoming of both above schemes for connectionless routing is that the metrics used to calculate the paths do not reflect the total delay a packet may experience in the network. The delay, which is composed of propagation, processing, queuing, and transmission delays, can vary greatly with the changes of the positions of the individual satellites and the network load.

A routing protocol for multi-layered satellite IP networks (MLSR) has been proposed in [1]. MLSR computes the routing tables of the satellites based on the delay measurements collected periodically. Under MLSR, the LEO satellites are grouped and managed by MEO satellites. LEO group topologies are hidden from other satellites by representing them as meta-nodes in the topology. The routing tables are calculated by GEO satellites based on this summarized information,

*Eylem Ekici was with the Broadband & Wireless Networking Laboratory, School of Electrical & Computer Engineering, Georgia Institute of Technology when this work was performed. This work is supported by the National Science Foundation under Grant ANI-0087762.

which are further refined by MEO managers for LEO satellites. In many cases, however, satellites are sparsely located in the MEO layer, LEO group abstraction cannot be restored in the MEO layer and MLSR cannot be implemented effectively. Furthermore, MLSR relies on periodic routing table calculations to handle ISL congestion and lacks a fast-reacting congestion resolution mechanism.

In this paper, we propose a new routing protocol: *Satellite Grouping and Routing Protocol* (*SGRP*), which operates on a two-layer satellite network consisting of LEO and MEO satellites. Collaboration between LEO and MEO satellite layers are utilized in SGRP: MEO satellites compute the routing tables for the LEO layer. The main idea of SGRP is to transmit packets in minimum-delay paths and distribute the routing table calculation for the LEO satellites to multiple MEO satellites. LEO satellites are divided into groups according to the footprint areas of the MEO satellites in each snapshot period. Snapshot periods are determined according to the predictable MEO trajectory and the changes in the LEO group memberships. The MEO satellite that covers a set of LEO satellites becomes the manager of that LEO group. Group managers are in charge of collecting and exchanging the link delay information of the LEO layer, and calculating the routing tables for their LEO members. LEO satellites receive routing tables from their group managers. Using this protocol, the calculation of the routing tables is shifted to MEO satellites, which effectively distributes the power consumption between LEO and MEO satellites. Since the signaling traffic is physically separated from the data traffic, link congestion does not affect the responsiveness of delay reporting and routing table calculation. Furthermore, responsive mechanisms to address link congestion and satellite failures are included in SGRP.

The remainder of this paper is organized as follows: In Section 2, the two-layer satellite network architecture is presented. The mobility modeling of the LEO/MEO joint constellation is detailed in Section 3. Section 4 introduces the definitions used in the paper. In Section 5, the new routing algorithm called Satellite Grouping and Routing Protocol is described in detail. The performance evaluation of SGRP is presented in Section 6. Finally, Section 7 concludes this paper.

2. Satellite network architecture

Routing complexity is a crucial issue in satellite networks. Since LEO satellites already have limited processing power, it is not desired to have all LEO satellites compute their own routing tables. The terrestrial gateways are constrained by geographical distribution of continents. Meanwhile, in order to reduce the system costs, the satellite coverage areas are usually not highly overlapped, which means that the terrestrial gateways do not have line-of-sight communications with many satellites (usually less than 5 satellites for gateways outside the polar regions). If we choose to use the terrestrial gateways for route computation, the majority of LEO satellites would be required to send their measurements to the gateways over several hops. Similarly, the routing tables calculated by the terrestrial gateways would be transmitted to the LEO satellites via several hops as well. Both directions of transmission result in an increase of the traffic load. On the other hand, if there is a MEO satellite constellation in operation and ISLs between LEO and MEO satellites can be set up, the LEO satellites can be partitioned into groups and the computation overhead can be distributed among the MEO satellites. Each LEO group would have line-of-sight communication a MEO satellite in the second layer. Transfer of link delay measurements and routing table distribution are reduced and the traffic load is not increased on the LEO satellite network.

We consider a two-layer satellite network and the terrestrial gateway stations. The grouping of LEO satellites is determined by the snapshot concept. In a snapshot period, LEO satellites are grouped according to the footprint areas of MEO satellites. The satellite members of a group are constant over this period. LEO satellites have direct links to their MEO group managers. Terrestrial gateways are fixed on the Earth, they have direct links to the LEO satellites within sight. They are in charge of address translation and the communication between the terrestrial autonomous systems and the satellite network. Terrestrial gateways together with LEO and MEO satellites form an autonomous system.

2.1. Satellite layers

The satellite network is composed of LEO and MEO satellite layers. We assume that both satellite layers provide global coverage.

1. **MEO layer**: The MEO layer is composed of all MEO satellites in the network. It has a total number of $N_M \times M_M$ satellites, where N_M is the number of planes in MEO constellation, and M_M is the number of satellites in a plane. MEO satellite is denoted by $M_{i,j}$, where $i = 1, \ldots, N_M$, $j = 1, \ldots, M_M$.
2. **LEO Layer**: The LEO layer consists of all LEO satellites in the network. The total number of satellites in this layer is $N_L \times M_L$, where N_L is the number of planes in LEO constellation, and M_L is the number of satellites in a plane. The LEO satellites are organized into a Walker constellation [13].

The *logical location* concept is used for the LEO layer [4]. In this paper, however, the position of a logical location is not fixed and the satellite that embodies a logical location varies with time. When the satellite assigned to a logical location changes, the successor satellite must take the necessary routing information from its predecessor. The links adjacent to the predecessor LEO satellite are also switched to the new LEO satellite. A logical location is referred to as (n, m), where n is the plane number, $1 \leq n \leq N_L$, and m is the satellite position

in the plane, $1 \leq m \leq M_L$. The LEO satellite representing the logical location (n, m) at time t is referred to as $L_{n,m}$.

The MEO satellite topology is captured by a series of *snapshots*. In every snapshot period, the logical locations covered by a MEO satellite are considered to be fixed although the LEO satellites that embody the logical locations may change. The snapshot period is determined according to the predictable MEO trajectories and the positions of the logical locations. The snapshot concept hides the mobility of the MEO satellites and is independent of the properties of the MEO constellation. The algorithm to determine the snapshot periods will be detailed later in Section 3.

2.2. Links in the network

There are three types of duplex links in the network.

1. **Inter-Satellite Links (ISLs)**: The communication within the same layer occurs through Inter-Satellite Links (ISLs). Satellites are connected to their immediate neighbors in the same layer via duplex ISLs. There are two kinds of ISLs in the network: *intra-plane ISL* and *inter-plane ISL*. Intra-plane ISLs can be maintained permanently as the relative positions of neighboring satellites in the same plane are fixed. Inter-plane ISLs are operated only outside the polar regions, and can be temporarily switched off due to changes in distance and viewing angle between satellites. $\text{ISL}_{s\to d}$ or $\text{ISL}_{d\to s}$ denotes an ISL that connects two satellites s and d in the same layer.
2. **Inter-Orbital Links (IOLs)**: The communication between MEO and LEO satellites occurs over Inter-Orbital Links (IOLs). If a LEO satellite s lies in the coverage area of a MEO satellite d, they are connected by an IOL, which is referred to as $\text{IOL}_{s\to d}$ or $\text{IOL}_{d\to s}$.
3. **User Data Links (UDLs)**: LEO satellites communicate with the terrestrial gateways via User Data Links (UDLs). Terrestrial gateways are directly connected to LEO satellites that cover them. Thus a terrestrial gateway can be connected to several LEO satellites, and a LEO satellite can maintain UDLs to multiple terrestrial gateways. The UDL between a LEO satellite s and a terrestrial gateway G is denoted by $\text{UDL}_{s\to G}$ or $\text{UDL}_{G\to s}$.

2.3. Satellite groups

In order to partition the LEO satellite network into administrative domains, the LEO satellites are grouped according to the footprint areas of the MEO satellites in each snapshot period.

A LEO group is defined as a set of logical locations that reside in the coverage area of the same MEO satellite. The members of a LEO group change as the MEO satellite moves. Hence, the groups must be redefined in each snapshot period. In a snapshot period, the MEO satellite that covers a set of logical locations becomes the group manager. Group managers are responsible for collecting and exchanging link delay information received from LEO layer, and calculating the routing tables for the LEO group members. A LEO group $\mathcal{L}_{i,j}$ is the collection of all LEO satellites that lie in the coverage area of the MEO satellite $M_{i,j}$, $\mathcal{L}_{i,j} = \{\mathcal{L}_{i,j}(k) \mid k = 1, \ldots, K_{i,j}\}$, where $K_{i,j}$ is the number of LEO members in group $\mathcal{L}_{i,j}$. The members of a LEO group are connected to the manager MEO satellite via IOLs. For example, in figure 1, the LEO group of MEO satellite $M_{i,j}$ is $\mathcal{L}_{i,j}$, which has six members $\mathcal{L}_{i,j}(1)$ through $\mathcal{L}_{i,j}(6)$.

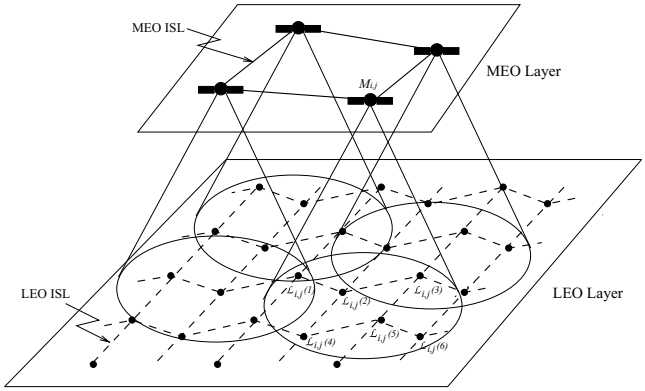


Figure 1. LEO/MEO joint constellation.

2.4. Gateway address translation

The terrestrial gateways are in charge of address translation and communication between the terrestrial autonomous systems and the satellite network. When a packet needs to be routed from gateway G_1 to gateway G_2 through satellite network, G_1 first finds the nearest LEO logical location for itself and G_2. Since the LEO logical locations are fixed with respect to the Earth, only the geographical location of the gateway is needed to determine the closest logical location. Assume that the logical location (n_1, m_1) is the nearest logical location to G_1, and (n_2, m_2) is the nearest to G_2. Then G_1 sends the packets to L_{n_1,m_1}, the LEO satellite which currently represents the logical location (n_1, m_1), through $\text{UDL}_{G_1\to L_{n_1,m_1}}$. The destination field of the packet is set as logical location (n_2, m_2), and is used for routing decisions inside the LEO network. After L_{n_2,m_2} receives the packet, it extracts the original destination G_2 from the data, then forwards the packet to gateway G_2 though $\text{UDL}_{L_{n_2,m_2}\to G_2}$.

3. Mobility modeling for LEO/MEO joint constellation

In order to create the snapshots of the satellite network, we must know the exact positions of the LEO and MEO satellites. Using the location information, the LEO groups and their group managers can be determined. In this section, we build a mobility model for LEO/MEO joint constellation. It gives the positions of LEO and MEO satellites at any time t, and the method to determine LEO groups and snapshot periods.

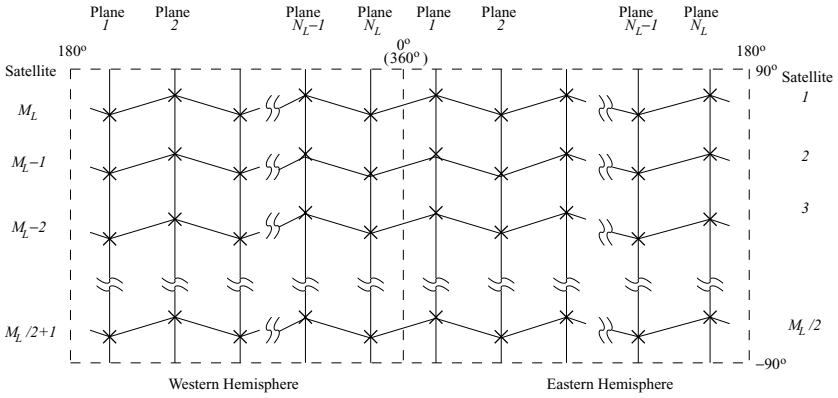


Figure 2. Logical locations in the LEO layer.

3.1. Modeling the LEO layer

The latitude $\phi(n, m)$ and longitude $\theta(n, m)$ of the LEO logical location (n, m) vary with time t and are calculated as follows:

$$\phi(n,m) = \begin{cases} \phi_0(n) - (m-1)\Delta\phi + \text{OFS}, \\ \quad \text{if } m < \lceil M_L/2 \rceil \\ -180^\circ - \phi_0(n) + (m-1)\Delta\phi + \text{OFS}, \\ \quad \text{if } m \geq \lceil M_L/2 \rceil \end{cases} \tag{1}$$

where $\text{OFS} = (w_L \times t)$ MOD $\Delta\phi$ is the offset within the latitude interval $\Delta\phi = 360^\circ/M_L$, w_L is the angular velocity of LEO satellite; $\phi_0(n)$ gives the latitude of the first satellite on the nth plane, and is defined as $\phi_0(n) = \{^{\phi_1,\ n \text{ odd}}_{\phi_2,\ n \text{ even}}$, where $\phi_1 \in (90^\circ, 90^\circ - \Delta\phi/2)$, $|\phi_2 - \phi_1| \leq \Delta\phi/2$. The first satellites in even-numbered planes have the same latitude ϕ_1, whereas the first satellites in odd planes are with latitude ϕ_2. Therefore, the satellites with same number m in all planes form a zigzag pattern, as shown in figure 2 if $\phi_1 \neq \phi_2$.

The longitude $\theta(n, m)$ of the logical location (n, m) is given by:

$$\theta(n,m) = \theta_0 + (n-1)\Delta\theta, \tag{2}$$

where θ_0 is the longitude of the first plane, and $\Delta\theta = 180^\circ/N_L$.

As satellites move, different satellites embody the same logical location at different time t.

3.2. Modeling the MEO layer

In this section, we illustrate the method to determine the satellite positions in a MEO constellation consisting of two planes. This model can also be modified to be used with any other MEO constellation.

In a MEO satellite constellation like ICO in figure 3, there are two crossing points for MEO planes 1 and 2, which are located on the equatorial plane, i.e, at latitude 0°. Assume that at time $t = 0$, MEO satellites $M_{1,1}$ and $M_{2,1}$ both move to northeast and are located at the first and second crossing points with longitude of 0° and 180°, respectively. The latitude Φ and longitude Θ of MEO satellite $M_{1,j}$ at any time t can be

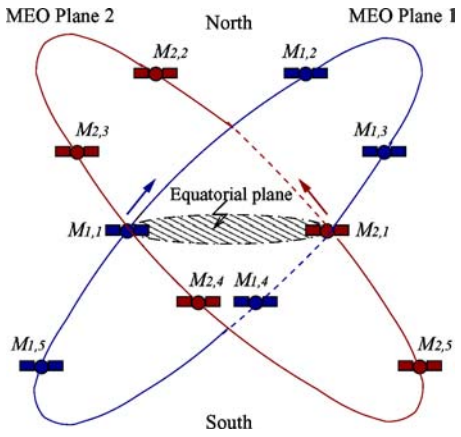


Figure 3. Initial positions of the MEO satellites.

computed by:

$$\Phi = \arcsin(\cos\alpha \cdot \sin r), \quad \Phi \in [-90^\circ, 90^\circ]$$
$$\Theta = \begin{cases} 360k_1 + \arccos(\cos\gamma/\cos\Phi) - w_E t, & \text{if } \Phi \geq 0 \\ 360k_2 - \arccos(\cos\gamma/\cos\Phi) - w_E t, & \text{if } \Phi < 0 \end{cases} \tag{3}$$

where α is the inclination angle for MEO plane; $\gamma = w_M t + (j-1)\Delta\Theta$, with w_M being the MEO satellite angular velocity, and $\Delta\Theta = 360/M_M$; k_1 and k_2 are independent integers to satisfy $\Theta \in [0^\circ, 360^\circ]$; w_E is the angular velocity of the Earth.

The latitude and longitude of the MEO satellites on plane 2 can be determined by:

$$\text{latitude}(M_{2,j}) = \text{latitude}(M_{1,j})$$
$$\text{longitude}(M_{2,j}) = (\text{longitude}(M_{1,j} + 180^\circ))\text{MOD } 360^\circ \tag{4}$$

3.3. Satellite groups and snapshot periods

Based on the exact positions of LEO and MEO satellites, and the footprint of every MEO satellite, we create the LEO satellite groups and determine the length of the snapshot periods at any time t.

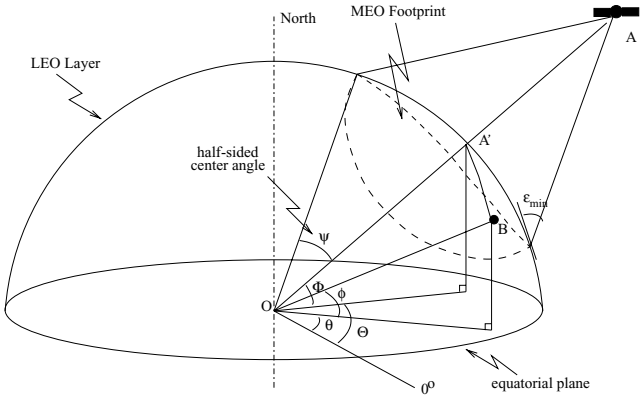


Figure 4. A MEO footprint.

3.3.1. Footprints of the MEO satellites on the LEO layer
The MEO footprints on the LEO layer are needed to determine the group membership of LEO satellites. The half-sided center angle of the MEO footprint on LEO the layer ψ is calculated as:

$$\psi = 90 - \epsilon_{\min} - \arcsin\left(\frac{R_E + h_L}{R_E + h_M} \cdot \cos\epsilon_{\min}\right), \qquad (5)$$

where R_E is the radius of the Earth, h_L and h_M are the plane altitude of LEO layer and MEO layer, respectively, and $\epsilon_{\min}$ is the minimum elevation angle of MEO satellites from the LEO layer.

Suppose that a LEO satellite $L_{n,m}$ is at (ϕ, θ), where ϕ and θ represent the latitude and the longitude of $L_{n,m}$, and a MEO satellite $M_{i,j}$ is at (Φ, Θ). For $L_{n,m}$ to lie in the footprint of $M_{i,j}$, the following condition must be satisfied:

$$\angle \mathrm{A'OB} = 2\arcsin\frac{|\mathrm{A'B}|}{2(R_E + h_L)} \le \psi \qquad (6)$$

where as shown in figure 4, A and B represent the positions of $M_{i,j}$ and $L_{n,m}$, respectively, A′ is the sub-satellite point of $M_{i,j}$ on the LEO orbit sphere.

3.3.2. Group definition and snapshot determination
Assume that the satellite network topology is periodic with T, where T is the least common multiple of the revolution periods of the Earth and the MEO satellites, and the time needed for any two satellites to be exactly on a given logical location. T is referred to as *system cycle*. The satellite topology can be considered as a periodically repeating series of P topology snapshots in the system cycle T. Over the interval $[t_i, t_{i+1}]$, $i = 0, 1, \ldots, P-1$, the LEO satellites' group membership is constant. Snapshot periods may have different lengths.

The snapshots and the LEO satellite groups are created according to the following criteria:

1. A LEO group is created according to the footprint of the MEO satellites on the LEO layer. Generally, the LEO satellites that lie in the same footprint of a MEO satellite form a group, and this MEO satellite becomes the group manager.
2. According to the definition, LEO satellite groups can be overlapped. If a LEO satellite lies in an overlapping region covered by several MEO satellites, it has more than one MEO group managers. To balance the management load, a *primary manager* is chosen among them. Primary manager takes care of the routing table calculation of a LEO satellite in a snapshot period. Since the trajectory of the MEO satellites is predictable, a LEO satellite chooses the MEO satellite with the longest remaining coverage time as its primary manager.[1] The members of a LEO group change as MEO satellites move, hence, the groups must be redefined for each snapshot period.
3. The snapshot period is further determined according to the changes in the LEO group memberships. Assume that at time $t = t_i$, at least one of the LEO satellites is no longer covered by its primary manager in snapshot i. In such a case, a new snapshot of the system must be created. Every LEO satellite chooses the group manager with the maximum predicted service time as its primary manager for snapshot $i + 1$. According to this criteria, new snapshots are created at times $t_1, t_2, \ldots, t_P = T$. The snapshots and the LEO groups repeat with a period of T.

Group information database can be uploaded to all satellites. The database information includes the start time of each snapshot period, LEO satellites' group membership and their MEO group managers in every snapshot.

4. Definitions

Definition 1 (Group Manager and Primary Manager)**.** Let $\mathcal{H}(x)$ refer to the MEO manager set of LEO satellite x, then $\mathcal{H}(x) = \{M_{i,j} \mid x \in \mathcal{L}_{i,j}\}$ includes all MEO satellites whose footprint covers x. The primary manager of x is written as $\mathcal{PH}(x)$. It is selected from $\mathcal{H}(x)$, and has the longest remaining coverage time for x, i.e, within all MEO satellites that currently cover x, $\mathcal{PH}(x)$ still covers x after all others exclude x in their footprints.

$$\mathcal{PH}(x) = \arg\max_{M_{i,j}}\{\text{remaining coverage time of } M_{i,j}, \text{w.r.t } x \mid M_{i,j} \in \mathcal{H}(x)\}. \qquad (7)$$

Definition 2 (Care-of Member List)**.** Every MEO satellite has a "care-of member" list in each snapshot period. The care-of member list $\mathcal{CM}(M_{i,j})$ of a MEO satellite $M_{i,j}$ is defined as

$$\mathcal{CM}(M_{i,j}) = \{x \mid \mathcal{PH}(x) = M_{i,j}\}. \qquad (8)$$

Hence $M_{i,j}$ is the primary manager of every LEO satellite in $\mathcal{CM}(M_{i,j})$.

[1] A mathematical method is explained in [2] to compute the remaining coverage time of a satellite over a ground terminal. The same method can be used to determine the remaining coverage time of a MEO satellite to a LEO satellite.

Definition 3 (Delay Function)**.** Let $l_{x\to y}$ be a direct ISL from node x to node y in LEO layer. The delay function $\mathcal{D}(l_{x\to y})$ is defined as follows:

$$\mathcal{D}(l_{x\to y}) = \begin{cases} \text{Delay from } x \text{ to } y, & \exists\, l_{x\to y} \\ \infty, & \text{otherwise} \end{cases}. \tag{9}$$

Definition 4 (Delay Report)**.** Delay report $\mathcal{DR}(x)$ of LEO satellite x is a set of tuples $\{y, \mathcal{D}(l_{x\to y})\}$, where y is a LEO satellite such that $\text{ISL}_{x\to y}$ exists between x and y.

Delay report $\mathcal{DR}(M_{i,j})$ of MEO satellite $M_{i,j}$ is a collection of the delay report of $M_{i,j}$'s care-of members.

$$\mathcal{DR}(M_{i,j}) = \{\mathcal{DR}(x) \mid x \in \mathcal{CM}(M_{i,j})\}. \tag{10}$$

Delay report $\mathcal{DR}(M_i)$ of MEO plane i is a collection of the delay report of $M_{i,j}$ in plane i.

$$\mathcal{DR}(M_i) = \{\mathcal{DR}(M_{i,j}),\ j = 1, \ldots, M_M\}. \tag{11}$$

Definition 5 (Plane Crossing Point)**.** Crossing points of plane i and plane l are referred to as $\mathcal{CP}(i, l)$, indicating where the two planes cross each other. There are two crossing points for each pair of i and l.

After collecting the delay information in LEO network, each MEO satellite has the same picture of the LEO network topology. MEO satellite $M_{i,j}$ computes the minimum delay paths from $\mathcal{CM}(M_{i,j})$ to all destinations. These paths are then used to create the routing tables. Before sending out routing tables to LEO satellites, $M_{i,j}$ tries to aggregate faraway LEO satellites into groups to reduce the size of the routing tables. To do this, the *remote groups* of a source satellite x are defined.

Definition 6 (Remote Group)**.** A remote group of LEO satellite x is a LEO group that is not covered by any satellite in $\mathcal{H}(x)$. The set of x's remote group is written as

$$\mathcal{RM}(x) = \{\mathcal{L}_{i,j} \mid M_{i,j} \notin \mathcal{H}(x)\}. \tag{12}$$

Definition 7 (Path)**.** $\mathcal{P}_{x\to y}$ is defined as the minimum delay path associated with source x and destination y. It is a sequential list of satellites on the path.

In our satellite network architecture, the routing tables are created by MEO satellites using the delay measurements in the LEO layer. MEO group managers prepare different routing tables for each of their care-of members. In our algorithm, two types of routing tables are needed: the *original routing table* and the *simplified routing table*.

Definition 8 (Original Routing Table)**.** Original routing table $\mathcal{ORT}_{M_{i,j}}$ is kept in the MEO satellite $M_{i,j}$. It provides an entry for each of its care-of members, and registers paths from $\mathcal{CM}(M_{i,j})$ to all destinations. The path from satellite x to a destination satellite y is defined as:

$$\mathcal{ORT}_{M_{i,j}}(x, y) = \mathcal{P}_{x\to y}, \quad \text{where } x \in \mathcal{CM}(M_{i,j}). \tag{13}$$

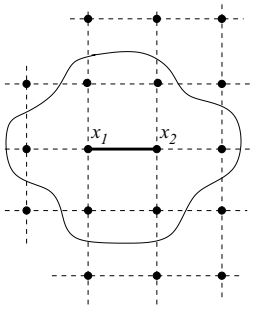


Figure 5. Congestion area of congested link $l_{x_1\to x_2}$ when $r = 1$.

Definition 9 (Simplified Routing Table)**.** Simplified Routing Table $\mathcal{SRT}_x$ of LEO satellite x is created by and sent from its MEO manager $M_{i,j}$. The construction of $\mathcal{SRT}_x$ is based on original routing table $\mathcal{ORT}_{M_{i,j}}$ and the group membership of destination satellites. Each entry of this routing table has a destination field and a next-hop field, where *next-hop* is the second node on $\mathcal{P}_{x\to Dest}$, and written as $\mathcal{SRT}_x(Dest)$. Here *Dest* can be any LEO satellite or a remote group. If the paths to all satellites in a remote group $\mathcal{L}_{i,j}$ have the same next-hop, the entries to all those LEO satellite destinations are replaced by a single entry in the simplified routing table. The destination field of this entry is set as $\mathcal{L}_{i,j}$.

Definition 10 (Congestion Area)**.** The congestion area of a congested link $l_{x_1\to x_2}$ is defined as:

$$\mathcal{CA}(l_{x_1\to x_2}) = \bigcup \left\{ l_{y_1\to y_2} \,\middle|\, \text{where } \mathcal{P}_{x_k\to y_i} \le r, k = 1 \text{ or } 2 \right\} \tag{14}$$

where r is the radius in the number of hops of the congestion area. Figure 5 shows the congestion area of a congested link when $r = 1$.

5. Satellite grouping and routing protocol

The goal of our new *Satellite Grouping and Routing Protocol* (*SGRP*) is to forward the packets on minimum delay paths in spite of the satellite mobility, and to distribute the routing table calculation for the LEO satellites to multiple MEO satellites. The delay metric used in the route computation is the sum of the processing, queuing, and transmission delays in the satellites and the propagation delays on the ISLs. Routing tables are calculated by MEO satellite group managers, transmitted to and stored in LEO satellites.

In this section, the detailed design of SGRP is presented. It includes three phases:

- Delay report from LEO satellite to MEO layer,
- Delay exchange in MEO layer,
- Routing table calculation,

The SGRP also has mechanisms to resolve congestion and satellite failures to avoid dropping packets.

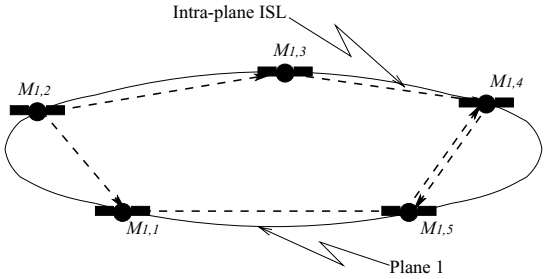


Figure 6. Intra-plane exchange.

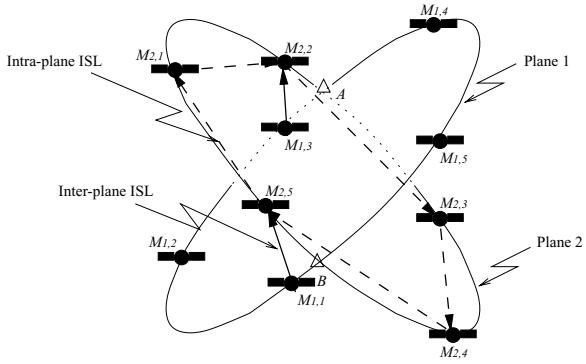


Figure 7. Inter-plane exchange.

5.1. Delay report

Delay information of LEO links needs to be reported to MEO satellites every T_c period, it is done as follows:

Initialization: At the beginning of a new snapshot period, MEO satellite $M_{i,j}$'s care-of member list $\mathcal{CM}(M_{i,j})$ is initialized as empty.

Step 1: Delay Reporting—At the end of every measurement interval of length T_c, a LEO satellite x monitors the delay on its outgoing links. A delay report $\mathcal{DR}(x)$ is created from the measured delay value and sent to x's primary manager $M_{i,j} = \mathcal{PH}(x)$ via $\text{IOL}_{x \to M_{i,j}}$.

Step 2: Delay Reception—After receiving a delay report $\mathcal{DR}(x)$, $M_{i,j}$ adds x into its own delay report $\mathcal{CM}(M_{i,j})$. $\mathcal{CM}(M_{i,j})$ is formed after all the delay reports from $M_{i,j}$'s care-of members have been received.

5.2. Delay exchange

After collecting link delay measurements from their group members, MEO group managers exchange the measurements inside the MEO layer to obtain a common picture of the LEO network topology. Our proposed exchange method includes two steps: *intra-plane exchange* and *inter-plane exchange*.

Step 1: Intra-plane Exchange

In MEO layer, the delay reports are first circulated in the same MEO plane.

1. MEO satellite $M_{i,j}$ sends its delay report $\mathcal{DR}(M_{i,j})$ to its two adjacent neighbors, $M_{i,p}$, in the same plane through $\text{ISL}_{M_{i,j} \to M_{i,p}}$, where $p = j - 1, j + 1$.
2. After receiving delay reports $\mathcal{DR}(M_{i,j})$, $M_{i,p}$ checks to see if it has been received before. If so, it is discarded.
3. $M_{i,p}$ forwards the new report $\mathcal{DR}(M_{i,j})$ on the other intra-plane ISL, which is different from the incoming one, i.e. $\text{ISL}_{M_{i,p} \to M_{i,p+1}}$ or $\text{ISL}_{M_{i,p} \to M_{i,p-1}}$.

Figure 6 shows the circulation of delay reports in MEO plane 1. $M_{1,2}$ sends out $\mathcal{DR}(M_{1,2})$ to its neighbors $M_{1,1}$ and $M_{1,3}$. Then the report follows the dashed lines in the direction of the arrows. In the end, $M_{1,4}$ and $M_{1,5}$ each receives a duplicate report, upon which the circulation of $\mathcal{DR}(M_{1,2})$ is terminated.

Step 2: Inter-plane Exchange

After the LEO delay information is exchanged within plane i, a copy of the same information must be sent out to plane $l, l = 1, \ldots, N_M, l \neq i$, and circulated there as well. The steps of the inter-plane delay report exchanging are as follows:

1. The two satellites on plane i nearest to plane crossing points $\mathcal{CP}(i, l)$ are chosen to be plane i's starting points. The two satellites nearest to $\mathcal{CP}(i, l)$ on plane l are selected as their reception satellites respectively. $\mathcal{DR}(M_i)$ is sent from plane i to plane l via the inter-plane ISLs.
2. The two reception satellites on plane l forward $\mathcal{DR}(M_i)$ clockwise via their intra-plane ISLs to the neighboring MEO satellites.
3. After receiving $\mathcal{DR}(M_i)$, $M_{l,m}$ first checks to see whether it has been received before. If so, the delay report is discarded, otherwise, it is forwarded clockwise to the next neighboring MEO satellite.

Figure 7 shows the transfer of $\mathcal{DR}(M_1)$ from plane 1 to plane 2. $\mathcal{CP}(1, 2) = \{A, B\}$. $M_{1,1}$ and $M_{1,3}$ are chosen as the starting points, their reception satellite are $M_{2,5}$ and $M_{2,2}$, respectively. Starting from $M_{2,5}$ and $M_{2,2}$, the report is circulated clockwise over the dashed lines. Note that the circulation of different plane's delay reports is processed in a parallel way, i.e., the delay report of one plane can be sent to different planes simultaneously.

5.3. Routing table calculation

Routing tables are prepared by the MEO satellites for their care-of members and updated every T_c period. There are two kinds of routing tables: the original routing tables register the detailed path and are kept in MEO satellites, whereas the simplified routing tables are sent to the LEO satellites.

Step 1: Original Routing Table Calculation—The MEO satellites perform routing table calculations after they received all the delay reports. The MEO satellite $M_{i,j}$ computes the minimum delay paths from the LEO satellites in $\mathcal{CM}(M_{i,j})$

to all LEO destinations, and adds them into original routing table $\mathcal{ORT}_{M_{i,j}}$.

Step 2: Simplified Routing Table Calculation—Based on $\mathcal{ORT}_{M_{i,j}}$, MEO group managers arrange the paths into destination and next-hop pairs for each of its care-of members. Before sending routing tables to LEO layer, $M_{i,j}$ tries to aggregate the destinations in remote groups to reduce the size of routing tables. The path aggregation is done as follows:

```
Let S = all satellites in LEO layer
for L_{i,j} ∈ RM(x)
  if the second node on P_{x→y} = t, ∀y ∈ L_{i,j}
    SRT_x(L_{i,j}) = t
    S = S − L_{i,j}
  end if
end for
for each y ∈ S
  SRT_x(y) = the second node on P_{x→y}
end for
```

When $\mathcal{SRT}_x$ is ready, it is sent from $\mathcal{PH}(x)$ to x via $\text{IOL}_{\mathcal{PH}(x)\to x}$.

5.4. Congestion avoidance

In our algorithm, data packets are routed according to the delay information gathered every T_c period. If traffic load changes fast, the routing decision cannot reflect the fluctuation of the real-time delay; congestion may occur. The congestion avoidance phase is introduced to deal with the congestion reactively and has three steps:

Step 1: Congestion Detection—To avoid congestion in the LEO network, every LEO satellite continuously monitors the queue lengths of the output buffers of their adjacent links. If the queue length associated with $l_{x_1\to x_2}$ is more than ξ packets, then "congestion" is said to have occurred on link $l_{x_1\to x_2}$. x_1 then promptly reports $\mathcal{D}(l_{x_1\to x_2}) = \infty$ to all its MEO managers in $\mathcal{H}(x_1)$.

Step 2: Information Propagation—Upon receiving a congestion warning of link $l_{x_1\to x_2}$, $M_{i,j}$ sets $\mathcal{D}(l_{x_1\to x_2}) = \infty$. Then, it propagates $\mathcal{D}(l_{x_1\to x_2}) = \infty$ in MEO layer using the same intra- and inter-plane exchange methods explained previously in Section 5.2.

Step 3: Path Recalculation—To reduce the computation overhead, MEO group managers only recalculate those paths affected by the congestion. Meanwhile, they try to lead the long routes away from entering the congestion area.

A MEO satellite M checks all paths in $\mathcal{ORT}_M$, and searches those affected by the congested link. If a path is either originated or destined within the congestion area $\mathcal{CA}(l_{x_1\to x_2})$, it will be kept. If a path goes through $\mathcal{CA}(l_{x_1\to x_2})$, then M "cuts" the congestion area when re-computating this path, i.e, set all delays associated with links in $\mathcal{CA}(l_{x_1\to x_2})$ to infinity, thus leads these paths away from entering the congestion area. The path recalculation in MEO satellite M for $x \in \mathcal{CM}(M)$ is summarized below.

```
Let S = all satellites in LEO layer
for each y ∈ S
  if l_{x1→x2} is on ORT_M(x, y) = P_{x→y}
    if l ∈ CA(l_{x1→x2}), ∀l ∈ P_{x→y}
      keep P_{x→y}, search next y
    end if
    if l ∉ CA(l_{x1→x2}), where l is the first and last link on
       path P_{x→y}
      set D(l_{y1→y2}) = ∞, ∀l_{y1→y2} in CA(l_{x1→x2})
    end if
    M recalculates P_{x→y}
    ORT_M(x, y) = P_{x→y}
  end if
end for
```

After the calculation, M updates affected parts in $\mathcal{ORT}_M$, aggregates the new paths, and sends packets to update the affected entries in simplified routing table $\mathcal{SRT}_x$ of its member x accordingly.

5.5. Satellite failures

A satellite may fail or be shut down temporarily for reasons such as maintenance and testing, or when crossing oceans or polar regions to save energy. When a satellite fails, all minimum delay paths passing through this satellite must be rerouted, so that the packets that normally pass through the failed satellite would not be dropped. In our algorithm, the rerouting is done in the following way: When a satellite fails, its direct neighbors are the first to sense this occurrence. They immediately send reports to MEO group managers. Upon receiving failure notification of a LEO satellite s, $M_{i,j}$ sets all link delays associated with s to infinity, then propagates the update delay report in the MEO layer.

To reduce the computation overhead, MEO group managers only recalculate those paths affected by the failure. A MEO satellite M checks the paths in $\mathcal{ORT}_M$, finds those affected by the failed satellite s. If the failed satellite lies on a path, M recalculates the path, updates the corresponding entry in $\mathcal{ORT}_M$, and performs group aggregation before changing into (Dest, next-hop) pairs for its care-of members.

If a packet arrives at the LEO satellite x and finds that the failed satellite is the next hop on its path, i.e., its routing table has not yet been updated, some special routing decision must be made to avoid dropping useful packets. Here we utilize the rerouting method in case of satellite failures in [4]. The packets destined to the failed satellite are deflected into orthogonal directions. The detailed rerouting algorithm can be found in the original paper.

Table 1
Parameters used for MEO and LEO satellite constellations.

	MEO	LEO
Altitude	$h_M = 10390$ km	$h_L = 700$ km
Number of planes	$N_M = 2$	$N_L = 12$
Number of satellite per plane	$M_M = 5$	$M_L = 24$
Angular velocity	$w_M = 1^\circ$/min	$w_L = 3.6^\circ$/min
Minimum elevation angle at LEO layer	$\epsilon_{\min} = 10^\circ$	–
Orbit inclination angle	45°	90°
Number of intra-plane ISLs	2	2
Number of inter-plane ISLs	0 or 1	0 or 2
Longitude of logical location $(n, 1)$	–	$\theta_0 = 7.5^\circ$
Latitude of logical location $(n, 1)$	–	$\phi_1 = 86.75^\circ, \phi_2 = 82.5^\circ$

6. Performance evaluation

Our simulation consists of three major parts: First, find the snapshot period and group membership information in each snapshot period according to the parameters of LEO and MEO satellite constellations. Secondly, using SGRP algorithm, keep track of the end-to-end delay between some terrestrial source-destination pairs, with the background traffic changing dynamically. Last, analytically show that the hierarchy in SGRP can reduce communication overhead compared to the centralized and distributed approaches.

6.1. Snapshot periods identification

In the two-layer satellite networks, the ICO network is chosen as the MEO satellite constellation, the LEO satellite constellation is a slightly modified version of the Teledesic network, where the orbital inclination is 90° instead of 98.2°. The system parameters are given in Table 1. The system cycle for these parameters is $T = 1440$ minutes, or one day.

Using our computation method in Section 3, there are a total of 93 snapshot periods in a system cycle. As expected, the snapshots repeat after time T. The mean duration time for all 93 snapshots is 15.5 minutes. The length distribution of the snapshot duration is given in figure 8, where the durations are in minutes. It can be seen that the lengths of snapshot periods are not fixed.

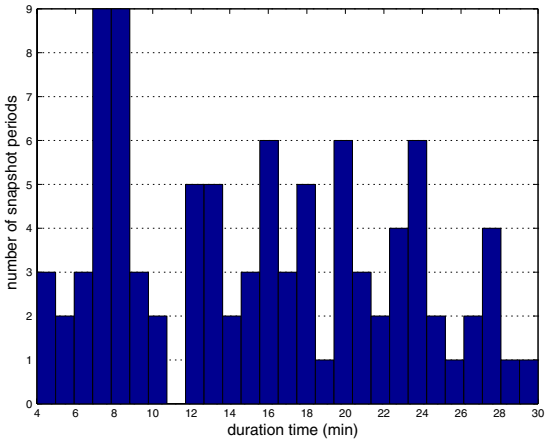


Figure 8. Distribution of snapshot duration.

Intensity level and Corresponding Expected Traffic (2005)

Intensity level	1	2	3	4	5	6	7	8
Traffic (million minutes/year)	1.6	6.4	16	32	95	191	239	318

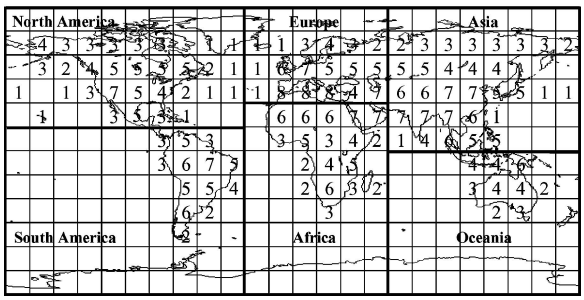


Figure 9. Earth zone division and user density levels [12].

6.2. Traffic modeling

We divide the Earth into $15^\circ \times 15^\circ$ geographical zones, and map each zone with a LEO logical location. Because of the asymmetry of the IP traffic, the user behavior and host behavior are different for each zone. For example, the source of *http* pages are more likely to be located in North America than in Central Africa. Hence, we build two databases for the user density level and host density level for each zone, where the user density level represents the amount of source requests in each zone and the host density level implies the host distributions over the geographic zones. The global background traffic can be generated using a traffic matrix model.

6.2.1. User density level

The forecasted voice traffic over LEO satellite systems for the year 2005 in [12] (as shown in figure 9) is referred to determine the user density levels. Here we assume that the potential requirement for satellite network IP traffic from each geographical zone is proportional to the expected volume of voice traffic. As users show different activities during different time of the day, to make the traffic model more accurate, we take the daily evolution of user density into consideration. Assuming that the daily evolution of traffic activity per user is the same for all users worldwide, and the local time of each traffic zone is equal to the solar time of the respective zone's center longitude. The daily user activity profile introduced in [8] is used. The user traffic distribution of each hour in percentage of the total traffic within a day is shown in figure 10.

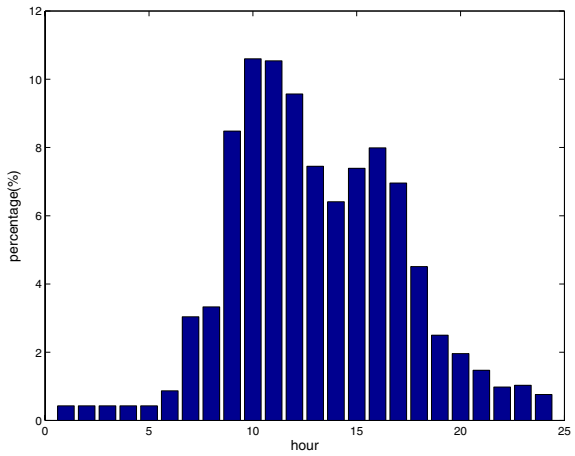


Figure 10. User activity in each hour (%) [8].

6.2.2. Host density level

The statistics of January 2001 in [10] is used to get the host density level for different terrestrial zones. The host density level gives the distribution of the Internet hosts in different continents, which is shown in Table 2. According to the data, we adjust the user density level to get the host density level of each zone by the following equation:

$$h_j = \frac{u_j}{\sum_i u(i)} \cdot N_h(k), \tag{15}$$

where h_j is the host density level of zone j, of which the user density level is u_j; $\sum_i u(i)$ is the sum of user density level of zones in continent k; $N_h(k)$ is the number of hosts in continent k. It can be seen that continent k's percentage share of host density on the Earth is equal to $p(k)$ in Table 2.

Table 2
Internet hosts distribution by continent in January 2001 [10].

Continent k	$N_h(k)$: # of hosts ($\times 10^3$)	$p(k)$: (%)
North America	71871.5	71.27
Europe	17698	17.55
Asia	7686.4	7.62
Oceania	1873.65	1.86
South America	1474.8	1.46
Africa	241.9	0.24

6.2.3. Traffic matrix

The inter-satellite traffic requirement between satellites i and j, i.e, T^{ij}, depends on the user density level u_i, the host density level h_j, and the distance $d(i, j)$ between the satellites.

$$T^{ij} = \frac{(u_i \cdot h_j)^\alpha}{(d(i, j))^\beta}. \tag{16}$$

Here i corresponds to the LEO logical location (n, m), where $n = \lceil i/M_L \rceil$, $m = i \bmod M_L$, M_L is the number of satellites in a LEO plane. Setting $\alpha = 0.5$, $\beta = 1.5$, we can get the traffic flow shares among the continents in Table 3.

In our satellite network, the links are modeled as finite capacity queues, the traffic requirements between satellites are mapped to the ISLs according to the shortest path the packets will take. They provide the arrival rates in the queuing model. We assume Poisson arrival rate and exponentially distributed service time, then the queuing delay of each link can be deduced by the $M/M/1/K$ queuing model.

The average packet arrival rate of each pair of satellites (packets/sec) is computed by:

$$\lambda_{ij} = \frac{T^{ij}}{\sum_{k=0}^{N_L \times M_L} \sum_{l=0}^{N_L \times M_L} T^{kl}} \times \text{(total offered traffic)}, \quad \text{where } i, j = 1, 2, \ldots, N_L \times M_L. \tag{17}$$

Here the "total offered traffic" represents the total traffic generated worldwide.

6.3. Simulation of delay performance

We developed our own simulator on C++. For each simulated routing protocol, the simulator outputs the corresponding end-to-end delay metric. In all simulations, the capacity of all UDLs and ISLs are chosen as 160 Mbps, and each outgoing link has been allocated a buffer size of 5 MB. If we assume an average packet size of 1000 bytes, the link capacity becomes 20000 packets per second and the buffer size becomes 5000 packets. The delay metric is sampled every 1 minute.

Three types of routing protocols are evaluated using our simulator: SGRP, the Datagram Routing Algorithm (DRA) [4], and the optimal routing computed by the Dijkstra algorithm [9]. Data packets are carried in LEO satellite layer. DRA forwards packets in the minimum propagation delay paths. Therefore, the queuing delay caused by the non-uniform traffic distribution is ignored. SGRP measures the link delay

Table 3
Continental traffic flow shares in %.

	Destination					
Source	North America	Europe	Asia	South America	Africa	Oceania
North America	86.18	6.74	4.18	1.76	0.45	0.70
Europe	25.10	55.88	13.52	1.62	2.84	1.04
Asia	24.04	20.89	47.74	1.15	1.75	4.43
South America	52.39	13.02	5.96	25.12	1.85	1.66
Africa	25.63	43.34	17.33	3.53	7.95	2.22
Oceania	26.48	10.58	29.22	2.11	1.49	30.12

values every T_c period and uses the delay values as a reference for computing the minimum delay paths. The Dijkstra algorithm [9] is used to calculate the routing tables in SGRP. Paths are adjusted when link congestion or satellite failures occur. The SGRP parameters used in the simulator are: the delay measurement interval $T_c = 4$ minutes unless specifically stated, the radius r of congestion area is set as 1. The optimal routing represents the ideal scenario that each satellite is assumed to be aware of the overall satellite topology and its knowledge of the link delays is updated in real time. Therefore, the optimal routing returns the best delay performance, which is hard to achieve in real systems and can only be approached at the cost of frequent delay measurement as well as heavy communication and computation overhead.

Our experiments are based on the observation of the end-to-end delay between certain terrestrial source-destination pairs. To evaluate the performance of the LEO/MEO satellite architecture and SGRP, three sets of experiments are conducted:

- **Path Optimality**: The first set of the simulations show the differences of end-to-end delay returned by SGRP, DRA, and the optimal routing.
- **Effect of Satellite Failures**: This set of simulations shows the effect of satellite failures on the performance of SGRP, with comparison with DRA and optimal routing.
- **Effect of Link Congestion**: Our routing algorithm SGRP has reaction mechanism when congestion occurs. This set of simulations shows the performance difference among SGRP, DRA, and optimal routing in case of link congestion.

6.3.1. Path optimality

The first set of experiments compares the end-to-end delay among the paths computed by SGRP, DRA, and the optimal routing. The experiments are based on the observation of the end-to-end delay between three terrestrial source-destination pairs. The first two pairs are with the same source node located at $(112.5^\circ E, 37.5^\circ N)$ in Asia. The destination nodes are at $(277.5^\circ W, 33.25^\circ N)$ in North America and $(52.5^\circ E, 52.5^\circ N)$ in Europe, respectively. The paths between these two pairs go through areas with traffic concentration. The path between source-destination pair 1 is with longer distance than that of pair 2. The third pair has the source located at $(142.5^\circ E, 37.5^\circ S)$ in Oceania and the destination at $(37.5^\circ E, 18.25^\circ S)$ in Africa. The path associated with the third pair does not travel through high traffic concentration areas. For each of the source-destination pair, the sender generates traffic with an average rate of 8 Mbps (1000 packets per second) for 100 minutes.

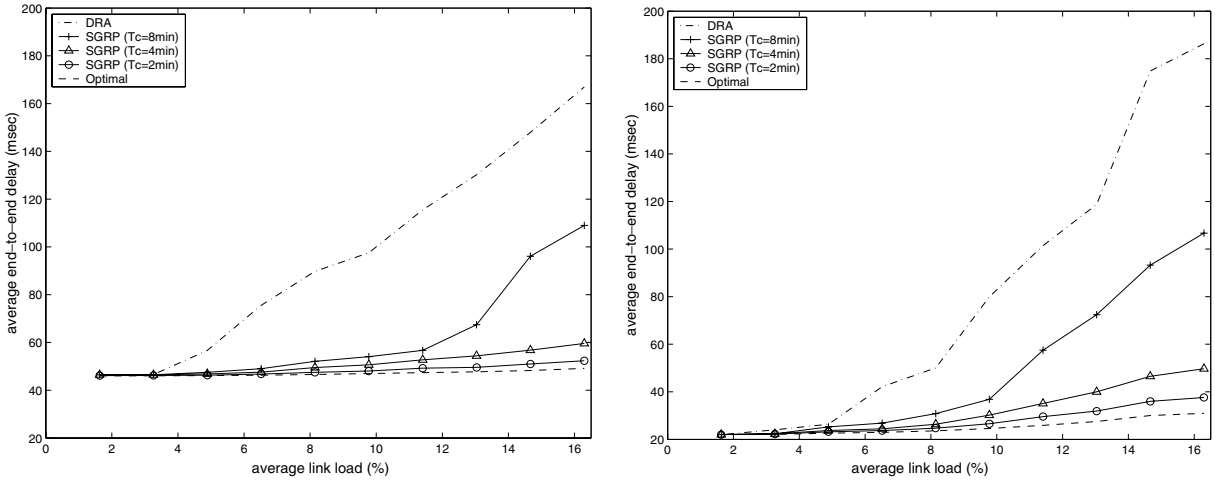
To compare the delays of different schemes under different link load, we increase the ISL utilization in the LEO layer gradually. It is done as follows:

- First, the packet arrival rate is generated by equation (17), it gives the average traffic rates of flows between each pair of satellites. Flows are generated with exponentially distributed rates with fixed means λ_{ij}.
- The rates are mapped to ISLs according to the minimum propagation delay paths the packets will take. The load of a link is the sum of all the rates of flows that pass through this link. Some ISLs are more heavily loaded than others according to the traffic distribution model.
- Assuming that the average load proportion across all the links keeps the same, hence to increase the ISL utilization statistically, the "total offered traffic" in equation (17) is increased, which affects the flow rates λ_{ij} and in turn changes the average load of each satellite link. The queuing delays of all the satellite links are calculated by the M/M/1/K queuing model.
- The delay of a link is the sum of its propagation delay and queuing delay at computation time.

In our simulation, each time a different value of the "total offered traffic" is chosen, the routes and end-to-end delays of certain flows are monitored for 100 minutes. The satellite link loads are changing dynamically with fixed nominal means. The end-to-end delay performance of the SGRP, DRA, and the optimal routing are depicted in figure 11. For each specific value of average link load, the end-to-end delay is averaged over the 100-minute monitoring time. Note that as the result of non-uniform traffic distribution shown in Table 3, the load of links varies greatly among different satellite links. Thus, some of the links may get congested even when the average link load is as low as 3%.

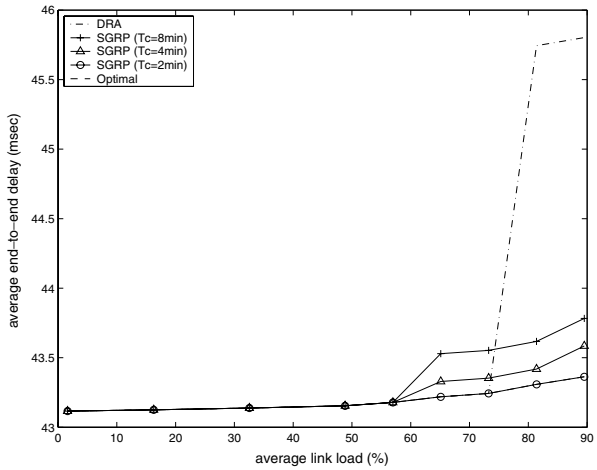
It can be seen that for paths that go through some high traffic concentration areas, e.g., source-destination pair 1 and 2, when the average link load is below 3%, the end-to-end delay performance of the three algorithms is similar. This is reasonable because when the traffic load is light, the propagation delay is the dominant factor in the end-to-end delay. However, as the average link load increases, the delay performance of SGRP and DRA deviates from the value returned by the optimal routing. The end-to-end delay of the path calculated by DRA increases dramatically when the average link load is greater than 8%. This is because when average link load increases, ISLs in areas with higher traffic density tend to be congested more easily. DRA reflects packets only when they approach or enter into the congestion area, whereas the routing scheme based on SGRP can have a big picture of the traffic distribution in LEO networks and reduce the traffic entering into the congested area. As SGRP leads long paths away to avoid even the vicinity of the congested links, however, these routes may experience longer delay compared to the paths calculated by the optimal routing algorithm.

For paths that travel only through areas with lower traffic concentration, e.g., source-destination pair 3, SGRP does not introduce higher delay than the optimal value until the average link load is high, e.g., 57% in figure 11(c). The delay performance of both DRA and SGRP is very close to the optimal value, e.g., the delay deviations from the optimal value for SGRP and DRA are within 0.5 msec and 2.5 msec, respectively (Note that the scale of the y-axis is different than those in figures 11(a) and (b)). Hence, for paths that do not travel through high traffic density areas, the performances of SGRP



(a) Source-destination pair 1.

(b) Source-destination pair 2.



(c) Source-destination pair 3.

Figure 11. Comparison of average end-to-end delay performance.

and DRA are not affected by congestion in other areas and are very close to the optimal value.

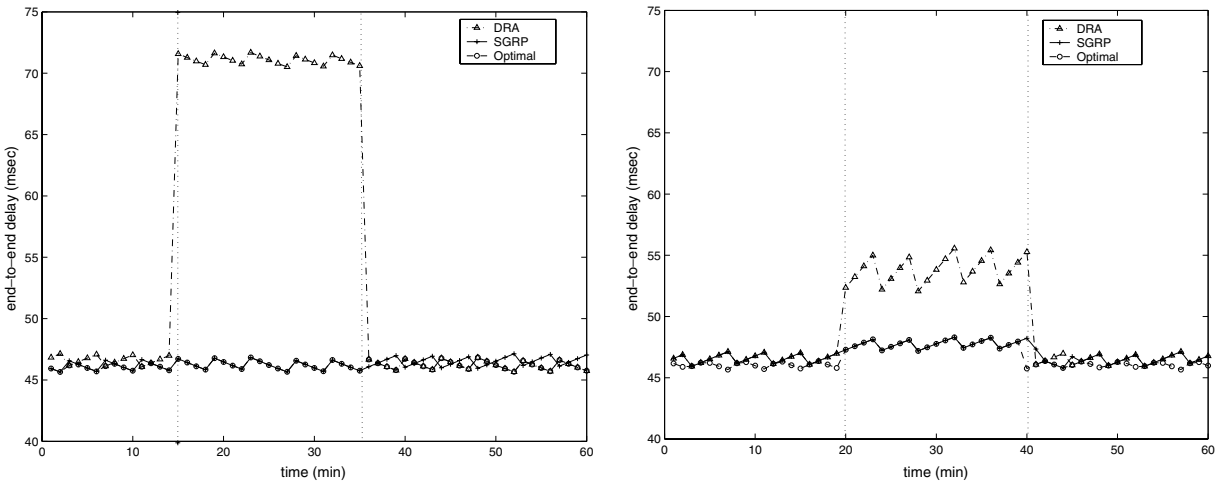
As explained in Section 5, the LEO satellites periodically measure delays of adjacent links. This delay information is then used to compute the routing tables for the coming measurement interval T_c. The length of T_c affects the delay performance of SGRP. If T_c is too large, the delay report obtained will not be able to capture the delay behavior in the next T_c period, which may cause the computed path to be sub-optimal. We have simulated SGRP with different T_c values of 8, 4, and 2 minutes, respectively. As seen in figure 11, with the decrease of measurement interval T_c, i.e., when routing tables are updated more frequently, the end-to-end delay values returned by SGRP approach the optimal value more closely. If the path does not travel through high traffic concentration areas, the delay difference of SGRP from the optimal value is ignorable. For example, in figure 11(c), when $T_c = 2$ min, the curve representing the path delay between source-destination pair 3 overlaps with that of the optimal delay. When T_c is large ($T_c = 8$ min in figure 11), the delay difference of SGRP from the optimal value grows largely under link congestion.

6.3.2. Effect of satellite failure

SGRP introduces a reaction mechanism against satellite failures and link congestion. In the following two sets of experiments, we compare the end-to-end delay of three different routing schemes mentioned previously under these events. To reflect the effect of real-time changes on delay performance, the background traffic is adjusted every hour according to the time of the day. All paths and link loads are updated after recalculation.

When a satellite fails, it effects the routing decision and the path delay. In this set of experiments, we keep track of the end-to-end delay of the first source-destination pair using these three algorithms, respectively. The sender generates traffic of 1000 packets per second for 60 minutes from 8 am to 9 am. The satellite representing the logical location of $(292.5^\circ W, 67.5^\circ N)$ is assumed to be out of service from 8:15 am to 8:35 am.

In figure 12(a), the instantaneous end-to-end delays associated with these three algorithms are depicted. DRA routes packets on the minimum propagation delay path, the satellites do not send delay reports to others. Thus, only the immediate



(a) Under satellite failure. (b) Under link congestion.

Figure 12. Comparison of instantaneous end-to-end delay performances.

neighbors know the satellite failure. When a packet is received by one of these neighbor satellites, and is destined to the failed one, it is deflected to one of the orthogonal directions. In SGRP, the satellite failure is reported to the MEO layer by its neighbors immediately. This failure report is then exchanged among all MEO satellites, causing them to update the routing tables of all the LEO satellites. Hence, we expect that SGRP have better performance than DRA under satellite failures. From the figure, we can see that the failure has minor effect on SGRP, yet in the satellite failure period, the path calculated by DRA undergoes higher end-to-end delay, which is about 55% higher than that of SGRP. On the other hand, the delays of SGRP and optimal routing are very close either under normal condition or when a satellite fails. Because when a satellite fails, the failure report packets are immediately received and passing around in MEO layer. New shortest paths are calculated and take effect after LEO satellites receive the new routing tables. This mechanism compensates the effect of satellite failures.

6.3.3. Effect of link congestion

Similarly, we depict the change of instantaneous end-to-end delay for source-destination pair 1 of the three algorithms when link congestion occurs. This congestion is created by injecting some heavy traffic into the satellite network in a certain area. In our experiment, the sender generates traffic of 1000 packets per second for 60 minutes in a peak hour from 10 am to 11 am. The congestion occurs at the link from LEO logical location ($277.5^\circ W$, $63.25^\circ N$) to ($277.5^\circ W$, $48.25^\circ N$) between 10:20 am and 10:40 am. To simplify the simulation, we confine the congestion to this link, and setting the load on this path to 100% of the link capacity.

From figure 12(b), the path calculated by DRA always undergoes higher delay within the congestion period. This delay is about 13% higher than that of the path calculated by SGRP. The average difference between the delays of SGRP and the optimal routing is about 0.5 msec. However, when congestion occurs, their delay performance is about the same. SGRP recalculates the routing tables right after congestion happens. The recalculation tries to keep the local traffic within the congestion area, but route the long path away from the congested area. Therefore, the effect of congestion will be compensated by enacting new routing tables.

6.4. Analysis of communication overhead

SGRP divides the LEO satellites into groups according to the snapshot periods and distributes the routing table calculation of all LEO satellites to several MEO satellites. Therefore, a hierarchy is introduced in the architecture. In order to demonstrate the efficiency of SGRP, we analytically compare the communication overhead of each round of routing table calculation in SGRP with the centralized and fully distributed routing table calculation approaches.

In the centralized routing table calculation scheme, all routing tables are calculated by a designated terrestrial gateway. The satellites in LEO layer create their delay reports and send them to the gateway through minimum hop paths. The terrestrial gateway calculates the individual routing tables for all the LEO satellites separately and sends these routing tables to the corresponding satellites again over minimum hop paths.

In the fully distributed routing table calculation approach, every satellite is responsible for calculating its own routing table. The delay reports are broadcast to all satellites. Once a satellite receives all delay reports, it calculates the shortest paths to all other nodes. Using the shortest paths, every satellite creates its own routing table that contains the next hop to reach all other nodes in the network.

In figure 13, the communication overhead of the three routing table calculation schemes are compared for a satellite network. The number of MEO satellites is set as 10, i.e., 2 planes with 5 satellites in each plane as in the ICO constellation. The total number of LEO satellites was changed and their effect on the communication overhead was recorded. The total

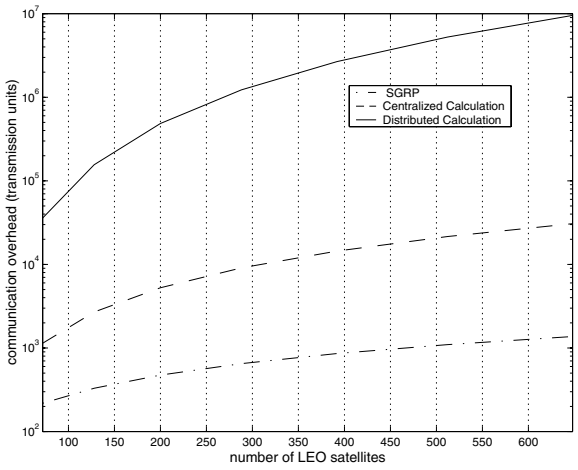


Figure 13. Communication overhead comparison.

communication overhead is expressed in terms of transmission units, which is an entry either in the delay report or in a routing table.

Among these three schemes, SGRP has the least amount of communication overhead. Central routing table calculation generates more communication overhead as the total number of satellites in the network increases. By introducing the hierarchy in SGRP, every LEO satellite only sends delay report to its MEO primary manager. Rather than broadcasting, delay reports are exchanged in MEO layer in an efficient way. After calculation, routing tables are sent back to corresponding LEO satellites through one hop from MEO primary manager to its care-of members. SGRP's communication overhead stays below that of the centralized calculation scheme in all cases. On the other hand, as the distributed calculation scheme requires broadcasting of delay reports to all LEO satellites, which boosts up its communication overhead, the distributed calculation scheme's communication overhead is the highest among the three.

7. Conclusions

In this paper, we introduce a satellite IP network consisting of LEO and MEO layers together with a new routing protocol: Satellite Grouping and Routing Protocol. Using this protocol, LEO satellites are dynamically divided into different groups, for each group a MEO satellite is assigned as the group manager. MEO group managers collect the link delay information from their LEO members, and compute the minimum-delay path for them.

In this paper, we assume that the traffic load on satellite system is moderate, that packets are routed within LEO layer. MEO satellites are used for routing table calculation and transmission of signaling and data control packets. Since the signaling traffic is physically separated from the data traffic, the congested links do not affect the transmission of the delay measurements. SGRP enables the collaboration between different satellite network constellations. MEO satellites are aware of the overall topology of LEO and MEO layers, which gives them the possibility of not to constrain the routing to LEO layers. Besides the management functions and route computation, MEO satellites can be used for other purposes as well, such as packet forwarding and navigation.

The performance of the SGRP algorithm has been assessed with simulations. The performance of SGRP is better than datagram routing algorithm. When satellite failures or link congestion occur, SGRP has mechanisms to reduce their effects on routing. We also showed that SGRP calculates the routing decisions with low communication overhead, it distributes the computational burden to multiple MEO satellites, thus balances the power consumption between LEO and MEO satellites.

References

[1] I.F. Akyildiz, E. Ekici and M.D. Bender, MLSR: A novel routing algorithm for multi-layered satellite IP networks, IEEE/ACM Transaction on Networking 10(3) (2002) 411–424.

[2] C. Chen, A QoS-based routing algorithm in multimedia satellite networks, in: *Proceedings of IEEE 58th Vehicular Technology Conference (VTC2003-Fall)* (Orlando, Florida, Oct. 6–9, 2003) vol. 4, pp. 2703–2707.

[3] H.S. Chang, B.W. Kim, C.G. Lee, S.L. Min, Y. Choi, H.S. Yang, D.N. Kim and C.S. Kim, FSA-based link assignment and routing in low-Earth orbit satellite networks, IEEE Transactions on Vehicular Technology, 47(3) (1998) 1037–1048.

[4] E. Ekici, I.F. Akyildiz and M.D. Bender, A distributed routing algorithm for datagram traffic in LEO satellite networks, IEEE/ACM Transaction on Networking 9(2) (2001) 137–147.

[5] T.R. Henderson and R.H. Katz, On distributed, geographic-based packet routing for LEO satellite networks, in: *Proceedings of IEEE GLOBECOM 2000* (2000) vol. 2, pp. 1119–1123.

[6] J.H. Hu and K.L. Yeung, Routing and re-routing in a LEO/MEO two-tier mobile satellite communications system with inter-satellite links, in: *Proceedings of IEEE ICC'2000*, (2000) vol. 1, pp. 134–138.

[7] J. Lee and S. Kang, Satellite over satellite (SOS) network: A novel architecture for satellite network, in: *Proceedings of IEEE INFOCOM'2000* (2000) vol. 1, pp. 315–321.

[8] J. Perdigues, M. Werner and N. Karafolas, Methodology for traffic analysis and ISL capacity dimensioning in broadband satellite constellations using optical WDM networking, in: *Proceedings of 19th AIAA International Communication Satellite Systems Conference (ICSSC'01)* (April, 2001).

[9] A. Tanenbaum, *Computer Networks*, 3rd edn. (Pretence Hall, Inc., 1996).

[10] Telcordia NetSizer, Internet hosts distribution by continent in January (2001) Available at http://www.infometre.cefrio.qc.ca/fiches/fiche275.asp

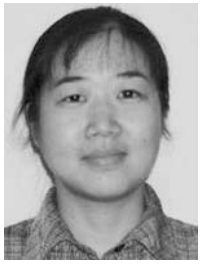
[11] H. Uzunalioglu, I.F. Akyildiz and M.D. Bender, A routing algorithm for LEO satellite networks with dynamic connectivity, ATM-Baltzer Journal of Wireless Networks (WINET) 6(3) (2000) 181–190.

[12] M.D. Voilet, The development and application of a cost per minute metric of the evaluation of mobile satellite systems in a limited-growth voice communications market, Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, USA (Sept. 1995) http://theses.mit.edu/Dienst/UI/2.0/Describe/0018.mit.theses%2f1995–189.

[13] C.J. Wang, Structual properties of a low Earth orbit satellite constellation—the Walker Delta network, in: *Proceedings of MILCOM'93* (1993) vol. 3, pp. 968–972.

[14] W. Werner, G. Berndl and B. Edmaier, Performance of optimized routing in LEO intersatellite link networks, in: *Proceedings of IEEE 47th Vehicular Technology Conference* (1997) vol. 1, pp. 246–250.

[15] W. Werner, A. Jahn, E. Lutz and A. Böttcher, Analysis of system parameters for LEO/ICO-satellite communication networks, IEEE Journal on Selected Areas in Communications 13(2) (1995) 371–381.
[16] L. Wood, A. Clerget, I. Andrikopoulos, G. Pavlou and W. Dabbous, IP routing issues in satellite constellation networks, International Journal of Satellite Communications 19 (2001) 69–92.



Chao Chen received the BE and ME degrees from Deparment of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China in 1998 and 2001, respectively. She is currently working toward her Ph.D. degree in the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA. She is a graduate research assistant in the Broadband and Wireless Networking Laboratory at Georgia Institute of Technology. Her current research interests include satellite and space networks, as well as wireless ad hoc and sensor networks.
E-mail: cchen@ece.gatech.edu

Eylem Ekici has received his BS and MS degrees in Computer Engineering from Bogazici University, Istanbul, Turkey, in 1997 and 1998, respectively. He received his PhD degree in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, GA, in 2002. Currently, he is an assistant professor in the Department of Electrical and Computer Engineering of the Ohio State University, Columbus, OH. Dr. Ekici's research interests include wireless sensor networks, space-based networks, and next generation wireless networks, with a focus on modeling, multiaccess control, routing and multicasting protocols, and resource management.
E-mail: ekici@ece.osu.edu