A QoS-based Routing Algorithm in Multimedia Satellite Networks

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

Abstract— **Real-time multimedia applications impose strict delay bounds and are sensitive to delay variations. Satellite link handover increases delay jitter and signaling overhead as well as the termination probability of ongoing connections. To satisfy the QoS requirements of multimedia applications, satellite routing protocols should consider link handovers and minimize their effect on the active connections. A new routing algorithm is proposed to reduce both the inter-satellite handover and ISL handover. Once a connection request arrives, the remaining coverage time of satellites is used in the deterministic UDL routing. In the probabilistic ISL routing, the propagation delay and existence probability of ISL links are considered to reduce the delay and ISL handover probability. The rerouting algorithm is called when link handover occurs. Experiments show that this routing algorithm results in small delay jitter, low rerouting frequency, and low rerouting processing overhead.**

I. Introduction

In recent years there has been a rapid growth of multimedia services in Internet. As an integral part of the global communication infrastructure, satellite networks will be faced with an increasing demand on real-time multimedia applications. Satellite systems can provide global coverage and constantly sustaining high bandwidth services. Non-GEO satellites have propagation delays comparative to terrestrial networks. The on-board processing capability and inter-satellite links (ISLs) introduced in many LEO satellite systems help to build a robust communication framework. However, satellite networks have different characteristics from terrestrial networks. Satellites are constantly moving, which causes the network connectivity and satellite link delays varying. When satellite links are switched off, handover is required to maintain the active connections. There are two types of handover in satellite networks [1]:

- *Inter-satellite handover:* When the sender or receiver leaves the coverage area of the initial satellites, the entire path should be re-created. The occurrence of inter-satellite handover depends on the time that the ground station remains in a satellite coverage area.
- *ISL handover:* Due to satellite mobility, some ISLs in the network are not always available. For example, when a satellite enters the polar regions, its adjacent inter-plane ISLs are turned off. Similarly, inter-plane ISLs through the seams have very short lifetime. Consequently, if a path contains such a link, it must be rerouted when the link is turned off. Since the link termination only depends on the movement of the satellites, the timing of link shutdowns can be calculated. If the connection duration is known, the ISL handover can be predicted.

This work is supported by the National Science Foundation under Grant ANI-0087762.

The link handover causes delay jitter and signaling overhead. Moreover, due to the inefficiency of network resources and rerouting delay, handover increases the forced termination probability of ongoing connections which is more annoying than blocking a new connection request. Real-time multimedia applications impose strict delay bounds and are sensitive to delay variations. As a result, routing protocols through satellite network should consider link handover and minimize its effect on each individual connection.

A dynamic routing concept is introduced for ATM-based satellite networks in [2]. The dynamic network topology is considered as a periodically repeated series of K topology snapshots. Using a sliding window, a set of k-ordered path sequences between satellite nodes are selected in each topology snapshot with an aim to minimize handover delay jitter and reduce link handover rate. However, the optimization is not done between end users and the inter-satellite handovers are not considered. The predictive routing protocol proposed in [3] provides guaranteed QoS services in satellite networks. It exploits the deterministic nature of the LEO satellite topology to predict the traffic load on the ISLs up to a short time in the future. k-ordered paths for a particular connection is computed for each staggered cell to maximize the minimum residual bandwidth. The optimal path is picked from the path set to reduce the link changes as well as to balance the user traffic. This protocol does not consider inter-satellite handovers and the computation overhead grows dramatically as k increases. The probabilistic routing protocol [4] utilizes the LEO satellite network dynamics and call statistics, tries to reduce the number of rerouting attempts due to link handover. An ISL is removed from route computation if it is expected to experience an ISL handover with a probability higher than a target probability (p) during the route establishment phase of a new call. The computation of the ISL handover probability is deduced from the hexagon effective footprints of satellites.

In this paper, a new routing protocol is proposed to reduce both the inter-satellite handover and the ISL handover

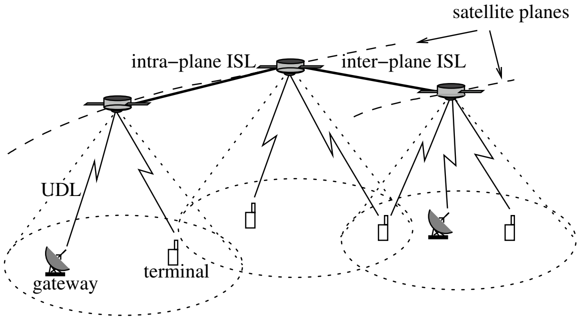


Fig. 1. Satellite network architecture.

probability. It is based on a general satellite constellation model in which satellite footprints may be overlapped. The predictability of link handover is utilized while computing the ISL path through the satellite constellation. Different from the probabilistic routing protocol [4], our routing protocol does not remove the links with a handover probability higher than p to avoid high new call blocking probability. A modified version of the footprint handover rerouting protocol (FHRP) [5] is used for rerouting when link handover occurs.

The remainder of this paper is organized as follows: The satellite network architecture is introduced in Section II. Section III describes our routing and rerouting algorithm in detail. The performance evaluation of the algorithm is presented in Section IV. Finally, Section V concludes the paper.

II. Satellite Network Architecture

The satellite network architecture considered in this paper is shown in Figure 1. It consists of satellites orbiting the Earth and ground stations on the Earth surface. Satellites may lie within one layer or in multiple layers, such as a combination of LEO and MEO layers. Ground stations may be fixed, which perform as gateways between satellite and terrestrial networks. Mobile ground stations are handheld terminals that move around with users. It is assumed that the movement of mobile ground stations can be ignored compared to the fast movement of satellites.

The satellites are connected through inter-satellite links (ISLs). There are two types of ISLs: intra-plane ISLs and inter-plane ISLs. The intra-plane ISLs connect satellites within the same plane and are maintained permanently since their relative positions are fixed. The inter-plane ISLs are between satellite in different planes. They are operated only outside the polar regions and need to be switched off temporarily due to the change of distance and viewing angle between them. The ISLs enable the routing of messages in satellite network without the requirement of terrestrial resources. Satellites communicate with the ground stations via user data links (UDLs). The footprints of satellites can be overlapped, thus a ground station can be connected to several satellites.

Since the satellite constellation is well-planned before deployment and does not change during operation, typically the location of any satellite (i.e., altitude, latitude, and longitude) at any time can be calculated according to the trajectory information. The location information of mobile ground stations (i.e., latitude and longitude) can be obtained using GPS technology and reported to the satellites. The location information of fixed and mobile ground stations is stored in databases either centrally or in a distributed manner, and can be retrieved whenever needed. The satellites periodically exchange their local information, which includes the available bandwidth on outgoing links and location information of ground stations within their footprints.

III. The New Routing Algorithm

Real-time multimedia applications impose strict delay bounds and are sensitive to delay variations. For a network to deliver QoS guarantees, it must reserve and control resources accordingly. The changing connectivity pattern of satellites calls for link handover to maintain the active connections. However, link handover increases delay jitter as well as signaling overhead. Excessive handovers also increase the blocking probability of ongoing connections.

The goal of our new routing and rerouting algorithm is to reduce the delay jitter while guarantee the bandwidth requirement. It incorporates the location information of satellites and ground stations to predict the lifetime of satellite links, trying to build stable paths for connection requests and reduce the probability of link handovers during connection lifetime. The routing problem considered in this paper is as follows: The connection establishment requests arrive on-line; A connection is established by allocating the required bandwidth along some path between the source and the destination nodes; The allocated bandwidth is released when the connection terminates.

It is assumed that the duration of a connection is exponentially distributed with known mean holding time. When a ground station issues a connection request, it should specify the following four parameters:

1) Location of source s: It is known to the ground station by GPS service;
2) ID of destination ground station d: It is used to retrieve the location of d in the location databases;
3) Expected connection duration $1/\mu$: It is specified through the distribution probability function;
4) Requested bandwidth bw: For CBR-type applications, the requested bandwidth is fixed through the connection duration. For VBR-type applications, the requested bandwidth can be described by maximum bandwidth and sustained bandwidth, or calculated using a token bucket model and the requested delay bound [3].

In this section, the detailed design of our new routing protocol is presented. It includes the following three parts:

- *Deterministic UDL routing* chooses the ingress and the egress satellites according to the locations of the source and the destination ground stations.
- *Probabilistic ISL routing* selects the path within satellite constellation between the ingress and the egress satellites.

- *Handover rerouting* in case of inter-satellite handover and ISL handover.

A. Deterministic UDL Routing

Most of the existing work assume that the minimum number of satellites is used to achieve global coverage. Thus, the overlapped area of the neighbor satellites' footprints does not constitute a significant portion of the overall coverage area. However, the overlapped area increases at higher latitude for polar-type satellite constellations. Furthermore, overlapped coverage areas can be utilized to increase the resources available to regions with dense population.

In this paper, it is assumed that a ground station may lie in the overlapping areas of several satellite footprints. Upon receiving the connection request, the source ground station selects the ingress and the egress satellites for the path between the source and the destination. Generally there are two types of metrics to select the access satellites:

- *Maximum coverage time (Max-Time)*: Select the satellites with sufficient bandwidth and the maximum remaining coverage time to the ground station. By doing so, the probability of inter-satellite handover is minimized. The computation of the remaining coverage time of satellites can be done with the knowledge of location information and will be explained later in this section.
- *Maximum received power (Max-Power)*: Select the satellite with sufficient bandwidth and the strongest received power. It is assumed that all the satellites have the same transmit power. Then, the selection of access satellite based the Max-Power metric equals to choosing the closest satellite to the ground station. This can be easily done using the satellites' location information.

To reduce the inter-satellite handover probability, we choose the Max-Time metric for UDL routing. Once the ingress and the egress satellites are chosen, the required bandwidth is allocated along the corresponding UDLs.

The remaining coverage time of a satellite S for a ground station E is calculated as follows.

[PARAMETERS]:

r_f: The radius of satellite S's footprint. $r_f = Re \cdot [\arccos(\frac{R_e}{r_s} \cdot \cos\theta_{min}) - \theta_{min}]$, where R_e is the Earth radius, r_s is the satellite orbit radius, θ_{min} is E's minimum elevation angle.
(v, ϵ_0): Satellite S's footprint velocity and movement direction.
$E(L_e, l_e)$: Ground station E's latitude and longitude.
$S(L_s, l_s)$: Satellite S's latitude and longitude.

Multimedia traffic has longer connection holding times compared to the traffic in the traditional circuit switched networks. For multimedia connections, Earth's rotation may no longer be ignored. Hence, the Earth's rotation is considered in computing the footprint velocity v and direction ϵ_0.

[STEPS]:

– *Compute azimuth angle of S toward E (AZ):* Using the law of sines and cosines for spherical triangles,

$$\begin{aligned} \sin\alpha &= \frac{\sin|l_e - l_s| \cdot \sin L_s}{\sin\gamma} \\ \cos\gamma &= \cos L_e \cos L_s \cos(l_s - l_e) + \sin L_e \sin L_s \end{aligned} \tag{1}$$

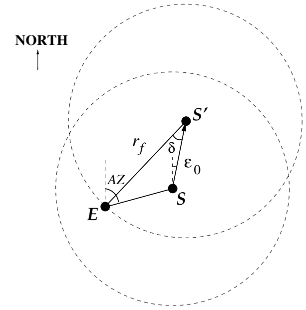


Fig. 2. Calculation of the remaining coverage time of S toward E.

The azimuth angle can be computed as:

$$AZ = \begin{cases} \alpha, & \text{if } S \text{ lies northeast of } E \\ \pi - \alpha, & \text{if } S \text{ lies southeast of } E \\ \pi + \alpha, & \text{if } S \text{ lies southwest of } E \\ 2\pi - \alpha, & \text{if } S \text{ lies northwest of } E \end{cases} \tag{2}$$

– *Compute the remaining coverage time of satellite S towards E:* Suppose after time T, the satellite moves from S to S' and the ground station E is on the edge of the satellite's footprint as shown in Figure 2, the trace of the sub-satellite point during time period T is represented by SS'.

First, if S, S', and E lie in a line,

$$SS' = \begin{cases} r_f - SE = r_f - \gamma R_e, & \text{if } \epsilon_0 = AZ \\ r_f + SE = r_f + \gamma R_e, & \text{if } |AZ - \epsilon_0| = \pi \end{cases} \tag{3}$$

Otherwise, $\angle ESS' = |\pi - |AZ - \epsilon_0||$. Using the law of sines in spherical triangle ESS',

$$\frac{\sin(r_f/R_e)}{\sin\angle ESS'} = \frac{\sin\gamma}{\sin\delta} = \frac{\sin(SS'/R_e)}{\sin(\angle ESS' + \delta)}$$

The remaining travel distance SS' can be written as:

$$SS' = R_e \cdot \arcsin\left[\frac{\sin(r_f/R_e)\sin(\angle ESS' + \delta)}{\sin\angle ESS'}\right] \tag{4}$$

where $\delta = \arcsin[\frac{\sin\angle ESS' \cdot \sin\gamma}{\sin(r_f/R_e)}]$.

Therefore, the remaining coverage time $T = SS'/v$. The satellite with the maximum coverage time T^* is selected as the access satellite. The ingress and the egress satellites are the satellites that provide the maximum coverage time (i.e., $T^*_{r,u}$ and $T^*_{r,d}$) to the source and the destination ground stations respectively. Assume exponential connection duration T_c with mean $1/\mu$, then,

$$Prob[\text{inter-satellite handover}] = P(T_c > T^*_r) = e^{-\mu T^*_r} \tag{5}$$

where $T^*_r = \min(T^*_{r,u}, T^*_{r,d})$.

B. Probabilistic ISL Routing

The ISL routing is done after the decision of the ingress and the egress satellites. Since the ISL handover operations cause high signaling overheads and long delays, two coefficients are assigned to each link i: propagation delay d_i and existence probability p_i.

- *Propagation delay (d_i)* the propagation delay of each satellite link at any specified time can be easily deduced by the satellite trajectory information. d_i for an inter-plane ISL is changing constantly with the satellite movement.
- *Existence probability (p_i)* is the probability for the ISL link not to be shut down either before the connection ends or an inter-satellite handover occurs.

Since the knowledge of the exact time that the inter-satellite handover would occur is known after the deterministic UDL routing, and the connection duration T_c conforms to exponential distribution with mean $1/\mu$, if we can predict the ISL handover time ($T_{i,lh}$) of link i, then

$$\begin{aligned} p_i &= Prob[T_{i,lh} > min(T_c, T_r^*)] \\ &= \begin{cases} 1, & \text{if } T_{i,lh} \geq T_r^* \\ 1 - e^{-\mu T_{i,lh}}, & \text{if } T_{i,lh} < T_r^* \end{cases} \end{aligned} \tag{6}$$

The cost of link i is computed as:

$$C_i = \begin{cases} d_i \cdot (1 - \ln p_i), & \text{if available bandwidth } \geq bw \\ \infty, & \text{if available bandwidth } < bw \end{cases} \tag{7}$$

As $p_i \to 0, C_i \to \infty$. Higher existence probability contributes to lower link cost. When $p_i = 1$, C_i is represented by the link propagation delay. Distributed Dijkstra or Bellman-Ford algorithm is applied to find the minimum cost path through the satellite constellation upon a connection request. Once a ISL path is found, the required bandwidth is allocated along the path.

C. Handover Rerouting

The handover rerouting algorithm is modified from the augmentation algorithm of FHRP [5]. Suppose if at time $t = t_e$, one of the ground station moves out of the footprint of its access satellite S. A new satellite S' with the maximum coverage time is selected as the new access satellite. Instead of computing the new ISL path immediately, the path augmentation algorithm is handled by S' as follows:

1) The satellite S' checks whether it is already on the ISL path. If so, the portion of the old path up to S' is deleted and the reserved bandwidth is released. The new ISL path starts from S'.
2) If S' is not on the ISL path, a direct link to one of the satellites on the path is searched starting from the other end of the path, If a direct link with sufficient bandwidth to support the connection is found, the link is augmented to the original path.
3) If a direct link between S' and the satellite nodes on the ISL path with required capacity is not found, the reserved bandwidth on the old path is released and a full rerouting (i.e., deterministic UDL routing and probabilistic ISL routing) is performed.
4) If the ingress and the egress satellites of the last computed route have both been updated, the probabilistic ISL routing between the new ingress and the egress satellites is called. This is to prevent frequent rerouting attempts due to non-optimal routes.

During connection time, if one of the satellite links along the ISL path needs to be switched off, full rerouting is called.

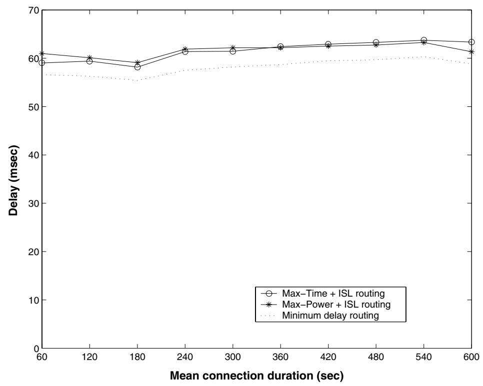
IV. Performance Evaluation

We have extended the VINT *network simulator (ns2.1b9a)* [6] by including modules of our routing and rerouting algorithms. A LEO satellite network with 288 satellites are considered in the simulation. There are 12 orbital planes with 24 satellites in each plane. Satellite orbits are 1375 km in altitude with an orbit inclination angle of 84.7^o. The minimum elevation angle of ground stations is 40^o. Each satellite has 2 intra-plane ISLs and 2 inter-plane ISLs.

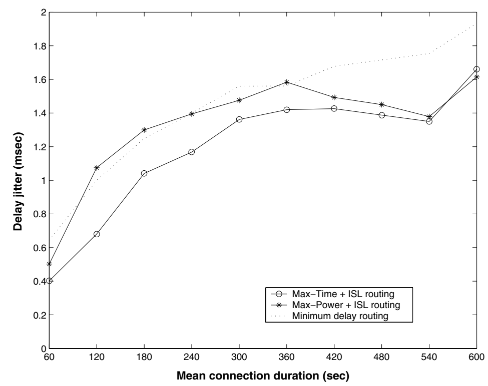
To evaluate our routing protocol, the path metrics (i.e., path delay, delay jitter, rerouting frequency, and rerouting overhead) of a connection between a source-destination pair are monitored. The source is located at $(33.39^oN, -84.26^oW)$ in Atlanta, the US, and the destination is at $(39.55^oN, 116.25^o)$ in Beijing, China. The sender generates connection requests with different mean connection duration $1/\mu$. The results are the average of 100 independent experiments. Performance comparisons are made among three different algorithms: Max-Time deterministic UDL routing with probabilistic ISL routing, Max-Power deterministic UDL routing with probabilistic ISL routing, and the Bellman's algorithm that returns the minimum delay path between the source and the destination.

Figure 3(a) and 3(b) give the delay metrics (i.e., end-to-end delay and delay jitter) of the above three routing algorithms. Apparently the minimum delay routing returns the path with shortest end-to-end delay, as the other two algorithms divide the end-to-end routing into UDL routing and ISL routing. However, their extra delay difference is within 5% of the delay of the shortest path. The delay jitter is represented by the variance of end-to-end delays. Among the three algorithms, the Max-Time with ISL routing has the minimum delay jitter. This is because that our rerouting algorithm tries to keep the original path and reduce the link handover probability. However, the frequent path updates generated by the minimum delay routing gives more opportunity to delay variance among different paths. Especially when the mean connection duration increases, the path updates more frequently, which in turn causes larger delay jitter.

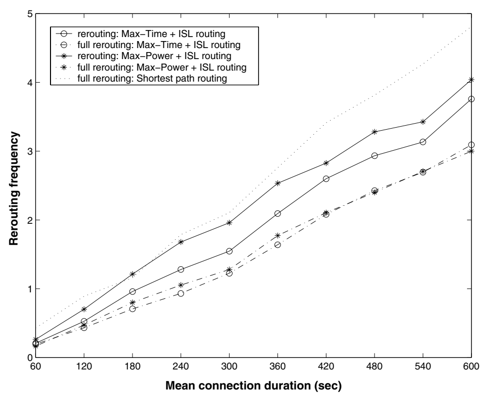
The average rerouting frequency and the rerouting overhead of the three routing algorithms are shown in Figure 3(c) and 3(d) respectively. The rerouting frequency is measured by the number of handover attempts in the connection duration. For Max-Time and Max-Power with ISL routing algorithms, we also plotted their full rerouting frequency, which stands for the average number of rerouting attempts with actual computation of a new ISL path. The number of rerouting is the sum of the full rerouting and the augmentation rerouting numbers. For minimum delay routing, all handover attempts call for full rerouting computation. Among the three algorithms, minimum delay routing has the highest rerouting frequency. Moreover, the full rerouting frequency of the Max-Time and Max-Power with ISL routing algorithms is much lower than that of the



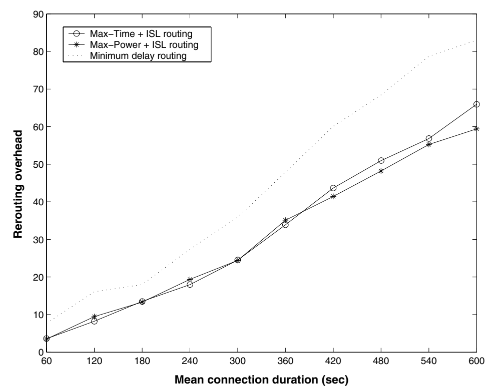
(a) Delay



(b) Delay jitter



(c) Rerouting frequency



(d) Rerouting overhead

Fig. 3. Comparison of path metrics.

minimum delay algorithm. The rerouting overhead is represented by the number of updates in the routing tables of all nodes, i.e., removing the record of a connection or recording a rerouted connection. The augmentation rerouting attempts to keep part of the original path by checking if the new access satellite is on or next to the old path. Therefore, less nodes are to be updated and processing overhead is reduced. This is also shown in the Figure 3(d) where minimum delay routing causes much higher rerouting overhead than the other two.

V. Conclusions

In this paper, the deterministic UDL routing based on the maximum coverage time together with the probabilistic ISL routing are introduced for routing between two ground stations via satellite networks. A rerouting algorithm is also described for link handover. The routing algorithm utilizes the satellite trajectory information and the connection statistics to reduce the probability of both the inter-satellite handover and the ISL handover while satisfy the user's bandwidth requirement. Experiment results have shown that this routing algorithm reduces delay jitter, rerouting frequency, and rerouting processing overhead compared with those of the minimum delay routing and the Max-Power UDL routing with probabilistic ISL routing.

References

[1] I. F. Akyildiz, H. Uzunalioglu, and M. D. Bender, "Handover management in low earth orbit (LEO) satellite network," *ACM-Baltzer Journal of Mobile Networks and Applications (MONET),* December, 1999.

[2] M. Werner, "A dynamic routing concept for ATM-based satellite personal communication networks," *IEEE Journal on Selected Areas in Communications,* vol. 15, no. 8, pp. 1636-1648, October, 1997.

[3] O. Ercetin, S. Krishnamurthy, S. Dao, and L. Tassiulas, "Provision of guaranteed services in broadband LEO satellite networks," *Computer Networks,* vol. 39, pp. 61-77, 2002.

[4] H. Uzunalioglu, I. F. Akyildiz, and M. D. Bender, "A routing algorithm for connection-oriented low earth orbit (LEO) satellite network with dynamic connectivity," *ACM-Baltzer Journal of Wireless Networks (WINET),* vol. 6, pp. 181-190, 2000.

[5] H. Uzunalioglu, I. F. Akyildiz, Y. Yesha, and W. Yen, "Footprint handover rerouting protocol for LEO satellite networks," *ACM-Baltzer Journal of Wireless Networks (WINET),* vol. 6, no. 3, June, 1999.

[6] The Network Simulator (*ns*), *version 2.1b9a*. http://www.isi.edu/nsnam/ns/