In VN, the logical position of a virtual node is used to represent the position of a satellite that virtual nodes and satellites correspond one to one at any time.

When a request arrives, we calculate k candidate paths (defined in (3)) using the k shortest paths algorithm.

The selected path should have enough available capacity to transmit the requested data

Here we define the transmission delay as the number of hops.

The satellite network can be modeled as a directed graph: g =(V,E), where V is the set of satellites and E is the set of inter-satellite links (ISLs).

In this paper, DeepQ-Network(DQN) [14] is employed to learn the optimal routing policy for the LEO satellite network, which is extended from Q-learning by utilizing neural networks.

InDQN, two neural networks Q(s,a;θ) and Q(s,a;θ-) are preserved, where the and are the network weight parameters. Q(s,a;θ) is the evaluation network for selecting an action while Q(s,a;θ-) is the target network for training.

The default aggregation function is set as the ∑ function and the readout function is approximated by a three-layer fully-connected neural network.

The Q-Network is a five-layer fully-connected neural network.

Because the input satellite network state is basically composed of topology and the links states, it will first be transformed into its corresponding line graph,which is able to convert edges into nodes while ensuring the equivalence.

We use the Iridium constellation as the target LEOsatellite network, where 66 satellites are distributed in 6 orbits and we select 6 of these satellites as the source and destination nodes.

The source and the destination satellites are chosen randomly from the selected 6 satellites and the required bandwidth is randomly generated from several kinds of packages(i.e.,16,32,64 bandwidth units).

The initial bandwidth for every link is set to be 200 bandwidth units. The exploration rate decays from 1 exponentially every 3 episodes and stops at 0.001

As The Update function, the hidden size of RNN is set to be 20

Besides, a link state contains a set of fields, i.e., ei = (f1 f2 f3 f4 f5), where f1 denotes the rest available capacity of the link, f2 indicates the occupied bandwidth of the link, f3 is the link betweenness (a measure of centrality inherited from graph theory that indicates how many paths may potentially traverse), f4 is an action vector indicating whether the requested data is transmitted over this link, and f5 is zero padding.

Method using DQN without GNN that are also based on Virtual Node are selected: ShortestPath(SP). SP chooses the shortest path for every coming request.

RandomPath(RP). RP decides the routing path randomly from the candidate paths.

RequestBalance(RB). RB evenly splits the required bandwidth into k parts and distributes them to all k candidate paths.

DQN without GNN(DWG)[15]. DWG utilizes DQN algorithm without GNN, i.e., it directly uses a fully connected neural network to approximate the Q-value of satellite networks state

Compared with SP algorithm, our algorithm outperforms amazingly in throughput with comparable delay. Because SP algorithm always chooses the shortest paths for requests, an episode is quick to end and the delay is short. The throughput has a 47.1% improvement with a 13.7% delay increase for our algorithm.

In comparison with RP, our algorithm outperforms not only in throughput but also in delay. Our algorithm improves the throughput by 33.4% and decreases the delay by 6.1%