Agent-Based Distributed Routing Algorithm with Traffic Prediction for LEO Satellite Network

Zhenyu Na[1], Zihe Gao[2], Yang Cui[3], Liming Chen[4] and Qing Guo[4]

[1]*School of Information Science and Technology,*
Dalian Maritime University, Dalian, China
[2]*Institute of Telecommunication Satellite,*
China Academy of Space Technology, Beijing, China
[3]*Institute of Science and Technology, Harbin Institute of Technology, Harbin, China*
[4]*Communication Research Center, Harbin Institute of Technology, Harbin, China*

nazhenyu@dlmu.edu.cn, zhg_hit@163.com, ycui@dlmu.edu.cn, qguo@hit.edu.cn

Abstract

Satellite network, especially low earth orbit (LEO) satellite network, which has advantages of global coverage and short round-trip time (RTT), has played an increasingly important role in the future generation of global communication network. Due to fast velocity, frequent handover and time-varying topology, designing effective routing algorithm for LEO satellite network has always been challenging. An Agent-based distributed Routing Algorithm with Traffic Prediction (TPARA) is proposed in this paper. TPARA algorithm, which fully takes traffic density of the surface of the Earth into account, consists of two parts: traffic prediction and routing decision. In the former, not only the heterogeneity of traffic on the ground is considered, but upcoming traffic from terrestrial terminals is predicted using an improved Kalman filtering as well. In the latter, mobile agents (MAs) are used to explore satellite network and collect routing information. The ultimate routing decision is determined by both current status and future status of satellite network. Simulation results show that, compared with the classical algorithm of ACO, TPARA not only has shorter delay, but alleviates congestion as well.

Keywords: *LEO, satellite network, routing algorithm, traffic prediction, mobile agent*

1. Introduction

The next generation of communication network will be featured by wireless mobility, global coverage and broadband high-speed transmission. Compared with geostationary orbit (GEO) and medium earth orbit (MEO) satellite systems, LEO satellite system has shorter transmission delay, low propagation loss and globally seamless coverage (especially, coverage of polar zones). In addition, as satellite network is able to connect terrestrial networks with aerospace communication platforms and different wireless terminals, LEO satellite network will be the indispensable infrastructure of the globally integrated communication network oriented to the future [1-2].

In LEO satellite system, satellites directly communicate with each other through inter-satellite links (ISLs) on which the data packets and signaling are simultaneously transmitted. However, as LEO satellite network is usually composed of tens of satellite on several orbits, its routing problem is more complicated than GEO and MEO satellite networks. On one hand, LEO satellite with high velocity leads to time-varying and dynamic topology, as result, ISLs will frequently break or recover. On the other hand, since broadband multimedia services will dominate the future LEO satellite network, drastically increasing traffic will overload or congest satellite nodes. Due to these factors, designing effectively adaptive routing algorithm for LEO satellite network is highly imperative [3-4].

Considering above characteristics of LEO satellite network, this paper proposes an agent-based distributed routing algorithm with traffic prediction (TPARA). In TPARA, based on the space-time distribution of traffic on the surface of the Earth, an improved Kalman filtering is presented to predict imminent traffic in satellite nodes, while mobile agents take charge of collecting network information. The routing decision is not only dependent on the current status, but the future status of the network. When traffic is sharply changes, it can be predicted by TPARA, and then a traffic notification can be received by satellite node. Consequently, the adjacent satellite nodes increase or decrease traffic to be sent to the node so that congestion (overload) and underload of node can be alleviated.

2. Related Work

In recent years, some routing algorithms for LEO satellite network have been proposed.

In early stage, inspired by ATM (Asynchronous Transfer Mode) technique, many researches were generally concentrated on connection-oriented routing algorithms [5-7]. Concretely, according to the topological periodicity of LEO satellite network, its continuous and time-varying topology can be decomposed into a series of static topological slices so that dynamic routing can be transformed into periodically static routing. These algorithms use offline computations, and satellites only serve as transponder, while the optimal paths from source node to destination node are only computed on the ground station. Satellite node only forwards packets according to routing tables. However, these algorithms are less robust more often than not due to frequent handover in LEO satellite network. In addition, centralized-processing in ground station for routing computation can not adapt to the overwhelming increase of multimedia and interactive traffic. Therefore, distributed and adaptive routing algorithms become highly indispensable and inevitable.

With the development of IP (Internet Protocol) technique oriented to non-connection, the routing algorithms based on satellite IP are widely developed. The key point of this type of routing algorithms lies in the routing on ISLs. Many algorithms were proposed such as [8-10]. They adopt logical address [8-9] or cell address [10], and then their routing strategy is converted into finding the shortest path, but the shortest path may not be the optimal path because the satellite node on the shortest path may be overloaded or congested. Therefore, balanced-load based routing algorithms draw researchers’ more attention.

[11] and [12] proposed routing algorithms based on load balance. The former mitigates congestion through bypassing the congested node, while in the latter satellite node continuously monitors its own status and notify its adjacent nodes to adjust traffic load. Both the algorithms only consider traffic optimization on transmission path and do not take the unevenness of traffic on the surface of the Earth into account at all. According to most of previous researches, balancing-load routing algorithms mainly focus on adjust of traffic in satellite node to avoid congestion. However, the root cause behind congestion is never been considered in literature. From the perspective of system, congestion in LEO satellite network is highly related to the distribution of traffic on the ground such that balancing-load routing algorithms is extremely

meaningful for the satellite node of which below sub-satellite point the traffic load changes sharply.

To cope with dynamic environment, artificial intelligent (AI) has been proposed which imitates behaviors of some animals by means of computation. In addition, with the explosive increase of multimedia traffic, QoS (Quality of Service) is another essential component for routing algorithms named QoS routing, which have multiple QoS constrains. Combined with multiple QoS constrains, many routing algorithms based on AI have been proposed, such as artificial neural networks [13-14], genetic algorithms [15-16], ant colony optimization (ACO) [17-18] and particle swarm optimization [19-20]. Compared with previous work, routing algorithms based on AI have the advantages of adaptation self-organization and self-learning. However, few current algorithms integrate AI with traffic prediction and balance. As described before, traffic balance is only focused on satellite node and ISLs by adopting passive traffic adjust without considerations of ground traffic distribution and traffic prediction. Especially, traffic prediction, as a sort of proactive scheme, can estimate the upcoming traffic and take adequate measure in advance to avoid congestion.

In view of above analyses, a novel distributed routing algorithm based on AI with traffic prediction and balance is proposed. Multi-agent is introduced into the algorithm to realize distributed and adaptive functions. Two improvements are highlighted in the algorithm: one is an improved adaptive filtering is adopted to predict traffic; another is the density of traffic on the ground being considered. Mobile agents explore and collect routing information, while the ultimate routing decision is determined by both current status and future status of satellite network.

3. Overview of Agent-Based Distributed Routing Algorithm with Traffic Prediction (TPARA)

3.1. Topology Model of LEO Satellite Network

As used in our previous papers [21-22], the Iridium constellation is utilized in this paper. The Iridium system is composed of six separate orbit planes and each plane has eleven satellites. Thus, each satellite has four neighbor satellites: two in the same orbit plane while the other two in two neighbor orbits. The links between two satellites in the same orbits are called intra-plane ISLs, while the links between satellites in different planes are called inter-plane ISLs. In other words, each satellite has four ISLs. When a satellite flies above the Polar Regions, due to the relative speed between satellites and the Doppler frequency shift, inter-plane ISLs will be broken or recovered.

3.2. Traffic Prediction in TPARA

Routing in satellite network is an optimization problem in essence that can be divided into two associated parts: path exploration and path maintenance. Each routing algorithm need to collect the information of network status for routing decision. However, collecting such information always lags behind the change of network status that routing decision strategy has to use hysteretic information. Therefore, if routing information can be forecast, the performance of the routing algorithm will be enhanced.

According to historical data, prediction is a proactive process using proper methods and techniques to estimate the future status. By now, statistical regression analysis, time series forecasting, neural networks and wavelet prediction have been commonly-used prediction algorithms. As satellite network covers entire surface of the Earth, the local time of sub-satellite point, population distribution and the level of economic development are more often than not so different that result in the unevenness of traffic on the surface of the Earth. In addition, the diversity of topography also influences traffic distribution in satellite network. For example, when satellite continuously flies over continents and oceans, traffic below sub-satellite point is often drastically time-varying. These factors and the nonlinear combination of these factors make traffic prediction for satellite network so complicated that a well-performed prediction algorithm is imperative. As our previous work, an improved Kalman filtering algorithm is proposed [23] and it has satisfactory prediction results for stochastic and time-varying environments. Therefore, the filtering algorithm is adopted in this paper for traffic prediction in satellite network.

3.3. Application of Multi-Agent System

Traditionally, previous routing algorithms are prone to a certain specific background, resulting in that they are difficult to be extended or applied to other conditions. Recent years, distributed AI has been widely adopted in routing algorithms because it is well adaptive to large-scale, open and dynamic environment. Thereinto, multi-agent is an important branch of distributed AI. In multi-agent system, different agents are responsible for different subtasks and work cooperatively to finish a complete task. The final goal of multi-agent system is that using different agents optimize resource utilization. In this paper, mobile agent is applied to routing algorithm which is divided into several subtasks being finished by different agents.

4. Traffic Prediction and Traffic Distribution

4.1. Traffic Prediction

Kalman filtering (KF) can be expressed by a state equation and an observation equation, *i.e.*,

$$\mathbf{X}(n+1)=\mathbf{F}(n+1,n)\mathbf{X}(n)+\mathbf{w}(n) \tag{1}$$

$$\mathbf{y}(n)=\mathbf{C}(n)\mathbf{X}(n)+\mathbf{v}(n) \tag{2}$$

where, $\mathbf{X}(n)$ is the $M\times 1$ state vector; $\mathbf{F}(n+1,n)$ is the $M\times M$ state transition matrix; $\mathbf{w}(n)$ is the $M\times 1$ state noise vector. $\mathbf{y}(n)$ is the $N\times 1$ observation vector; $\mathbf{C}(n)$ is the $N\times M$ observation matrix; $\mathbf{v}(n)$ is the observation noise.

It is supposed that prediction is based on measurement being taken every interval τ. Let $x(n)$ denote the traffic at time n, then $x'(n)$ denotes the first-order differential of $x(n)$. If τ is small enough, the following formula can be obtained

$$x(n)=x(n-1)+Tx'(n-1) \tag{3}$$

Similarly, the change rate of the traffic rate $x''(n)$ can be expressed as

$$x'(n)=x'(n-1)+Tx''(n-1) \tag{4}$$

As many stochastic factors influence the change the traffic velocity, noise is modeled as

$$x''(n-1)=w(n) \tag{5}$$

Therefore, according to (1)-(5), $\mathbf{F}(n+1,n)$, $\mathbf{w}(n)$ and $\mathbf{C}(n)$ can be given by

$$\mathbf{F}(n+1,n)=\begin{bmatrix}1 & T & 0\\ 0 & 1 & T\\ 0 & 0 & 0\end{bmatrix} \tag{6}$$

$$\mathbf{w}(n)=\begin{bmatrix}0 & 0 & w(n)\end{bmatrix}^{\mathrm{T}} \tag{7}$$

$$\mathbf{C}(n)=\begin{bmatrix}1 & 0 & 0\end{bmatrix} \tag{8}$$

Consequently, the observation equation (2) becomes the following form

$$y(n)=\mathbf{C}(n)\mathbf{X}(n)+v(n) \tag{9}$$

Let $\hat{\mathbf{X}}_1(n)$ and $\hat{\mathbf{X}}_2(n)$ denote the estimates of $\mathbf{X}(n)$ derived from (1) and (2), i.e.

$$\hat{\mathbf{X}}_1(n)=\mathbf{F}\hat{\mathbf{X}}_1(n-1) \tag{10}$$

$$y(n)=\mathbf{C}\hat{\mathbf{X}}_2(n) \tag{11}$$

Considering a linear combination of the two estimates (10) and (11), shown by

$$\hat{\mathbf{X}}(n)=[\mathbf{I}-\mathbf{A}(n)]\hat{\mathbf{X}}_1(n)+\mathbf{A}(n)\hat{\mathbf{X}}_2(n) \tag{12}$$

where, $\mathbf{A}(n)$ is a $M\times M$ weight matrix that $\mathbf{A}(n)=\mathbf{m}(n)\mathbf{C}(n)$. The estimated error vector $\boldsymbol{\xi}(n)$ of the state vector $\mathbf{X}(n)$ can be expressed as

$$\boldsymbol{\xi}(n)=\mathbf{X}(n)-\hat{\mathbf{X}}(n)=[\mathbf{I}-\mathbf{m}(n)\mathbf{C}(n)][\mathbf{X}(n)-\hat{\mathbf{X}}_1(n)]-\mathbf{m}(n)v(n) \tag{13}$$

Then, the correlation matrix $\mathbf{R}(n)$ of the error vector $\boldsymbol{\xi}(n)$ can be gotten

$$\mathbf{R}(n)=\mathrm{E}[\boldsymbol{\xi}(n)\boldsymbol{\xi}^{\mathrm{T}}(n)]=[\mathbf{I}-\mathbf{m}(n)\mathbf{C}(n)]\cdot\mathbf{R}_1(n)\cdot[\mathbf{I}-\mathbf{m}(n)\mathbf{C}(n)]^{\mathrm{T}}+\mathbf{m}(n)\sigma_v^{\mathrm{T}}\mathbf{m}^{\mathrm{T}}(n) \tag{14}$$

where,

$$\mathbf{R}_1(n)=\mathrm{E}\left[\left(\mathbf{X}(n)-\hat{\mathbf{X}}_1(n)\right)\left(\mathbf{X}(n)-\hat{\mathbf{X}}_1(n)\right)^{\mathrm{T}}\right] \tag{15}$$

Let $\dfrac{\partial \mathrm{tr}\mathbf{R}(n)}{\partial \mathbf{m}(n)}=0$, then $\mathbf{m}(n)$ can be solved as

$$\mathbf{m}(n)=\mathbf{R}_1(n)\mathbf{C}(n)^{\mathrm{T}}\cdot\left[\mathbf{C}(n)\mathbf{R}_1(n)\mathbf{C}(n)^{\mathrm{T}}+\sigma_v^2\right]^{-1} \tag{16}$$

Therefore,

$$\mathbf{R}(n)=[\mathbf{I}-\mathbf{m}(n)\mathbf{C}(n)]\cdot\mathbf{R}_1(n) \tag{17}$$

To solve $\mathbf{R}_1(n)$, considering

$$\mathbf{X}(n)-\hat{\mathbf{X}}_1(n)=\mathbf{F}[\mathbf{X}(n-1)-\hat{\mathbf{X}}_1(n-1)]+\mathbf{w}(n) \tag{18}$$

Note that the expectation operation is independent of the state transition matrix $\mathbf{F}$, while $\mathbf{w}(n)$ and $\mathbf{X}(n)$ are uncorrelated, formula (15) can be rearranged as

$$\mathbf{R}_1(n) = \mathbf{F}\mathbf{R}_1(n-1)\mathbf{F}^T + \mathbf{Q}_w \tag{19}$$

where, $\mathbf{Q}_w$ is the correlation matrix of the state noise. Obviously, the calculation of $\mathbf{R}_1(n)$ can be performed recursively. Due to noise is unknown, it should be dynamically estimated as [23]

$$\hat{\mathbf{Q}}_w(n+1) = (1-d_n)\hat{\mathbf{Q}}_w(n) + \frac{d_n}{n+1}[\varepsilon^2(n+1)\mathbf{K}(n+1)\mathbf{K}^{\mathrm{T}}(n+1) + \mathbf{P}(n+1) - \mathbf{F}\mathbf{P}(n)\mathbf{F}^{\mathrm{T}}] \tag{20}$$

$$\hat{\sigma}_v^2(n+1) = (1-d_n)\hat{\sigma}_v^2(n) + d_n[\varepsilon^2(n+1) - \mathbf{C}P(n+1|n)\mathbf{C}^{\mathrm{T}}] \tag{21}$$

Where, d_n is forgetting factor.

The complete traffic prediction process is given below. For n = 1, 2, …, formulas (10), (16), (17), and (19)-(21) should be calculated recursively.

Where,

$$\mathbf{K}(n+1) = \mathbf{P}(n+1|n)\mathbf{C}^{\mathrm{T}}(n)(\mathbf{C}(n)\mathbf{P}(n+1|n)\mathbf{C}^{\mathrm{T}}(n) - \hat{\sigma}^2(n))^{-1} \tag{22}$$

$$\varepsilon(n+1) = y(n+1) - \mathbf{C}(n)\mathbf{F}\hat{\mathbf{X}}(n) - \hat{r}(n) \tag{23}$$

$$\mathbf{P}(n+1) = [\mathbf{I} - \mathbf{K}(n+1)\mathbf{C}(n)]\mathbf{P}(n+1|n) \tag{24}$$

$$\mathbf{P}(n+1|n) = \mathbf{F}\mathbf{P}(n)\mathbf{F}^{\mathrm{T}} + \hat{\mathbf{Q}}_w(n) \tag{25}$$

$$\hat{r}(n+1) = (1-d_n)\hat{r}(n) + d_n[y(n+1) - \mathbf{C}(n)\hat{\mathbf{X}}(n+1)] \tag{26}$$

4.2. Traffic Distribution

In general, ground stations and terminals communicate with the satellite of LoS (Line of Sight). In view of traffic load, four factors should be taken into account. The former two is natural factor, while the latter two is human factor. First, land and ocean are the most important natural factor that influences traffic of satellites. The second natural factor is the rotation of the Earth, because daytime traffic is often dominant compared with night traffic. The first human factor is the development of economics, because the developed region has higher traffic than the under-developed region. The last factor is emergency and assembly, such as Shanghai World Expo. Due to above factors, the traffic of satellite node will be time-varying that because of high-speed of LEO satellites, the topology of LEO satellite network has always been time-varying, resulting in traffic of satellite changing with the change of sub-satellite point. Traffic change highly affects the performances of routing algorithms.

Therefore, two hypotheses are made in TPARA. One is traffic density of satellite node is only related to its queue length and ISL capacities. Specifically, if current queue length is close to buffer size, the satellite is busy and vice versa. Another hypothesis is that traffic source of satellite node is originated from the ground and its neighbor satellites. Traffic come from neighbor satellites is mainly dependent on packet routing scheme. In contrast, traffic come from the ground, as described above, is determined by natural and human factors. The key idea of TPARA is that routing decision is updated by predicting traffic from ground and traffic sent to the satellite node is reduced to maximize utilization of network resources.

It is evident that if potential traffic from ground could be quantified, traffic prediction can be more accurate. According to coverage of LEO satellite and levels of regional economics, the surface of the Earth is divided into some regions as shown in Figure1 in which the numbers from 1 to 10 stand for the potential traffic dependent on the natural and human factors. The length and the width of each region are 15° in longitude and latitude. It can be seen that the number 1 meaning the least traffic occupies the most of the surface of the Earth because ocean covers 71% of the Earth, while the numbers 9 and 10 only exists in North America, Europe and East Asia, because these regions are often developed in economy. NOEKF algorithm is used to predict traffic for every region. TPARA includes two parts: traffic prediction and routing decision. The latter will be discussed in the next section.

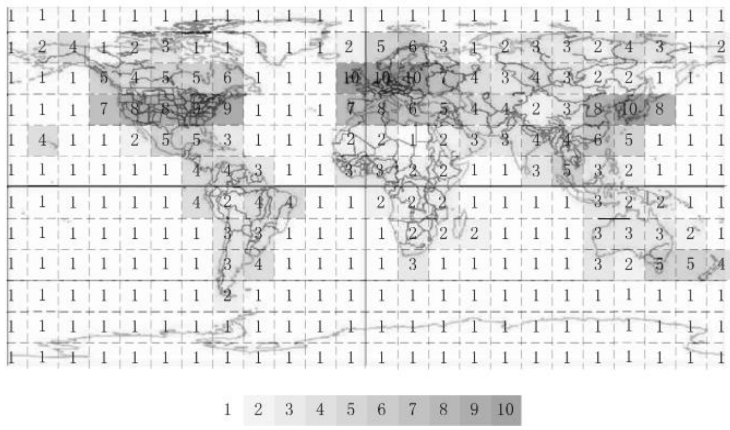


Figure 1. Potential Traffic Distribution on the Surface of the Earth

5. Descriptions of Agent-based distributed Routing Algorithm with Traffic Prediction (TPARA)

5.1. Routing Decision of TPARA

When it comes to routing decision, the concept of multi-agent is still adopted to implement distributed routing. The behaviors of TPARA can be summarized as follows.

(1) From each satellite node s, forward MAs are generated and sent to specific destination node d in regular interval. They are indicated with $F^i_{s\to d}$, where i is the identifier of forward MA.

(2) Each forward MA collects non-local routing information about path and traffic. Forward MAs are snet hop-by-hop from source node to destination node with the same priority to data packets. All forward MAs explore satellite network concurrently, independently and asynchronously. For each intermediate node, random decision policy is used to select the next hop. In the course of movement, forward MAs record information about traversal time and node identifier along the paths they passed.

(3) When forward MA arrives at destination node, it is destroyed and then backward MA $B^i_{s\to d}$ is generated, going back to the source node along the same path

$$P^i_{\mathrm{s\to d}} = [s, v_1, v_2, \cdots, d]$$

as before but in the opposite direction, where $v_j \in N(v_j) \cap P^i_{s \to d}$ is intermediate node. With higher priority than data packets, backward MA arrives at v_j and updates local routing information.

(4) When backward MA returns to the source node, it is deleted from the network. Each MA has a maximum TTL (time-to-live). If the life of MA exceeds TTL, it will also be deleted.

(5) Data packets are routed according to random decision policy based on the information contained in the routing tables. These tables are derived from the pheromone tables which are used to route MAs. In the meantime, the pheromone tables are determined and modified by MAs. Only the optimal paths are stored in the data-routing tables which guarantee transmission quality of data packets.

5.2. Data Structure

It is supposed that there are N nodes in satellite network and each node has L neighbor nodes.

In TPARA, each satellite node k comprises five components including pheromone matrix T^k, data-routing table R^k, link queue L^k, predicted traffic P^k and statistical parametric model M^k, as demonstrated in Figure 2.

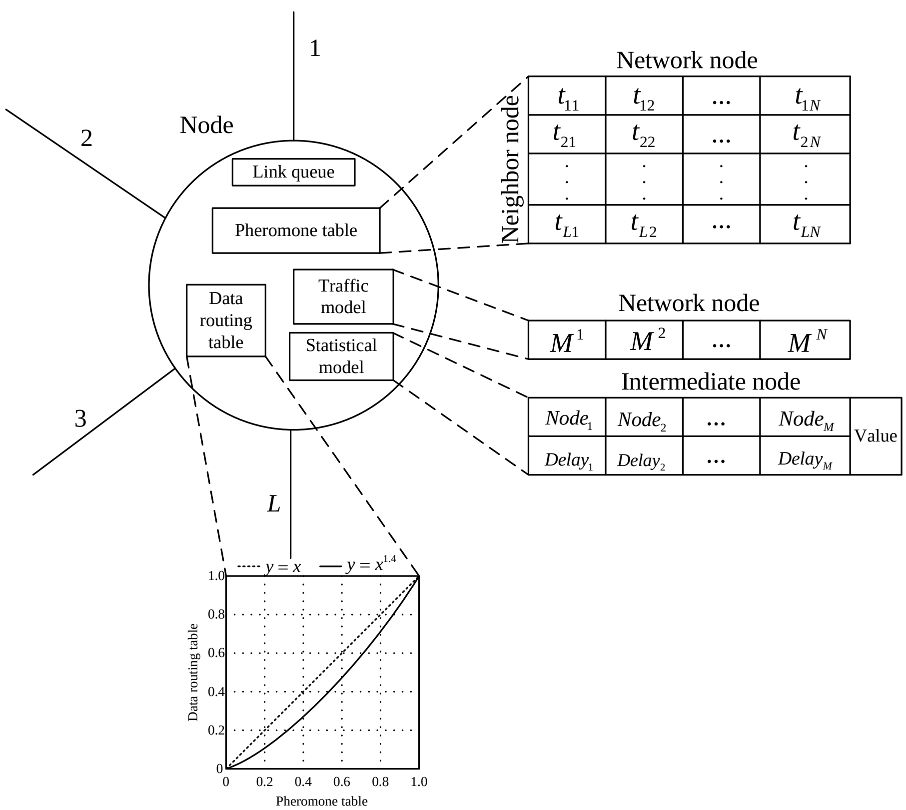


Figure 2. Data Structure Stored in Satellite Node

Pheromone matrix T^k is similar to routing tables in distance-vector algorithms and the element t_{nd} of T^k is the probability from current node to one of neighbor nodes. The value of t_{nd} is in the range of [0, 1] and following expression is held

$$\sum_{n \in N_k} t_{nd} = 1 \quad d \in [1, N], \; N_k = \{\text{neighbors}(k)\} \tag{27}$$

where, N is the number of nodes in satellite network.

Data routing table R^k is a routing table used to forward data packets. R^k has the same structure to T^k. The elements of R^k are obtained by exponential transformation and re-normalization to 1 of the corresponding elements of T^k. Data packets are sent to neighbor nodes in probabilities according to the elements of R^k.

Link queue L^k is used to describe the status of the queue in local satellite node. L^k gives snapshots of queue status, while T^k is the path information gotten by MAs. In other words, T^k records long-term global information of satellite network and L^k reflects short-term traffic information in local node.

P^k records the predicted traffic from the surface of the Earth. Its elements are obtained from NOEKF algorithm and determine the future status of satellite along with queue L^k.

Statistical model M^k is a vector of $(N-1)$-dimension with data structures $(\mu_{\mathrm{d}}\sigma_{\mathrm{d}}^2, W_{\mathrm{d}})$, where μ_{d} and σ_{d}^2 stand for the mean and the variance of the traversal time in view of the destination node d, respectively. W_{d} is the optimal traversal time to node d of the latest W observed values.

5.3. Implementation of TPARA

(1) Forward MA behavior

Every interval Δt, forward MA $F_{s \to d}$ is generated to establish a feasible and low-cost path from source node s to destination node d. As forward MA shares the same priority and queue to data packets, MA is able accurately record network congestion. The destination d is chosen in a probability p_d which is defined as

$$p_d = \frac{f_{sd}}{\sum_{d'=1}^{N} f_{sd'}} \tag{28}$$

where, f_{sd} is the number of packets sent from node s to node d.

While moving to destination node, forward MA keeps a private memory H in which the number and the time of nodes it passed are stored. The list

$$V_{v_0 \to v_m} = [v_0, v_1, \cdots, v_m]$$

maintains the order of the nodes visited by forward MA and the list

$$T_{v_0 \to v_m} = [T_{v_0 \to v_1}, T_{v_1 \to v_2}, \cdots, T_{v_{m-1} \to v_m}]$$

records the traversal time from v_{k-1} to v_k of forward MA.

At each intermediate node k, forward MA chooses its next hop according to the following rule: If $n \in V_{s\to k}$, $\forall n \in N_k$, all the neighbors have been visited by the forward MA, then it chooses the next hop in probability of p_{nd}, but excluding the node from which the forward MA coming.

$$\begin{cases} p_{nd} = \dfrac{1}{|N_k|-1} & \forall n \in N_k \wedge (n \neq v_{i-1} \vee |N_k| = 1) \\ p_{nd} = 0 & \text{otherwise} \end{cases} \tag{29}$$

Otherwise, random decision policy is applied as follows:

$$\begin{cases} p_{nd} = \dfrac{t_{nd} + \alpha l_n}{\sum_{n=1}^{|N_k|-1} t_{nd} + \alpha(|N_k|-2)} & \forall n \in N_k \wedge n \notin V_{s\to k} \\ p_{nd} = 0 & \text{otherwise} \end{cases} \tag{30}$$

where, $l_n \in [0,1]$ represents queue status which is defined as

$$l_n = 1 - \frac{q_n}{\sum_{n'=1}^{|N_k|} q_{n'}} \tag{31}$$

q_n means the data packets to be sent from node k to neighbor node n. $\alpha \in [0,1]$ is a weight factor. Formula (30) determines the probabilities of selecting the next hop nodes.

(2) Backward MA behavior

Once forward MA $F_{d\to s}$ arrives at the destination node, it is destroyed and then backward MA $B_{d\to s}$ is generated which contains all information in $F_{d\to s}$. Path update is started from backward MA going back to node s along the original path in the opposite direction. Along the path, all data structures at each node will be updated. A difference between forward and backward MAs is that

(3) Local models updates

When backward MA arrives at intermediate node k, three operations are carried out by backward MA: M^k is updated, path from k to d is evaluated and path to d is strengthened. μ_d and σ_d^2 are updated according to following expressions

$$\mu_d \leftarrow \mu_d + \eta(T_{k\to d} - \mu_d) \tag{32}$$

$$\sigma_d^2 \leftarrow \sigma_d^2 + \eta\left((T_{k\to d} - \mu_d)^2 - \sigma_d^2\right) \tag{33}$$

$$W_d = T_{k\to d} \quad (\text{if } T_{k\to d} < W_d) \tag{34}$$

where, η weights the number of the most recent samples.

After updating M^k, the path visited by the forward MA must be evaluated as below

$$r = c_1\left(\frac{W}{T}\right) + c_2\left(\frac{I_{\sup} - I_{\inf}}{(I_{\sup} - I_{\inf}) + (T - I_{\inf})}\right) \tag{35}$$

where, r is the reinforcement factor , W is the optimal traversal time of the latest w observed samples experienced by forward MA. $I_{\sup}$ and $I_{\inf}$ are estimates of the supremum and the infimum of μ in a certain confidence interval. The coefficients c_1 and c_2 are weights of two terms.

r stands for the current status of satellite node. When satellite moves from low-traffic region to high-traffic region, it becomes idle because the amount of data packets is small. In other words, r should be set larger such that routing decision policy encourages data to be sent to this node. But in fact, traffic of this satellite node would sharply increase to leading to congestion. Therefore, reinforcement factor r should be revised by predicted traffic P^k.

Suppose that the prediction interval is Δt and $\hat{P}$ is the predicted traffic at time $(t+\Delta t)$. Let Q and $q(t)$ denote the queue length and queue occupancy at time t. The ISL capacity is C, the average packet size is d_{avg} and the buffer size is B. The congestion index $\xi(t)$ of satellite node is defined as below:

$$\xi(t) = \frac{(Q - q(t)) \cdot d_{\text{avg}} + C \cdot \Delta t - \hat{P}}{B} \tag{36}$$

$\xi(t)$ shows the future status of satellite node and the revised factor γ is derived from $\xi(t)$

$$\gamma = \frac{1}{1 + \exp(-a\xi)} \tag{37}$$

where, a is the slope parameter.

Then, the final reinforcement factor r is calculated as below:

$$r = r \times \gamma \tag{38}$$

(4) Pheromone table update

After revision of reinforcement factor r, the pheromone table T^k is updated as below:

$$t_{fd} \leftarrow t_{fd} + r(1 - t_{fd}) \tag{39}$$

t_{fd} represents the probability if choosing neighbor node f when destination node is d. The probability t_{nd} receives an implicitly negative feedback because all the probabilities sum up to 1 for the same destination d:

$$t_{nd} \leftarrow t_{nd} - rt_{nd} \quad \forall n \in N_k, n \neq f \tag{40}$$

(5) Data-routing table update

The data-routing table R^k is updated after every update in the pheromone table. Data packets are forwarded according to random policy whose parameters which are dependent on the elements of the data-routing table. The elements of R^k are the exponential transformation of elements of pheromone table. Compared with the path of low probability, the path with higher probability is inclined to be selected.

$$R_{nd}^{k} = (t_{nd})^{\varepsilon} \tag{41}$$

$$R_{nd}^{k} = \frac{R_{nd}^{k}}{\sum R_{id}^{k}} \quad i \in N_k \tag{42}$$

ε is set to 1.4 after many tests because it not only effectively avoids 'bad' path, but also guarantees multi-path transmission.

(6) Other discussions

When forward MA behavior is discussed, it is obvious that vast MAs are able to enhance performance of routing algorithm, but decrease its efficiency. Therefore, MA is generated with medium and constant speed In TPARA.

In addition, link switch is less discussed in previous LEO satellite routing algorithms [24]. There are two cases considered in this paper: link break and link recovery. When ISL connecting two satellites breaks, these two satellites will change original routing that one satellite can not take the other one satellite as the destination node or the next-hop node. Likewise, their neighbor satellites will also change their routing. Thus, two satellites update local routing table and then broadcast failure notification agent (FNA). As FNA records a list of destination node on whose path the link is broken, all neighbor nodes that receive FNA update their routing tables. Further, if one of above neighbor nodes also changes original routing due to link break, the node sends FNA to its neighbor nodes until all nodes of satellite network know and adapt to link break. When ISL recovers, two separate satellites can be connected by their ISL. These two satellites generate recovery notification agent (RNA) to explore the recovered link. Specifically, each of two satellites updates its routing table according to the information detected by RNA and then broadcasts the update to its neighbor nodes until all nodes of satellite network know and adapt to link recovery.

6. Simulation Results

In this paper, Iridium satellite constellation is introduced to evaluate the performance of TPARA. As illustrated before, the traffic density on the surface of the Earth is considered in TPARA. The potential traffic density is plotted in Figure 3 in which the surface of the Earth is divided in to 288 grids by longitude 15° and latitude 15°. Every grid stands for relative traffic density ranging from 1 to 10 according to economic development.

Other simulation parameters of TPARA are set as follows: $\alpha = 0.3$, $\eta = 0.2$, $c_1 = 0.7$, $c_2 = 0.3$ and $a = 1$. Ant Colony Optimization (ACO) in [25] and TPARA are compared through simulation.

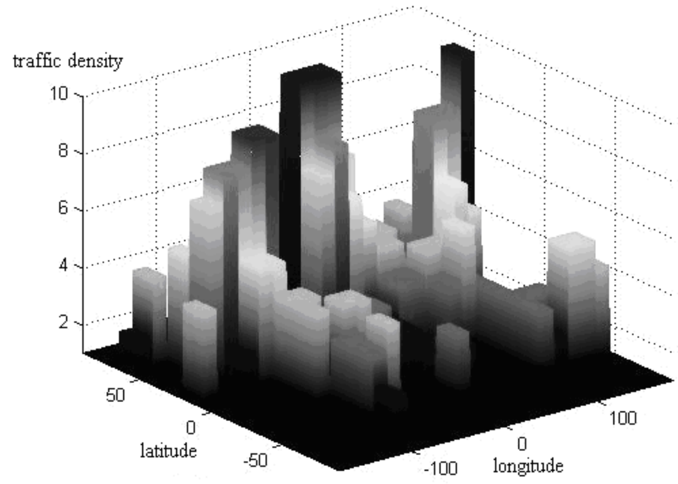


Figure 3. Potential Traffic Density on the Surface of the Earth

Figure 4 and Figure 5 show prediction performances of TPARA (*i.e.* NOEKF prediction algorithm). It is apparent that NOEKF algorithm outperforms CKF (classical Kalman filtering) and FARIMA algorithm, which is a well-performed prediction algorithm [26]. Though CKF algorithm has large error with real traffic, the trend of change is correct. As the premise of FARIMA model is that the observed time series must be a generalized stationary process, but in most cases, the real traffic does not accord with the condition. It can be seen from Figure 5 that when real traffic changes drastically, FARIMA model cannot track the sharp change of traffic. In contrast, since NOEKF algorithm makes accurate estimation of time-variant noise, it can exactly reflect the change trend of the real traffic in time, so it is able to accurately predict traffic.

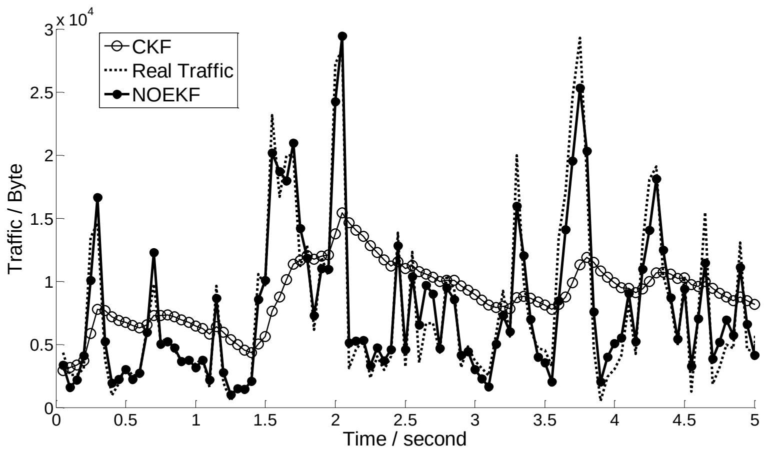


Figure 4. Traffic Prediction of NOEKF and CKF Algorithms

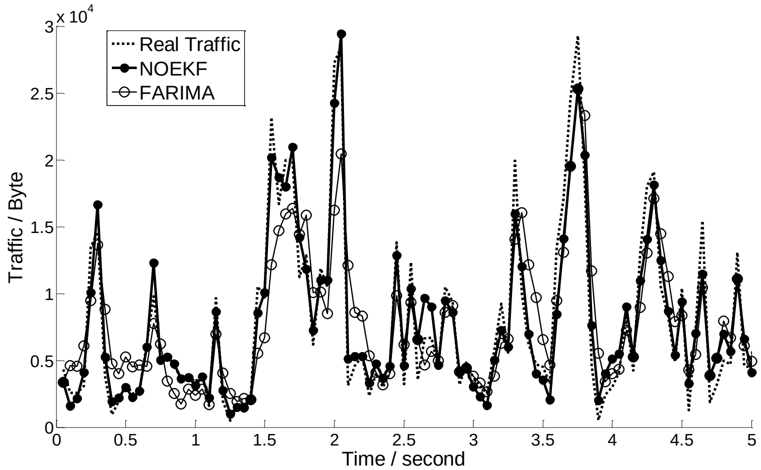


Figure 5. Traffic Prediction of NOEKF and FARIMA Algorithms

Figure 6 presents the link congestion probability of satellite node in ACO and TPARA. As potential traffic density is different on the surface of the Earth, when satellite moves to heavy traffic region, the congestion probability of satellite links increases due to a huge number of data packets from ground. Compared with ACO, traffic prediction which is an important component of TPARA revises the routing strategy and effectively avoids packets being sent to the impending congested satellite nodes.

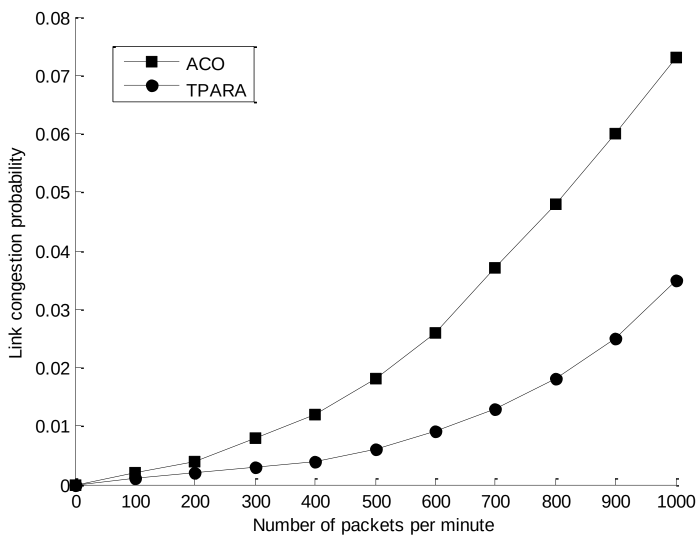


Figure 6. Link Congestion Probability of Satellite Node

Figure 7 depicts the average packet delay for ACO and TPARA. When traffic load sharply changes at 20 s, packet delay of ACO drastically increases to jam satellite node. In contrast, TPARA has gentle increase of packet delay mainly because it well deals with bursty traffic through predicting traffic and sending packets to neighbor nodes. Therefore, TPARA has better robustness than ACO.

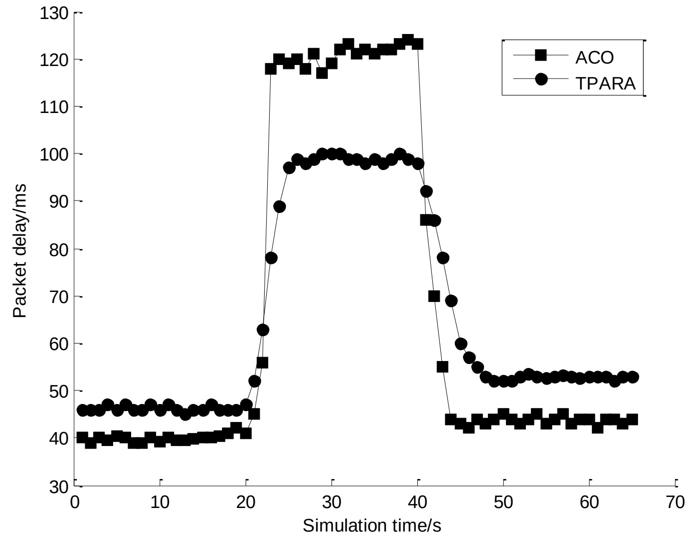


Figure 7. Average Packet Delay for ACO and TPARA

From above results, it can be seen that taking both traffic density on ground and traffic prediction into account, TPARA shows better load balance because it not passively dealing to bursty traffic but actively predicting traffic to avoid node congestion.

7. Conclusions

In this paper, an Agent-based distributed Routing Algorithm with Traffic Prediction TPARA is proposed. TPARA includes two modules of traffic prediction and routing decision. The former predicts traffic from the surface of the Earth and then revises routing decision. The latter uses mobile agent to collect the information about network status. As a result, routing decision for satellite node is not only dependent on current status, but related to the future status. Because of both traffic prediction and load balance, TPARA improves robustness of satellite network. Simulation results also demonstrate that TPARA outperforms typical ACO algorithm. As the algorithm is distributed and adaptive, it is easily deployed in the future satellite network.

Acknowledgements

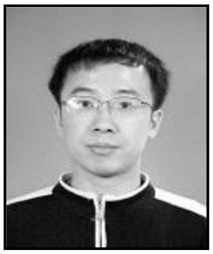
This paper is supported by the Fundamental Research Funds for the Central Universities and Major Program of National Natural Science Foundation of China under grant No. 60532030. The Institute of Telecommunication Satellite of China Academy of Space Technology (CAST) of China and Communication Research Center of Harbin Institute of Technology give constructive suggestions to the algorithm designs and simulations.

References

[1] V. V. Darnopykh, "Application of Unified Methodical Approach to Online Planning of Target Operation of Satellite Monitoring and Communication Systems", Journal of Computer and Systems Sciences International, vol. 49, no. 1, **(2006)**.
[2] P. Thompson, B. Evans, L. Castenet, M. Bousquet and T. Mathiopoulos, "Concepts and Technologies for a Terabit/s Satellite - Supporting Future Broadband Services via Satellite", Proceedings of the 3-rd International Conference on Advances in Satellite and Space Communications (SPACOMM), Budapest, Hungary, **(2011)** April 17-22.
[3] E. Ekici, I. F. Akyildiz and M. D. Bender, "A Distributed Routing Algorithm for Datagram Traffic in LEO Satellite Networks", IEEE/ACM Transactions on Networking, vol. 9, no.2, **(2001)**.
[4] J. Wang, J. Yan and Z. Cao, "Optimization of Sequent Snapshots Routing Algorithm in LEO Satellite Networks", Journal of Astronautics, vol. 30, no. 5, **(2009)**.
[5] H. Uzunalioglu, I. F. Akyildiz and M. D. Bender, "A Routing Algorithm for Connection-Oriented Low Earth Orbit (LEO) Satellite Networks with Dynamic Connectivity", Wireless Networks, vol. 6, no. 3, **(2000)**.
[6] M. Werner, "A Dynamic Routing Concept for ATM-Based Satellite Personal Communication Network", IEEE Journal on Selected Areas in Communications, vol. 15, no. 8, **(1997)**.
[7] H. S. Chang, B. W. Kim, C. G. Lee, S. L. Min, Y. Choi, H. S. Yang, D. N. Kim and C. S. Kim, "FSA-Based Link Assignment and Routing in Low-Earth Orbit Satellite Networks", IEEE Transactions on Vehicular Technology, vol. 47, no. 3, **(1998)**.
[8] E. Ekici, I. F. Akyildiz and M. D. Bender, "A Distributed Routing Algorithm for Datagram Traffic in LEO Satellite Networks", IEEE/ACM Transactions on Networking, vol. 9, no. 2, **(2001)**.
[9] E. Ekici, I. F. Akyildiz and M. D. Bender, "Datagram Routing Algorithm for LEO Satellite Networks", Proceedings of IEEE INFOCOM' 00, 19th Annual Joint Conference of the IEEE Computer and Communications Societies, Tel Aviv, Israel, **(2000)** March 26-30.
[10] D. S. Mauro, E. Cianca and M. Ruggieri, "Improved Algorithms for Internet Routing in Low Earth Orbit Satellite Networks", Space Communications, vol. 20, **(2005)**, pp. 3-4.
[11] Y. S. Kim and W. J. Park, "Adaptive Routing in LEO Satellite Networks", Proceedings of IEEE 51st Vehicular Technology Conference (VTC 2000-Spring), Tokyo, Japan, **(2000)** May 15-18.
[12] T. Taleb, Z. M. Fadlullah, T. Takahashi, R. Wang, Y. Nemoto and N. Kato, "Tailoring ELB for Multi-Layered Satellite Networks", Proceedings of IEEE International Conference on Communications (ICC' 09), Dresden, Germany, **(2009)** June 14-18.
[13] A. Jankowska, M. C. Schut and N. Ferreira-Schut, "A Wireless Actuator-Sensor Neural Network for Evacuation Routing", Proceedings of 3rd International Conference on Sensor Technologies and Applications, Athens/Glyfada, Greece, **(2009)** June 18-23.
[14] S.-Z. Huang and L.-y. Huang, "The Optimal Routing Algorithm of Communication Networks Based on Neural Network", Proceedings of International Conference on Intelligent Computation Technology and Automation, Changsha, China, **(2010)** May 11-12.
[15] P. P. Repoussisa, C. D. Tarantilisa, O. Bräysyb and G. Ioannoua, "A Hybrid Evolution Strategy for the Open Vehicle Routing Problem", Computers & Operations Research, vol. 37, no. 3, **(2010)**.
[16] X. -D. Zhang, L. -H. Ding and Y. Rao, "QoS Routing by Genetic Algorithm for LEO Satellite Networks", Proceedings of 2nd International Symposium on Computational Intelligence and Design, Changsha, Hunan, China, **(2009)** December 12- 14.
[17] B. -B. Jiang, H. -M. Chen, L. -N. Ma and L. Deng, "Time-Dependent Pheromones and Electric-Field Model: a New ACO Algorithm for Dynamic Traffic Routing", International Journal of Modelling, Identification and Control, vol. 12, **(2011)**, pp. 1-2.
[18] X.-M. Wang and X.-M. An, "An Energy and Location Aware ACO Based Routing Algorithm for Wireless Sensor Networks", Acta Electronica Sinika, vol. 38, no. 8, **(2010)**.
[19] S. Masrom, Z. Z. A. Siti, A. M. Nasir and A. S. A. Rahman, "Hybrid Particle Swarm Optimization for Vehicle Routing Problem with Time Windows", Proceedings of 13th WSEAS International Conference on Mathematical Methods, Computational Techniques and Intelligent Systems, Wisconsin, USA, **(2011)**.
[20] H. Liu, Q.-S. Zhang and W.-P. Wang, "The Vehicle Routing Problem of Reverse Logistics with Grey Recycling Demand Based on Hybrid PSO", Proceedings of IEEE 17th International Conference on Industrial Engineering and Engineering Management (IE&EM), Xiamen, China, **(2010)** October 29-31.
[21] Z. -H. Gao, Q. Guo and Z. -Y. Na, "A Distributed Multipath Routing Strategy for LEO Satellite Networks", Tamkang Journal of Science and Engineering, vol. 14, no. 2, **(2011)**.
[22] Z. -H. Gao, Q. Guo and Z. -Y. Na, "A Novel Optimized Routing Algorithm for LEO Satellite IP Networks", Journal of Systems Engineering and Electronics, vol. 22, no. 6, **(2011)**.
[23] Z. -Y. Na, Q. Guo, Z. -H. Gao, J. -Q. Zhen and C. -Y. Wang, "A Novel Adaptive Traffic Prediction AQM Algorithm", Telecommunication Systems, vol. 49, no. 1, **(2012)**.

[24] Z. -Y. Na, Z. -H. Gao, Y. Cui, L. -M. Chen and Q. Guo, "QoS Routing Algorithm Applying Multi-Agent System for LEO Satellite Network", International Journal of Future Generation Communication and Networking, 6, in Press, **(2013)**.
[25] H. -B. Xia, W. -B. Xu and Y. Liu. "New Ant Colony Optimization for Load-Balancing and Routing Mechanism", Application Research of Computers, vol. 26, no. 7, **(2009)**.
[26] M. Jiang, C. -M. Wu, M. Zhang and D. -M. Hu, "Research on the Comparison of Time Series Models for Network Traffic Prediction", Acta Electronica Sinica, vol. 37, no. 11, **(2009)**.

Authors



Zhenyu Na, was born in 1981, Harbin, P.R.China. He received the B.S. degree in communication engineering, M.S. degree in information and communication engineering from Harbin Institute of Technology (HIT) in 2004 and 2007, respectively. In 2010, he received doctoral degree in information and communication engineering at Communication Research Center of HIT.

Now, he is a lecturer at School of Information Science and Technology of Dalian Maritime University (DLMU) with research interests including satellite communications and networks, wireless networks, MIMO-OFDM communications. Until now, he has published more than 10 papers in international journals and conferences.

Zihe Gao, was born in 1983, Yingkou, P.R.China. He received B.S. degree in communication engineering from HIT in 2005, and received M.S. and PH.D. degrees in information and communication engineering at CRC from HIT in 2007 and 2011, respectively.

Now, he is a senior engineer of the Institute of Telecommunication Satellite of China Academy of Space Technology. His main interests cover satellite communication system design, satellite link modeling, intelligent routing and QoS routing algorithm. Up to now, he has published 8 journal papers.

Yang Cui, was born in 1982, Harbin, P.R.China. He received the B.S. degree in communication engineering, M.S. degree in information and communication engineering from Harbin Institute of Technology (HIT) in 2005 and 2007, respectively. In 2012, he received doctoral degree in information and communication engineering at Communication Research Center of HIT.

Now, he is an assistant professor at Institute of Science and Industry Technology of HIT. His research interests cover wireless communications and networks, QoS, radio resource management for heterogeneous wireless networks. Until now, he has published more than 10 papers in international journals and conferences.

Qing Guo, was born in 1964, Harbin, P.R.China. He received B.S. Degree in radio engineering from Beijing University of Posts and Telecommunications (BUPT), M.S. and Ph.D. Degrees from HIT in 1985, 1990 and 1998, respectively. Currently, he is the professor, deputy director and Ph.D. supervisor of Communication Research Center at HIT.

His research interests are in the area of broadband satellite communication, multimedia communication, wireless transmission protocol, QoS and broadband access technology. He has been a reviewer for several famous journals and international conferences including SCIENCE IN CHINA Series E and F, Chinese Journal of Electronics, Chinese Journal of Aeronautics, ICC, *etc*.

Until now, Prof. Guo has won the 2nd Prize of National Science and Technology Progress of China and 4 science and technology awards of provincial and ministerial level, published more than 10 patents and 60 research papers in domestic and international journals and international conferences.