

Research Article

Dynamic Routing Strategy for Directed Transmissions of High-Valued Contents in NGSO Satellite-Based Internet

Zeyu Hu,¹ Wei Li,² Tianjia Wang,¹ and Yong Li ¹

¹Key Laboratory of Universal Wireless Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

²The State Radio Monitoring Center, Beijing 100037, China

Correspondence should be addressed to Yong Li; liyong@bupt.edu.cn

Received 9 March 2022; Accepted 3 May 2022; Published 10 June 2022

Academic Editor: Huaming Wu

Copyright © 2022 Zeyu Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nongeostationary orbit (NGSO) satellite-based Internet is an important part of the future network. One single NGSO satellite has the characteristics of high-speed movement, small coverage area, and short connection time; therefore, the method of multisatellite collaboration is adopted in NGSO satellite-based Internet. Caching is an effective technique to reduce the transmission load and delay, especially for high-value content. However, the traditional caching cannot be applied to the NGSO satellite network directly due to the frequent switching of service satellites. As a result, some high-valued cached data need to be transmitted among satellites, such that an effective routing between the source and destination node needs to be established orderly in the satellite network. This paper proposes a dynamic routing strategy based on multilayer satellite architecture for unicast and multicast scenarios, respectively. For the unicast, a routing strategy based on Q-learning is proposed to optimize bandwidth and delay simultaneously, and the effectiveness is verified by the comparison against the traditional routing strategy. For the multicast, an advanced Q-learning multicast routing algorithm based on spatial directionality is proposed to optimize the overall transmission success rate and resource cost, and the superiority of the proposed routing strategy is verified by simulation results.

1. Introduction

Nongeostationary orbit (NGSO) satellite-based Internet is regarded as an important part of the sixth generation (6G) mobile networks due to its extended capability in the coverage area [1]. Satellites can be divided into several categories according to the orbit altitudes. Among all types of satellites, low earth orbit (LEO) satellites have attracted more attention in both the academia and the industry because of their advantages in link delay and pathloss [2]. Since Release14, the 3rd generation partnership project (3GPP) has been interested in the integrated communication between satellite and terrestrial systems. At present, they are promoting the integrating of NGSO satellite network into 5G [3].

The LEO satellite has the characteristics of small coverage and frequent switching for the ground users. Therefore, multiple LEO satellites are needed to form a constellation to

satisfy communication requests by intersatellite cooperation [4]. The intersatellite link is designed to realize the interaction of signaling and data transfer between satellites, which makes it possible for the resources (such as caching resources and computing resources) to be jointly allocated in the NGSO satellite network [5]. On the other hand, some requests are repeated multiple times in the communication network, which can be considered high-value content [6]. Therefore, caching high-valued contents is an effective technique to reduce duplicate transmissions and network overheads in the communication system. Na et al. designed a caching method using the buffer in the satellite, and the performance of the NGSO satellite network system was improved [7]. Furthermore, Liu et al. proposed a novel caching algorithm by optimizing content placement in the NGSO satellite network, and the effectiveness of caching in the LEO satellite was further demonstrated in [2]. In recent years, some creative work has been proposed. Li et al. proposed a

creative model called Temporal Netgrid Model to portray the time-varying topology of large-scale NGSO satellite networks, and the routing strategy based on this method is instructive [8]. Furthermore, some work on the efficient content management and the integrated satellite/terrestrial cooperative transmission scheme has been improved in subsequent studies [9, 10].

The LEO satellite is in relative high motion with respect to the ground. Limited by the requirement of the minimum elevation of the link between each LEO satellite and the user on the ground can only be maintained for a period of time. Consequently, multiple satellites are required to relay the connection when an user needs to be served by the satellite network for a long time, but it is unrealistic to cache the high-valued contents in all satellites. Therefore, the high-valued contents need to be transferred in the satellite network directionally. It is possible for multiple LEO satellites to request the same high-value content at the same time, because each LEO satellite covers a different area. In other words, the route between one source node and one or more destination nodes needs to be established for directional transmissions of the high-value cached data. The directional transmission means that the LEO satellite serving the ground users transmits the data to the next LEO satellite in an successive route. This is a valuable routing problem in NGSO satellite-based Internet. To solve this problem, Burleigh proposed a method of contact graph routing in 2008 [11]; the central idea is to calculate one routing in advance according to the fixed orbit of satellites, so as to realize the transfer of data orderly in the satellite network. This method has been further studied in the literature [12], in which Wang et al. proved the feasibility of this routing strategy, and showed its advantages in delay and cache consumption. However, this method only calculates the orbit data and ignores the situations of the cache space and link status in the satellite network.

This paper proposes a routing strategy for directional transmission of high-valued cached data by intersatellite links, and the routing strategy is effective and suitable for the network of satellites in space. The main contributions of our work are summarized as follows:

- (i) An optimization problem of routing strategy considering both delay and bandwidth is proposed in NGSO satellite system, and it is suitable for the mobility of NGSO satellite
- (ii) A unicast routing strategy based on reinforcement learning is effectively applied to the solution with only one receiving node. In addition, the effectiveness of proposed unicast algorithm is proved in comparison with various traditional algorithms
- (iii) A multicast routing strategy combining spatial location and transmission efficiency is proposed, which effectively solves the problem that reinforcement learning is difficult to converge in multicast routing strategy
- (iv) The effectiveness of the proposed dynamic routing strategies is simulated in NGSO system, and the results show that the proposed unicast and multi-

cast routing strategies are better than the traditional ones in terms of end-to-end delay and bandwidth consumption in various random scenarios

The rest of this paper is organized as follows. In Section 2, the overall system model is introduced, and the multilayer satellite network is presented which uses the clustering method to limit the dynamics of the LEO satellite. In Section 3, the mathematical model of the unicast routing problem in the NGSO satellite network is formulated, and the routing strategy suitable for this model is proved by comparing with other heuristic algorithms. In Section 4, this paper compares the success rate between the proposed unicast routing strategy and the traditional contact graph routing, and the effectiveness of the proposed strategy is verified by system-level simulator. In Section 5, the problem of multicast routing applicable to space scenarios is represented, and a routing strategy based on directionality is proposed. Moreover, the effectiveness of proposed algorithm is verified in Section 6. Finally, we conclude our work in Section 7.

2. System Model

The orbit altitude determines the function in the NGSO satellite system. In this case, the LEO satellite has the lower orbit which has greater advantages in transmission delay for the ground equipment, and it is suitable to be used as an executor in the network; the medium earth orbit (MEO) satellite has the higher orbital which has a larger coverage and the more information, and it is suitable to be the manager in the network. The MEO and LEO can be converged into a multilayer satellite architecture to achieve effective complementarity of these characteristics. Based on the multilayer satellite architecture [13], this paper mainly studies the routing problem in the NGSO satellite network. In this architecture, the requests, link status, storage space, and other related information are transmitted to the cluster head using interorbital links. Then, the cluster head makes routing strategy based on these information and sends them to each LEO satellite participating in the routing. Intersatellite link is used for information and data exchange between LEO satellites, which is an important research direction in NGSO satellite system. Meanwhile, some common NGSO satellite systems, such as LeoSat, use lasers as intersatellite links, which ensures the extremely high throughput. For this propose, the MEO satellites are regarded as the control center nodes using interorbital link, and the LEO satellites are regarded as the edge caching and data transmission nodes, as shown in Figure 1.

A number of LEO satellites are within the coverage area of the same MEO satellite for certain period of time. Then, these LEO satellites and the MEO satellite form a cluster by the interorbital link, and the MEO is the cluster-head node of this cluster. The cluster-head node caches the status information of each satellite in this cluster, such as current topology and available bandwidth. Due to the relative motion between the MEO and the LEO, the relative position in the cluster is approximately considered to be fixed only in a short period of time. As a result, time needs to be divided

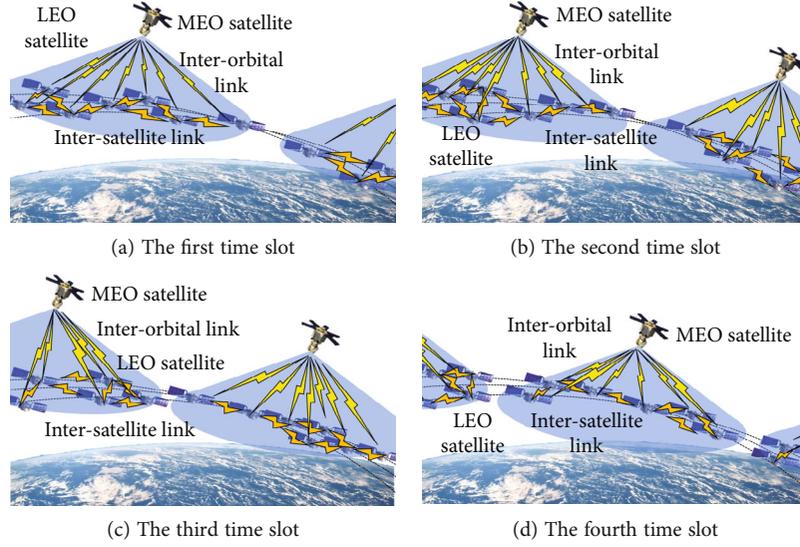


FIGURE 1: There is relative motion between MEO satellite and LEO satellite, and the clusters with MEO satellite as the cluster head are different at various time slots. In each independent time slot, the MEO satellite controls LEO satellites to make an efficient routing strategy.

into smaller time slots due to the change of relative positions between these satellites. The time slot is denoted as T . At the beginning of each T , the LEO satellites in this cluster report their status information to the cluster head, and the cluster-head node establishes the topology structure according to the information. It is assumed that the topology of this cluster does not change until the end of T , and the topology needs to be rebuilt in another different time slot as shown in Figure 1.

The model for this paper is summarized as follows. In a certain time slot T , a number of LEO satellites are covered by one MEO satellite to form a cluster by intersatellite link, and at least one LEO satellite in this cluster needs to send high-valued contents to one or more other LEO satellites for caching. The core of this problem is how to make a reasonable dynamic routing strategy in the cluster head, which uses the intersatellite link for transmission of the high-valued content. The cases of one requesting LEO satellite and multiple requesting LEO satellites are divided into unicast and multicast to discuss separately due to the difference between the two problem models.

3. Problem Formulation for Unicast Routing

In this section, the mathematical model of unicast routing strategy is proposed for one time slot T .

3.1. Dynamic Unicast Routing in the NGSO Satellite Network. Assuming that at the beginning of the time slot T , the MEO satellite receives a routing request from the LEO satellite, where the packet size of the data requested to be transmitted is set to F . The channel capacity required to transmit data can be expressed as $C_{\text{request}} = F/T - t_{\text{delay}}$, where t_{delay} denotes the delay, which is composed of four parts: the delay generated by information of LEO satellites reporting to the MEO satellite and information initialization ($t_{\text{delay-init}}$), the delay generated by uploading of traffic request

($t_{\text{delay-req}}$), the delay generated by routing strategy generation ($t_{\text{delay-route}}$), and distribution ($t_{\text{delay-dis}}$). All LEO satellites in the cluster can be directly connected to the MEO satellite by the interorbital link as illustrated in Figure 1.

As shown in Figure 2, a complete transmission route is limited by both the total delay and the total bandwidth of the link. However, the two resources are limited in different ways, due to the fact that the link delay is limited by the sum of the delay of all links and the link bandwidth is limited by the smallest one among all links. The conditions of an eligible routing in this cluster are as follows:

$$C_{\text{request}} = \frac{F}{T - t_{\text{delay}} - \sum_{l=1}^L D_l}, \quad (1)$$

$$C_{\text{provide}} = B_l \log \left(1 + \frac{PH}{N_0 B_l} \right), \quad (2)$$

$$C_{\text{request}} \leq \min(C_{\text{provide}}), \quad (3)$$

where B_l represents the bandwidth of the l^{th} hop link of the route. H is the channel gain, P is the power, N_0 is the background Gaussian noise, and they are all treated as constants here, because the time slot is short.

$$T - t_{\text{delay}} = \sum_{l=1}^L D_l + D_{\text{idle}}, \quad (4)$$

$$D_{\text{idle}} \geq 0, \quad (5)$$

where D_l represents the transmission delay of the l^{th} hop link of the route, and L is the total number of hops in the routing strategy. D_{idle} is a positive number, which denotes the time duration when the link is idle.

For the routing strategy, not only the requirement of available bandwidth needs to be satisfied, but also the

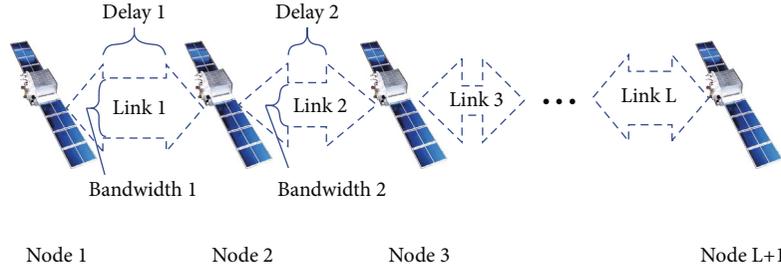


FIGURE 2: Both delay and bandwidth need to be considered in the formulation of unicast routing strategy. The selection of different intersatellite links is the key of this optimization problem.

usability needs to be guaranteed. The condition of reasonable in routing strategy is that:

$$\sum_{i=1}^N R_{i,j} \leq 1, \forall i, j = 1, \dots, N, \quad (6)$$

$$\sum_{i=1}^N R_{i,i} = 0, \forall i = 1, \dots, N, \quad (7)$$

$$\sum_{i=1}^N R_{i,j} = \begin{cases} 1, & R_{i,j} \in \text{RoutingStrategy} \\ 0, & \text{otherwise} \end{cases}, \quad (8)$$

where the number of LEO satellites in the cluster is N , and R is a matrix whose size is $N \times N$. i and j represent the index of the LEO satellite node. Assuming $R_{i,j} = 1$, the data transferred is from i_{th} node to j_{th} node in the routing strategy. The constraint condition is that each node can only have one input and one output at most, which is to ensure that there is no loop in the routing strategy.

When there is more than one routing strategy to satisfy the traffic requirement in the cluster, it is necessary to compare the optimization results of the routing strategies. Eqs. (1)–(5) ensure that the link set on the route can provide sufficient bandwidth resources, while Eqs. (6)–(8) ensure the feasibility of the routing strategy and eliminate loop routes. The rationality and connectivity are guaranteed in the routing strategy, and it is optimized by adjusting the delay and the remaining bandwidth of the link. Then, the routing problem can be reduced to the following optimization problem:

$$\underset{D}{\text{minimize}} \sum_{j=1}^N \sum_{i=1}^N (R_{i,j} D_{\text{route}}), \quad (9)$$

$$\underset{B}{\text{minimize}} \sum_{j=1}^N \sum_{i=1}^N (R_{i,j} B_{i,j} - R_{i,j} B_{\text{route}}). \quad (10)$$

Eq. (9) is used to optimize transmission time, and unnecessary hops will be eliminated in the routing strategy. Eq. (10) is used to optimize the remaining bandwidth to avoid network congestion.

The combination of (9) and (10) is a multiobjective optimization problem, which can be normalized by parameter γ :

$$\underset{M}{\text{minimize}} \sum_{j=1}^N \sum_{i=1}^N (R_{i,j} M_{i,j}), \quad (11)$$

$$M_{i,j} = \gamma D_{\text{route}} - (1 - \gamma)(B_{i,j} - B_{\text{route}}).$$

For optimization function (11), the constraints should include Eqs. (1)–(8). The variable to be optimized is the multiobjective optimization $M_{i,j}$ of the routing strategy, and the constraints are applied to ensure the sufficient bandwidth, reasonable delay, and reasonable routing by (1)–(8), respectively. From the conditions of the multiple optimization objectives in Eq. (11) and the complex constraints in (Eqs. (1)–(8)), this problem involves integer programming, and the solution space is relatively large; therefore, it is a typical NP-hard dynamic programming problem. In this case, the heuristic algorithm will be employed.

3.2. Unicast Routing Strategy Using Q-Learning. Based on the dynamic transition characteristics of the states for this problem, the machine learning algorithm is expected to solve the routing problem with its excellent strategy optimization ability in dynamic decision problems. For this problem, the complete environmental information is unknown. Therefore, the reinforcement learning algorithm can obtain the routing strategy by feedback reward from the environment interaction. The temporal-difference learning algorithm is applied to perform optimization, owing to its advantage of high learning efficiency and fast convergence [14].

In this problem, the formulation of R -table is the most worthy of concern. For every two nodes i and j , $R_{i,j}$ has two states: connected and disconnected, which can be selected. In this case, 0 and 1 are chosen to represent the two states to simplify, and the R -table is limited to $R_{i,j} = 0, 1$. The selection of $R_{i,j}$ is only related to the current node index i and the next node index j , and it can be approximately regarded as a Markov decision process (MDP). Admittedly, there are some hidden constraints in the modeling. First of all, R -table itself has certain constraints, that is, link relationship. Namely, the $R_{i,j}$ can only be set as disconnected forcibly if there is no one-hop-link to connected the satellite i and j directly. However, this does not affect the

Markov property of the state transition. Secondly, the accumulation of transmission delay may cause the problem that Eq. (5) constraint is not satisfied, but in fact, this restriction condition is loose due to the sufficient bandwidth and huge throughput of the intersatellite link. And the delay as the optimization objective, the optimization process would reduce the delay gradually when the initial routing strategy have been satisfied. Therefore, the cumulative delay of multihop links does not affect the property much. Finally, in the rate limit of Eq. (3), the rate limit of Eq. (3) seems to be obtained only after all multihop routes are obtained, but in fact, it only needs to be determined during each state transition because C_{request} is fixed. In this way, it is equivalent to adding a condition similar to link relationship that will set as disconnected forcibly, which does not affect the Markov property. In this case, the optimization model is casted to a MDP, and the four variables need to be determined, namely, states set, actions set, probability set, and reward set. All LEO satellites in the cluster form a set for the State. The data transition between two connected LEO satellites is regarded as an action function, and all the available actions form the action set Action. The transition probability matrix is positively correlated with the value depends on the probability table, which is optimized by the Reward. The reward function Reward is positively correlated with the selected routing with Eq. (11).

Q-learning is an effective heuristic algorithm used for routing problems, and the idea of temporal-difference is introduced into the algorithm. The State-Action table has been introduced as the core of Q-learning, and the selecting of neighbor is the transfer of state. The ϵ -greedy algorithm is introduced, which is helpful to speed up the convergence of Q-table. Meanwhile, the counter mechanism is introduced in order to avoid the problem of local optimum in algorithm convergence. In addition, other details are shown in Algorithm 1.

3.3. Algorithm Performance in Unicast Routing. The well-known heuristic algorithms include genetic algorithm, simulated annealing algorithm, tabu search algorithm, and ant colony algorithm. In these algorithms, ant colony algorithm [15] and genetic algorithm [16] are widely used in solving the routing problem.

A simulator is built to simulate the scenario in which the MEO satellite has obtained the information of all links and caching space in the cluster and formulates routing strategies for traffic requests in a time slot T .

In order to compare the difference of convergence effect and speed of several heuristic algorithms, the optimization problem (11) is solved by ant colony algorithm, genetic algorithm, and the proposed reinforcement learning algorithm, respectively. The Monte Carlo method is used to simulate those algorithms. The parameters are given randomly, and the simulation results of 1, 10 and 100 snapshots are given by Figures 3(a)–3(c), respectively. The convergence speed and optimization effect are concerned in the algorithm selection. In this case, the x -axis is set as the iterations times, and the y -axis is set as the optimization effect. However, for the optimization effect, different data requests and other factors

```

1:  $\forall Q(S,A) = \text{Connect Matrix}$ 
2: for  $episode = 1, 2, \dots$  do
3:    $S(1) \leftarrow \text{Begin State.}$ 
4:   for  $l=1,2,\dots$  end do
5:     if  $\text{Random} < \epsilon$  then
6:        $A(l+1) \leftarrow \text{Max}(Q(S(l), A(l))).$ 
7:     else.
8:        $A(l+1) \leftarrow \text{Random}(Q(S(l), A(l))).$ 
9:     end if.
10:     $S^* \leftarrow A^{(l)}S(l).$ 
11:    if  $S^*$  satisfy the constraints for (1), (2) and (3) then
12:       $\text{Reward}(l+1) \leftarrow \text{Constant} - M(l+1)$ 
13:      N.B.:  $\forall M(l) < \text{Constant}$ 
14:    else
15:       $\text{Reward}(l+1) \leftarrow 0$ 
16:    end if
17:     $Q(S, A) \leftarrow Q(S, A) + \alpha[\text{Reward} + \beta$ 
18:       $\text{Max}(Q(S^*, A)) - Q(S, A)]$ 
19:    if  $S^*$  is equal to End State then
20:      break
21:    else
22:       $S(l+1) \leftarrow S^*$ 
23:    end if
24:    if  $Q - \text{Table}$  converge then
25:      Counter = Counter+1
26:      if Counter  $\geq$  Counter_Threshold then
27:        break all
28:      end if
29:    else
30:      Counter clear
31:    end if

```

ALGORITHM 1: Q-learning in unicast routing.

such as topology state have a directly impact on the results. In this case, the delay and bandwidth generated by the firstly routing strategy are set to the maximum M value of the same conditions. Then, the normalization M is optimized gradually with the iteration of the algorithm. The convergence times and final convergence values are shown in Figure 3.

The simulation results indicate that the proposed reinforcement learning algorithm has the smallest occupied bandwidth and delay convergence compared with the ant colony algorithm and the genetic algorithm, because the ability of ϵ -greedy search strategy to get out of the semilocal convergence is applied in the proposed algorithm. In addition, Figure 3 also proves the usability of the algorithm in terms of time complexity, and the algorithm only needs less than 20 iterations to get convergence results in most cases.

4. Performance Evaluation for Unicast Routing

For the multilayer satellite system model proposed in Section 2, the model is used to build the system-level simulator. It is mainly used to simulate the formulation of routing strategy in a time slot T for complex and variable states. For the same network status and traffic requests, two

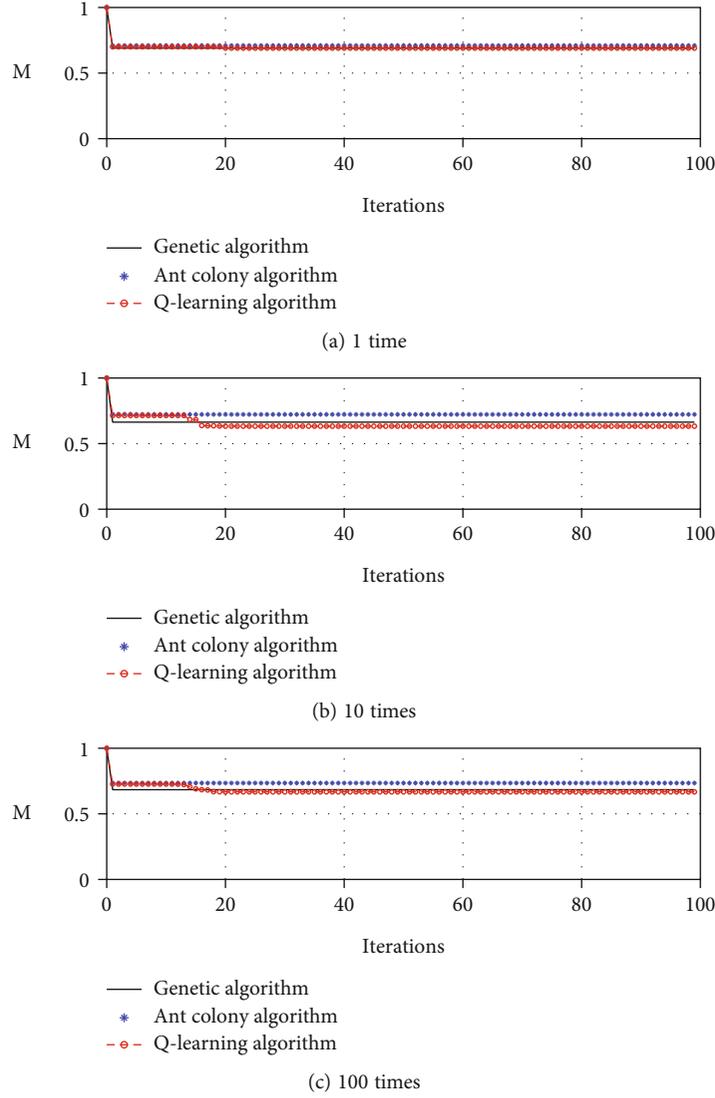


FIGURE 3: Three heuristic algorithms were applied to solve the routing problem, and the comparison results showed that the Q-learning algorithm was better than the genetic algorithm and the ant colony algorithm for the optimization effect.

algorithms, the dynamic routing strategy using Q-learning and the traditional contact graph routing strategy, are applied, respectively. The success rate of the routing strategy is used as the output of the statistical result of the system-level simulator.

In this simulation, a complete cluster is composed of several LEO satellites and one MEO satellite. The number of LEO satellites is determined by the coverage of the MEO satellite, which is explained in Section 2. The network relationship in this cluster conforms to the characteristics of small world network. In this way, all LEO satellites in the cluster are not connected directly by one hop, but there is the possibility of establishing links between adjacent satellites. Besides, the connection relationship that a satellite can maintain at one time slot T is limited, which is due to the limited intersatellite link resources of each LEO satellite. It is assumed that the statistical characteristics, the location, link status, and available buffer space of LEO satellites are randomly distributed in each T . It is worth noting that there

is a minimum gap any two LEO satellites, which is more suitable for the actual security requirements of the NGSO satellite system.

For the routing algorithm in this simulation system, the mathematical model and solution are proposed in Section 3. The current network status has been obtained by the MEO satellite because it is the cluster head, and the routing strategy is formulated dynamically. In other words, this is a dynamic decision-making problem of multiobjective optimization using the reinforcement learning algorithm in the cluster head. The traditional contact graph routing is calculated in advance according to satellite orbit information, but it cannot get the current link and cache space information. In this simulation, due to the randomization of satellite location information, the shortest distance is used to make routing strategy. In this case, the shortest distance means that this route is the shortest of all routes from the source node to the destination node. This method was chosen for improving the success rate of the contact graph routing.

TABLE 1: Simulation parameters in unicast routing.

Symbol	Describe	Value
T	Time slot	100 (s)
$t_{\text{delay-req}}$	The delay of request	10 (ms)
$t_{\text{delay-route}}$	The delay of the routing strategy	0 ~ 1 (ms)
D	Cached data size	1 ~ 25 (MB)
B_l	The bandwidth of the link	0 ~ 1 (Mbps)
Node_Number	Number of LEO satellites	30
Minimum_Gap	Minimum distance between two satellites	1 (km)
Maximum_Gap	Maximum distance between two satellites	1000 (km)
λ	Normalization parameter	0.5
ε	Greedy coefficient	0.3
α	Learning rate	0.7
β	Discount factor	0.8
Counter_Threshold	Counter threshold	10
Constant	Maximum reward value	100

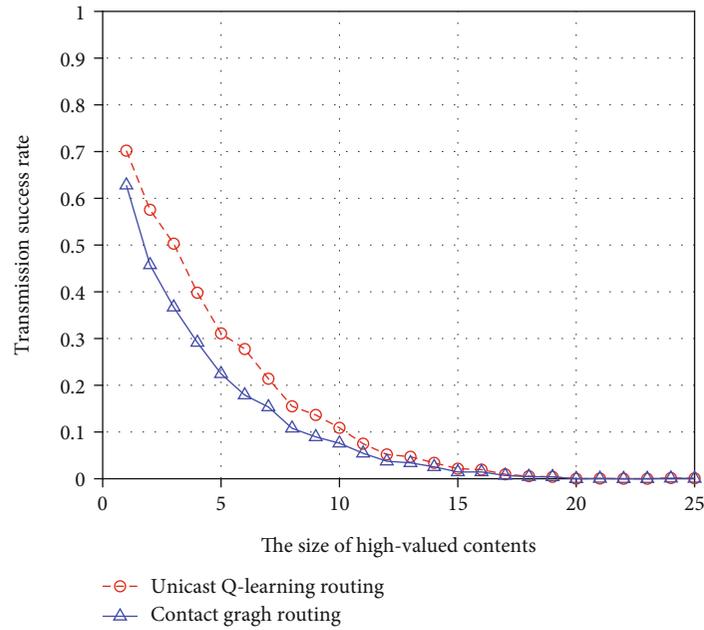


FIGURE 4: The success rate of dynamic routing based on the Q-learning algorithm is higher than the contact graph routing at all of high-value content sizes.

The bandwidth B_l of all links in the system is uniformly distributed on 0 ~ 2 Mbps. The available cache space of each LEO satellites is randomly distributed, which is considered as sufficient or insufficient. Other important simulation parameters are given in Table 1.

In the NGSO satellite network, the capacity of link is an important parameter, which is the key factor to determine the success of directional transmission of high-valued content. However, in order to adapt to the complex and changeable working environment, the link bandwidth is set to be randomly distributed in the system simulation. According

to parameters given in Table 1, the mathematical expectation of all link bandwidth in the cluster is $E(B_l) = 1$ Mbps. Besides, the maximum value $\text{Max}(B_l)$ in all links is not exceed 2 Mbps. The size of high-valued content data D and link capacity B_l can be approximately considered to have a positive proportion relationship, because a larger bandwidth can transmit a larger amount of data in a certain period of time without changing other conditions. Therefore, in this simulation, the adjustable parameter D is used instead of the capacity of all links. The value of D is limited between 1.25 MB and 25 MB, which is based on the maximum

```

1:  $\forall Q(S,A) = \text{Connect Matrix}$ 
2:  $Travel \leftarrow \text{Destination Node}$ 
3:  $Arrive \leftarrow \text{Source Node}$ 
4: for  $episode = 1, 2, \dots$  do
5:    $Begin\ State \leftarrow Travel(1)$ 
6:    $End\ State \leftarrow Arrive$ 
7:   for  $p=1,2,\dots$  end do
8:     if  $Travel$  is empty then
9:       break
10:    end if
11:    for  $l=1,2,\dots$  end do
12:      if  $Random < \epsilon$  then
13:         $A(l+1) \leftarrow \text{Max}(Q(S(l), A(l)))$ 
14:      else
15:         $A(l+1) \leftarrow \text{Random}(Q(S(l), A(l)))$ 
16:      end if
17:       $S^* \leftarrow A^{(l)}S(l)$ 
18:       $Q_p \leftarrow Q(S^*, A)$ 
19:      if  $S^*$  is empty then
20:         $Q_p \leftarrow \text{Reward} \leftarrow \text{Constant}$ 
21:      else
22:        if  $S^*$  is equal to  $End\ State$  then
23:           $Q_p \leftarrow \text{Reward} \leftarrow \text{Constant}$ 
24:        else
25:           $\text{Reward} \leftarrow 0$ 
26:           $Q_t \leftarrow \text{Reward} + \beta \text{Max}(Q(S^*, \text{all } A))$ 
27:           $S(l+1) \leftarrow S^*$ 
28:        end if
29:      end if
30:       $Q(S, A) \leftarrow Q(S, A) + \alpha(Q_t - Q_p)$ 
31:      if  $Q - Table$  converge then
32:         $Counter = Counter + 1$ 
33:        if  $Counter \geq Counter\_Threshold$  then
34:          break
35:        end if
36:      else
37:         $Counter$  clear
38:      end if
39:    end for
40:     $Arrive \leftarrow Travel \cap Route$ 
41:     $Travel \leftarrow Travel \notin Arrive$ 
42:  end for
43:end for

```

ALGORITHM 2: Q-learning in multicast routing.

bandwidth limited. The Monte Carlo method has been applied to eliminate individual differences, and the number of the snapshots is 100. Simulation results are shown in Figure 4.

According to the simulation results, the proposed dynamic routing strategy is superior to the traditional contact graph routing strategy for the transmission success rate of the high-valued content. When the link has a large carrying capacity, that is, the size of cached data is small, and the transmission success rate is relatively high, which can reach about 70% in this simulation. At this time, the failure rate is mainly due to the lack of available links, which is because some LEO satellites do not have enough cache space. In contrast, the routing strategy proposed in this paper has a higher

success rate, because the accurate link and cache space information are reported in this time slot. This routing strategy is more suitable for the complex cluster environment, compared with the contact graph routing only calculated by the orbit information of LEO satellites. With the decline of the capacity of all links, the transmission success rate shows the downward trend in both routing strategies. Interestingly, the success rate of the proposed routing is always no less than the contact graph routing.

5. Problem Formulation for Multicast Routing

In NGSO satellite-based internet, there may be multiple receiving satellites of the high-valued contents. In this section, the mathematical model of multicast routing strategy is proposed for one time slot T .

5.1. Mathematical Model of Routing. The design method of the multicast routing is quite different from the unicast routing. For a unicast routing strategy, there are one source node and one destination node in this network, and only one reasonable path between them is needed to be found. However, there are several destination nodes in the network for the multicast routing strategy, and mutual constraints exist between multiple destination nodes. It is the most challenge and difficulty to make reasonable and efficient multicast routing strategy.

The problem model also needs to be further revised for multicast routing. The effect of delay can be modeled in unicast routing, but in multicast routing, only the transmission rate can be assigned large enough to meet the transmission rate requirements of high-value content. The original problem of Eqs. (1) and (2) is reduced to a bandwidth-only expression:

$$B_{\text{request}} \leq \min(B_{i,j}), \forall i, j \in \text{RoutingStrategy}. \quad (12)$$

The reasonableness condition of the routing strategy is revised to

$$\sum_{i=1}^N R_{i,j} \leq 1, \forall i = 1, \dots, N, \quad (13)$$

$$\sum_{j=1}^N R_{i,j} \leq N, \forall j = 1, \dots, N, \quad (14)$$

$$\sum_{i=1}^N R_{i,i} = 0, \forall j = 1, \dots, N, \quad (15)$$

$$\sum_{i=1}^N R_{i,j} = \begin{cases} 1, & R_{i,j} \in \text{RoutingStrategy} \\ 0, & \text{otherwise} \end{cases}. \quad (16)$$

In this multicast problem, the change of the equations from (6) to (14) indicates that the route has changed from a single broken line to a tree structure. At the same time, the consistency of (6) and (13) limits the occurrence of loops, because all nodes have at most one input neighbor. Moreover, the optimization goal has been changed to consume less bandwidth to meet transmission requirements,

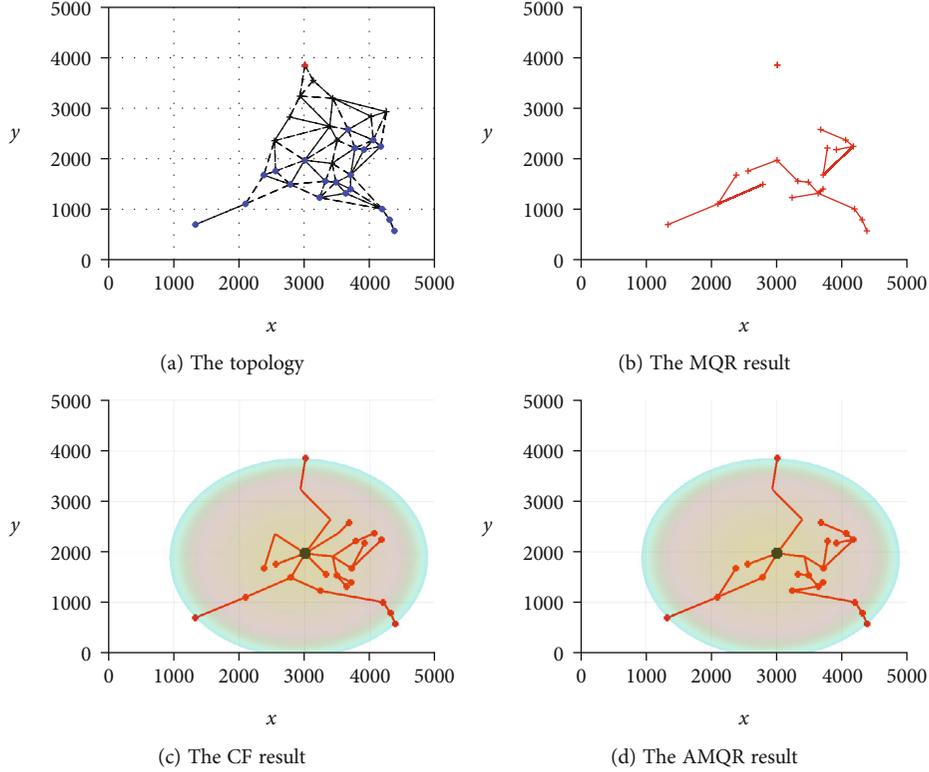


FIGURE 5: Typical LEO connections in a cluster are simulated, and the routing results generated by different algorithms are visually plotted.

$$\underset{B}{\text{maximize}} \sum_{j=1}^N \sum_{i=1}^N (R_{i,j} B_{\text{request}}). \quad (17)$$

For this, a dynamic multicast routing strategy is designed to solve the new optimization problem.

5.2. Intelligent Optimization of Routing. For Eqs. (12), (13), and (17), Algorithm 1 is modified, which is mainly based on the multinode combination and reward condition. The detailed is shown in Algorithm 2.

To observe the performance of the Q-learning in the multicast routing (QMR) algorithm, a typical simulation scenario was constructed. In this scenario, there are 30 nodes randomly located within a square with sides of 5,000 kilometers. The prerequisite for the existence of a link between two nodes is that the distance is less than 1000 km. Both the source and destination nodes in the cluster are randomly determined. The final topology is shown in Figure 5(a). The triangle represents the location of the source node, the large circle represents the location of the destination node, and the dotted line represents the available intersatellite links.

However, the conclusion that the QMR algorithm has poor performance in terms of routing success rate was found, as shown in Figure 5(b), a typical failure case. Due to random exploration, it is difficult for the QMR algorithm to quit after falling into the trap, and it becomes a new trap in the exploration process of other nodes after the exploration times run out. For this problem, the QMR algorithm

needs to be further improved, and a reasonable solution is to preset initial value. Reasonable initial value can help the QMR algorithm avoid some illogical situations and eliminate these potential traps in advance.

5.3. Directionality of Routing. In the traditional multihop network, the study of the multicast algorithm mainly focuses on the generation scheme of multicast tree [17]. Compared with the method of generating multicast structure directly, the method of eliciting multicast route based on one core node has more extensive applicability. However, most of the multicast tree generation algorithms based on the core node are obtained by repeated iteration of the Dijkstra algorithm, and the time complexity is a disadvantage of these methods. For these problems, the applicable working scenarios of the multicast algorithm need to be further analyzed. Firstly, the position and attitude information of LEO can be obtained by MEO. Secondly, the satellite operates in space, and the space between the two satellites is relatively empty. To sum up, the physical location and logical location in this network can be further unified.

For this reason, an algorithm named center forward (CF) is proposed. The core idea of the CF algorithm is to give each node a reasonable forward direction. The first step of CF is to construct a minimum sphere which can cover all the useful nodes, and the position of the node closest to the center of this sphere is determined. Second, all paths in this cluster are assigned a value which is used to determine whether they are near the center, and the value is recorded by F -table. Finally, all destination and source nodes select the maximum

```

1: The center of the minimum sphere enclosing for the
   source node and all the destination nodes is calcu
   lated, and the nearest node  $Center(x_c, y_c, z_c)$  is ob
   tained.
2: for Row = 1, 2, ... end do
3:   for Column = 1, 2, ... end do
4:     if F(Row, Column)  $\in$  Connect Matrix then
5:        $(x_s, y_s, z_s) \leftarrow Start\ Column$ 
6:        $(x_e, y_e, z_e) \leftarrow End\ Row$ 
7:       if Start is equal to Center then
8:          $F \leftarrow -inf$ 
9:       else
10:         $\vec{V}_{hop} = (x_e, y_e, z_e) - (x_s, y_s, z_s)$ 
11:         $\vec{V}_{forward} = (x_c, y_c, z_c) - (x_s, y_s, z_s)$ 
12:         $Projection \leftarrow |\vec{V}_{hop}| \times \vec{V}_{hop} \cdot \vec{V}_{forward} / (|\vec{V}_{hop}| |\vec{V}_{forward}|)$ 
13:        if  $abs(Projection - |\vec{V}_{forward}|) \leq$ 
           Threshold then
14:           $F \leftarrow F_{max}$ 
15:        else
16:           $F \leftarrow -\log(abs(\vec{V}_{forward} - |\vec{V}_{forward}|) / |\vec{V}_{forward}|)$ 
17:        end if
18:        end if
19:      else
20:         $F \leftarrow -inf$ 
21:      end if
22:       $F(Row, Column) \leftarrow F$ 
23:    end for
24:  end for

```

ALGORITHM 3: Center forward.

value in F – table until they connect to the center node. A typical routing for CF is shown in Figure 5(c). The detail is shown in Algorithm 3.

Projection is an important concept in solid geometry, which can directly reflect the effect of a vector in another direction. Here, the goal of each node is to move toward the center. Therefore, the direction of the line from each node to the center node is the optimal direction of travel, and the projection of each hop in this direction needs to be recorded. In addition to direction, distance is also a key factor that needs to be considered. For the CF algorithm, too long or too short is not a favoring choice, and only the one consistent with the required forward value is the most efficient choice. Therefore, logarithm is a reasonable calculation method to be selected. In addition, the upper limit is set to avoid the excessive value in the F – table.

For the topology in Figure 5(a), the route obtained by the CF algorithm is shown in Figure 5(c). The CF algorithm can find a reasonable forward direction for each node within the sphere coverage, which is close to the center node, and the CF algorithm has a higher rate of reasonable routing generation by comparison with the QMR algorithm. However, some redundant and illegal links exist, and the CF algorithm cannot avoid this problem. For this, the multicast routing algorithm needs to be further improved.

5.4. Integrated Routing. The Advanced Multicast Q-learning Routing (AQMR) is an improved routing algorithm, which combines CF and QMR algorithm. Firstly, the CF algorithm generates an F – table. Secondly, the F – table is assigned as the initial value of the Q – table in the QMR algorithm. Finally, the MQR algorithm is used to optimize the multicast routing.

Similarly, the AQMR algorithm is simulated, and the result is shown in Figure 5(d). The AQMR algorithm can find a reasonable multicast routing, which is similar to the results of the CF algorithm. The difference is that it can effectively eliminate some irrational routing.

6. Performance Evaluation for Multicast Routing

Further, the multicast routing in a time period T is simulated. Here, 30 LEO satellites under the coverage of 1 MEO satellite are randomly generated, and simulation parameters are given in Table 2.

The research focuses on the transmission of high quality content over NGSO networks by multicast. In this scenario, the successful transmission rate and transmission cost of high-value content are the key point. Therefore, the number of destination nodes that receive the high-value content successfully and the average bandwidth usage per transmission

TABLE 2: Simulation parameters in multicast routing.

Symbol	Describe	Value
B_l	The bandwidth of the link	0 ~ 10 (Mbps)
Node_Number	Number of LEO satellites	30
Destination-Nodes_Number	Number of destination nodes	10
Minimum_Gap	Minimum distance between two satellites	1 (km)
Maximum_Gap	Maximum distance between two satellites	1000 (km)
ε	Greedy coefficient	0.5
α	Learning rate	0.2
β	Discount factor	0.8
Counter_Threshold	Counter threshold	10
Constant	Maximum reward value	10
F_{\max}	Upper limit of F	5
Threshold	Threshold of difference in F	0.1

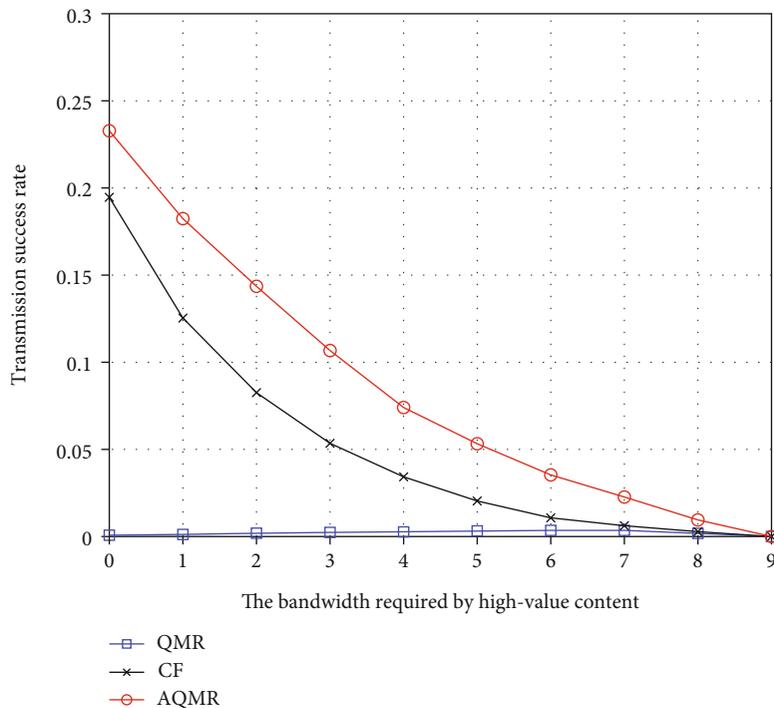


FIGURE 6: As bandwidth requests increase, the success rate of the routing strategy decreases due to the decrease in the number of links available. But at all times, the performs of AQMR is higher than MQR and CF.

routing are the indicators of this simulation. To verify the applicability of the algorithm, these parameters are set as random values which are the location of the LEO satellite, the identity of the source node, the identity of the destination node, and the bandwidth between any two interconnected satellites. The bandwidth required by high-value content is the independent variable, which increases from 1 (MB) to 10 (MB). The statistical results of the successful rate and the average bandwidth occupied by different algorithms are shown in Figures 6 and 7.

The simulation results in Figure 6 are in line with expectations. The QMR is clearly the worst in terms of transmis-

sion success rate due to an initial exploration that was too random, and the AQMR has a higher success rate than the CF because the latter cannot avoid undesirable routing. The simulation results show that the proposed AQMR is more effective than the CF algorithm which only applies the traditional multicast routing idea and the QMR algorithm which only applies Q-learning. Meanwhile, the low success rate of all algorithms is limited by the problem that all nodes in this cluster may not be fully connected.

Figure 7 shows the average occupied bandwidth for each successful routing. In the simulation results, the QMR algorithm occupies the least bandwidth, followed by the AQMR

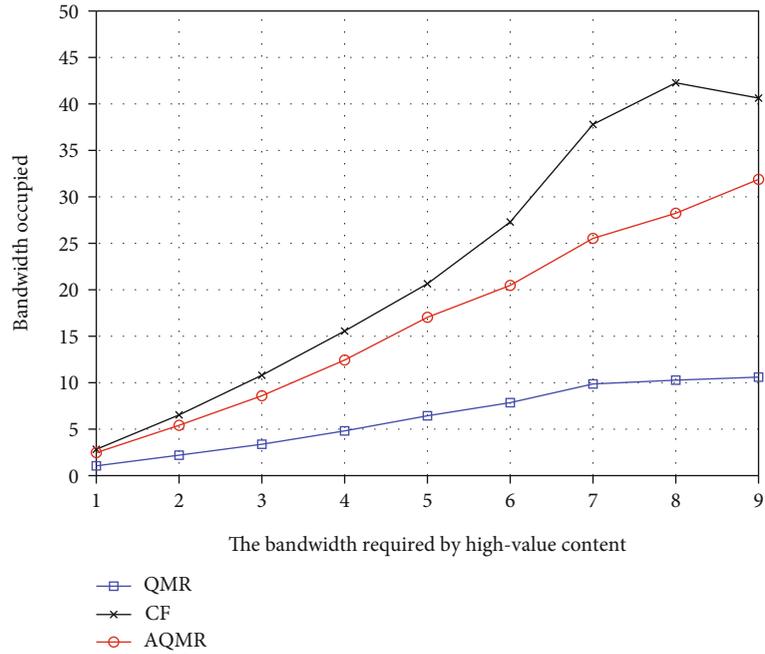


FIGURE 7: The bandwidth occupancy is calculated for each path that successfully transmits the high-value content. The optimization effect of AQMR is higher than CF without considering the QMR algorithm due to the lower success rate.

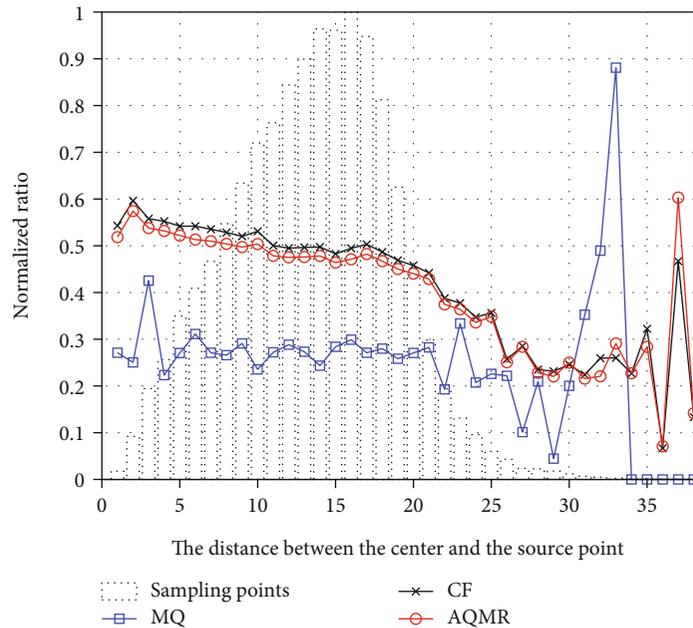


FIGURE 8: The bandwidth occupancy is calculated for the center-source distance, and the statistical results proved the relevant speculation.

algorithm, and the CF algorithm has the worst performance. However, this result is biased because the success rate of the QMR algorithm is relatively low, which is proved in Figure 6. In many cases, only a small part of the multiple destination nodes can generate routing successfully, which is easy to access the source node relatively. The simulation results are consistent with the expectation. Compared with the CF algorithm, some redundant routing is eliminated in the AQMR algorithm effectively, which reduces the bandwidth occupancy of successful routing.

Furthermore, a characteristic was found in the simulation that the relative distance between the source node and the center node had an effect on the bandwidth occupancy. As shown in Figure 5(a), this is a typical scene with the large distance. Most of the destination nodes are on one side and the source nodes are on the other side, and the direction of multicast routing is uniform relatively. On the other hand, the smaller distances will result in scattered routing directions. For this, the parameter named center-source distance is defined, which can reflect the multicast situation effectively.

In the simulation, the effect of the number destination nodes on the results needs to be eliminated, and the number in 19 groups from 2 to 20 is set. It was repeated 2500 times in each group of simulation, and 47,500 groups of simulation results were obtained. Among them, 13 sets of data were removed due to the large distance, and the remaining data accounting for 99.97% were shown in Figure 8. The maximum center-source distance is divided into 1000 equally spaced parts as the X -axis, and the normalized bandwidth proportion is taken as the Y -axis.

The result of these algorithms is consistent with the prediction without considering the oscillation data caused partly by too few sampling points, the AQMR is superior to CF in effect. Furthermore, the bandwidth occupancy of AQMR and CF tends to decrease with the increase of center-source distance, but the QMR has no significant difference. This shows that AQMR and CF are affected by the direction of multicast routing, while QMR is not. It proves the inference about the bandwidth occupation of the QMR again.

7. Conclusion

In this paper, the directional transmission of high-valued contents over the intersatellite links is discussed in the NGSO satellite-based Internet. Based on the multilayer architecture, the problem of cached data transmission is formulated as a dynamic routing problem. The case of one destination node is modeled as a dynamic decision-making problem, and the unicast Q -learning routing strategy is proposed. Moreover, a simulation is built to verify the effectiveness for the proposed method, and the results show that the transmission success rate of the proposed routing strategy is always better than the traditional contact graph routing strategy, because more accurate network information rather than the simple satellite orbit information is acquired. On the other hand, this paper establishes an adaptive solution for the problem of high-value content directional transmission in multiple destination nodes, and the simulation results show that compared with the QMR algorithm which only considered Q -learning and the CF algorithm which only considered directional, the proposed AQMR algorithm has more advantages for this problem.

This paper has proposed an effective solution for the directional transmission in the NGSO network, and the next step is to study the impact of different network structures on the delivery of high-value content. The NGSO satellite changes link relationship due to movement. In this case, it would be interesting to have a method of mapping from network topology and high-value content transmission without specific link relationships. As a result, we will look at methods to deliver high-value content that are more commonly applicable in the NGSO satellite system.

Data Availability

In the database of Key Laboratory of Universal Wireless Communications, Beijing University of Posts and Telecommunications, if necessary, the data and code will be uploaded to the designated location of the journal after the article is published.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by Beijing University of Posts and Telecommunications Excellent Ph.D. Students Foundation under Grant No. CX2020209, in part by Project No. A01B02C01-202015D0.

References

- [1] Q. Bi, "Ten trends in the cellular industry and an outlook on 6G," *IEEE Communications Magazine*, vol. 57, no. 12, pp. 31–36, 2019.
- [2] S. Liu, X. Hu, Y. Wang, G. Cui, and W. Wang, "Distributed caching based on matching game in LEO satellite constellation networks," *IEEE Communications Letters*, vol. 22, no. 2, pp. 300–303, 2018.
- [3] 3GPP NR, "Solutions for NR to support non-terrestrial networks (NTN) (Release 16)," *3GPP TR*, vol. 38, p. 821, 2019.
- [4] P. Chini, G. Giambene, and S. Kota, "A survey on mobile satellite systems," *International Journal of Satellite Communications & Networking*, vol. 28, no. 1, pp. 29–57, 2010.
- [5] Y. Zhu, L. Qian, L. Ding, F. Yang, C. Zhi, and T. Song, "Software defined routing algorithm in LEO satellite networks," in *2017 International Conference on Electrical Engineering and Informatics (ICELTICs)*, pp. 257–262, Banda Aceh, Indonesia, 2017.
- [6] J. Yao, T. Han, and N. Ansari, "On mobile edge caching," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2525–2553, 2019.
- [7] Z. Na, Z. Pan, X. Liu, Z. Deng, Z. Gao, and Q. Guo, "Distributed routing strategy based on machine learning for LEO satellite network," *Wireless Communications & Mobile Computing*, vol. 2018, article 3026405, pp. 1–10, 2018.
- [8] J. Li, H. Lu, K. Xue, and Y. Zhang, "Temporal netgrid model-based dynamic routing in large-scale small satellite networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 6009–6021, 2019.
- [9] J. Li, K. Xue, J. Liu, Y. Zhang, and Y. Fang, "An ICN/SDN-based network architecture and efficient content retrieval for future satellite-terrestrial integrated networks," *IEEE Network*, vol. 34, no. 1, pp. 188–195, 2020.
- [10] J. Li, K. Xue, D. Wei, J. Liu, and Y. Zhang, "Energy efficiency and traffic offloading optimization in integrated satellite/terrestrial radio access networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 4, pp. 2367–2381, 2020.
- [11] S. Burleigh, "Dynamic routing for delay-tolerant networking in space flight operations," in *SpaceOps 2008 Conference*, American Institute of Aeronautics and Astronautics, 2008.
- [12] G. Wang, S. Burleigh, R. Wang, L. Shi, and Y. Qian, "Scoping contact graph-routing scalability: investigating the system's usability in space-vehicle communication networks," *IEEE Vehicular Technology Magazine*, vol. 11, no. 4, pp. 46–52, 2016.
- [13] S. Li and F. Tang, "Load-balanced cooperative transmission in MEO-LEO satellite network," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pp. 564–571, Krakow, Poland, 2018.

- [14] Y. Sun, M. Peng, Y. Zhou, Y. Huang, and S. Mao, "Application of machine learning in wireless networks: Key techniques and open issues," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3072–3108, 2019.
- [15] M. Dorigo, G. Caro, and L. Gambardella, "Ant algorithms for discrete optimization," *Artificial Life*, vol. 5, no. 2, pp. 137–172, 1999.
- [16] Q. Zhang, L. Ding, and Z. Liao, "A novel genetic algorithm for stable multicast routing in mobile ad hoc networks," *China Communications*, vol. 16, no. 8, pp. 24–37, 2019.
- [17] S. Yousefi and M. Maleki, "Multicast scheduling algorithm supporting spatial mini-slot reuse for IEEE 802.16 mesh networks," *China Communications*, vol. 10, no. 1, pp. 116–133, 2013.