A Load Balancing Algorithm Based on Satellite Delay Tolerance Network

Li Yang [1,2, a] Jinyuan Liu [1,2, b] Chengsheng Pan [1, c] Debin Wei [1,2, d]

[1] Information and Engineering College, Dalian University, Dalian 116622, China

[2] Communication and Networks Key Laboratory, Dalian 116622, China

[a] Yangli945@126.com, [b] atlantis918@hotmail.com, [c] pancs@sohu.com,

[d] weidebin@163.com

Abstract

In order to solve the problem of high dynamic topology and high transmission delay in DTN satellite networks, the unbalanced data flow between two inter-satellite links is unbalanced. In this paper, a BLO load balancing algorithm based on LEO satellites is proposed. By establishing the topology model of DTN satellite network, modifying the format of bundle package, considering the residual capacity of satellite nodes for congestion avoidance, applying the multi-objective mixed integer linear programming to increase the minimum connection time of satellite nodes and reduce the maximum connection time of satellite nodes, making DTN satellite Network flow is more balanced. Simulation results show that BLO-DTN algorithm can reduce the data packet loss rate and average delay, further improve the resource utilization rate of DTN satellite network and effectively improve the performance of DTN satellite network.

Keywords

Delay/Tolerance network, Load balancing, Congestion avoidance, Linear programming.

1. Introduction

DTN is a protocol for adding a bundle layer above the transport layer above the transport layer for the characteristics of intermittent wireless network communication links, dynamic changes of the network topology and large link communication delay[1] [2] , In order to achieve storage - carrying - forwarding mechanism. Due to the large delay, frequent interruption and asymmetric bandwidth of the DTN satellite network, routing algorithms and congestion policies in traditional networks are no longer suitable for DTN satellite networks. In the DTN satellite network, when congestion occurs, there is generally no available communication opportunity to relieve congestion. Therefore, it is more necessary to perform congestion avoidance before congestion occurs.

Rango et al. [3] proposed a hop-by-hop method of local flow control in interstellar network, which can send some information of local buffer space back to hop by explicit feedback. However, because the satellite network can not guarantee that there is an end-to-end path between any pair of nodes at any time, the delay will usually be large even if there is an end-to-end path. The time required for the message to be transmitted from the source node to the target node is very long, Even before it reaches the target node, congestion has taken place. Seligman et al. [4] proposed a storage-based routing scheme for DTN networks. Before node congestion occurs, nodes store messages to avoid congestion. However, in the DTN satellite network, because of sparse nodes, large delay and limited communication opportunities, nodes can not store information at any time. This makes the above storage scheme reduce the probability of successful message transmission. Therefore, the best way to avoid congestion in a DTN satellite network is to predict the remaining capacity in advance.

Literature [5] proposed a dynamic equalization algorithm based on load balancing for three-layer satellite networks with high, medium and low trunnions. By setting different congestion thresholds for different satellites, the transmission reliability of multi-layer satellites is improved. Literature [6] modified the classical shortest path algorithm and proposed a dynamic routing algorithm based on satellite topological transformation to reduce the routing handover probability and computation. Literature [7] proposed routing weight adaptive algorithm, after flow adjustment, making the average number of hops in satellite transmission, the average transmission delay is optimized. However, none of the above documents consider the flow control in satellite communications, which makes the load of the entire satellite network unbalanced. In addition, the current Contact Graph Route algorithm [8] and Prophet algorithm [9] do not consider the remaining capacity of network nodes.

In this paper, the topology model of 3D network is established based on the characteristics of DTN satellite network. The method of congestion avoidance is modified by modifying the format of bundle package. Finally, BLO-DTN load balancing algorithm based on multi-objective mixed integer linear programming is used to increase the minimum connection time of satellite nodes, The maximum connection time of satellite nodes and the satellite routing algorithm are optimized to make the flow between the stars more balanced so as to improve the performance of the DTN satellite network.

2. Algorithm design

2.1 Contact topology model

In the later satellite network, let the set of nodes be V ,which represents the set of all satellites and ground stations in the network. The edge set is $\mathrm{E} \subseteq \mathrm{V} \times \mathrm{V}$,which represents the collection of inter-satellite links and star links in the network. Define $\theta : E \times R^{+} \rightarrow \{0,1\}$ as the existence function,which indicates whether edge e exists at time t.With $A(e)$ to represent the existence of side e,namely

$$A(e) = \{\mathrm{t} \in \mathrm{R}^{+} : \theta(e,t) = 1\} \tag{1}$$

Since the existence of side e is a time period, equation (1) can be written as:

$$A(e) = \{[t_1, t_2) \cup [t_3, t_4) \cup \cdots\} \tag{2}$$

For satellite networks, each satellite node periodically moves in accordance with a set orbit, so the link's existence time is also usually periodic The period of existence function e is defined as T ,So for any $t \in R^{+}$,any $k \in N$,有 $\theta(e,t) = \theta(e, t + kT)$. Then equation(2)can be further expressed as:

$$A(e) = \{[t_1 + kT, t_2 + kT) \cup [t_3 + kT, t_4 + kT) \cup \cdots\} (k \in N) \tag{3}$$

In this paper, based on the discrete virtual topology, the system is divided into time slices, and then static routes are calculated. Divide the time slice according to formula (3), guaranteeing in every time slice, $\theta \equiv 1$or $\theta \equiv 0$.

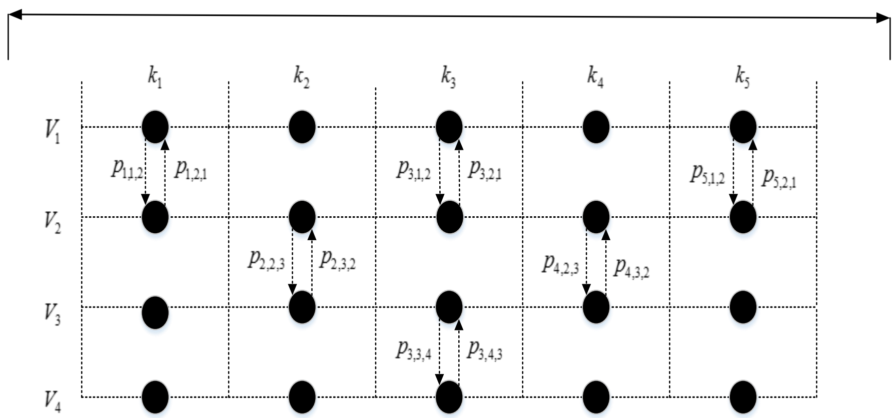

Figure 1 Contact topology

Figure 1 shows a network of four satellite nodes,node collection $V=\{V_1,V_2,V_3,V_4\}$,According to the link between the four satellite nodes, the topology is divided into time slices $\{k_1,k_2,k_3,k_4,\cdots\}$. $p_{k,i,j}$ represents the connection between node V_i and node V_j in the k time slot,when the connection does not exist, $p_{k,i,j}=0$;when the connection exists, $p_{k,i,j}=1$.Then the entire network connection topology can be represented by a three-dimensional adjacency matrix $[P]_{k,i,j}$.

2.2 Congestion Avoidance mechanism

In the delay tolerance network, the storage resources and communication resources of the nodes are generally very limited. Moreover, the late satellite networks generally can not guarantee the existence of an end-to-end connection. Therefore, the messages need to be stored in the relay nodes for a long time, The node's persistent storage resources are depleted, leading to congestion and network performance degradation. The best solution to the congestion problem is to take action before congestion occurs. In the process of calculating the delay of the network routing, usually only consider the local node size of the message stored and the capacity of the connection, but did not take into account the rest of the network node cache size. This will result in the receiver rejecting to save the message, that is, congestion, when sending messages according to the calculated path due to insufficient buffer capacity. Therefore, this paper carries out congestion avoidance by modifying the format of the bundle package defined by Bundle protocol. In the late satellite network, the connection information can be obtained from Eq. (3), and the capacity of the connection can be calculated accordingly. However, it is difficult for nodes in the network to obtain cache information. Therefore, the idea of this paper is to exchange information between nodes through the communication between nodes. Specific steps are as follows:

Modify the bundle package format. Table 1 shows the Bundle protocol defined standard data unit format, contains some of the Bundle process control information and routing information required. This article adds a 4-byte field in the packet, to record the current node routing information for the message, the remaining capacity of the connection RC (Residual Contact).

Table 1 Bundle protocol standard data unit format

1Byte	3Byte
version number	Bundle Processing control sign
Block length	
Destination node scheme Offset	Destination node SSP Offset
Source node scheme Offset	Source node SSP Offset
Report endpoint scheme Offset	Managed endpoints SSP Offset
Create time	
Serial number	
Survival	
Dictionary length	
Address Dictionary Array	
Load block	

After receiving the message, the next-hop node parses Bundle basic block information, reads the value written by the last-hop node, and compares the value with the remaining buffer capacity (Residual Buffer) of the current node:If $RC \le RB$,when the last hop node has less remaining capacity than the remaining cache capacity of the node in the route calculation, the connection is still considered in the topology used by the route calculation.;If $RC > RB$,that is, the last hop node is in the route If the remaining capacity of the connection is less than the remaining buffer capacity of the local node, some messages may be rejected due to insufficient buffer storage during subsequent sending. Therefore, before the local node caches the buffer, notify Hop node, recalculate the path to the following message, and remove the connection in the topology used in the calculation.

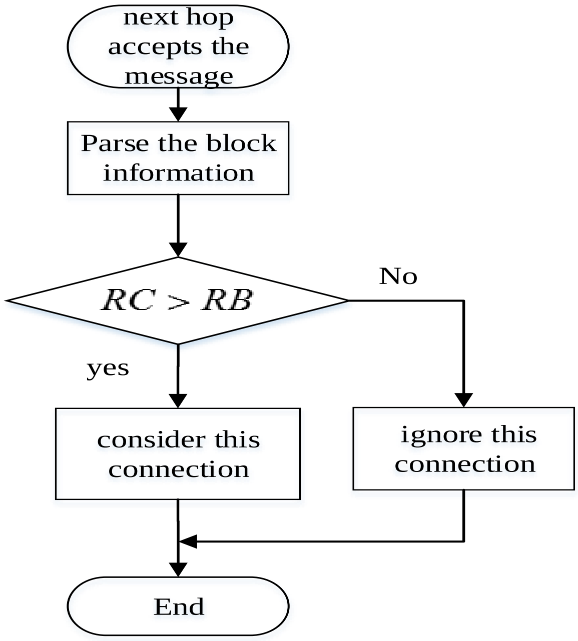

Figure 2 Process flow chart

3. During the Bundle hosting phase, the next-hop node sends the Bundle hosting request message to the hop-up node and sends the comparison result with the previous hop node for corresponding processing.

2.3 Load Balancing Problem Definition

In the Delay tolerance satellite network, the number of satellite nodes is smaller, the distribution is not uniform, and the connection between some nodes is high, which leads to larger flow generated by these nodes and more node loss, further reducing the capacity Satellite network communication opportunities decrease. Therefore, it is very important for the message to be forwarded by all nodes as far as possible, which is of lasting significance to the lasting and reliable operation of the entire network. Find a solution to make the entire network flow more balanced. This article uses matrix $[L]_{k,i,j}$ to represent this suitable solution,The elements of the matrix $l_{k,i,j}$ indicates the path actually selected during the message forwarding,in the time slice k send from node i to node j ,therefore, $[L]$ is subset of $[P]$.

The goal of this article is to connect to topology $[P]$,solve a $[L]$ can make flow balanced, This paper uses a multi-objective mixed integer linear programming to solve the problem。Suppose there are N nodes in the whole network,The entire network topology is divided into K time slices,use t_{max} to represent the maximum time in any connection in the entire network can use, t_{min} represent the minimum amount of time that any connection in the entire network can use。K is the status of the node,N is the number of nodes.

In the satellite nodes under the premise of the communication link bandwidth is constant, in order to make the entire network flow more balanced, that is, to make each connection using the total time closer, so, on the one hand, to increase the original use of time The use of smaller connections, namely:

$$max\{t_{min}+\alpha(\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}l_{k,i,j}*t_k)=t_{min}+\alpha*t_{obj}\} \quad (4)$$

The constraint is:

$$\begin{cases}\sum_{k=1}^{K}l_{k,i,j}*t_k\geq t_{min}\\ l_{k,i,j}=l_{k,j,i}\\ l_{k,i,j}\in\{0,1\}\end{cases} \quad (5)$$

Among them, $t_{obj}=\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{N}l_{k,i,j}*t_k$ represents the total time the connection is used throughout the network, α indicates its weight coefficient。Equation (4) increases the usage time of the connection by adding an $\alpha*t_{obj}$ to the shortest time t_{min} used for the connection. The three formulas of Eq. (5) indicate that there is a lower boundary at the time when each connection of the whole network is used t_{min} ; For the same link, the forward link is equivalent to the reverse link; the link in the network either exists or does not exist, only the two mutually exclusive state. On the other hand, to reduce the use of time-consuming connection, that is：

$$min\{t_{max}\} \quad (6)$$

The constraint is:

$$\begin{cases}\sum_{k=1}^{K}l_{k,i,j}*t_k\leq t_{max}\\ \sum_{k=1}^{K}l_{k,i,j}*t_k\geq t_{min}\\ l_{k,i,j}=l_{k,j,i}\\ l_{k,i,j}\in\{0,1\}\end{cases} \quad (7)$$

Wherein, the constraint of formula (7) indicates that there is an upper boundary for each connection usage time of the entire network t_{max} and a lower boundary t_{min} ;For the same link, the forward link is equivalent to the reverse link; the link in the network either exists or does not exist, only the two mutually exclusive state.

2.4 Solve the formula

Each state can be related to a connected graph by means of solving the Blossom algorithm of the general matching problem. Over time providing a uniform contact, the contact arc can be weighted as a function of the time a node does not have time to be assigned a link between them. Find the algorithm for maximum edge number matching so that each vertex matches at most one edge. In particular, given a topological matrix [P],Blossom algorithm found [L] maximum arcs is limited to each node's interface.

Match by augmenting the path by iterating over an initial null match. The enhancement path is the continuity between the two edges. The isolated vertices match at edges that match and do not alternate. It can be proved that the matching is the largest iff and only if there is no matching source path for the extended path. Effectively finding an extension path is a core part of the Blossom algorithm and is used to achieve flowering and contraction. Although the Blossom algorithm successfully enables the best arc combination with a given interface limit of 1, it searches for perfect matches only, meaning that all vertices must be overwritten. In particular, our problem is to pursue maximum weight under normal circumstances and then reduce the maximum (imperfect) match to the perfect match.

Input: Contact Topology[P] Set [P] size K×N×N Start State Time $[T]_K$

Output: Contact Plan [L] size K×N×N Set Invalid Contact Time(ICT)

for each node

$ICT_{i,j} \leftarrow 0 \quad \forall i, j$;

for each node

for $k \leftarrow 0$ to K

Set Weights for each invalid time

$[\mathrm{W}]_{k,i,j} \leftarrow ICT_{i,j} \quad \forall i, j$;

If $[\mathrm{L}]_{k,i,j} = 0$

$ICT_{i,j} \leftarrow ICT_{i,j} + \mathrm{t}_k$

end if

Solve State by Blossom

Blossom($[P]_k, [\mathrm{L}]_k, [\mathrm{W}]_k$)

end for

Figure 3 BLO-DTN algorithm

The more chances a connection joins without planning for the final contact time [L], the more Blossom algorithm has with such a selection as the subroutine contains the appropriate maximal match reduction. The algorithm is named BLO-DTN, the algorithm is described below. Arc contact time metrics provide a given contact i, j. The specific algorithm is as follows.

3. Performance evaluation

3.1 Simulation environment

In order to verify the performance of the algorithm, the network simulation software OPNET Modeler14.5 and STK were used to simulate and analyze the protocol. The LEO satellite network constructed in this paper consists of 9 walker-distributed LEO satellites and 3 ground stations. The main parameters of the satellite constellation are shown in Table 2.

Table2 Satellite constellation main parameters

parameter name	Parameter value
Track height	1680km
Number of orbital planes	3
Orbital satellite number	3
Orbital inclination	60 degrees
Link rate	1-10Mbps
Ground station	Kashi、Beijing、Sanya

With OPNET Modeler14.5 built simulation view shown in Figure 4

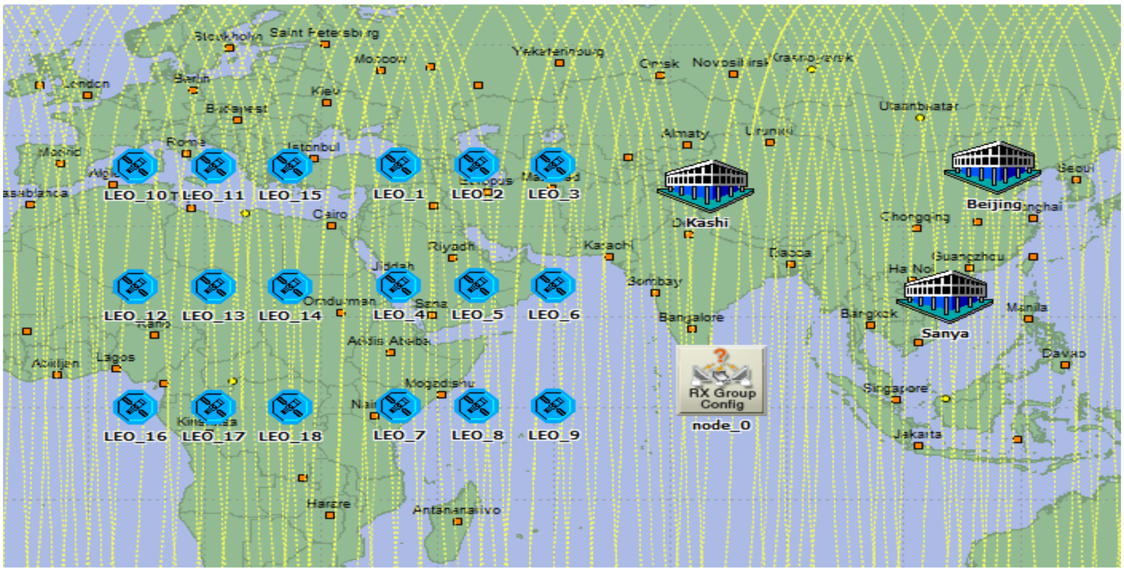

Figure 4 build low-orbit satellite two-dimensional map OPNET

3.2 Simulation results

(1)Load statistics and equalization coefficients

In order to reflect the load of satellites under different routing algorithms, the number of satellites processing data packets in the statistical simulation. Figure 5 shows the statistics of satellite processing data packets involved in transmission. In order to better describe the load balancing effect, the load distribution index is introduced [10]:

$$f=(\sum_{i=1}^{n}x_i)^2/(n\sum_{i=1}^{n}x_i^2) \tag{8}$$

Where n is the total number of LEO satellite ISLs; indicates the number of link packets that are numbered;; $f\in[0,1]$, f value indicates a more balanced load.

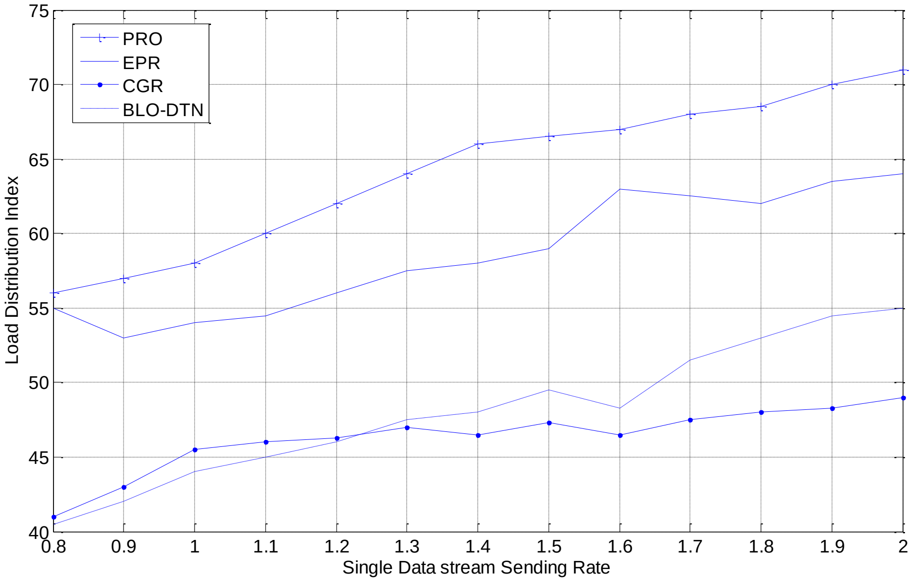

Figure 5. Load distribution index diagram

(2)Data loss rate

As shown in FIG 6, the BLO-DTN algorithm maintains a lower packet loss rate because the introduced load balancing mechanism increases the minimum flow and reduces the maximum flow, and the packet loss rate is obviously smaller than those of the other three algorithms. When network load aggravates, the packet loss ratio of PRO algorithm and EPR algorithm is significantly higher than CRG algorithm and BLO-DTN algorithm..

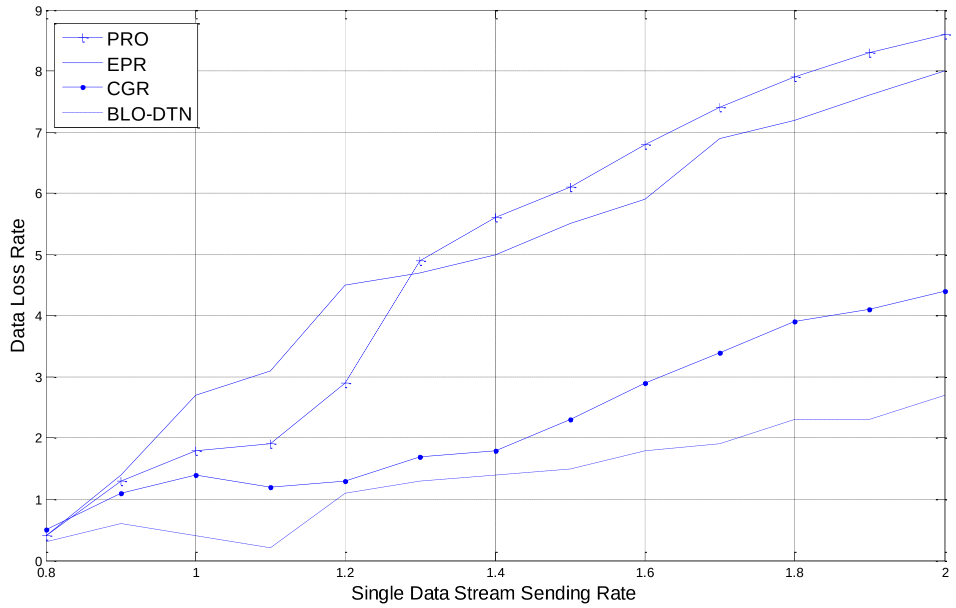

Figure 6. Data loss rate

(3)Data transmission delay

The load balancing mechanism of BLO-DTN algorithm increases the complexity of the algorithm without sacrificing the delay, and the average of BLO algorithm is obviously lower than the other three algorithms. When buffering 3-5MB, the EPR algorithm increases the delay greatly due to a large amount of message redundancy can not be processed, while the priority queue scheduling mechanism in CGR algorithm can reduce the delay well. The BLO-DTN algorithm in this paper is effective So that flow is more balanced, to avoid a large number of messages in the queue or transmission redundancy, effectively reducing the delay.

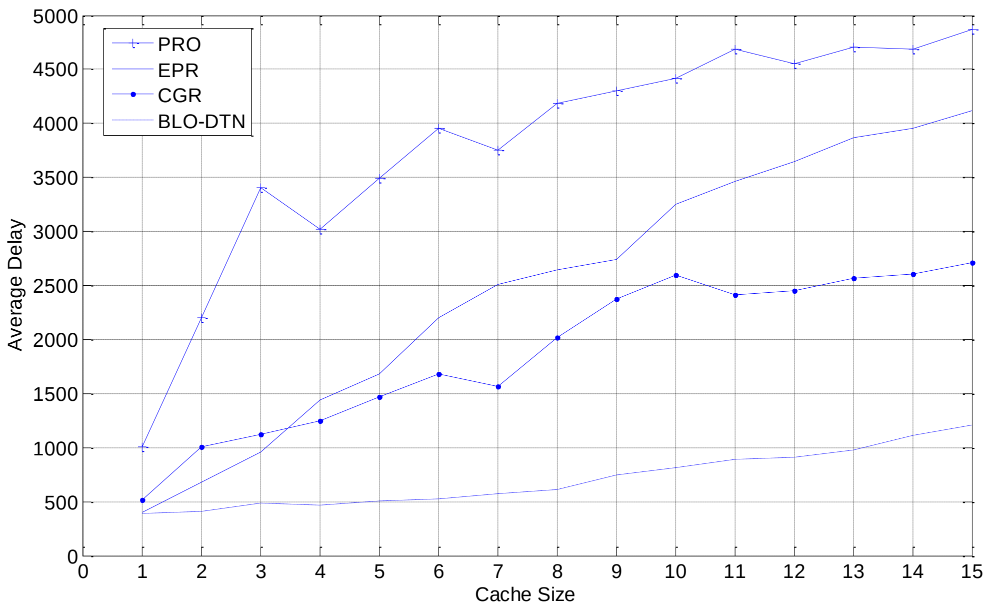

Figure 7. Transmission delay comparison chart

4. Conclusion

In this paper, aiming at the problem of high dynamic topological change and long transmission delay in DTN satellite networks, which leads to unbalanced data flow in the inter-satellite link, the paper first modifies the format of the bundle to consider the remaining capacity of satellite nodes for congestion avoidance, The BLO-DTN algorithm based on mixed integer linear programming is more effective in that the interstellar flow is more balanced, the routing efficiency is improved, and the average transmission delay and packet loss rate are effectively reduced. Therefore, the BLO-DTN algorithm combined with the remaining capacity management mechanism is more suitable for the DTN satellite network with high dynamic, easy to interrupt and long delay.

Acknowledgements

The National Natural Science Foundation of China（61722105）.

References

[1] S. Farrell and V. Cahill, Delay- and Disruption-Tolerant Networking[J] Artech House, 2006.
[2] V. Cerf , A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss . Delay-Tolerant Networking Architecture[J], IETF RFC 4838, Apr. 2007:89-95.
[3] RANGO F D. Hop-by-hop local flow control over interplanetarynetworks based on DTN architecture[C]//2008 IEEE InternationalConference on Communications (ICC 2008). Beijing, China, c2008:1920-1924.
[4] Seligman M, Fall K, Mundur P. Storage routing for DTN congestion control[J]. Wireless Communications and Mobile Compu-ting, 2007, 7(10): 1183-1196.
[5] Adaptive routing for packet-oriented intersatellite link networks: Performance in various flow scenarios. Mohorcic, Mihael,Werner, Markus,Svigelj, Ales,Kandus, Gorazd. IEEE Transactions on Wireless Communications . 2002
[6] Analysis and Comparison of LEO and MEO Satellite Networks. P. Truchly,P. Buran. 49th International Symposium ELMAR-2007 . 2007
[7] Datagram routing algorithmfor LEO satellite networks. Ekici E,,A kyildiz I F,Bender M D. Proc IEEE IN FOCOM 2000 .
[8] Scott.Burleigh,“ContactGraphRouting”,Internet-Draft,July2010.work-in-progresshttp://tools.ietf.org/html/draft-burleigh-dtnrg-cgr.
[9] Vahdat A, Becker D. Epidemic routing for partially-connected Ad-Hoc network. Duke University Technical Report Cs-2000-06 Tech. Rep, 2000: 1-14.
[10] Taleb T, Mashimo D, Jamalipour A, et al. Explicit load balancing technique for NGEO satellite IP networks with on-board processing capabilities[J]. IEEE/ACM transactions on Networking, 2009, 17(1):281-291.