



## Article

# Caching-Aware Intelligent Handover Strategy for LEO Satellite Networks

Tao Leng <sup>1,2</sup>, Yuanyuan Xu <sup>1,2</sup> , Gaofeng Cui <sup>1,2,\*</sup> and Weidong Wang <sup>1,2</sup>

<sup>1</sup> School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; 2016010101@bupt.edu.cn (T.L.); xyy1223@bupt.edu.cn (Y.X.); wangweidong@bupt.edu.cn (W.W.)

<sup>2</sup> Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China

\* Correspondence: cuigaofeng@bupt.edu.cn; Tel.: +86-10-6119-8609

**Abstract:** Recently, many Low Earth Orbit (LEO) satellite networks are being implemented to provide seamless communication services for global users. Since the high mobility of LEO satellites, handover strategy has become one of the most important topics for LEO satellite systems. However, the limited on-board caching resource of satellites make it difficult to guarantee the handover performance. In this paper, we propose a multiple attributes decision handover strategy jointly considering three factors, which are caching capacity, remaining service time and the remaining idle channels of the satellites. Furthermore, a caching-aware intelligent handover strategy is given based on the deep reinforcement learning (DRL) to maximize the long-term benefits of the system. Compared with the traditional strategies, the proposed strategy reduces the handover failure rate by up to nearly 81% when the system caching occupancy reaches 90%, and it has a lower call blocking rate in high user arrival scenarios. Simulation results show that this strategy can effectively mitigate handover failure rate due to caching resource occupation, as well as flexibly allocate channel resources to reduce call blocking.

**Keywords:** LEO handover strategy; multiple attributes; caching; deep reinforcement learning



**Citation:** Leng, T.; Xu, Y.; Cui, G.; Wang, W. Caching-Aware Intelligent Handover Strategy for LEO Satellite Networks. *Remote Sens.* **2021**, *13*, 2230. <https://doi.org/10.3390/rs13112230>

Academic Editors: Sooyoung Kim and Giovanni Giambene

Received: 27 April 2021

Accepted: 1 June 2021

Published: 7 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, the 5th generation mobile communication system (5G), which aims to provide high-speed wireless services [1] for global users, has developed rapidly. However, due to the impacts of terrain and cost on infrastructure construction, terrestrial cellular networks can only cover densely populated areas [2], and it fails to provide communication services for mountainous terrain, oceans, and air areas. The advantages of satellite communication, which are wide coverage, strong resistance to destruction, and insensitive to terrain factors, can compensate for the limitations of the terrestrial mobile communication networks. Therefore, satellite communication has become one of the key technical components for systems beyond 5G to achieve global coverage [3]. It is widely implemented in many fields [4], such as military, disaster emergency, digital broadcasting and television, and mobile communication. Some scholars also proposed that the 6th generation (6G) wireless communication will be “5G + satellite network” [5]. Thus, many countries and companies are actively engaging in the research and implementation of satellite communication systems, especially the Low Earth Orbit (LEO) communication systems with lower propagation delay [6], such as Starlink, OneWeb, etc.

Generally, the LEO satellite system has a dynamic topology, which leads to frequent handover between terrestrial terminals and satellites [7]. Unfortunately, frequent handover not only makes the LEO satellite systems difficult to guarantee the quality of service (QoS) of users but also leads to the waste of radio resources. Moreover, the existing studies on satellite handover strategies, such as inter-beam handover and inter-satellite

handover [8], mainly focus on handover based on the receiving signal quality, remaining service time, and so on.

Papapetrou et al. [9] proposed three different handover criteria: the maximum remaining service time, the maximum number of idle channels, and the minimum distance, which could be applied to new coming or handover calls. The strategy based on the maximum remaining service time can greatly reduce the handover times, delay, and signaling cost. On basis of this strategy, Hu et al. [10] proposed a velocity-aware handover prediction method to find the shortest path of the time-expanded graph dynamically. Duan et al. [11] proposed a distributed handover method by taking into account the impact of routing, and it can reduce the propagation delay while keeping the handover times acceptable. Seyedi et al. [12] proposed a simple real-time handover strategy, which exploited both the global positioning system (GPS) infrastructure and multiple satellites, to minimize the expected handover times. The strategy based on the maximum number of idle channels selected the candidate satellite with the largest number of available channels. Thus, the load distribution of satellite systems can be balanced, and the limited network resources can be efficiently utilized. Zhou et al. [13] proposed a dynamic channel reservation scheme based on priorities. The traffic, which is predicted based on the deterministic movement property of LEO satellites, was used to obtain the thresholds for reserved channels. It can effectively reduce handover failure rate and improve channel utilization by dynamically adjusting the thresholds according to the traffic conditions. The strategy based on the minimum distance selected the candidate satellite by considering the distance between the satellites to avoid link interruption. Wu et al. [14] proposed a graph theory-based inter-satellite handover strategy, which adopted the shortest path algorithm to obtain the optimal handover scheme. Furthermore, other single-attribute handover criteria can be achieved with this handover model by changing the path weights. Since the above studies only consider the effect of a single attribute of the candidate satellite, they cannot achieve a good trade-off between handover times, system load, and success rates of handover.

Li et al. [15] proposed a multi-layer handover management framework and different handover procedures based on handover prediction which can reduce handover delay and signaling cost. Furthermore, they also proposed a dynamic handover optimization method, which takes traffic, rate demand, and channel gain into account, aiming at reducing the dropping rate and guaranteeing the QoS of mobile terminals. Li et al. [16] proposed a user-centric handover scheme for ultra-dense LEO satellite networks, and it can realize seamless handover by buffering user's downlink data in multiple satellites simultaneously. Wu et al. [17] proposed a handover algorithm based on the potential game, and the strategy considered the remaining service time and the satellite elevation angle which can be utilized to minimize the average satellite handover times and decrease call-dropping probability. They also proposed a terminal random-access algorithm aiming at balancing the network load. He et al. [18] proposed a load-aware satellite handover strategy based on multi-agent reinforcement learning, and it can balance satellite load to avoid network congestion, while, at the same time, maintain low signaling overhead. Miao et al. [19] proposed an LEO satellite handover strategy based on the multi-attribute decision. The strategy used the technique for order preference by similarity to an ideal solution (TOPSIS) evaluation method to calculate the weighted values of three attributes, which include signal strength, remaining service time, and remaining idle channels. And then, the strategy selected the candidate satellite with the best overall performance. Zhang et al. [20] considered the impacts of channel quality, remaining service time, and the number of service users on handover strategy, and they used the entropy method to weight each factor and transform it into a single objective optimization problem. Xu et al. [21] analyzed a quality of experience (QoE)-driven handover strategy, which considered routing delay, remaining service times, and remaining idle channels with high-speed mobile users for LEO satellite networks. Table 1 shows the summary and comparison of the above studies.

**Table 1.** Comparison of existing handover strategies.

Strategy Types	Authors	Handover Factors				Performance
		Service Times	Channels	Distance	Others	
Single attribute	Papapetrou et al. [9]	✓				reduce the handover times
			✓			balance load, reduce handover failure rate
					✓	avoid link interruption
	He et al. [10]	✓			velocity-aware	handover prediction, find the shortest path
	Duan et al. [11]	✓			routing delay	reduce the propagation delay
	Seyedi et al. [12]	✓			GPS, multiple satellite	minimize the handover times
	Zhou et al. [13]		✓		traffic prediction	reduces handover failures rate, improves channel utilization
Multiple attribute	Wu et al. [14]			✓		optimal handover strategies for end-to-end communication
	Li et al. [15]		✓		traffic, rate demand	reducing the dropping rate, guarantee the QoS of mobile users.
	Wu et al. [17]	✓		✓		minimize the handover times, decrease call-dropping probability
	He et al. [18]	✓	✓		load-aware	balance load, maintain low signaling overhead
	Miao et al. [19]	✓	✓		single strength	reduce handover times, balance load and guarantee QoS
	Zhang et al. [20]	✓	✓		number of users, satellite power	reduce handover times, balance load and guarantee SNR
	Xu et al. [21]	✓	✓		routing delay	reduce handover times, failure rate and transition delay

Besides factors related to the communication metrics, resource management by considering the integration of communication, computing and caching (3C) is also important for future mobile edge computing (MEC) enhanced satellite systems [22]. Caching data at satellite nodes can improve communication efficiency by avoiding duplicate transmissions [23,24]. Liu et al. [23] proposed a novel caching algorithm by optimizing contents placement in LEO satellite constellation networks, which was used to minimize user terminals content access delay. Zhang et al. [25] analyzed the caching restricted resource allocation with joint optimization of satisfaction index and spectrum efficiency for multi-beam satellite systems. Since the satellite on-board caching resource is limited, we focus on evaluating the inter-satellite handover strategies for the LEO satellite systems with the caching aware strategy.

Moreover, we aim to tackle the following problems encountered by the existing handover strategies:

- (1) Although the existing handover strategies analyze several factors that affect the performance of handover, the effect of limited on-board caching is not considered. Moreover, the joint-effect of multiple attributes, which are on-board caching, remaining service time, and idle channels, is not considered either.
- (2) The existing handover strategies make the handover decisions with the snap shot-based topology. However, the topology of LEO satellite networks is time varying, and

the snap shot-based handover strategies cannot guarantee the long term performance of the dynamic system.

To solve these problems, the effects of on-board caching and joint-effect of multiple attributes on handover strategies are analyzed in this paper. Furthermore, an intelligent handover strategy based on deep reinforcement learning (DRL) is proposed to reduce the dropping probability and call blocking rate. The main contributions are listed as below.

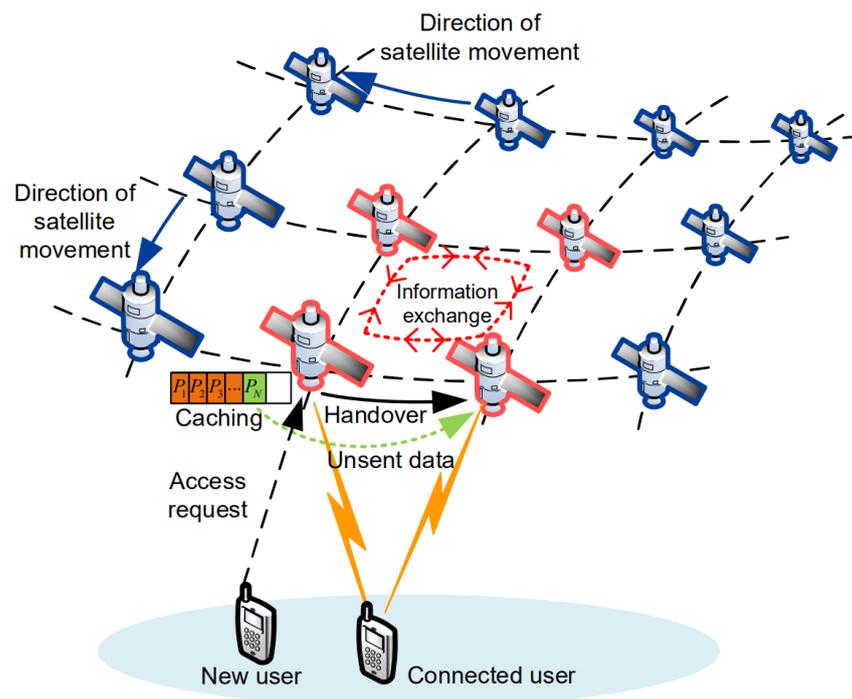
- (1) A novel framework for caching-aware intelligent handover strategies is proposed for LEO satellite networks. Different from existing handover strategies, the joint-effect of multiple attributes, including remaining service time, remaining idle channels, and remaining caching capacity, on handover performance are investigated with dynamic network topology.
- (2) To adapt to the dynamic topology of satellite systems, the inter-satellite handover process is modeled as a Markov decision process, and the process for the intelligent handover strategy is provided in detail.
- (3) An intelligent handover algorithm based on DRL is proposed. The algorithm can make decisions on when will the handover be activated and select the target satellite in each time slot. Moreover, the DRL algorithm can make continuous handover decisions, which makes the whole system obtain the maximum long-term benefits. Simulation results demonstrate the effectiveness of the proposed handover strategy.

## 2. System Model

### 2.1. System Architecture and Handover Factors

LEO satellites are typically deployed at low altitude, such as 500 km to 1500 km, and the LEO satellites move with a high speed relative to the users on the ground. Thus, frequent handover may occur during the service time of users. This paper considers a constellation of LEO satellites in the sun-synchronous orbit. There are 12 orbits within the constellation, and 9 satellites are located at each orbit. The satellites are located at an altitude of 1000 km and have an orbital inclination of 99.4843 deg. A typical handover scenario in the LEO satellite system is presented in Figure 1. Each terrestrial user terminal establishes a communication link with an LEO satellite for transmitting data. Due to the high mobility of the LEO satellite, the terrestrial users will move out of the coverage of the serving satellite after a period of connecting, and handover is required to ensure continuous communication. In addition to this case, the access of new coming users can also lead to handover within the system. And the adjacent satellites will continuously exchange the remaining resource information through the inter-satellite link. When the source satellite detects that the handover is needed for the connected user links, it will analyze the resource information and select the best satellite from the candidate satellite list. Before the handover, the data that has not been sent will be cached, and then the cached data should be sent to the optimal candidate satellite after the new link is established.

In the  $P$  time slot, the serving satellite refers to the satellite currently connected to the subscriber, which provides communication services to the subscriber during the time slot. Candidate satellites refer to satellites that are available for connection in the  $P$  time slot other than the serving satellite. The candidate satellites can be selected when the handover is activated for the user's communication link. There are 2 to 4 visual satellites for a user in this system. In addition to the serving satellites, there are 1 to 3 candidate satellites. Once the handover is activated, the optimal candidate satellite is selected from the candidate satellites according to the intelligent policy.



**Figure 1.** Handover scenario in LEO satellite system.

## 2.2. Remaining Service Time

Due to the high mobility, a LEO satellite stays in the visible range of the user for about 10 min generally, during which the satellite can communicate with the user. This period is called the maximum service time. The maximum remaining service time refers to the time that the communication link can maintain before the serving satellite moves out of the user's visual range. In Figure 2, due to the user's speed is far less than the speed of the satellite, we assume that the user terminal is stationary relative to the satellite. For every user and its serving satellite, the elevation angle between the user and the satellite will gradually increase from the minimum value to the maximum value and then, finally, decrease to the minimum value. Specifically, at  $T_0$ , the user terminal enters the satellite coverage area, and the elevation angle from the user terminal to the satellite is the smallest. At this point, the satellite begins to provide communication services for users, and the remaining service time is the longest. With the movement of the satellite, the elevation angle increases gradually, and the user terminal gets the maximum elevation angle at  $T_1$ . The position of the satellite footprint  $Q_1$  coincides with the point  $H$ , and  $H$  stands for the closest point from the location of the user terminal to the satellite footprint trajectory. After  $T_1 - T_0$ , the elevation angle of the user terminal is minimum again in  $T_2$  slot, and the user leaves the satellite service coverage area. According to the geometric relationship, we can obtain the maximum service time  $T_{\max}$  [20] as follows:

$$T_{\max} = \frac{2\Gamma(t_0)}{\omega}, \quad (1)$$

where  $\Gamma(t_0)$  is the radian of  $\angle Q_0OH$ ,  $O$  is the center of the Earth, and  $\omega$  is the angular velocity of the satellite.  $\Gamma(t_0)$  can be calculated with  $Y(t_0)$  and  $Y_{\min}$  [20] as follows:

$$\cos \Gamma(t_0) \cos(Y_{\min}) = \cos Y(t_0), \quad (2)$$

where  $Y_{\min}$  is the radian of  $\angle Q_1OU$ ,  $O$  is the center of the Earth, and  $Q_1U$  represents the shortest distance between the user terminal and the satellite footprint trajectory.  $Y(t_0)$  is the radian of  $\angle Q_0OU$ , and  $Q_0U$  represents the longest distance between the user terminal

and the satellite footprint trajectory. According to the geometric relationship, we can calculate the maximum service time with the total radius of the satellite footprint trajectory  $2\Gamma(t_0)$ .

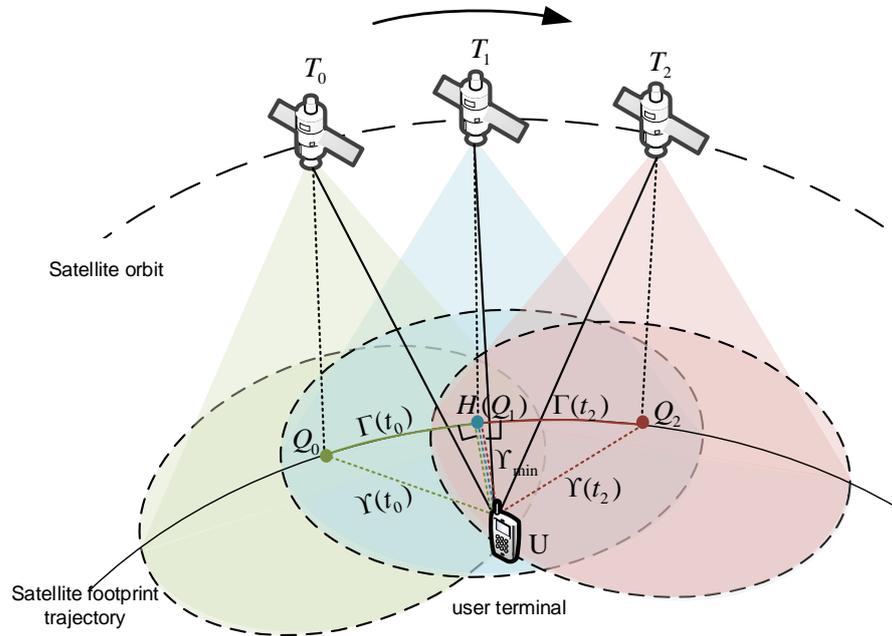


Figure 2. Satellite coverage at different times.

First,  $Y(t_0)$  [20] can be obtained as shown in the Figure 3. It can be expressed as

$$(R_e + h) \cos(\theta_{\min} + Y(t_0)) = R_e \cos \theta_{\min}, \tag{3}$$

where  $R_e$  is the radius of Earth,  $h$  is the altitude of satellite orbit, and  $\theta_{\min}$  is the user's minimum elevation angle. Then,  $Y(t_0)$  can be expressed as

$$Y(t_0) = \arccos\left(\frac{R_e}{R_e + h} \cos \theta_{\min}\right) - \theta_{\min}, \tag{4}$$

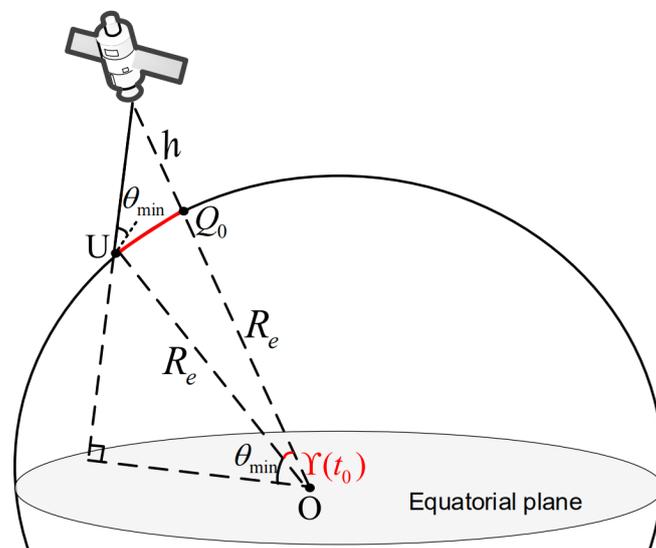


Figure 3. The geometric relationship of  $Y(t_0)$ .

As show in Figure 4,  $Y_{\min}$  [20] can be obtain as follows:

$$Y_{\min} = \arccos \sqrt{1 - \frac{d_{\min}^2}{R_e^2}}, \tag{5}$$

where  $d_{\min}$  is the shortest distance of user terminal to the satellite footprint trajectory plane. In the actual system, the latitude and longitude of the satellite footprint point  $(\phi_t^s, \varphi_t^s, 0)$  and user  $(\phi_t^u, \varphi_t^u, \zeta_t^u)$  for time slot  $t$  can be obtained according to the GPS positioning system. Converting latitude, longitude and altitude coordinates  $(\phi, \varphi, \zeta)$  to Earth-Centered Earth-Fixed (ECEF) coordinates  $(x, y, z)$ . The conversion formula is:

$$\begin{cases} x = (\frac{a}{\sqrt{1 - e^2 \sin^2 \phi}} + \zeta) \cos \phi \cos \varphi \\ y = (\frac{a}{\sqrt{1 - e^2 \sin^2 \phi}} + \zeta) \cos \phi \sin \varphi \\ z = ((1 - e^2) \frac{a}{\sqrt{1 - e^2 \sin^2 \phi}} + \zeta) \sin \phi \end{cases}, \tag{6}$$

where  $a$  is the length of the Earth’s semimajor axis, and  $e$  is the eccentricity of the Earth. Use three coordinates of the satellite footprint  $\{(x_t^s, y_t^s, z_t^s) | t = 0, 1, 2\}$  to determine the satellite footprint trajectory plane equation

$$Ax + By + Cz + D = 0, \tag{7}$$

where

$$\begin{cases} A = (y_2 - y_1)(z_3 - z_1) - (y_3 - y_1)(z_2 - z_1) \\ B = (z_2 - z_1)(x_3 - x_1) - (z_3 - z_1)(x_2 - x_1) \\ C = (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1) \\ D = -Ax_1 - By_1 - Cz_1 \end{cases} \tag{8}$$

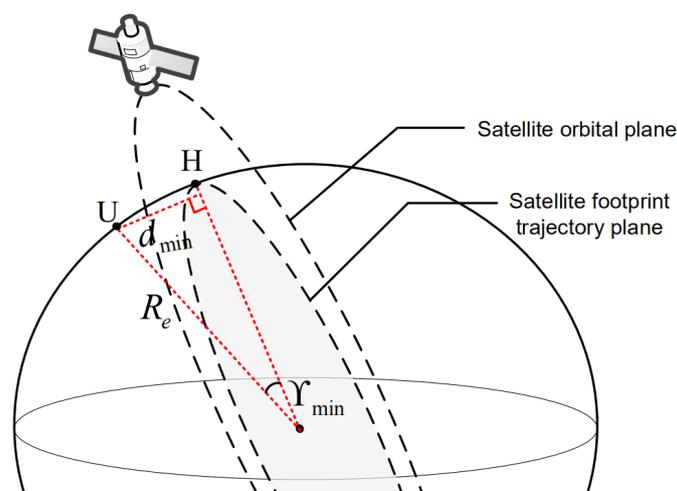


Figure 4. The geometric relationship of  $Y_{\min}$ .

According to the plane Equation (7),  $d_{\min}$  can be expressed as

$$d_{\min} = \frac{Ax^u + By^u + Cz^u + D}{\sqrt{A^2 + B^2 + C^2}}. \quad (9)$$

Finally, the angular velocity  $\omega$  of the satellite in the ECEF coordinate system can be got according to the geometric relationship in Figure 5, which can be expressed as

$$\omega \approx \omega_s + \omega_e \cos(\pi - \sigma), \quad (10)$$

where  $\omega_s$  and  $\omega_e$  are, respectively, the angular velocity of the satellite and the Earth in the Earth-Centered Inertial (ECI) coordinate system,  $\sigma$  is the satellite orbit inclination,  $\omega_s$  can be expressed as

$$\omega_s = \sqrt{\frac{\mu}{(R_e + h)^3}}, \quad (11)$$

where  $\mu$  is the Kepler constant,  $R_e$  is the Earth radius, and  $h$  is the altitude of the satellite orbit. With Equations (4) and (5),  $\Gamma(t_0)$  can be written as

$$\Gamma(t_0) = \arccos\left(\frac{\cos Y(t_0)}{\cos(Y_{\min})}\right). \quad (12)$$

Then, the maximum service time of satellite can be obtained by Equation (1), and the remaining service time at the current slot  $T$  can be expressed as

$$T_{rem} = T_{\max} - (T - T_0), \quad (13)$$

where  $T_0$  is the recorded service start time.

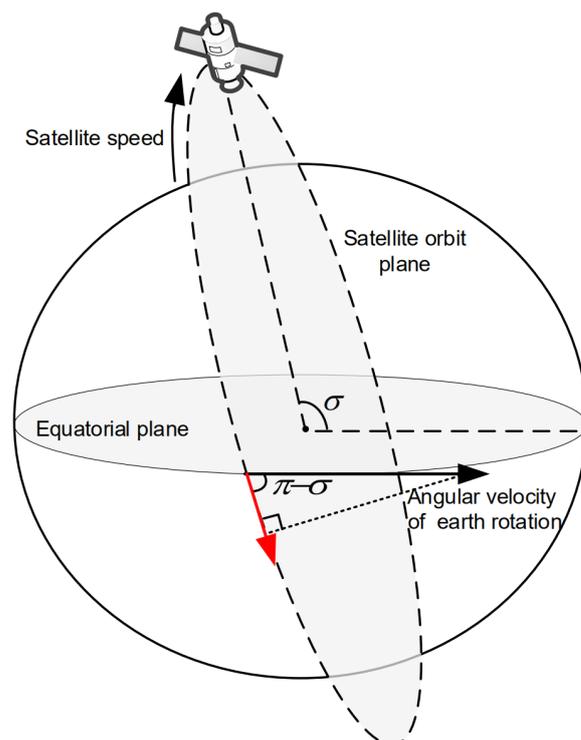


Figure 5. Angular velocity of satellite.

### 2.3. Remaining Idle Channels

Each satellite has a fixed number of channels. Channels that are occupied will not be allocated to other users, while unoccupied channels are idle and waiting to be allocated. The number of remaining idle channels reflects the satellite load. Unbalanced load distribution will lead to a large number of idle channels in the light-load satellite coverage area, and the dropped calls and call blocking rate will increase in the overload areas. This situation will greatly reduce the overall performance of the system. In this paper, the number of idle channels is obtained in real-time through information interaction between satellites and utilized as a factor to make the handover decisions. For example, the handover strategy can dynamically reset the links of connected users to the adjacent satellites with idle channels to reduce the blocking rate of new calls.

### 2.4. Remaining Caching Capacity

During the handover period, the user's unsent data will be cached to the serving satellite, and they will not be sent to the target satellite until the handover is completed. The data will be cached and utilized to ensure the integrity of user data and avoid packet loss caused by handover. Apart from the occupied caching resources, the remaining caching resources that can be used for caching the data during the handover process are called the remaining caching capacity. However, the limited satellite on-board caching resource is difficult to meet the requirement of a large amount of data caching caused by multi-user simultaneous handover processes, and it will lead to handover failure and data packet loss. Therefore, the on-board caching resource is regarded as one of the decision factors of handover, and the information can also be obtained from the information interaction between adjacent satellites.

## 3. Caching-Aware Intelligent Handover Strategy

### 3.1. Handover Flow

This paper proposes an inter-satellite handover strategy, where the handover decisions are made by the serving satellite, and handover flow is shown in Figure 6. Firstly, the user periodically reports the user's location to the serving satellite, and the satellite calculates the remaining service time according to the user's location. At the same time, each satellite will periodically send its resource information to adjacent satellites, mainly including the number of idle channels and the remaining caching capacity. For example, at a certain time slot, the serving satellite will receive resource information from adjacent satellites and make handover decisions according to the user location information and the satellites resource information. In detail, the trained DRL network is adopted to decide when will the handover be executed and the candidate satellite. If the handover procedure is activated, the serving satellite will send the handover request to the optimal candidate satellite. Then, the optimal candidate satellite applies for the resources. If the resources are sufficient, it will send the handover response to the serving satellite, who is caching the unsent user's data and sends the handover notification to the user. From then on, the user and the candidate satellite can establish the communication link, and the handover can be completed. If the candidate satellite has no remaining resources, or the serving satellite's caching is overflow or the link is interrupted, handover failure occurs.

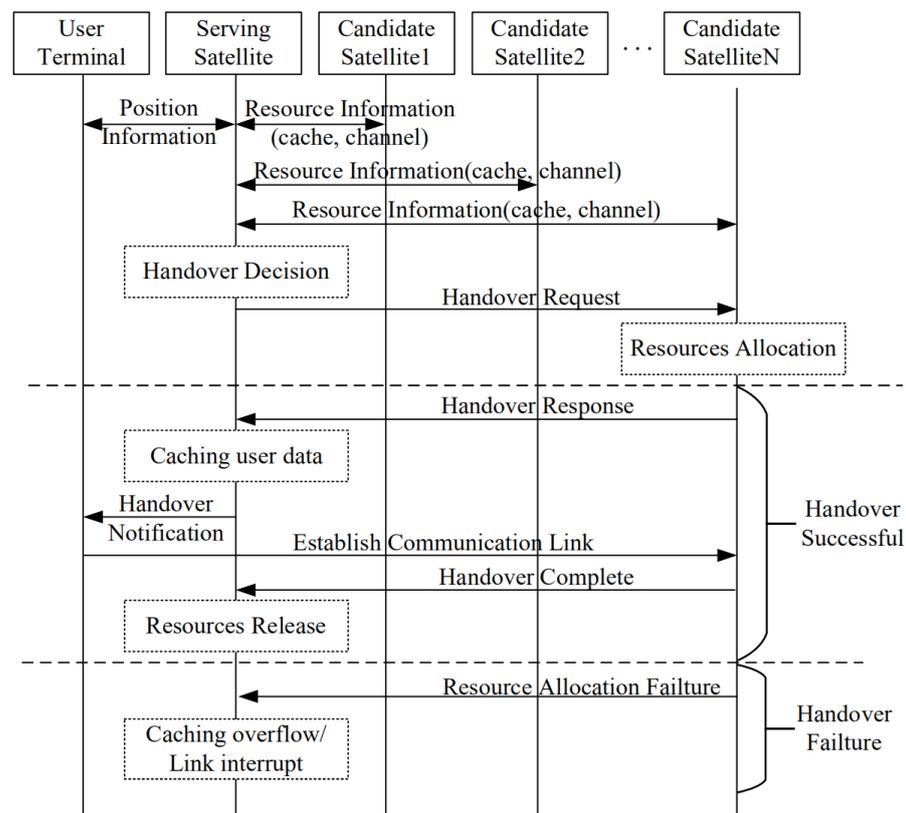


Figure 6. Handover flow chart based on serving satellite decision.

Based on the handover signaling flow chart, we will analyze the handover latency. The latency of this process is composed of propagation latency and data transmission latency, which is ignored here as the signaling transmission latency is much smaller than the propagation latency. Firstly, the signaling interaction from user to satellite for reporting location information and information exchange among adjacent satellites occurs periodically, so this part of the signaling overhead is not counted as part of the handover delay. When the serving satellite decides to activate handover based on the state of the environment, it interacts with the optimal candidate satellite for the handover request signaling. The request and response signaling is carried out on the inter-satellite link, and the maximum signaling time between two satellites at a time is approximately 17 ms, as can be calculated from the STK simulation environment. When the serving satellite finishes caching the unsent data, it sends a handover command to the ground user, and the propagation delay from satellite to ground user is about 3.3 ms. The user receives the handover command and establishes a connection with the optimal candidate satellite. Then, the handover process is finished. This process involves two Earth-satellite link transmission and three inter-satellite link transmissions. The total propagation delay of the process is 57.6 ms.

Here, we should note that there will be several reasons that will cause satellite handover. Besides the satellite movement, we also focus on the following two situations in this paper:

- When the new coming user asks for access, the communication links of the connected users may be reset from the serving satellite that has no idle channels to another candidate satellite. Thus, the channels can be released for the new coming users.
- If the remaining caching capacity is less than the amount of data that will be sent by the users, the handover cannot be carried out. Otherwise, the handover will fail, and it will result in packet loss and a sharp decline in user experience.

Moreover, the handover decisions will be made by considering the joint-effects of several attributes, such as remaining service time, remaining idle channels, and remaining caching capacity, to obtain the best overall system performance.

### 3.2. Intelligent Handover Strategy with Multiple Attributes

The handover decisions will affect the resource utilization of the connected satellite resources, so the handover decisions at each time slot will also be affected by the decisions at the previous time slots; thus, the process can be modeled as the Markov decision process. In order to achieve a dynamic and continuous handover decision that takes into account multiple factors in a continuous environmental state, we propose an intelligent handover strategy based on the DRL. The strategy can intelligently decide when will the handover be activated and select the optimal satellite based on the resource utilization information of the candidate satellite. It fully takes into account the joint-effect of multiple attributes which include the remaining service time, the number of idle channels, and the remaining caching capacity. Among these attributes, two attributes are the state of the candidate satellite (the remaining service time and the number of idle channels), and another attribute is the state of the serving satellite (the remaining caching capacity). Moreover, the handover strategy can make continuous decisions for the dynamic LEO network. The network structure is shown in Figure 7, and the detailed explanation for the DQN training process can be found in Reference [26]. We need to train this decision network in advance and use the parameter-stabilized DRL network to output the decision actions. During the actual handover process, the satellite simply feeds the environment information from the mobile terminal periodically into the DRL network, and the corresponding handover decision will be obtained.

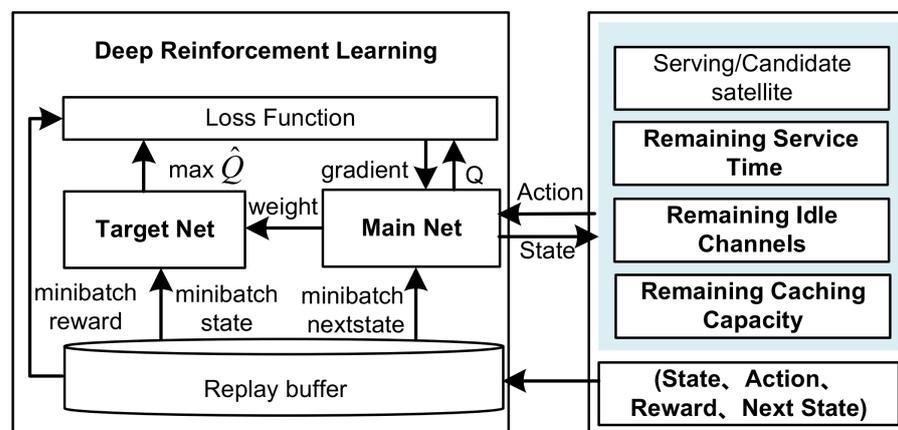


Figure 7. The intelligent handover network based on DQN.

The specific design of an intelligent handover decision network based on DRL, which is shown in Figure 7, can be denoted as a tuple  $(S, A, P, R)$ .  $S$  denotes the state space of the LEO system.  $A$  is the handover action space.  $P$  is the space of the state transition probability.  $R$  denotes the reward of state and action.

**State space (S):** the state space is derived from the environment of the LEO satellite communication system. There are four environment state quantities in this paper: the satellite label  $\Omega$ , the number of remaining idle channels  $\Theta$ , the remaining cache capacity  $C$ , and the remaining service time  $T$ . Firstly, the satellite label is used to express which satellite is currently connected to the user, and which satellites are candidates. The number of remaining idle channels, remaining cache capacity, and the remaining service time are the factors to be taken into account for handover. And each state consists of the corresponding state of the serving satellite and the handover candidate satellites. The state at time slot  $p$  can be defined as

$$s_p = \{\Omega_p, \Theta_p, C_p, T_p\}. \quad (14)$$

$\Omega_p$  is the set of satellite label, which can be expressed as  $[\Omega_p^*, \Omega_p^1, \Omega_p^2, \dots, \Omega_p^K]$ .  $\Omega_p^*$  is defined as  $[\Omega_{1,p}^*, \Omega_{2,p}^*, \dots, \Omega_{i,p}^*]$ . The upper corner of the status variable symbol is used to refer to the serving satellite or the candidate satellite. “\*” represents the relevant status of the serving satellite which connected to the user. The lower corner “ $i, p$ ” indicates the user number  $i$  and the time slot  $p$ . So,  $\Omega_{i,p}^*$  is used to express the serving satellite label of user  $i$ .  $\Omega_{i,p}^* \in \{1, 2, 3, \dots, M\}$ ,  $M$  is the total number of the LEO satellites.  $\Omega_p^K$  is defined as  $[\Omega_{1,p}^K, \Omega_{2,p}^K, \dots, \Omega_{i,p}^K]$ . The upper corner “ $k$ ” indicates the relevant status of the candidate satellite  $k$ . For example, the maximum number of visible candidate satellites for the user in this paper is 3, then the upper corner is 1 or 2 or 3. So,  $\Omega_{i,p}^K$  is used to express the  $k$ th candidate satellite label of user  $i$ .  $\Theta_p$  refers to the information of satellite channels, which can be expressed as  $\Theta_p = [\Theta_p^*, \Theta_p^1, \Theta_p^2, \dots, \Theta_p^K]$ .  $\Theta_p^*$  is defined as  $[\Theta_{1,p}^*, \Theta_{2,p}^*, \dots, \Theta_{i,p}^*]$ , and  $\Theta_{i,p}^*$  is the number of idle channels of the serving satellite of user  $i$ .  $\Theta_{i,p}^* \in \{1, 2, 3, \dots, I\}$ ,  $I$  is the total channel number of a single satellite.  $\Theta_p^K$  is defined as  $[\Theta_{1,p}^K, \Theta_{2,p}^K, \dots, \Theta_{i,p}^K]$ , and  $\Theta_{i,p}^K$  is the number of idle channels for user  $i$ 's  $K$ th candidate satellite.  $C_p = [C_p^*, C_p^1, C_p^2, \dots, C_p^K]$  and  $T_p = [T_p^*, T_p^1, T_p^2, \dots, T_p^K]$  is the remaining caching capacity and service time.

**Action space (A):** For each time slot, the agent will decide whether the handover be activated or not and select the optimal target satellite for every user. The action of time slot  $p$  can be expressed as

$$a_p = [\Lambda_{1,p}, \Lambda_{1,p}, \dots, \Lambda_{i,p}], \quad (15)$$

where  $\Lambda_{i,p} \in \{0, 1, 2, 3, \dots, K\}$ . The value ‘0’ means that the handover for user  $i$  will not be activated. Other values mean that handover will be activated, and the corresponding number is the label of the candidate satellites.  $K$  is the maximum number of candidate satellites.

**Transition probability (P):** Since the system state in this paper is continuous and the state is affected by the handover decision, the transition probability from  $s_p$  to  $s_{p+1}$  with the action  $a_p$  is difficult to be obtain. Hence, a model-free DRL framework based on a deep Q-learning network (DQN) is adopted.

**Reward function (R):** We divide the reward function  $R(s_p, a_p)$  into two parts: gain function  $g_p$  and cost function  $l_p$ .

$$R(s_p, a_p) = \sum_{i=1}^N g_{i,p} - l_p, \quad (16)$$

$$g_{i,p} = \begin{cases} w_1 \Theta_{i,p}^* + w_2 C_{i,p}^* + w_3 T_{i,p}^*, & \text{if } a_p = 0, \text{ handover is not activated} \\ w_1 \Theta_{i,p}^{a_p} + w_2 C_{i,p}^{a_p} + w_3 T_{i,p}^{a_p}, & \text{if } a_p \neq 0, \text{ handover is activated} \end{cases}, \quad (17)$$

$$l_p = w_4 \alpha_p + w_5 \beta_p + w_6 \delta_p. \quad (18)$$

$g_{i,p}$  is the gain of remaining communication resources. Three attributes, which are the number of idle channels, remaining caching capacity, and remaining service time, are normalized.  $w$  is the weight of each attribute, and  $w_1 + w_2 + w_3 = 1$ .  $l_p$  is the cost function.  $\alpha_p$  is the number of successful handover times in time slots  $p$ , and  $\beta_p$  is the number of

failed handover times. The failed handover defined here refers to the handover failure caused by several reasons. For example, the candidate satellite does not have idle channels, or the remaining service time is reduced to zero, or the serving satellite does not have enough caching capacity.  $\delta_p$  is the number of dropped calls caused by insufficient resources (remaining service time, channel, etc.).

According to the definitions of state, action, and the reward, we can calculate the target Q value:

$$y_p = R(s_p, a_p) + \gamma \max_{a'} \hat{Q}(s_p, a'; \theta^-). \quad (19)$$

Then, the mean square error loss function is calculated as

$$L(\theta) = E[(R(s_p, a_p) + \gamma \max_{a'} \hat{Q}(s_p, a'; \theta^-) - Q(s_p, a; \theta))^2], \quad (20)$$

where  $\theta$  is the main network parameters, and  $\theta^-$  is the target network parameters. After calculating the loss function, the gradient descent strategy is used to update the main network parameters. To break the time correlation between the sequences, the replay buffer is used to store the experience, and a random sample of minibatch is used for learning. The detailed algorithm is described as Algorithm 1.

---

**Algorithm 1** Intelligent handover algorithm based on DRL.

---

**Initialize** replay buffer  $D$

**Initialize** action-value function  $Q$  with random weights  $\theta$

**Initialize** target action-value function  $\hat{Q}$  with weights  $\theta^- = \theta$

**Repeat**

**Initialize**  $episode = 1$  and  $p = 1$

**Initialize** the start state  $s_1$  with connection relationship.

**for**  $p = 1, T$  **do**

Select a random action  $a_1$  With probability  $\varepsilon$  otherwise select  $a_p = \arg \max_a Q(s_p, a; \theta)$ .

Execute action  $a_p$  in LEO environment emulator.

Receive a reward  $R(s_p, a_p)$  and next state  $s_{p+1}$

Store transition  $(s_p, a_p, r_p, s_{p+1})$  in  $D$ .

Sample minibatch of transition  $(s_p, a_p, r_p, s_{p+1})$  randomly from  $D$  when  $D$  is full.

Update the main network parameters  $\theta$  using gradient descent with the goal of minimizing the loss function  $L(\theta)$  defined in Equation (20)

Every  $C$  steps reset  $\hat{Q} = Q$ .

**end for**

**Until**  $episode > episode\_max$ .

---

#### 4. Results

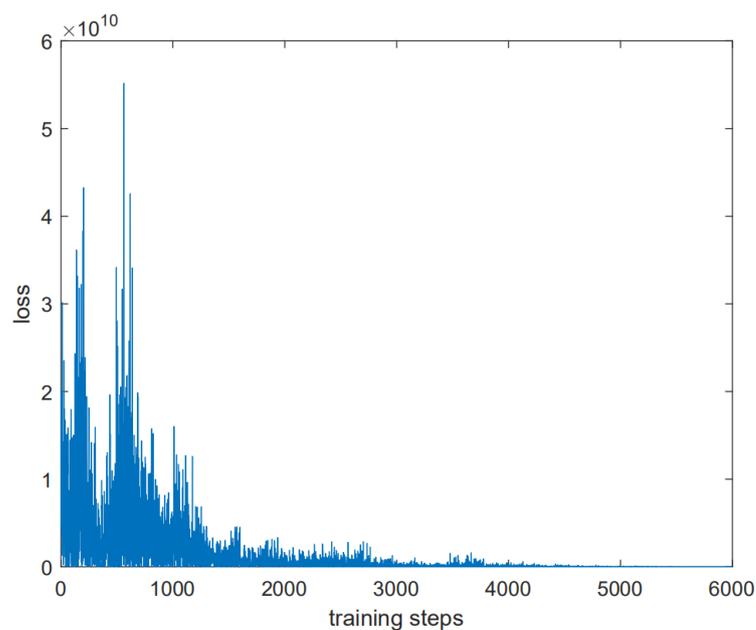
Satellite tool kit (STK) is used to obtain the topology of the satellite network. The simulation parameters are shown in Table 2. And the number of visible satellites for each terrestrial user is from 2 to 4, which means that each user has up to 3 candidate satellites except the serving satellite. Five user terminals are set up in Beijing with the locations following a uniform distribution. We use STK software to simulate the scene and obtain the geographical coordinates of each user. And STK is used to simulate LEO satellite constellation to obtain satellite operation data.

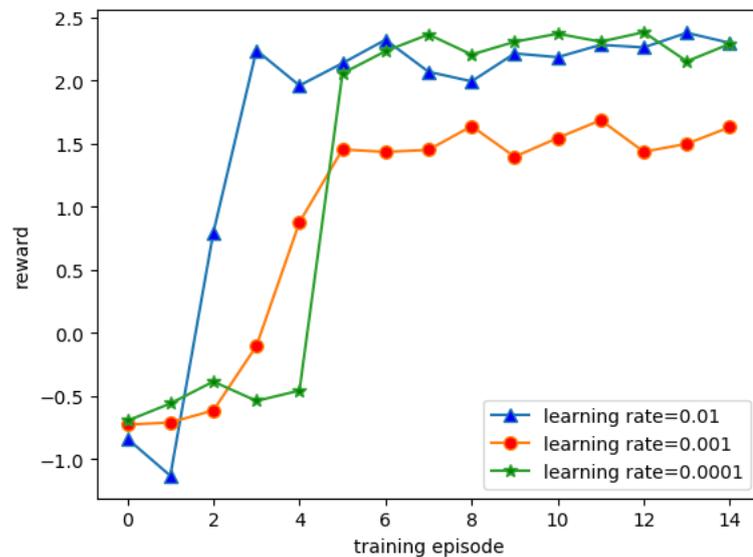
**Table 2.** Simulation parameters.

Parameter	Value
<b>Scene parameters</b>	
Constellation type	Sun-synchronous orbit
Orbital altitude	1000 km
Orbital inclination	99.4843 deg
Number of planes	12
Number of satellites per plane	9
Minimum elevation angle of user terminal	12 deg
Number of channels per satellite	200
On-board caching capacity for handover	1 GB
Time slot length	100 ms
Data rate	[500, 800] Mbit/s
<b>DRL Network parameters</b>	
Replay buffer	10,000
Observation size	3000
Minibatch size	200
Activation function	ReLU
Learning rate	0.01
Discount factor	0.9

In this paper, DQN is adopted, and Figure 8 shows the convergence of the DQN network loss function with the increase of training steps. It shows that the loss function converges when the number of training steps reaches about 4500.

In the simulation, the convergence characteristics of the reward value with different learning rates, which are 0.0001, 0.001, and 0.01, are compared and shown in Figure 9. Simulation results show that different learning rates will achieve different performances, and the convergence speed will also vary. It can be seen from Figure 9 that the convergence speed with learning rate 0.01 is the fastest, and the reward value after convergence has a small fluctuation range and high stability. The reward value with a learning rate of 0.01 is much higher than that, with a learning rate of 0.001. Therefore, the learning rate in this paper is set as 0.01.

**Figure 8.** Convergence of the DQN loss function.



**Figure 9.** Comparison of reward values under different learning rates.

Figure 10 compares the performance of four handover strategies in terms of handover failure rate. RST represents the traditional handover strategy based on the remaining service time (RST), which takes the maximum remaining service time as the selection criteria for selecting optimal candidate satellites, and handover is activated when the current serving satellite has no remaining service time. NIC is a traditional handover strategy based on the number of remaining idle channels (NIC), which takes the maximum number of remaining idle channels as the selection criteria for selecting candidate satellites, and handover is activated when there is no remaining service time. MAF is a multi-attribute fusion (MAF) decision based on the TOPSIS evaluation strategy. The strategy considers the remaining service time and the number of idle channels to select the optimal candidate satellite, and handover is activated when there is no remaining service time. IMF is the handover strategy of Intelligent multi-attributes fusion (IMF) decisions based on the DRL, which is proposed in this paper. It can be seen from Figure 10 that the handover failure rate increase with the raise of caching occupancy. Among the strategies, the RST strategy has the highest handover failure rate because it does not consider the remaining channel resources and the remaining caching capacity. Moreover, the growth trends of the handover failure rate of RST, NIC, and MAF are significantly improved when the caching occupancy is higher than 50%. When the caching occupancy is 90%, the handover failure rate is close to 100%. When the caching occupancy is 90%, the handover failure rate of IMF is close to 20%, which shows its performance gain over the referred strategies. Therefore, the proposed IMF handover strategy has the best performance with high caching occupancy.

Moreover, the performance of these strategies on call blocking rate is also compared. Here, the call blocking rate is defined as:

$$P_{block} = \frac{\sum_{slot=1}^T \frac{N_{block}}{N_{new}}}{T}. \quad (21)$$

Here,  $N_{new}$  new coming users arrive in the satellite coverage area and send an access request to the satellite. However, because the satellite does not have enough remaining channel resources for serving all the new coming users,  $N_{block}$  new coming users' access requests are rejected. In this paper, we consider the average call blocking rate in one hour under a high load rate (95% channel occupancy). The call blocking rate can reflect the flex-

ibility of the handover algorithm and the QoS of users in the coverage area. As shown in Figure 11, the call blocking rate of every handover strategy is increasing with the rise of user arrival rate. Among them, the new call blocking rate of RST is the highest because the channel state is not considered. The call blocking rate of the NIC is lower than that of the RST strategy. The MAF strategy considers the remaining service time and the number of idle channels, but the blocking rate is still high because it cannot flexibly adjust the connections of connected users. The IMF strategy can not only consider multiple factors but also flexibly configure the connected user link. If necessary, the connected user link can be reset to the other satellites with the idle channel resources, and it can greatly reduce the call blocking rate. The simulation results show that the call blocking rate of IMF is the lowest. Especially when the user arrival rate is 10, the call blocking rate of IMF is about 23% lower than that of the RST strategy.

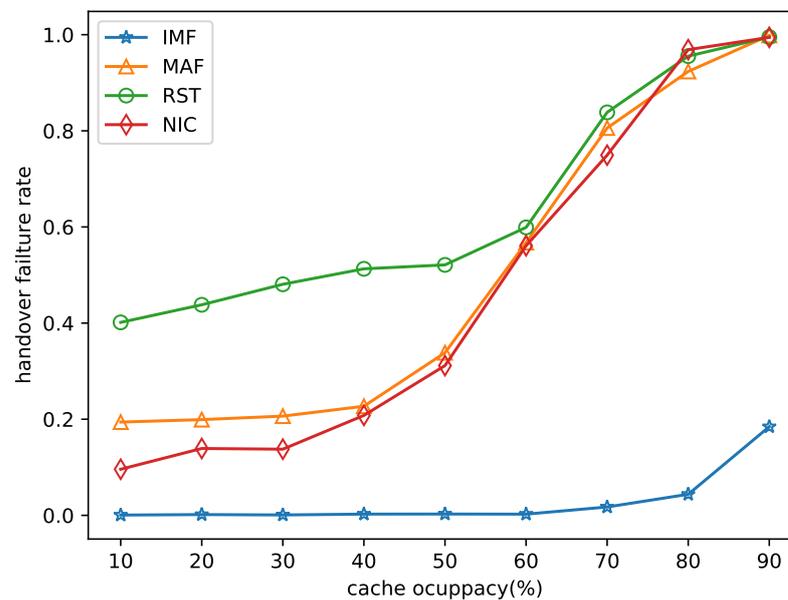


Figure 10. Performance comparison of handover failure rate.

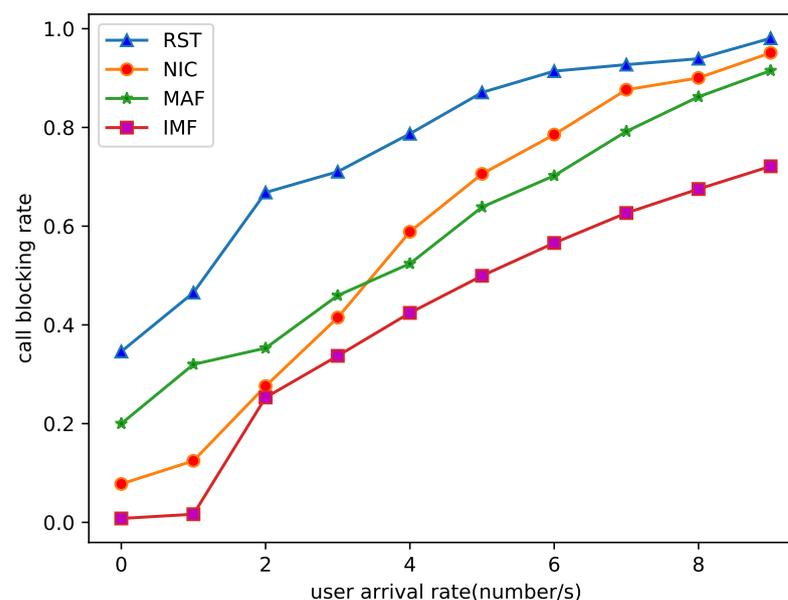


Figure 11. Performance comparison of call blocking rate.

## 5. Discussion and Conclusions

In this paper, a caching-aware intelligent handover strategy is proposed for the LEO satellite network. The strategy is different from existing handover strategies. First, the handover strategy focuses not only on the selection of the optimal candidate satellite for handover but also on the handover moment when it will be activated. That is to say the intelligent handover strategy needs to judge whether the handover can be activated at this moment and to select the candidate satellite. Secondly, the effects of multiple factors, including remaining service time, remaining idle channels, and remaining caching capacity, on the handover strategy are jointly considered. Then, DRL is adopted to make continuous intelligent handover judgment, which enables sequential decision to maximize long-term gains by interacting with the environment.

Via simulation, it is verified that the caching-aware intelligent handover strategy has a significant performance improvement in both the handover failure rate and call blocking rate. The performance of this strategy is compared with typical RST, NIC single-attribute handover, and MAF multi-attribute strategies. When the system caching occupancy is 10%, this strategy reduces the handover failure rate by nearly 40% compared to the RST handover strategy and nearly 20% compared to the MAF strategy. With the increasing of caching occupancy, the handover failure rate of the RST, NIC, and MAF handover strategies increase rapidly, while the caching-aware handover strategy proposed in this paper only has a failure probability of 18.3%. This shows that the handover strategy proposed in this paper can improve the quality of service for users in high caching occupancy scenarios. In addition, the system call blocking rates of each policy are compared for different user arrival rates. When the user arrival rate is high, the proposed intelligent handover strategy reduces the call blocking rate by 25% compared to the RST handover strategy and by 18.5% compared to the MAF handover strategy. Therefore, it can be concluded that the proposed strategy can effectively balance the system load, relieve the network pressure, reduce packet loss, and improve the quality of service for users.

Moreover, the DRL used in this strategy can make intelligent and continuous handover decisions to obtain maximum long-term gains. The complexity of the DRL mainly depends on the dimensions of state and action space, which is affected by the number of users, satellite and handover factors considered. Generally, the training process of DRL will cost long time to converge. Once the training process is finished, the complexity of decision process is low. Therefore, the training process can be executed on ground gateways with a large amount of computing resources, and the decision process can be implemented on-board. Moreover, artificial intelligence applied in satellite communication systems is still an open topic, and we will also investigate other light-weight models that could be utilized for resource management in satellite networks in our future works.

**Author Contributions:** Conceptualization, T.L., Y.X., G.C., and W.W.; methodology, T.L., Y.X., and G.C.; software, T.L. and Y.X.; validation, T.L., Y.X., and G.C.; formal analysis, T.L., G.C., and W.W.; investigation, T.L., G.C. and W.W.; resources, T.L., Y.X., and G.C.; data curation, T.L.; writing—original draft preparation, T.L., Y.X.; writing—review and editing, G.C.; visualization, T.L., Y.X., and G.C.; supervision, T.L., G.C., and W.W.; project administration, G.C. and W.W.; funding acquisition, G.C. and W.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China [No.61971054].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fourati, H.; Maaloul, R.; Chaari, L. A survey of 5G network systems: Challenges and machine learning approaches. *Int. J. Mach. Learn. Cybern.* **2020**, *12*, 1–47. [\[CrossRef\]](#)
2. Giambene, G.; Kota, S.; Pillai, P. Satellite-5G integration: A network perspective. *IEEE Netw.* **2018**, *32*, 25–31. [\[CrossRef\]](#)
3. Chen, S.; Sun, S.; Kang, S. System integration of terrestrial mobile communication and satellite communication—The trends, challenges and key technologies in B5G and 6G. *China Commun.* **2020**, *17*, 156–171. [\[CrossRef\]](#)
4. Kodheli, O.; Lagunas, E.; Maturo, N.; Sharma, S.; Shankar, B.; Montoya, J.; Duncan, J.C.M.D.; Spano, D.; Chatzinotas, S.; Kisseleff, S.; et al. Satellite Communications in the New Space Era: A Survey and Future Challenges. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 70–109. [\[CrossRef\]](#)
5. You, X.; Wang, C.; Huang, J.; Gao, X.; Zhang, Z.; Wang, M.; Huang, Y.; Zhang, C.; Jiang, Y.; Wang, J.; et al. Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts. *Sci. China Inf. Sci.* **2021**, *64*, 1–74. [\[CrossRef\]](#)
6. Portillo, I.D.; Cameron, B.G.; Crawley, E.F. A technical comparison of three low Earth orbit satellite constellation systems to provide global broadband. *Acta Astronaut.* **2019**, *159*, 123–135. [\[CrossRef\]](#)
7. Kaushal, H.; Kaddoum, G. Optical communication in space: challenges and mitigation techniques. *IEEE Commun. Surv. Tutor.* **2017**, *19*, 57–96. [\[CrossRef\]](#)
8. Musumpuka, R.; Walingo, T.M.; Smith, J.M. Performance analysis of correlated handover service in LEO mobile satellite systems. *IEEE Commun. Lett.* **2016**, *20*, 2213–2216. [\[CrossRef\]](#)
9. Papapetrou, E.; Karapantazis, S.; Dimitriadis, G.; Pavlidou, F.N. Satellite handover techniques for LEO networks. *Int. J. Satell. Commun. Netw.* **2004**, *22*, 231–245. [\[CrossRef\]](#)
10. Hu, X.; Song, H.; Liu, S.; Wang, W. Velocity-aware handover prediction in LEO satellite communication networks. *Int. J. Satell. Commun. Netw.* **2018**, *36*, 451–459. [\[CrossRef\]](#)
11. Duan, C.; Jing, F.; Chang, H.; Song, B.; Xu, Z. A novel handover control strategy combined with multi-hop routing in LEO Satellite Networks. In Proceedings of the 2018 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), Vancouver, BC, Canada, 21–25 May 2018.
12. Seyedi, Y.; Wu, Z. A simple real-time handover management in the mobile satellite communication networks. In Proceedings of the IEEE 2015 17th Asia-Pacific Network Operations and Management Symposium, Busan, Korea, 19–21 August 2015.
13. Zhou, J.; Ye, X.; Pan, Y.; Xiao, F.; Sun, L. Dynamic channel reservation scheme based on priorities in LEO satellite systems. *J. Syst. Eng. Electr.* **2015**, *26*, 1–9. [\[CrossRef\]](#)
14. Wu, Z.; Jin, F.; Luo, J.; Fu, Y.; Shan, J.; Hu, G. A graph-based satellite handover framework for LEO satellite communication networks. *IEEE Commun. Lett.* **2016**, *20*, 1547–1550. [\[CrossRef\]](#)
15. Li, Y.; Zhou, W.; Zhou, S. Forecast based handover in an extensible multi-layer LEO mobile satellite system. *IEEE Access* **2020**, *8*, 42768–42782. [\[CrossRef\]](#)
16. Li, J.; Xue, K.; Liu, J.; Zhang, Y. A user-centric handover scheme for ultra-dense LEO satellite networks. *IEEE Wireless Commun. Lett.* **2020**, *9*, 1904–1908. [\[CrossRef\]](#)
17. Wu, Y.; Hu, G.; Jin, F.; Zu, J. A satellite handover strategy based on the potential game in LEO satellite networks. *IEEE Access* **2019**, *7*, 133641–133652. [\[CrossRef\]](#)
18. He, S.; Wang, T.; Wang, S. Load-Aware satellite handover strategy based on multi-Agent reinforcement learning. In Proceedings of the 2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020.
19. Miao, J.; Wang, P.; Yin, H.; Chen, N.; Wang, X. A multi-attribute decision handover scheme for LEO mobile satellite networks. In Proceedings of the IEEE 5th International Conference on Computer and Communications, Chengdu, China, 6–9 December 2019.
20. Zhang, S.B.; Liu, A.J.; Liang, X.H. A multi-objective satellite handover strategy based on entropy in LEO satellite communications. In Proceedings of the IEEE 6th International Conference on Computer and Communications, Chengdu, China, 11–14 December 2020.
21. Xu, H.; Li, D.; Liu, M.; Han, G.; Huang, W.; Xu, C. QoE-driven intelligent handover for user-centric mobile satellite networks. *IEEE Trans. Veh. Technol.* **2020**, *69*, 10127–10139. [\[CrossRef\]](#)
22. Wang, S.; Zhang, X.; Zhang, Y.; Wang, L.; Yang, J.; Wang, W. A survey on mobile edge networks: Convergence of computing, caching and communications. *IEEE Access* **2017**, *5*, 6757–6779. [\[CrossRef\]](#)
23. Liu, S.; Hu, X.; Wang, Y.; Cui, G.; Wang, W. Distributed caching based on matching game in LEO satellite constellation networks. *IEEE Commun. Lett.* **2018**, *22*, 300–303. [\[CrossRef\]](#)
24. Qiu, C.; Yao, H.; Yu, F.R.; Xu, F.; Zhao, C. Deep Q-learning aided networking, caching, and computing resources allocation in software-defined satellite-terrestrial networks. *IEEE Trans. Veh. Technol.* **2019**, *68*, 5871–5883. [\[CrossRef\]](#)
25. Zhang, P.; Wang, X.; Ma, Z.; Song, J. Joint optimization of satisfaction index and spectrum efficiency with cache restricted for resource allocation in multi-beam satellite systems. *China Commun.* **2019**, *16*, 189–201.
26. Hasselt, H.V.; Guez, A.; Silver, D. Deep Reinforcement Learning with Double Q-learning. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016.