Load-aware and Position-aided Routing in Satellite IP Networks

L. Wang, L. Liu, X. Hu

Wang Lu
1. Graduate University of Chinese Academy of Sciences
Beijing, China, 100049, and
2. Institute of Software, Chinese Academy of Sciences
Beijing, China, 100190
E-mail: wlu810@gmail.com

Liu Lixiang, Hu Xiaohui
Institute of Software, Chinese Academy of Sciences
Beijing, China, 100190

Abstract: Satellite communication is regarded as a promising way to achieve seamless global coverage with an array of advantages over the traditional terrestrial network. However, the non-uniform distribution of users and the highly topological dynamics have been a matter of facts. The traditional load balancing technique is able to guarantee a better traffic distribution around the Inter Satellite Links with heavy loads after considering the non-uniform distribution of users, but it transmits load information passively and fails to provide a fair traffic distribution over the entire constellation. Given the highly topological dynamics, traditional routing protocols may experience the convergence processes frequently or use a large amount of storage to keep the snapshots. This paper proposes a novel routing protocol named LPR for the Satellite IP Networks, which transmits the bottleneck link load information to upstream satellites in a Hop-by-Hop manner actively, to guarantee a better distribution of traffic among satellites. In response, upstream satellites select less congested paths and communicate a portion of data according to the priority of different classes of traffic. In addition, LPR applies the geographical position when selecting the routing path to eliminate the convergence process which may cause the routing tables invalid. Finally, various aspects of performance of LPR have compared with other mechanisms. Our evaluations show that a routing protocol with load-aware and position-aided leads to higher throughput and better traffic distribution compared to the traditional routing protocol and load balancing technique.
Keywords: satellite networks, load balancing, position, routing.

1 Introduction

Satellite networks exhibit unique features and offer an array of advantages over traditional terrestrial networks. They are able to provide coverage to extensive geographic areas and to interconnect among remote terrestrial networks. A satellite communicates with other satellites via inter-satellite links(ISLs). The links between satellites in the same plane are called intra-plane ISLs. The links between satellites in different planes are called inter-plane ISLs. Satellites communicate with the ground stations over user data links(UDL).

LEO satellites provide short round-trip delays and have been the focus of several researches for a period of time [1], [2]. Given recent advancements in satellite communications, researchers find that a combination of different layers of satellites would yield a better performance than

these layers individually. Researchers have proposed severals routing protocols for multi-layered satellite networks, such as MLSR [3], SGRP [4], SOS [5] et. al. MLSR is a link state routing protocol for multi-layered satellite networks. SGRP is a multi-layered satellite network routing protocol based on snapshots. In these multi-layered satellite network routing protocols, LEO satellites transmit data, and MEO/GEO satellites are managers and compute the routing paths. Another type of satellite network routing protocol is based on geographical location, such as LPDB [6]. LPDB utilizes satellites' inherent broadcast capabilities, and predict transmission direction according to the geographical location.

The majority of the routing protocols have convergence processes to collect information and update routing tables when the topology changes, however satellites move rapidly relative to the surface of the Earth and to the ground stations(LEO satellites speed at over 2500km/hour [7]) and a lot of packets may be lost during the convergence process, due to routing tables becoming invalid. Furthermore, satellites' rapid movements cause frequent convergence processes, which increases the algorithms' overhead. When satellites use routing protocols based on snapshots, it takes a large amount of storage. If satellites' links interrupted, routing tables of the current snapshot and the following snapshots would be invalid, and in turn it takes a long interval for the central control nodes, such as ground stations, GEO/MEO, to calculate new routing tables and update routing tables for the satellites.

In addition, an important missing point in the previous designs consists in their focus on searching for the paths without any consideration of the total traffic distribution over the entire constellation, which leads to unfair distribution of traffic. Recently, researchers propose some load balance mechanisms, such as CEMR [8], ELB [9] et. al. CEMR is based on a cost metric that involves both propagation and queuing delays, however it does not reflect the congestion state of the next hop, nor does it estimate the queuing delay a packet may experience there. ELB is an explicit load balancing mechanism that classifies users into three classes. Neighboring satellites exchange load information and reduce data forwarding rates when the queue occupancies exceed a pre-determined threshold. However, ELB only exchanges information among neighboring satellites actively and it takes a long interval to notify sources of the congestion state. Furthermore, although ELB applies a routing metric that instantly reflects both the one-way propagation delay and the instant queuing delay to avoid loops, the traffic forwarding may still cause loops.

The objective of this paper is to design a novel satellite network routing protocol named LPR (Load-aware and Position-aided Routing), which is able to guarantee fair distribution of traffic among the entire constellation and eliminate the convergence process. LPR calculates every link's load information and transmits it to the upstream satellites actively. LPR calculates the next hop set based on geographical position, and the next hop set includes a main path and several alternative paths. LPR transmits traffic on the main path when the load on the main path is low, and begins to communicate a portion of traffic on the alternative paths when the load on the main path becomes high. The main path is the routing path with the shortest distance.

2 Dissemination and Estimation of Load Information

2.1 Load Transmission

In traditional load balancing technique, the load information is exchanged among neighboring satellites. It is able to achieve better traffic distribution than the routing protocols which do not take traffic distribution into account, but the ISLs are expected to be heavily loaded around the overload region and others remain underutilized, and in turn it fails to achieve fair traffic

distribution over the entire constellation, as in Figure 1.

In LPR, each satellite calculates the load information and compares its own link load with the received link load , and updates the link load, if the local link load is larger than the received link load. Then satellites transmit the load information in a Hop-by-Hop manner. In other words, each satellite transmits the bottleneck link load information to upstream satellites actively. The upstream satellites receive the bottleneck link load, and begin to detour traffic when the bottleneck link load is larger than the threshold, resulting that all upstream satellites on the transmission path are able to communicate a portion of data when there is a link congested. In turn, the traffic is distributed fairly over the entire constellation and satellites are able to return back to free state quickly, as in Figure 2.

In addition to the fair traffic distribution, LPR is able to disseminate the load information instantly to all upstream satellites. In traditional load balancing technique, it takes a relatively long interval to transmit the overload information to upstream satellites, for the traditional load balancing technique only transmits load information actively to neighboring satellites, and load information is transmitted passively to the upstream satellites. Let I and O denote the total input and output traffic at a given satellite. A satellite is considered to be overloaded if the load exceeds the *threshold*. Let t_n denote the time for satellite n becomes overloaded when I is larger than O. Let d_n denote the transmission delay between satellite n and satellite $n + 1$. In traditional load balancing technique, it takes $\sum_{n=i}^{j-1} (t_n + d_n)$ for satellite i to know the overload sate originated from satellite j. But in LPR, it only takes $\sum_{n=i}^{j-1} d_n$ for satellite i to know the load information of satellite j, due to the active transmission of load information in a Hop-by-Hop manner.
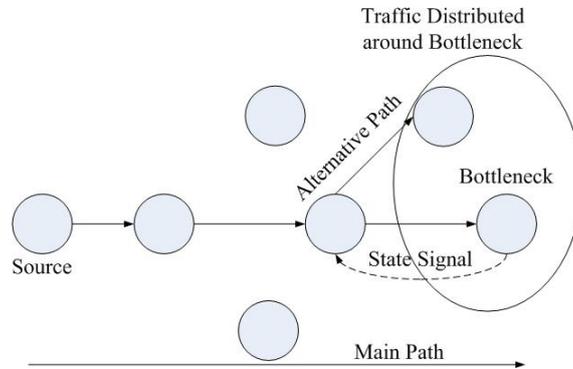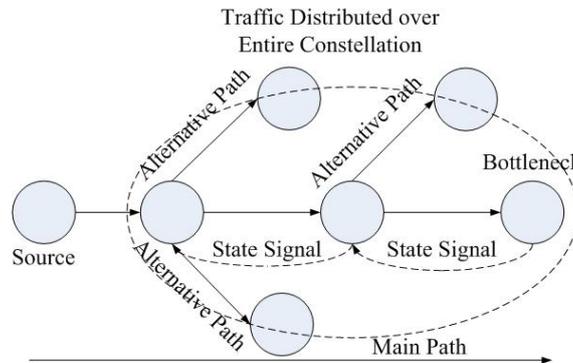

Figure 1: Traditional Load Balancing

Figure 2: LPR's Load Balancing

2.2 Load Estimation and Setting of Thresholds

In LPR, satellites keep a virtual queue for every link and estimate link load R_l over an appropriate interval time t_R, and t_R is estimated according to RTT in one hop. R_l is computed similar to [10]:

$$R_l = \frac{\lambda_l + k_q \cdot \overline{q}_l}{\gamma_l \cdot C_l \cdot t_R} \tag{1}$$

Here the definitions of λ_l and $\overline{q}_l$ are different from [10]. λ_l is the total amount of input data in a virtual queue for one link during t_R, and $\overline{q}_l$ is the persistent queue length of a virtual queue for one link during t_R. k_q controls how fast the persistent queue drains, γ_l is the target utilization that sets close to 1, and C_l is the link capacity. The input traffic λ_l and the persistent queue length $\overline{q}_l$ are measured using bytes [10]. LPR sets k_q to 0.8 to make full use of the limited memory space. The persistent queue length $\overline{q}_l$ is measured by using a low-pass filter that samples the instantaneous queue size, $q(t)$, every $t_R/20$.

2.3 Setting of Detouring Ratio

The objective of setting the detouring ratio is to allow the satellites' overloaded links to return back to light load state and keep this state for at least a period of θ.

Let I_t denote the total rate of traffic coming from terminals to the bottleneck link within the coverage area of a satellite, I_n denote the total rate of traffic coming from neighboring satellites to the bottleneck link, O denote the total rate of output traffic of the bottleneck link, and l_d denote the ISL delay. A satellite receives link load $R_l(t)$ from neighboring satellites at time $(t + l_d)$. $R_l(t)$ is the neighboring satellite's link load at time t, and the satellite's link load at time $(t + l_d)$ is $R_l(t + l_d)$:

$$R_l(t + l_d) = \frac{\gamma_l + k_q \cdot (\overline{q}_l + (I_t + I_n - O) \cdot I_d)}{\gamma_l \cdot C_l \cdot t_R} \tag{2}$$

In order to ensure a prompt recovery and a residual time in the light load for at least θ, the new rate of traffic coming from neighboring satellites I_n^{new} should satisfy:

$$R_l(t + I_d + \theta) = \frac{\gamma_l + k_q \cdot (\overline{q}_l + (I_t + I_n - O) \cdot I_d + (I_t + I_n^{new} - O) \cdot \theta)}{\gamma_l \cdot C_l \cdot t_R} < threshold \tag{3}$$

$threshold$ is the link load's threshold. If the link load is larger than $threshold$, the link is overloaded and satellites begin to detour a portion of traffic according to the detouring ratio on the alternative paths. The traffic detouring ratio φ can be computed as:

$$\varphi = 1 - min(max(0, \frac{I_n^{new}}{I_n}), 1) \tag{4}$$

3 Load-aware and Position-aided Routing

3.1 Path Selectioin

The key technologies required to support positioning over satellite systems have been developed, such as GPS, Galileo and GLONASS. With these advancements, satellites are able to get the high precision orbit determination. In LPR, path selection is performed by using the position. LPR calculates the routing path for every arriving packet, instead of maintaining routing tables to avoid the convergence process, however, it takes a large amount of computing resources. So

satellites cache the routing results until the topology state changes, to reduce the computation overhead.

Step 1: Predicting transmission direction-The polar constellation networks have orbital seams, and LPR assumes that there are no cross-seam links, for the regular hand over as satellites pass. Satellites would transmit receiving packets cross seam, if the destination and current satellites located on different side of the orbital seam. The orbital seams are located between the fixed planes of ascending and descending satellites in the same layer and in turn LPR is able to utilize the longitude difference between the orbital plane of current satellites and the orbital seam to judge whether packets should cross the seams or not. Assuming that satellite A's geographical position is $(long_a, latitude_a)$. The points of the intersections of the orbital seams and the equator are $Left(long_{left\_seam}, 0)$ and $Right$ $(long_{right\_seam}, 0)$. The point of the intersection of the plane of satellite A and the equator is $A'(plong, 0)$:

$$plong = long_a + \arcsin(\tan(latitude_a)/\tan(inclination)) \tag{5}$$

where $inclination$ is the constellation's inclination angle, and the constellation is composed of satellites in one layer. Let $left\_seam\_diff$ denote the longitude difference between A' and $Left$ and $right\_seam\_diff$ denote the longitude difference between A' and $Right$.

$$left\_seam\_diff = plong - long_{left\_seam} \tag{6}$$

$$right\_seam\_diff = long_{right\_seam} - plong \tag{7}$$

$left\_seam\_diff$ and $right\_seam\_diff$ are fixed in the same layer. Assuming that the point of intersection of the destination and the equator is $(long_{dest}, 0)$. If the difference between $plong$ and $long_{dest}$ is less than $left\_seam\_diff$ or $right\_seam\_diff$, packets will not be transmitted across the orbital seams and A's neighboring satellites set $Nei(A)$ includes all satellites having ISLs with A, otherwise packets will be transmitted across the orbital seams and $Nei(A)$ only includes satellites which are able to transmit packets across the seams by means of passing through the poles or another satellites layer. After estimating neighboring satellites set $Nei(A)$, satellite A calculates the distance from x , where $x \in Nei(A)$, to the destination to get next hop set $N(x)$:

$$N(x) = \{x|D(x, dest) < D(A, dest), x \in Nei(A)\} \tag{8}$$

In multi-layered satellite networks, all satellites are projected to the surface of the Earth when calculating the distance. If several satellites in different layers are projected in one point on the surface of the Earth, the satellite in the same layer with the sender has the highest priority.

Step 2: Selecting next hop-After predicting transmission direction, satellite A will choose next hop from $N(x)$ depending on the load information collecting from the transmission paths. LPR transmits traffic on the main path when the load is low, and begin to detour traffic when the load of the main path becomes high. LPR classifies traffic into three types, similar to [11], to guarantee the QoS requirements. Class A is delay-sensitive traffic, Class B is through-put sensitive traffic, and Class C is best effort traffic. When the load increases, LPR starts detouring traffic of Class C. If the requested detouring ratio of traffic is larger than the traffic percentage of Class C, the traffic of Class B is detoured as well. If the requested detouring ratio of traffic φ is larger than the traffic percentage of Class C and Class B, the traffic of A begins to detour. Let a, b and c denote the predicted traffic ratio of class A, B and C, the traffic detouring ratio is shown in Table 1.

3.2 Loop Avoidance

In LPR, forwarding traffic depending on the geographical position and the load information may cause loops. To cope with this issue, LPR adds the routing path into the packet header

Table 1: Traffic Detouring Ratio

	Class A	Class B	Class C
$\varphi < c$	0	0	$\frac{\varphi}{c}$
$c \leq \varphi < (b+c)$	0	$\frac{\varphi-c}{b}$	1
$\varphi \geq (b+c)$	$\frac{\varphi-b-c}{a}$	1	1

after selecting the next hop, to form the traversed link set LS. Consider satellite A sending the set of traversed link LS to satellite B, and satellite B will not select links in the LS. However it would take a lot of bandwidth to transmit the traversed link set information. LPR proposes a mechanism to reduce this overhead at the expense of local mapping. Considering that satellite A associates a traversed link set LS to $label_{LS}$, and sends the traversed link set LS and the mapping label $label_{LS}$ to satellite B. Satellite B will establish a mapping $label_{LS} \rightarrow LS$ and send an acknowledgement to Satellite A. After satellite A receives the acknowledgement of the mapping from satellite B, satellite A only includes $label_{LS}$ rather than the entire traversed link set LS in the packet header. Labels exchange among neighboring satellites do not have global meaning. Furthermore, a time-to-live value T is associated with a mapping at the receiver and the mapping is deleted, if no packet with the label associated with this mapping is received for T. Then the receiver notify the sender of deleting the timeout mapping $label_{LS} \rightarrow LS$. Considering that satellite B finds that $label_{LS} \rightarrow LS$ is timeout, satellite B will delete the mapping and notify satellite A that the mapping is timeout. Satellite A will delete the mapping after receiving the notification.

4   Simulation Setup

In this section, we describe the simulation topology and parameters that have been used to compare the performance of LPR, Dijkstra, ELB on Dijkstra and Snapshot [12]. We have used OMNeT++ simulator [13] and developed a constellation of Iridium satellites and the main parameters are shown in Table 2.

Table 2: Main Parameters of Satellite Topology

Parameters	Values
Satellite Number	66
Orbit Altitude	780km
Inclination Angle	84.4degrees
ISLs	2 Intra-SLs, 2Inter-SLs
Polar Area Border	70degrees
Plane Number	6
Seam Links	no
Constellation Type	Polar

It is assumed that each satellite maintains four ISLs with its neighboring satellites. Uplinks, downlinks and ISLs are each given a capacity equal to 25Mbps. The average packet size is set to 1KB. Lengths of drop-Tail based buffers equal to 200 KB. Simulations are all run for 2 hours, for

LEO satellite constellation operates at the altitude of 780 km with the system period of about 100 minutes.

In the simulation, we consider 100 data flows. The source and destination end-terminals are dispersed all over the world and the distribution of end-terminals is shown in Table 3. The traffic distribution is identical to [14]. The sources send data at constant rates from 0.8 Mbps to 1.5 Mbps. The traffic percentages of delay-sensitive traffic, throughput-sensitive traffic and best effort traffic are set to 20%,30% and 50%.

Table 3: Distribution of End Users

	North America	South America	Europe	Africa	Asia	Oceania
End Users Number	35	10	25	5	20	5

5   Simulation Results

In this section, we will show the performance results of LPR in the presence of different traffic classes in terms of average transmission delay and average throughput, and will compare the load balancing index among LPR, ELB on Dijkstra, Snapshot and Dijkstra.

5.1   Multiple Traffic Classes

The performance of LPR with multiple traffic classes is compared with Dijkstra algorithm. Figure 3 shows the average packet transmission delay. The delay of packets belonging to the delay-sensitive traffic is smaller compared to other types of traffic This is because the delay-sensitive data is always sent via the paths with the shortest distance. When the traffic load gets higher (the individual transmission rate is larger than 1.3Mbps), Dijkstra exhibits the minimum transmission delay, but at the expense of significant packet drops, as shown in Figure 4. LPR's transmission delay increases higher, due to the traffic detouring which increases the communication delay.
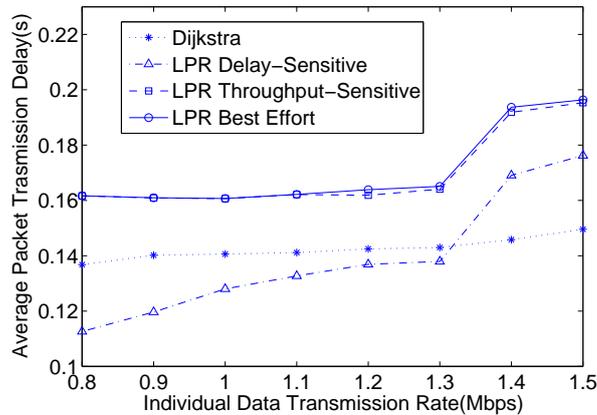

Figure 3: Packets Transmission Delay

As shown in Figure 5, the average throughput of Dijkstra is always lower than other schemes, due to the overflow of the drop-tail buffers, as shown in Figure 4. Throughput of the delay-sensitive traffic outperforms the other two types of traffic, because of the detouring priority.
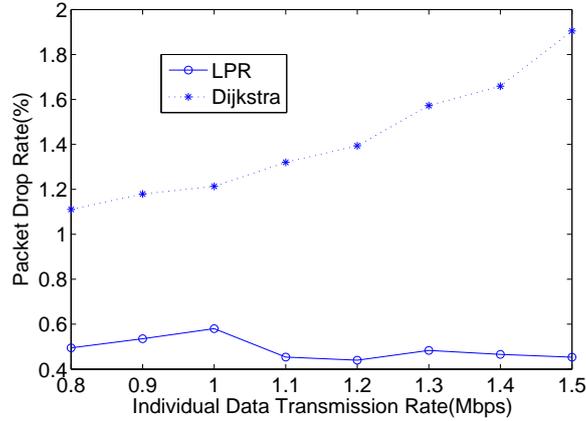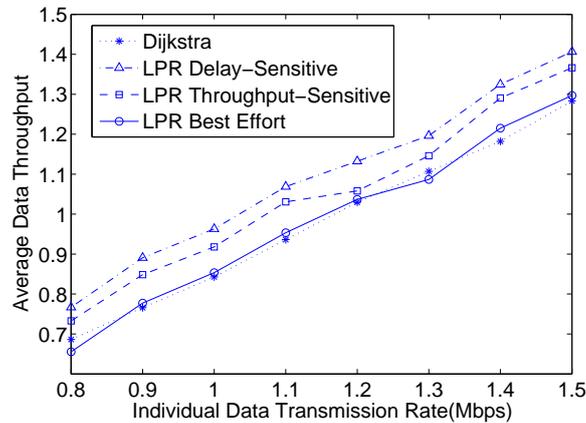
Figure 4: Packets Drop Rate

Figure 5: Throughput

These results demonstrate that LPR is able to guarantee QoS requirements by the effective load balance mechanism and a routing method without the convergence process.

5.2   Load Balancing Index

To evaluate the traffic distribution of LPR, this paper compares LPR with Dijkstra, Snapshot and ELB with the following traffic distribution index:

$$f = \frac{(\sum_{i=1}^{n} x_i)^2}{n \cdot \sum_{i=1}^{n} x_i^2} \tag{9}$$

Where n is the total number of ISLs and UDLs, x_i denotes the actual number of packets that traversed the ith link. The index ranges from zero to one and high value of this traffic distribution index represents good distribution of traffic over the entire constellation. In the simulation, LPR's threshold ranges from 0.85 to 0.95 and the interval is 0.5, and LPR's distribution index is the average of the results of these different thresholds, in order to evaluate the performance of LPR under different thresholds.

In Dijkstra scheme, the traffic distribution index is low (Figure 6), for packets always traverse through paths with the smallest number of hops without taking the traffic distribution into account. LPR scheme always try to distribute the traffic on the transmission paths when the link

is overloaded, so the traffic distribution is higher than the plots of Dijkstra and Snapshot. ELB on Dijkstra is higher than the plots of Dijkstra, Snapshot, for ELB searches for less congested paths, after receiving a Busy Signal Advertisement from neighboring satellites. However, in LPR scheme, not only neighboring satellites are able to communicate a portion of data via less congested paths, but also the upstream satellites. Furthermore the load estimation of LPR is more accurate than ELB, in that LPR's granularity of load estimation is the link, but ELB is the node. In turn, the distribution index of LPR is much higher than ELB on Dijkstra.
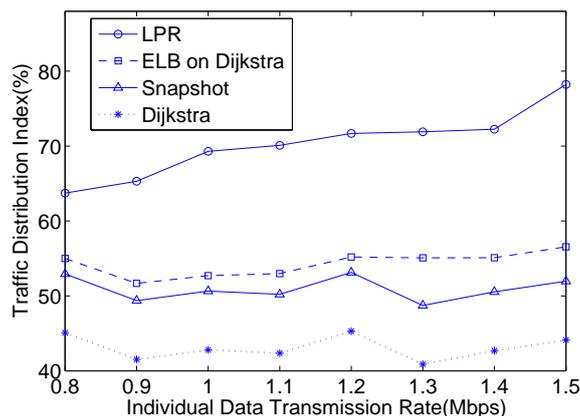

Figure 6: Traditional Load Balancing

6   Conclusions

In this paper, we propose a Load-aware and Position-aided Routing protocol named LPR, which transmits load information in a Hop-by-Hop manner and enables all the upstream satellites to know the load information instantly and to communicate a portion of traffic on the less congested paths when the load becomes high. Furthermore, LPR utilizes the geographical positions of satellites and select the routing path to eliminate the convergence process. Finally, we evaluate the performance of LPR in the Iridium constellation. The simulation results indicates that LPR is capable of guaranteeing a good distribution of traffic among satellites and supporting different QoS requirements. As future work, we intend to improve the ability of providing different QoS by differentiating traffic, and analyze LPR's impacts on packets reordering. Finally, we will implement LPR in Linux kernel to assess its strengths and limitations in practice.

Bibliography

[1] E. Ekici, I.F. Akyildiz, M.D. Bender, A Distributed Routing Algorithm for Datagram Traffic in LEO Satellite Network, *IEEE/ACM Transactions on Networking*, 9(2):137-147, 2001.

[2] T. Henderson, R.H. Katz, On Distributed, Geograhpic-based Packet Routing for LEO Satellite Networks, *Proceedings of IEEE GLOBECOM*, 2000.

[3] I.F. Akyildiz, E. Ekici, M.D. Bender, MLSR:A Novel Routing Algorithm for Multilayered Satellite IP Networks, *IEEE/ACM Transactions on Networking*, 10(3):411-424,2002.

[4] C. Chen, E. Ekici, A routing protocol for hierarchical LEO/MEO satellite IP networks, *Wireless Networks*, 11(4):507-521, 2005.

[5] J. Lee, S. Kang, Satellite over Satellite (SOS) Networks: A Novel Architecture for Satellite Network, *Proceeding of IEEE INFOCOM*, 2000.

[6] C. Chen, C. ZeSheng, Towards a routing framework in ad hoc space network, *Internation Journal of Ad Hoc and Ubiquitour Computing*, 5(1):44-55, 2010.

[7] L. Wood, Internetworking with Satellite Constellations, School of Electronics, Computing and Mathematics, University of Surrey, Guildford, 2001.

[8] B. Jianjun, L. Xicheng, L. Zexin, Compact Explicit Multipath Routing for LEO Satellite Networks, *Proceedings of IEEE Workshop on High Performance Switching and Routing*, HongKong, China, 2005.

[9] T. Taleb, D. Mashimo, A. Jamalipour, et. al., Explicit Load Balancing Technique for NGEO Satellite Networks with On-Board Processing Capabilities, *IEEE/ACM Transactions on Networking*, 17(1):281-291,2009.

[10] Y. Xia, L. Subramanian, I. Stoica, One more bit is enough, *IEEE/ACM Transactions on Networking*, 16(6):1281-1294, 2008.

[11] A. Svigelj, M. Mohorcic, G. Kandus, et. al., Routing in ISL networks considering empirical IP traffic, *IEEE Journal on Selective Areas Communication*, 22(2):261-272,2004.

[12] V. Gounder, R. Prakash, H. Abu-Amara, Routing in LEO-based satellite networks, *Emerging Technologies Symposium on Wireless Communications and Systems*,221-226,1999.

[13] A. Varga, OMNeT++ Discrete Event Simulation System, http://www.omnetpp.org.

[14] M. Mohorcic, M. Werner, A. Svigelj, et. al., Adaptive Routing for Packet-oriented intersatellite link networks: Performance in various traffic scenarios, *IEEE Transaction on Wireless Communications*, 1(4):808-818,2002.