

High Performance Routing in a LEO Satellite Network

Roy Kucukates

Cem Ersoy

Computer Networks Research Laboratory, NETLAB

Bogazici University, Istanbul, Turkey

kucukate@boun.edu.tr

ersoy@boun.edu.tr

Abstract

Satellite systems are designed as a backup system to the existing land based wireless systems. Because of the global coverage of the system it becomes more economic for the rural areas where there is no wired infrastructure of communication. Besides being more economic to the rural areas the satellite networks are not affected by the natural disasters and continue to give the service in such situations. In this research a routing method in a routing set concept is investigated. The system tries to minimize the maximum flow over a given set of shortest paths from the source to destination and aims to leave additional capacity over the inter-satellite links to be able to utilize better the limited capacity and, also, to avoid the congestion.

1. Introduction

The operators of new satellite systems give more importance over communication satellite networks than broadcasting systems since, they believe that on the average 50% of the world population will be living in rural areas in 21st century [1]. Along with the population need it will be preferred for the services that could be given on satellite networks such as health care, education, crisis management, environmental monitoring, electronic commerce and Internet [2, 3]. Also, the satellite networks are thought as a backup communication system where land based communication networks exist, because the satellite systems do not suffer from natural disasters (earthquake, flood, etc) and continue to give service globally.

The system is convenient for various QoS requirements for different communication needs as bulk transfers, video broadcasts, video conferencing, interactive computing, voice calls, distance learning, etc. [4, 5] and is intended to augment the existing terrestrial and cellular networks [6].

This paper will introduce a new packet routing mechanism over interconnected satellite networks that will be based on Routing Set (RS) concept and will give simulation results of some scenarios. It is known that the continents where communication can occur are located around the equator, namely, below the 60th latitude. This means the communication path distances between the satellites (inter satellite links) over the communicating regions are approximately equal. This last brings the result that a communication's propagation delay is related to the number of satellite nodes (hops) it passes. The RS method proves that all of the alternative shortest paths between

source and destination pairs have the same length. This means the propagation delay is the same for all path alternatives of the communication pairs. Our path finding algorithm tries to minimize the maximum flow over the RS. This method implies knowledge of the flows over the RS (not all of the network). The resulting network is a better-utilized and balanced flow network. This last reduces at the same time the processing delay and minimizes the buffer requirements of the switching nodes.

The second section categorizes and describes the satellite networks. Section 3, introduces the Routing Set concept. Section 4 will deal with Min Flow Max Residual (MFMR) routing algorithm. Section 5 will describe the test network, the related scenarios that will be simulated over the network using the RS and MFMR and will give the performance evaluation. Section 6 will conclude the paper.

2. Satellite networks

The satellite network suffers from the propagation delay. It is known that in an interactive communication the delays above 400 msec become "annoying" [1]. The delay is related to various factors. The most important is the propagation delay. Since the propagation path is long compared with land based systems the satellite networks are inherently slower than other networks. The altitude of the satellite constellation affects the delay. Due to Van Allen Radiation Belts the satellite networks can be at 3 different range of altitudes (orbits): Low Earth Orbit (LEO) satellite networks, below the first belt (500-1500 km); Medium Earth Orbit (MEO) satellite networks, between the first and second belt (5000-12000 km); and Geosynchronous Earth Orbit (GEO) above the second belt (20000-...km). The resulting end-to-end propagation delays from ground to ground are 20-25msec for a LEO, 110-130msec for MEO and 250-280msec for GEO systems [1, 7].

Thus, satellite systems can be categorized according to their orbits. The Geosynchronous orbit satellites are located over the equator and have the same angular speed as the earth. This means that a geosynchronous satellite keeps its place at the same point in the sky for a reference point on the earth. Its service area is constant and is called a footprint. The MEO and LEO orbits being closer to the earth in order to keep their satellites at the same altitude have to do a bigger angular motion compared to earth. This results in a motion of a satellite seen from a reference point over earth. That's why the service area of a LEO or MEO satellite changes during a tour. In order to give a

continuous service with a LEO or MEO satellite the system has more than one satellite nodes and they are moving on the same orbit with the same direction and angular speed. In such constellations the system can have two types of network architectures. In the first architecture bent pipe satellites are used and the system sends the messages to a ground switching center and the center routes the message over wired networks and broadcasts the packet at the destination satellite which, in its turn broadcasts the message to its footprint area and the destination terminal receives the packet. In the second type of architecture the satellites have switching capabilities and inter-satellite links and can route the packets in the air. The routing information can be static or dynamic. In static routing all of the route information are decided and uploaded to the satellite nodes. In dynamic routing the paths are rediscovered every time a new call (in circuit switching) or packet arrives (in datagram routing). The routing information can be discovered by a ground switching center or by the satellite itself. In both cases the satellites have to forward the packets along their inter-satellite links. The satellites in the same orbit do not change their position relative to the others in the orbit. Thus, they keep the same inter-satellite links (ISL) with their neighbors. These links are called intra-orbital links. They may have other links to the neighboring orbits. These links are called inter-orbital links [8]. The number of required satellite to cover the intended service area, ISL and required channels are described in [9]. The orbits used by these satellites are given in [10].

Besides the satellite networks, high-altitude aeronautical platforms also are being designed in order to give metropolitan area coverage with low-delay communication [11].

The satellite network architecture looks like a Twisted Manhattan network given in [12]. In fact, the real network has small differences that affect the optimal routing in this mesh architecture. The real architecture is given in Figure 1. The satellite network has two seams. Each seam is presented with a different shading on the figure. The upper right and the down left sides of the network are neighbors. However, since they move in opposite directions, they continuously break their inter-orbital links and reconnect with the next arriving satellite.

In Figure 1, the ground relative positions can be thought as fixed and the routing can be done accordingly. This means that the routing can be done according to source-destination pair positions on earth. If we consider an instant and think the network as static, the routing can be done easily. In our assumption, we associate the routing information with the current coordinates of the node it is found. The coordinate of the satellite is given by (s, o) where s is its position of satellite in the orbit counting from North Pole and o is the orbit number. For the next instant of time where a new satellite comes to the position, the recent satellite can forward the routing and active connection information over the ISL. Polar orbit satellite networks are special case topologies. All of the satellite nodes pass through the poles. In such a topology we can specify the network with the number of orbits and the number of satellites per orbit. Because of the spherical shape of the network, there exist more than one path to the destination node. The discovery of the shortest path depends on the direction where the destination is searched. In order to find the direction, a virtual network is introduced in the next section.

The performance of the system is a key issue in networks. As the satellites are moving faster than mobile nodes on earth the

transmitter and receiver will do handover very often. New nodes and also the whole network should guarantee the QoS especially to multimedia service users. Critical QoS parameters are given in [13, 14]. Besides the QoS, the communication protocols such as TCP/IP [15] or medium access control protocols [16] should be redesigned or modified in order to be used over satellite networks.

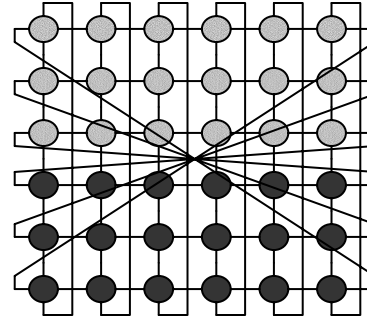


Figure 1. The satellite network topology

3. Routing Set

Since the satellites cover smaller areas in LEO systems, the traffic requirements become unbalanced [17] due to high population in cities and low at rural areas. The communication requirements between high population areas are larger. This problem can be resolved by distributing the flow in a balanced way over all possible ISLs between the communicating nodes. Recent research includes Dogleg, Parallel Highway and Polar Hop routing methods [18]. In Dogleg routing, the path is first horizontal and vertical or the reverse, but never both at the same time. In Parallel Highway routing, the flow is routed over a parallel path to the shortest paths. This is used to distribute the high flow outside of the crowded regions. The last one is the Polar Hop. In this method, all of the connections are routed over the path that pass through the polar region.

The footprints of satellites move faster than any terminal on earth. This causes to frequent handovers of terminals from a satellite to the other. This is called satellite-fixed cell system. The alternative of this is the earth-fixed cell system [19]. In this method the cell on earth is fixed and the satellite's antenna beams are steered so as to point toward this fixed cell during some interval of time. In either way the source-destination pair and the path between them are transferred to new satellites coming to the positions where the path was passing through. In a network where the intermediate nodes are changing all the time in a constant way the path handover becomes important. However, if the orbit period is divided into convenient small time intervals, the network can be thought as fixed during these intervals. The routing can be done according to this static network, since the newly arriving satellite will take the place of its ancestor. In some recent work, this network is considered as having finite states (Finite State Automation FSA) and all the routing decisions are pre-calculated for all states (static routing) in [20] it is shown that static routing gives better results than dynamic routing. In some work, it is believed that the FSA assumption is not realistic and the path has to do handover in every moment where the underlying network architecture changes due to the satellite motion and the routing must be done according to the probabilistic model that chooses the least

probable handover ISLs as the most convenient path. This is called the Probabilistic Routing Protocol (PRP) [21]. But this method assumes that the link cost is a function of the time and location homogeneous traffic load [22]. Whereas this assumption is not realistic since the traffic is not time-homogeneous. New traffic can easily congest the PRP path and leave the other paths underutilized. The maximal resource requirement of a path is determined by the hop count and by using a minimum hop count path will reduce the maximum resource consumption [23]. In our work we will use this assumption. If we consider the network in Figure 1, we can see that there exists more than one shortest path from the source to the destination. We call all the nodes in the rectangle, where the source is a corner and the destination is the other on the diagonal, a Routing Set (RS). All the directions toward the destination are located on a shortest path from the source to destination.

In Figure 2, all possible paths are shown. It can be proven that all of the paths using any one of the links with the specified directions are equal and are shortest paths. It can also be proven that the paths using these directions are loop-free. Thus, the routing problem for a satellite system becomes the “shortest path” discovery problem. However, since the network is spherical and there exist many RS between the source and destination and most of them pass through the polar region or through the horizontal plane a virtual network has to be considered while finding the right RS.

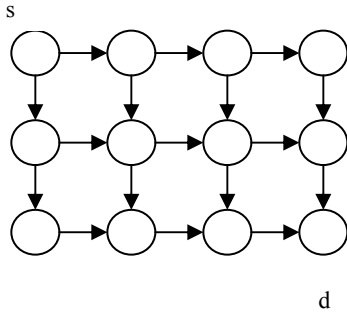


Figure 2. Paths from the source s to destination d .

We assume N orbits and S satellites per orbit. Because of the spherical shape of the earth, the first and the N^{th} orbits are neighbors in both sides. The nodes between the borders of these orbits are called seams. The border orbit satellites of different seams that are neighbors are counter-rotating. The satellite network will look like the topology shown in Figure 1, if we ignore the spherical shape of the earth. The real satellite network a virtual routing model of the whole network is presented in Figure 3.

In this model, all the real source and destination pairs are located in the center rectangle, namely in $Net 1$ and/or $Net 2$. The routes will be calculated on this virtual network and mapped back to the real connections. Suppose we have the source in $Net 1$ and the destination in $Net 2$. The destination node will be placed, according to a mapping function, virtually in all $Net 2$ and $Net 2^{-1}$. All the routing sets will be formed according to single source and multiple virtual destinations pairs. The minimum number of hop RS will be selected as primary RS since it includes the shortest path candidate. In this virtual network the maximum hop count for routing sets L_R can be calculated by the following formula:

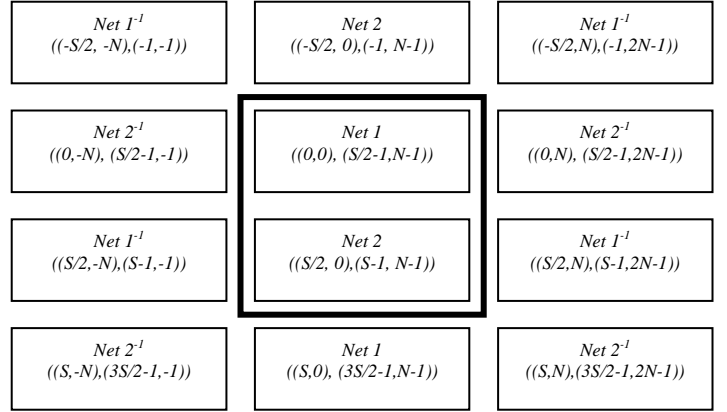


Figure 3. Virtual routing model.

$$L_R = |x_d - x_s| + |y_d - y_s|$$

We assume that satellite positions are given by the pair (x, y) where x represents the axis and y the ordinate on the coordinate plane. All the routing are calculated based on this node positions. For the shortest path the RS with the minimum length is chosen and the path is found in this RS. In this network, the direction of the RS is calculated by evaluation of formula (1) for six RS candidates (namely six addition operations). The path is selected within this RS and mapped back to the real nodes. The shortest path in our network is chosen by a flow distribution mechanism that will be described in the next section.

4. Minimum flow maximum residual (MFMR) path

We have chosen the boundaries of the communication path up to now. Every alternative path has the same length. During this step a routing algorithm has to be chosen. The algorithm has to [24]:

- Be of an average complexity,
- Do loop-free routing,
- Converge,
- Have minimum computational overhead,
- Have storage overhead.

In [8], the minimum propagation delay is taken as basis for routing and the number of hops is not considered. In the same work, the length of ISLs near the pole are assumed as shorter than those near the equator. Moreover, routing over the polar ISLs seems to have less propagation delay. However, in this method the polar ISLs will suffer from congestion. Also, the population percentage is higher close to the equator than to the Polar Regions. Which means that most of the ISLs within a RS have approximately the same length and choosing the least propagation delay path will congest the link whereas the other links will remain unused. This method will minimize at the same time the call blocking probability that is one of the main constraints in [20]. In our method, it is assumed that all the ISLs have the same length in order to give the same priority to all the paths.

In [22], the satellite network topology is considered as changing during time because of the constant motion of satellites. However, as the number of satellites grow (e.g. 66 for Iridium and 288 for Teledesic, etc.) with the FSA assumption, at a given instant, all of the satellites have their four ISLs

Table 2. Parameters for Scenario 1.

Grp	Transmit	Rcvng	Call (sec)	Load (10 ³ pk/sec)
1	(0,2), (1,1), (2,1), (1,2), (2,2)	(0,2), (1,1), (2,1), (1,2), (2,2)	40	H. flow: 10
2	(1,2), (2,2), (1,3), (2,4)	(1,2), (2,2), (1,3), (2,4)	40	L. flow: 4
3	(4,1), (3,2), (4,2)	(4,1), (3,2), (4,2)	40	N. flow: 7
4	(0,9), (5,3), (5,4)	(0,9), (5,3), (5,4)	40	H. flow 10
5	(3,10), (3,9), (3,9), (4,9), (5,8), (5,7)	(3,10), (3,9), (4,9), (5,8), (5,7)	40	L. flow: 4

Table 3. Parameters for Scenario 2.

Grp	Transmit	Rcvng	Call (sec)	Load (10 ³ pk/sec)
1	(3,10), (3,9), (4,9), (5,7), (5,8), (0,2), (1,1), (2,1), (1,2), (3,2), (1,3), (2,4), (4,1), (4,2), (0,9), (5,4), (5,3)	(2,2)	40	L.flow: 5

5.2. Result of the simulations

Since there is no operational satellite network yet, the routing is evaluated on a simulator, Opnet Modeler 7.0.B. As the objective of the routing was to minimize the maximum flows on the links, the scenarios are evaluated for their link flows. The maximum flow of the network is represented by the link with the maximum flow during the overall simulation time. The simulation durations are chosen as 1000 sec for the first scenario and 2000 sec for the second and the third scenarios. Each scenario is run 10 times and the maximum of the result is chosen. In the first scenario, besides the max flow, the end-to-end delay for various number of hops are evaluated. In the second scenario, the surrounding nodes of the hot-spot are evaluated for their max flow and buffer requirements. In the third simulation, the max flows of critical links around the intra-continental hot-spot are evaluated.

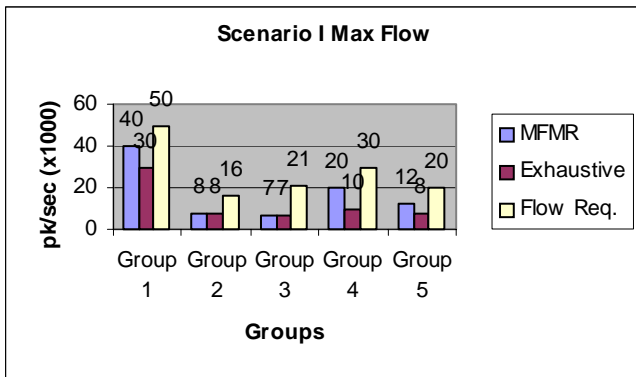


Figure 5. Comparison of the maximum flows

The maximum flows in each group are close or equal to the ideal case. The ideal case is evaluated by using exhaustive search. The flow requirement of the group is the sum of individual flows of the nodes in the group. The MFMR is below the total flow requirement of the network. This last means that

the flows are distributed over the RS and no additional flow from the other groups is routed over the specified RS. In Figure 6, it can be seen that even for the worst case (5 hop) the end-to-end delay from ground to ground is below 400msec.

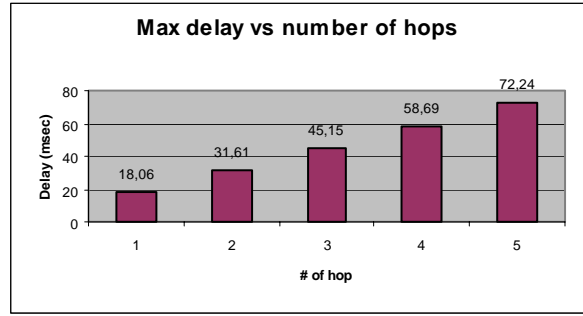


Figure 6. Effect of number of hops to the end-to-end delay.

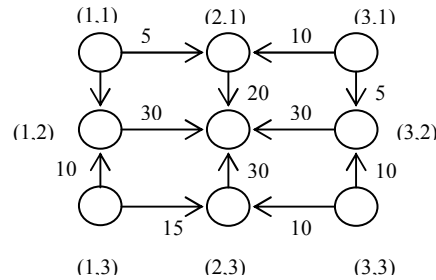


Figure 7. The max flows around the global hot-spot.

In the second scenario, the max buffer sizes of the surrounding satellite nodes are recorded. It is seen that, during the simulation, all the surrounding nodes had at most 3000 packets in their buffer at the same time. This results in that for the case studied distributing the flow over the RS reduces the processor utilization on the satellites. The resulting max flows around the hot-spot is given in Figure 7.

In the global hot-spot case the flows are evenly distributed around the satellite node. However, it should be noted that the hot-spot is chosen right in the upper middle of the communicating nodes. This position is a place where the nodes are distributed evenly in both directions (to the left and right.) The maximum flow of the overall network is 30,000 pk/sec.

The route calculation algorithm does not reserve channels before the connection establishment and while the path is chosen as most promising for a source, some others may be chosen the path and sent the channel reservation message. In this case, the system waits until one of the connections finishes, but the maximum value is captured by the probes and represented as a bad case on the graph.

The RS routing method gives better results when the RS is bigger. In the global communication case, the system can chose as a RS all satellites, which results in better distribution of the total flow. As the flows are distributed over the RS, the satellite nodes are less utilized and the buffer size requirements are limited to 3000 packets for the studied case. The RS algorithm intends to distribute the flows over all the available shortest path routes in order to reduce the link utilization and prevent the congestion. The algorithm tries to use the underlying network better than the dogleg, parallel highway and polar hop since it uses the shortest path route and the minimum flow link as the most convenient link.

6. Conclusion

The satellite networks are designed to give global coverage and communication services. The system, being far from the mobile terminals located on earth suffers from the high delay compared to land based wired networks. The satellite system is a backup network to the land based systems. Moreover, the network is not affected by the natural disasters and is more economical to reach the rural areas where land based communication systems become expensive. Since the system globally covers the earth, the potential number of subscribers is much more than land based wireless systems. This number of potential subscribers causes high network flow on the satellite system. Some research assumes that the link propagation delays are the bottleneck of the satellite systems and route accordingly. In this work, the Routing Set concept and the minimization of the maximum flow for satellite systems are designed and simulated. The system assumes that most of the population lives below the 60th latitude on earth. This assumption gives the right to have all the inter satellite links between the neighboring orbits equal. This results in the fact that choosing the minimum hop path between the source and the destination pair will reduce the processing delay and, the most important will distribute the maximum flow over underutilized links. This method will keep additional capacity on the links and on the satellite processors.

Three scenarios are implemented and simulated on Opnet. The scenarios are designed for representing the normal operation (Scenario I) and the natural disaster case where all of the nodes try to reach the hot-spot area with the simplest service (Scenario II). In the last scenario (Scenario III), the network simulates the continental hot-spot case. For the studied cases, it is seen that the system tries to minimize the maximum flow and reduces the buffer sizes at the satellite nodes and also keeps the end-to-end delay from ground to ground under the upper boundary of interactive communication (400 msec).

References

- [1] B. Miller, "Satellites Free the Mobile Phone," *IEEE Spectrum*, pp. 26-35, March 1998.
- [2] P. Chitre and F. Yegenoglu, "Next-Generation Satellite Networks: Architecture and Implementations," *IEEE Communications Magazine*, pp. 30-36, March 1999.
- [3] N. Ghani, and S. Dixit, "TCP/IP Enhancements for Satellite Networks," *IEEE Communications Magazine*, pp. 64-72, July 1999.
- [4] D. P. Connors, B. Ryu, and S. Dao, "Modeling and Simulation of Broadband Satellite Networks Part I: Medium Access Control for QoS Provisioning," *IEEE Communications Magazine*, pp. 72-79, March 1999.
- [5] S. Yoshida, H. Kimura, Y. Inoue, T. Masamura, and N. Yamauchi, "Interactive Multimedia Communication Systems for Next-Generation Education Using Asymmetrical Satellite and Terrestrial Networks," *IEEE Communications Magazine*, pp. 102-106, March 1999.
- [6] C. E. Fossa, R. A. Raines, G. H. Gansch, M. A. Temple, "A Performance Analysis of the IRIDIUM Low Earth Orbit Satellite System with a Degraded Satellite Constellation," *Mobile Computing and Communications Review*, Vol. 2, No. 4, pp. 54-61.
- [7] I. F. Akyildiz, S. H. Jeong, "Satellite ATM Networks: A Survey," *IEEE Communications Magazine*, pp. 30-43, July 1997.
- [8] E. Ekici, I. F. Akyildiz, and M. D. Bender, "Datagram Routing Algorithm for LEO Satellite Networks," *Proceedings of INFOCOM 2000*, Tel Aviv, Israel, 2000.
- [9] M. 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*, vol. 13, no. 2, pp. 371-381, Feb. 1995.
- [10] J. P. Hatton, "Orbits Used by High Altitude Satellites", <http://www.geocities.com/CapeCanaveral/Hangar/1668/orbits.htm>
- [11] G. M. Djuknic, J. Freidenfelds, and Y. Okunev, "Establishing Wireless Communications Services via High-Altitude Aeronautical Platforms: A Concept Whose Time Has Come?" *IEEE Communications Magazine*, pp. 128-135, Sept. 1997.
- [12] L. Wood, A. Clerget, I. Andrikopoulos, G. Pavlou, and W. Dabbous, "IP Routing Issues in Satellite Constellation Networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, Dec. 2000.
- [13] E. G. Cuevas, "The Development of Performance and Availability Standards for Satellite ATM Networks," *IEEE Communications Magazine*, pp. 74-79, July 1999.
- [14] R. Goyal, S. Kota, R. Jain, S. Fahmy, B. Vandalore, and J. Kallaus, "Analysis and Simulation of Delay and Buffer Requirements of Satellite-ATM Networks for TCP/IP Traffic", <ftp://ftp.netlab.ohio-state.edu/pub/jain/papers/jsac98/jsacfinal.htm>.
- [15] C. Partridge, and T. J. Shepard, "TCP/IP Performance over Satellite Links," *IEEE Networks*, pp. 44-49, Sept./Oct. 1997
- [16] H. Peyravi, "Medium Access Control Protocols Performance in Satellite Communications," *IEEE Communications Magazine*, pp. 62-71, March 1999.
- [17] A. Jamalipour, M. Katayama, and A. Ogawa, "Traffic Characteristics of LEOS-Based Global Personal Communications Networks," *IEEE Communications Magazine*, pp. 118-122, Feb. 1997.
- [18] R. Mauger, and C. Rosenberg, "QoS Guarantees for Multimedia Services on a TDMA-Based Satellite Networks," *IEEE Communications Magazine*, pp. 56-65, July 1997.
- [19] G. Maral, J. Restrepo, E. Del Re, R. Fantacci, and G. Giambene, "Performance Analysis for a Guaranteed Handover Service in an LEO Constellation with a 'Satellite-Fixed Cell' System," *IEEE Trans. on Vehicular Technology*, vol. 47, no. 4, Nov. 1998.
- [20] H. S. Chang, B. W. Kim, C. G. Lee, S. L. Min, Y. Choi, H. S. Kim, C. S. Kim, "Performance Comparison of Static Routing and Dynamic Routing in Low-Earth Orbit Satellite Networks," *Proceedings of IEEE VTC'96*, pp. 1240-1243, 1996.
- [21] H. Uzunalioglu, "Probabilistic Routing Protocol for Low Earth Orbit Satellite Networks," *Proceedings of IEEE ICC'98*, pp. 89-93, 1998.
- [22] H. Uzunalioglu, I. F. Akyildiz, Y. Yesha, and W. Yen, "Footprint Handover Rerouting Protocol for Low Earth Orbit Satellite Networks," *Wireless Networks*, pp. 327-337.
- [23] J. Lee, S. Kang, "Satellite over Satellite (SOS) Network: A Novel Architecture for Satellite Network," *Proceedings of IEEE INFOCOM 2000*, Tel Aviv, Israel, 2000.
- [24] S. R. Pratt, R. A. Raines, C. E. Fossa Jr, and M. A. Temple, "An Operational and Performance Overview of the IRIDIUM Low Earth Orbit Satellite System," *IEEE Communications Surveys*, <http://www.Comsoc.org/pubs/surveys>, Second Quarter, 1999.
- [25] R. K. Ahuja, T. L. Magnanti, J. B. Orlin, *Network Flows*, Ch. 4, pp. 108-113, Prentice-Hall Inc., 1993.
- [26] R. Kucukates, "A Minimum Flow Maximum Residual Method for Routing in LEO Satellite Networks" PhD Thesis Dissertation, Bogazici University, 2002.