electronics

MDPI

Article

Machine Learning-Based Solutions for Handover Decisions in Non-Terrestrial Networks

Mwamba Kasongo Dahouda , Sihwa Jin and Inwhee Joe *

Department of Computer Science, Hanyang University, Seoul 04763, Republic of Korea
* Correspondence: iwjoe@hanyang.ac.kr

Abstract: The non-terrestrial network (NTN) is a network that uses radio frequency (RF) resources mounted on satellites and includes satellite-based communications networks, high altitude platform systems (HAPS), and air-to-ground networks. The fifth generation (5G) and NTN may be crucial in utilizing communication infrastructure to provide 5G services in the future, anytime and anywhere. Based on the outcome of the Rel-16 study, the 3rd generation partnership project (3GPP) decided to start a work item on an NTN in 5G new radio (NR) Rel-17, and the focus of the study was on mobility management procedures, due to the movements of NTN platforms; especially, low earth orbit (LEO) satellites. Handover enhancements were discussed to tackle the frequent handover due to the fast satellite movement. Therefore, the major problem of handover in LEO satellite systems was the signaling storm created by handing over all user equipment (UE) in a cell to a new cell because when the UE crosses the boundary between the neighboring cell of a satellite, an intra-satellite or cell handover occurs; thus, all users in a cell are expected to experience a change of cell due to handover every few seconds. In addition, UE location is not easy to define due to moving cell/beam situations. In this study, we propose machine learning-based solutions for handover decisions in non-terrestrial networks for cell handovers or intra-satellite handovers to reduce signaling storms during handovers where the handover requests will be executed by clustered users. First, the dataset was generated by the simulator that simulates communication between users and satellites. Second, we preprocessed the data, and also used the feature creation technique to create the distance feature using the Haversine formula, and then applied clustering and classification algorithms. The experimental results show that the distance between a user and its cell center is an important parameter for handover decisions in NTN, and the random forest outperforms all models with a higher accuracy of 99% along with a better F1-score of 0.9961.

Keywords: non-terrestrial network (NTN); conditional handover (CHO); machine learning (ML); clustering

Citation: Dahouda, M.K.; Jin, S.; Joe, I. Machine Learning-Based Solutions for Handover Decisions in Non-Terrestrial Networks. *Electronics* **2023**, *12*, 1759. https://doi.org/10.3390/electronics12081759

Academic Editors: Ping-Feng Pai and Paul Mitchell

Received: 27 February 2023
Revised: 1 April 2023
Accepted: 5 April 2023
Published: 7 April 2023

1. Introduction

A traditional mobile network is a communication network where the link to and from end nodes is wireless, and the network is distributed over land areas called "cells", each served by at least one fixed-location transceiver. Currently, it is getting harder to offer the necessary quality of service (QoS) criteria [1–3], especially for users with high mobility, due to the state of technology and the volume of traffic consumed by users [4]. Mobility management is one of the major functions of network communication in traditional cellular networks or terrestrial networks (TN) and non-terrestrial networks (NTN), and mobility management aims to track where the users are, allowing calls and other user services to be delivered to them. Therefore, one of the main problems for these users is the proper initiation of the handover. Many wireless system issues, including decision-making, resource optimization, and network administration, can be presented in a way that is ideal for machine learning (ML) [5–8]; through intelligent adaptive learning and decision-making, ML can be utilized to improve radio communication. Since applications for NTNs include telecommunications, navigation, and remote sensing, these networks can provide coverage to remote or hard-to-reach areas and can also be used as a backup to terrestrial networks in case of failure or disaster. In Release 16 of the 3rd generation partnership project (3GPP) specification, the first 5G wireless technology in the world was standardized. In Release 17, work is being conducted to create 5G wireless technology to support non-terrestrial satellite networks [9]. Massive amounts of data are being produced while 5G networks are growing more sophisticated. To effectively manage the 5G networks, data-driven management, and artificial intelligence (AI) solutions are essential. Research on AI-enabled NG-RAN was carried out by 3GPP for Version 17; this study examined the high-level principles, functional framework, potential use cases, and related solutions for AI-enabled RAN intelligence [10]. For a selection of AI-based use cases, including network energy conservation, load balancing, and mobility optimization, 3GPP Release 18 will specify data collection upgrades and signalling support. Enabling 5G systems to support non-terrestrial networks (NTN) is one of the research areas of the 3GPP, with a focus on common scenarios where non-terrestrial networks provide access to user equipment (UE). Spaceborne platforms such as low earth orbiting (LEO), medium earth orbiting (MEO), geosynchronous earth orbiting (GEO), and high elliptical orbit (HEO) satellites are used by satellite communication networks [11]. Another unmanned aerial system (UAS) is the high-altitude platform station (HAPS) that can potentially be used to provide both fixed broadband connectivity for end-users and transmission links between the mobile and core networks used for backhauling traffic.

1.1. Non-Terrestrial Networks Overview

This section is very important because it gives a deep understanding of the topic studied, which is how to include satellite communication in new radio 5G communication (NR 5G). Since the goal of the 3GPP in Rel-17 was to specify the improvements found for NR NTN, a focus was placed on LEO and GEO as well as implied compatibility to accommodate high-altitude platform stations and air-to-ground scenarios. In addition to enhancing current data collection features, 3GPP Release 18 will look at how AI approaches might increase air interface capabilities. Moreover, 3GPP Release 18 will improve NR data collecting in the context of the self-organizing network (SON) by automating RAN planning, configuration, management, optimization, and healing, SON reduces the need for human involvement [12]. Minimization of drive testing (MDT) gives operators the ability to set up typical UEs to gather and submit measurement data, reducing the need for traditional drive testing. The work on Release-18 will take care of SON features left over from Release 17 and data gathering for random access channel (RACH) optimization. Therefore, the process in this research includes problem analysis, which means identifying some main challenges in non-terrestrial networks and deep comprehension of mobility management. A non-terrestrial network is a network or network segment that uses radio frequency (RF) resources mounted on satellites and includes satellite communications networks, high altitude platform systems (HAPS), and air-to-ground networks as described in Figure 1. Typically, a non-terrestrial network includes the following components: A GEO satellite is supplied by one or more sat-gateways that are installed throughout the satellite targeted coverage; one or more sat-gateways that connect the non-terrestrial network to a public data network (e.g., regional or even continental coverage) [11]. We assume that each UE in a cell only has access to one satellite gateway. A non-GEO satellite is subsequently connected to one or more satellite gateways at once. The system makes sure that there is enough time between each serving satellite to complete mobility anchoring and handover while maintaining service and feeder connection continuity. Both transparent and regenerative (with onboard processing) payloads are possible for satellite implementation. The satellite often produces many beams over a service region that is bounded by its field of view [13]. The beam imprints are typically elliptical. In NTN, we have three different sorts of links: a feeder connection, also known as a radio link, connects a satellite to a ground station, while a service link connects user equipment to the satellite. Inter-satellite links (ISL), which are additional links if there is a constellation of satellites, are optional. Regenerative payloads on the satellites are necessary for this. ISL may function in optical or RF frequency bands.

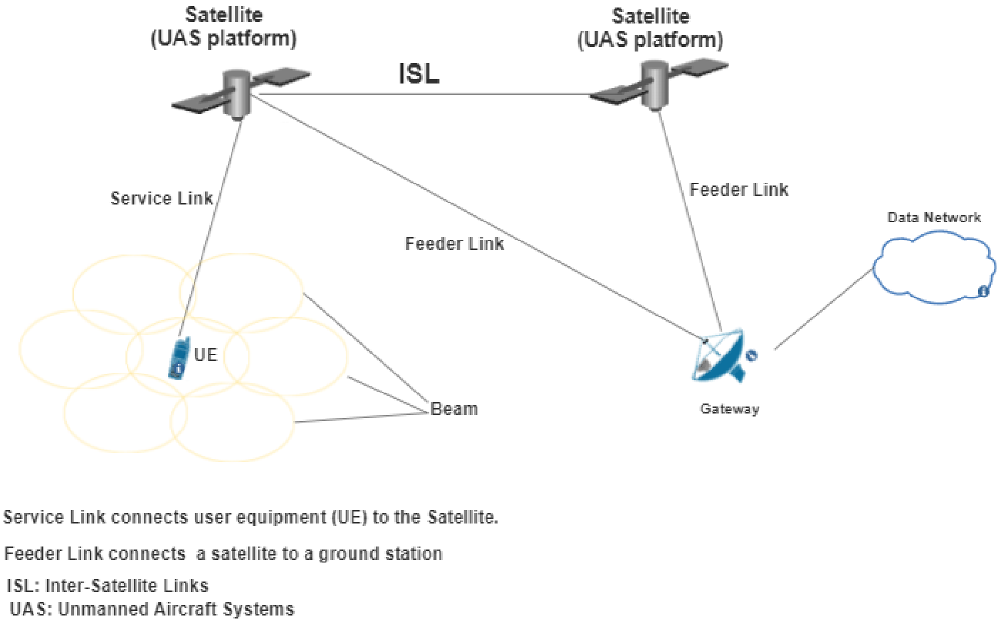


Figure 1. Non-terrestrial network typical scenario based on regenerative payload.

While a constellation of LEO and MEO satellites is utilized to provide services in both the Northern and Southern hemispheres, GEO satellites and UAS are employed to provide continental or regional service. In certain instances, the constellation may even offer coverage of the entire world, including the polar regions. In contrast to the GEO system, the LEO satellite system has many advantages such as efficient bandwidth usage, low propagation delay, and low power consumption at user terminals and satellites. However, in contrast to the GEO satellite system, the coverage area of LEO satellites is not constant. This is because satellites move asynchronously to the earth, handing them off between ground stations as they pass through different regions of the earth. Therefore, mobility management for LEO satellite systems is more difficult than for GEO systems. In some LEO satellite systems, the satellites communicate with each other using inter-satellite links (ISL).

Mobility Management

In mobile communication systems, mobility management allows the control and coordination of the movement of mobile devices in a wireless network. It involves a set of protocols and mechanisms that enable mobile devices to connect and disconnect from the network, to move between different network cells or access points, and to maintain ongoing communication sessions while in motion. In addition to handling handovers, mobility management also entails controlling the identity and location of mobile devices. Mobility management seeks to maintain communication between mobile devices as they move in the network, while simultaneously making effective use of available resources and limiting the impact on overall network performance. According to the 3GPP, for LEO NTN, mobility management procedures should be improved to take into account satellite movement-related factors such as measurement validity, user equipment (UE) velocity, movement direction, large and varying propagation delay, and dynamic neighbor cell set [13]. For GEO NTN, mobility management procedures need to be adjusted to accommodate large propagation delays. The mobility management in the non-terrestrial network is very different from the traditional network or terrestrial; In NTN, based on the architecture, we have many types of handover such as inter-satellite handover, intra-satellite handover, and inter-access network handover. However, in our study, we focus on an intra-satellite handover; in our proposed architecture, both cells/beams are served by the same satellite, and no other satellite is involved in the handover process. In LEO satellite systems, intra-satellite handovers are the most typical kind of handovers encountered because of the small area covered by beams and the rapid satellite speed. Thus, we can consider the user mobility negligible compared to high satellite speed. Mobility management of LEO satellites is therefore much more challenging than GEO or MEO systems. With a few exceptions, terrestrial network systems and LEO satellite systems have somewhat comparable mobility. In both systems, the relative position between the cells and the UE changes continuously, requiring the handover of the UE between adjacent cells. In terrestrial network systems, the UE moves through the cells, while in LEO systems the cells move through the UE. The cell size of LEO satellite systems is larger compared to terrestrial network systems. Moreover, the speed of the UE can be ignored in LEO satellite systems, since that speed is negligible compared to the rotational speed of the LEO satellite.

Conditional handover (also known as “soft handover”) is a type of handover process that occurs in cellular networks. In traditional (or "hard") handover, a mobile device is disconnected from one cell and connected to another cell before the call is resumed. In conditional handover, the mobile device simultaneously maintains connections to both the current and the target cells for a short period of time, allowing the system to evaluate the quality of the new connection before committing to the handover. This can help to minimize the interruption of the call and increase the chances of a successful handover.

Conditional handover is typically used in cellular systems, but the studies on solutions for adapting NR to support NTN documented in 3GPP TR 38.821 [13] defined 5 types of conditional Handover triggering methods such as (1) measurement-based triggering, (2) location (UE and Satellite) triggering, (3) time(r)-based triggering, (4) timing advance value-based triggering, and (5) elevation angles of source and target cells-based triggering. In non-terrestrial networks (NTNs), conditional handover is a technique used to manage the handover of mobile devices between different network nodes while maintaining a seamless communication session. In an NTN, the network architecture and infrastructure are different from traditional terrestrial networks, which can make handover more challenging. For example, in a satellite-based NTN, a mobile device may need to be handed over between different satellites or between a satellite and a ground station. Moreover, cells in terrestrial networks are stationary, although UE may be movable along various trajectories. As a result, to choose the appropriate target cell for each UE, the network needs the measurement report from the UE. The situation is very different in NTN, particularly for LEO satellites where the cells/beams are moving over time, and the high speed of LEO satellites will result in frequent handovers and high handover rates of a large number of UE. We considered the simulation in the earth-moving beams scenario; in NTN we have earth-fixed cells/beams and earth-moving cells; however, in our study, we did not consider earth-fixed beams. Therefore, we considered earth-moving cells where cells are moving on the ground; therefore, the NTN platforms and the beams are all moving but not at the same speed. One of the most important steps for the Handover is the measurement report from UE before the handover as shown in Figure 2. In NTN, the propagation delay is orders of magnitude more than in terrestrial systems, adding extra latency to mobility signaling activities including measurement reporting, HO command receiving, and HO request/ACK. As shown in Figure 2, the interruption time for the downlink can be calculated as the period of time between the network transmitting RRCReconfiguration with sync (Step 3) and the Target gNB receiving RRCReconfigurationComplete (Step 6). After step 3, the gNB is unable to communicate any further data; however, it can proceed once it has received the signal RRCReconfigurationComplete. The UE may continue sending data to the source gNB until RRCReconfiguration with sync is received (Step 6). Note that, the gNB is a 3GPP 5G Next Generation base station that supports the 5G New Radio.

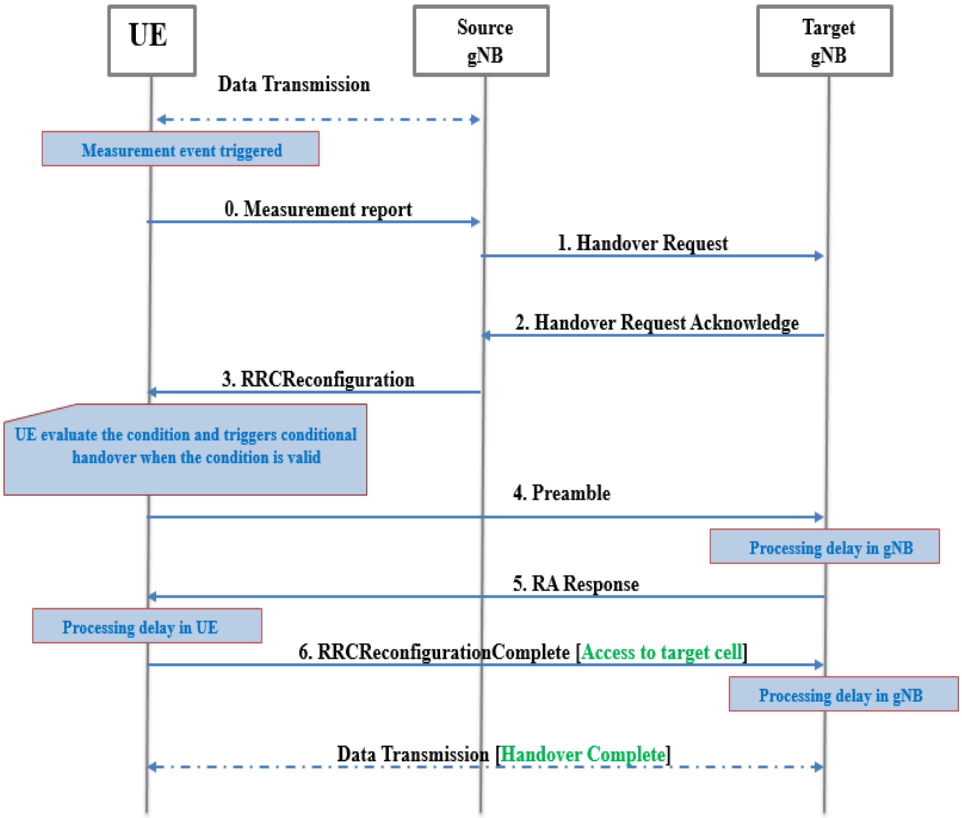


Figure 2. Procedure for Conditional Handover.

As a kind of solution, 3GPP defines several sets of the predefined set of measurement report mechanisms to be performed by UE. This predefined measurement report type is called Event, such as Event A3, which means the neighbor becomes offset better than serving, or A4, the neighbor becomes better than the threshold [13]. When certain conditions are met, the new handover procedure introduced in Release-16 by 3GPP allows the user equipment (UE) to decide whether to perform the handover or not. The previous handover procedure, which was reactive and prone to handover failures, left it up to the network to decide whether to perform the handover. Therefore, the 3GPP introduced the conditional handover (CHO) feature in 5G-NR Release-16, allowing UE to choose whether to conduct a handover when specific criteria are satisfied. Ideally, the fundamental handover process is still the same as it was in the basic handover procedure: UE transmits the measurement report to the source cell along with the neighbor cell PCI and signal strength (usually the reference signals received power), source cell decides to begin the handover procedure to best target cell, and target cell completes the handover operation. The conditional handover (CHO), as shown in Figure 2, is a handover that the UE performs when one or more handover execution conditions are satisfied. When receiving the CHO configuration, the UE begins evaluating the execution condition(s), and after a handover is carried out, the UE ceases analyzing the execution condition(s). In a conditional handover, the mobile device will maintain connections to both the current and the target network nodes for a short period of time, allowing the system to evaluate the quality of the new connection before committing to the handover. This can help to minimize the interruption of the call and increase the chances of a successful handover. Thus, machine learning (ML) algorithms can also be used to optimize the handover process by adapting to the specific characteristics of the network and mobile devices. For example, ML algorithms can learn to adjust the handover threshold and the timing of the handover based on the type of application and the mobility of the mobile device.

2. Related Work

Machine learning (ML) can be used to improve the performance and efficiency of the handover process in wireless networks by providing more accurate and efficient handover decisions. ML-based handover algorithms can be trained to analyze data on the network conditions and the requirements of the mobile device to make informed decisions about when and how to perform a handover. For instance, ML algorithms can be trained to predict the quality of the new connection before committing to the handover, which can help to minimize the interruption of the call and increase the chances of a successful handover. Many researchers have studied handover technique in new generation wireless networks [14–20], ranging from traditional techniques such as multi-attribute decision making (MADM) [21] to deep learning techniques such as the SINR change of a UE in the handover problem in 5G networks [22]. In [22–26], the authors combined their efforts to address two issues and proposed a learning-based load balancing handover for multi-user mobile mmWave networks where they characterized the user association as a non-convex optimization problem, and then they attempted to approximate the optimization solution of the problem by using the deep deterministic policy gradient (DDPG) method. Traditional cell-to-cell handovers rely on measurements of target cell radio strength, which necessitates frequent measurement gaps. Therefore, a handover solution based on machine learning was proposed in [27–29]; also, prediction-based handover strategies are recommended to reduce the number of measurement windows as in [30], an ultrafast and effective XGBoost-based predictive handover technique for next-generation mobile communications was presented. Another fascinating study about conditional handover was proposed in [31]. In this study, the author presents a novel prediction-based CHO (PCHO) scheme that employs deep-learning technology to address the drawbacks of CHO and make better informed preparatory decisions. All of the above-mentioned handover strategies are related to terrestrial networks (TNs); however, in NTNs, the network architecture and infrastructure are dissimilar from those of conventional terrestrial networks, which might make handover more challenging. Moreover, an in-depth analysis of the current status of onboard artificial intelligence and machine learning (AI/ML) processing is provided in [32], along with a list of important factors to take into account while implementing it in communication satellites. Also, it offers a helpful framework for comparing and assessing the feasibility of various AI chipsets for onboard AI/ML processing. However, our proposed architecture is for a regenerative onboard satellite payload; therefore, the machine-learning model is being trained offline. On-device ML can be utilized to execute inference with models directly on a device.

3. Proposed Method

3.1. NTN Simulator and Data Description

The simulator was designed to simulate the communication between satellites and UE through service links. It was developed using the C-sharp Net Framework and STK 11.3 engines. The STK engine is software for satellite orbit generation and communication analysis between satellites and UEs. The satellite is an LEO satellite with an altitude of 300 km and operates at a speed of 7.56 km/s.

The LEO satellite operated by the simulator owns three cells, and the diameter of each cell is set at 250 km; in addition, 62,908 users were placed randomly in the three cells. Due to the distinctive features of non-terrestrial networks, such as high mobility, wide coverage areas, and constrained network resources, handover procedures might be more complicated in an NTN. Additionally, the 3GPP study on NR support for non-terrestrial networks did not yet include a dataset for handover; as a result, in this study, we developed an NTN simulator for handover to design and simulate the behavior of user equipment and satellite in non-terrestrial communication networks. As shown in Figure 3, we designed an intra-satellite handover scenario between overlapped satellite beams; the satellite and UE are connected directly through a service link. Note that, here the $C1_{UE}$, $C2_{UE}$, and $C3_{UE}$ denote different clusters and each cluster contains many users.

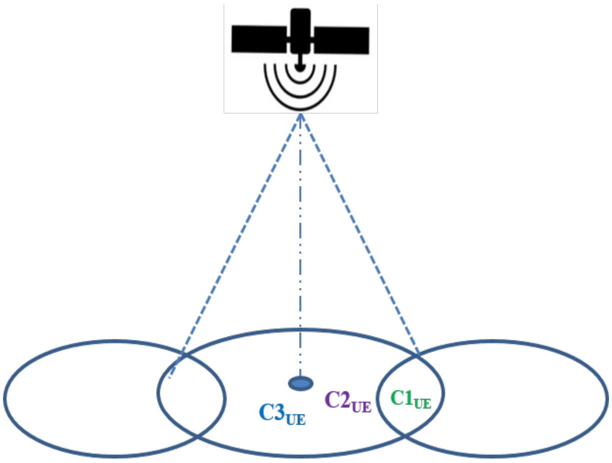
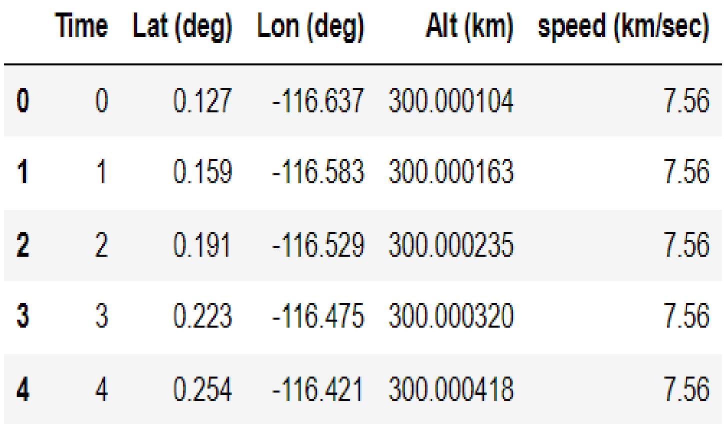


Figure 3. The proposed intra-satellite handover architecture.

3.1.1. LEO-Satellite Data Collection

We have gathered details about UE and LEO satellites. As seen in Figure 4, we selected the latitude, longitude, and altitude (LLA) data from the start (0 s) and end (120 s = 2 min) of the simulation for the LEO satellite. The speed of the LEO-Satellite was 7.56km/s, whereas the altitude was 300 km.



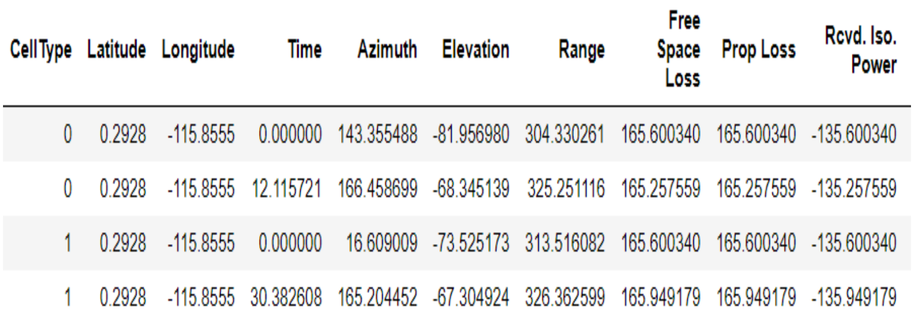
	Time	Lat (deg)	Lon (deg)	Alt (km)	speed (km/sec)
0	0	0.127	-116.637	300.000104	7.56
1	1	0.159	-116.583	300.000163	7.56
2	2	0.191	-116.529	300.000235	7.56
3	3	0.223	-116.475	300.000320	7.56
4	4	0.254	-116.421	300.000418	7.56

Figure 4. LEO-Satellite Features: latitude, longitude, and altitude (LLA) data from the NTN Simulator.

3.1.2. User Equipment Data Collection

To better understand the simulation of NTN communication between UE and Satellite, we collected important data about the user used to track all the UE in different cells.

Figure 5 shows the first ten features of the user data after 120 s of simulation. The Dataset has 62,908 rows and 20 columns; Table 1 describes the details of each feature.



CellType	Latitude	Longitude	Time	Azimuth	Elevation	Range	Free Space Loss	Prop Loss	Rcvd. Iso. Power
0	0.2928	-115.8555	0.000000	143.355488	-81.956980	304.330261	165.600340	165.600340	-135.600340
0	0.2928	-115.8555	12.115721	166.458699	-68.345139	325.251116	165.257559	165.257559	-135.257559
1	0.2928	-115.8555	0.000000	16.609009	-73.525173	313.516082	165.600340	165.600340	-135.600340
1	0.2928	-115.8555	30.382608	165.204452	-67.304924	326.362599	165.949179	165.949179	-135.949179

Figure 5. User Equipment Data Sample Description.

Table 1. User Equipment Features.

Feature	Description
CellType	Serving Cell or Target Cell, or Candidate Cell
Latitude [deg]	Latitude of a specific UE in degrees
Longitude [deg]	Longitude of a specific UE in degrees
Time [s]	Time the UE stays in the cell
Azimuth [deg]	Azimuth between satellite and UE
Elevation [deg]	Elevation between satellite and UE
Range [km]	Distance between satellite and UE
Free Space Loss [dB]	Loss due to propagation through free space
Prop Loss [dB]	The total propagation loss computed across all enabled propagation models
Rcvd. Iso. Power [dBW]	Received isotropic power is the power at the receiver before the pre-receive gains/losses and the receiver antenna gain added (in dBW).
Carrier Power at Rcvr Input [dBW]	Carrier Power at Rcvr Input is the power at the receiver after the receiver antenna gain is added (in dBW).
Flux Density [dBW/m^2]	The power from the desired transmitter crossing a unit area normal to the direction of wave propagation.
g/T: G/T = (Receiver Gain)/(System Temperature at the Receiver) [dB/K]	The ratio of the receive antenna gain G to the total system temperature T is the "figure of merit" for the receiver (in dB/K).
C/No [dB∗MHZ]	The carrier-to-noise density ratio (C/No) where C is the carrier power and No = kT (Boltzmann's constant × system temperature) is the noise density. It is equivalent to C/N with a normalized Bandwidth (B = 1).
C/N [dB]	The carrier-to-noise ratio (C/N) where C is the carrier power and N = kTB (Boltzmann's constant × system temperature × bandwidth) is the noise power
Eb/No [dB]	The energy per bit to noise ratio (Eb/No) where Eb is the energy per bit and No = kT (Boltzmann's constant × system temperature)
Propagation Delay [s]	The amount of time required for a signal to propagate through the physical link medium. This will vary depending on the propagation distance and the type of medium
Propagation Distance [km]	The distance of the physical link medium between a transmitter and a receiver for which a signal will travel
Spectral Flux Density [dBW $\times$ m^2 $\times$ Hz^{-1}]	The dimensions are watts/(m^2 $\times$ Hz). The power is computed across the receiver's bandwidth (as seen by the receiver's RF front end). The bandwidth is the receiver's total bandwidth

Thus, one of the crucial features of the UE is the location provided during the simulation, and we also provided details of each feature about the information of the UE.

3.2. Proposed Clustering Algorithm

Machine learning algorithms improve the performance and efficiency of the handover process in wireless networks by providing more accurate and efficient handover decisions, optimizing the handover process, and mobility management. However, it is worth noting that the implementation and deployment of ML-based handover algorithms in NTN can be complex and require a significant amount of data and computational resources. Therefore, we propose ML-based solutions for Handover decision-making using the K-Means clustering algorithm, an unsupervised machine learning algorithm, so that UE that has been grouped due to particular similarities are in the same cluster. This algorithm calculates the distance between each UE and its cluster center in each iteration [33]. Moreover, ML-based algorithms can also be used to improve mobility management in NTN by predicting the location of the mobile and providing accurate handover decisions. In this study, the major clustering algorithm component uses a two-step procedure known as expectation maximization. Consequently, each UE can be paired with its closest centroid in the expectation stage; then the new centroid is set after the maximizing step computes the mean of all the UE for each cluster. The k-means algorithm is described below in Algorithm 1.

Algorithm 1: UE k-means algorithm.

```
Read: k                    // Specify the number k of clusters to assign
Randomly initialize k centroids
repeat
    expectation                 // Assign each UE to its closest centroid
    maximization            // Compute the new centroid of each cluster
until The centroid positions do not change
```

As shown in Algorithm 1, specifying the number of clusters is very important; using the elbow method, we can find the optimal number for the clustering algorithm.

Optimization of the Number of Clusters with the Elbow Method

The elbow technique compares the percentage of the number of clusters that will form an elbow at a given moment to the appropriate number for the clustering algorithm. This technique uses visual analysis to evaluate the consistency of the optimal number of clusters [34]. As presented in Figure 6, the elbow approach plots within-cluster sum squares (WCSS) values on the Y-axis and the number of clusters on the X-axis to create the graphic using the WCSS concept. Therefore, we calculated the value of WCSS for various k values between 1 and 10, as shown in Equation (1).

$$WCSS = \sum_{C_k}^{C_n} \left(\sum_{d_i in C_i}^{d_m} distance(d_i, C_k)^2 \right), \quad (1)$$

where C is the cluster centroids and d is the UE in each cluster.

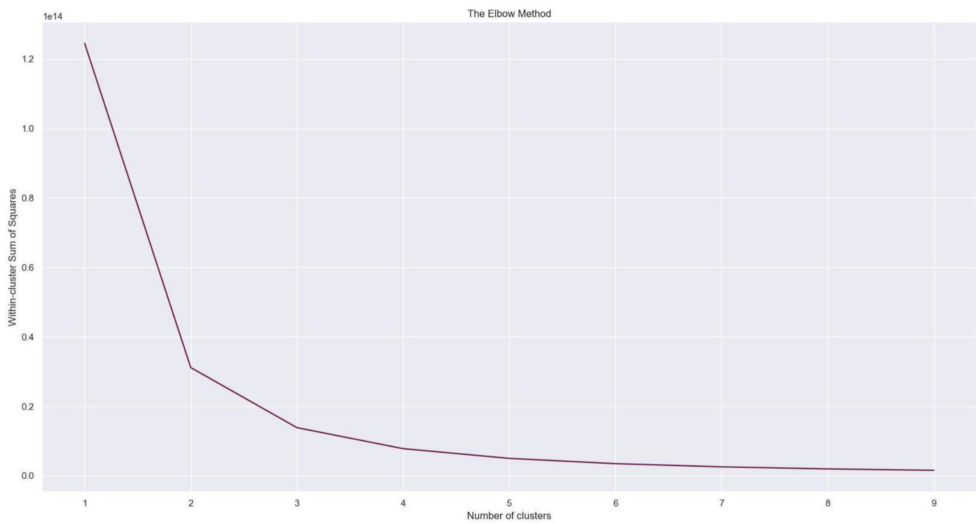


Figure 6. Elbow Method Graph.

3.3. Proposed Classification Algorithm

The triggering of measurement reporting is used based on the location of the UE, which means that in a specific cell, users that are close to the cell center will not have the same received signal strength (RSS) as those who are at the edge of the cell, as described in Figure 7.

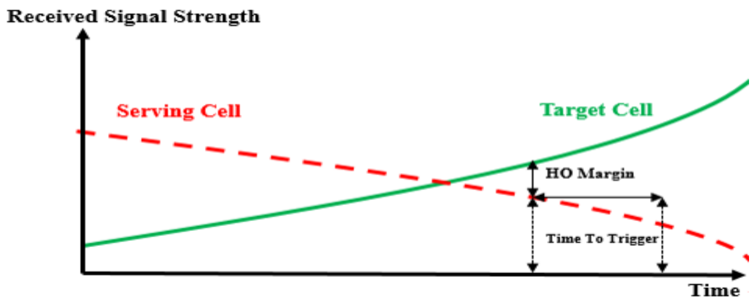


Figure 7. A sketch of near-far effect in different scenarios in NTN.

In NTN, there is just a simple variation in signal strength between two beams in an overlapped region. The users have trouble telling which cell has a better RSS because the handover procedure is subject to measurement events. In NTN, due to the movement of the satellite and the beams, a potentially large number of UE must perform HO at a given time [13], which could result in significant signaling overhead and service continuity issues depending on the speed of the satellite, propagation latency, and UE density.

As shown in Figure 8, while the satellite and the beams are moving, some users are handing out from the area no longer served by the cell, and others are also handing in. In that specific scenario, we have two types of users: the ones who are ready to hand over; and those who are not to perform the handover. During the feature engineering, we created two additional features called distance and handover-ready based on the dataset produced. Since we know the centroid of each cluster, we can use the Haversine formula to determine how far the UE is from the center of the cell. As described below, the haversine formula determines the great-circle distance between two points on a sphere knowing their longitudes and latitudes. Let the central angle θ between any two points on a sphere be:

$$\theta = \frac{d}{r}, \tag{2}$$

where d is the distance between the two UE along a great circle of the sphere, and r is the radius of the sphere. With the help of the haversine formula, it is possible to calculate the haversine of θ directly from the latitude and longitude of the two points:

$$hav(\theta) = hav(\varphi_2 - \varphi_1) + \cos\varphi_1 \cos\varphi_2 hav(\lambda_2 - \lambda_1), \tag{3}$$

Alternatively, to avoid employing cosines, which degrade resolution at small angles.

$$hav(\theta) = hav(\varphi_2 - \varphi_1) + (1 - hav(\varphi_1 - \varphi_2) - hav(\varphi_1 + \varphi_2)) \cdot hav(\lambda_2 - \lambda_1) \tag{4}$$

where (φ_1, φ_2) are the latitude of point 1 and latitude of point 2 whereas (λ_1, λ_2) are the longitude of point 1 and longitude of point 2. Here, we consider the coordinate of point 1 as the cell center coordinate, which is the centroid of the cluster, whereas point 2 is the coordinate of a specific UE. Finally, the haversine function $\mathrm{hav}(\theta)$, applied above to both the central angle θ and the differences in latitude and longitude, is as follows:

$$hav(\theta) = \sin^2(\frac{\theta}{2}) = \frac{1 - \cos\theta}{2} \tag{5}$$

The haversine function computes half a versine of the angle θ. To solve for the distance d, apply the archaversine (inverse haversine) to $h = hav(\theta)$ or use the arcsine (inverse sine) function:

$$d = r \times archav(h) = 2r \arcsin(\sqrt{h}), \tag{6}$$

or more explicitly:

$$d = 2r \arcsin(\sqrt{hav(\varphi_2 - \varphi_1) + (1 - hav(\varphi_1 - \varphi_2) - hav(\varphi_1 + \varphi_2)) \cdot hav(\lambda_2 - \lambda_1)}) \tag{7}$$

$$= 2r \arcsin(\sqrt{\sin^2(\frac{\varphi_2 - \varphi_1}{2}) + (1 - \sin^2(\frac{\varphi_2 - \varphi_1}{2}) - \sin^2(\frac{\varphi_2 + \varphi_1}{2})) \cdot \sin^2(\frac{\lambda_2 - \lambda_1}{2})}) \tag{8}$$

$$= 2r \arcsin(\sqrt{\sin^2(\frac{\varphi_2 - \varphi_1}{2}) + \cos\varphi_1 . \cos\varphi_2 \cdot \sin^2(\frac{\lambda_2 - \lambda_1}{2})}) \tag{9}$$

Based on the generated dataset, during the feature engineering process, we created two more features, the distance, and handover-ready. Here, the distance is the distance between two points on a sphere given their longitudes and latitudes; having the centroid of each cluster or cell, we calculated the distance between all the UE and their cell center. The handover ready feature is a binary feature (1 or 0); we assumed that UE close to the cell center ($C3_{UE}$ *and* $C2_{UE}$) is not ready for handover (we assigned zero as a value to them) those who are close to the edge of the cell $C1_{UE}$ are ready for handover (we assigned one as value to them) as shown in Figure 8. As a result, we have a binary classification problem, and then we applied a machine learning algorithm to predict if a specific UE is ready for handover or not in Non-Terrestrial Networks.

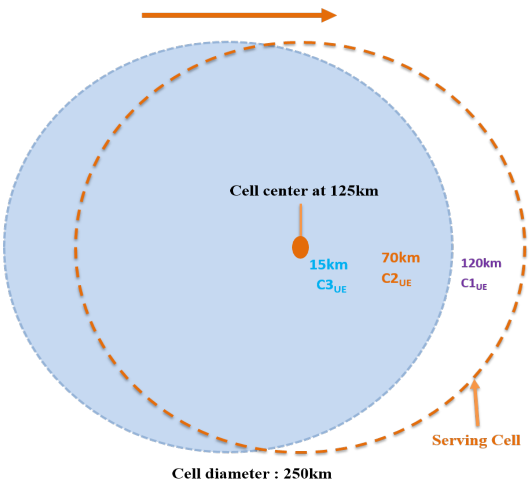


Figure 8. At a given time: UE ready for Handover are far from the cell center ($C1_{UE}$), and UE not ready for Handover are close to the cell center ($C3_{UE}$ and $C2_{UE}$).

4. Experimental Results

4.1. Performance Result of the Proposed Clustering Algorithm

To visualize the location of all the 62,908 UE, we selected three features, the type of cell, the latitude, and the longitude of each UE, as depicted in Figure 9.

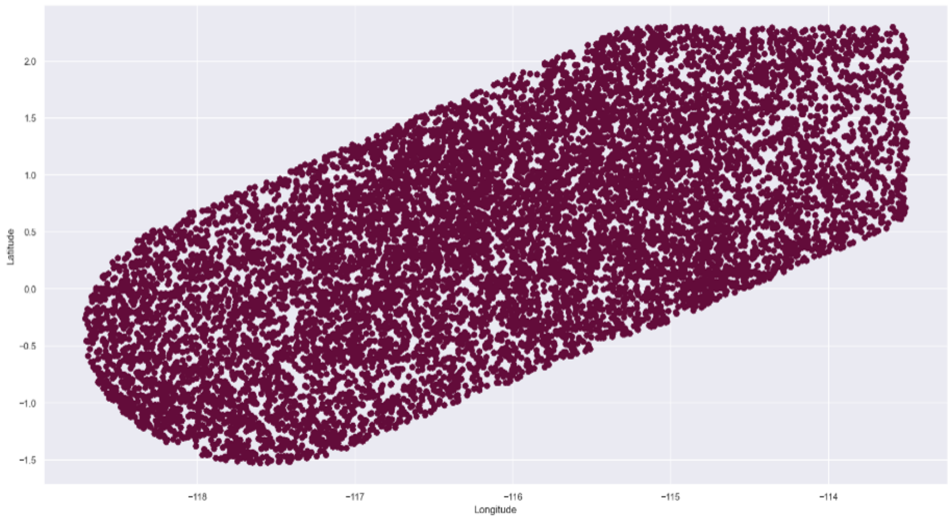


Figure 9. Location of User Equipment based on their latitude and longitude.

Moreover, the elbow technique shown in Figure 6 explained the number of clusters on the X-axis and the WCSS values on the Y-axis to produce the visual utilizing the WCSS idea.

The elbow point is at 3, as shown in Figure 6; thus, we selected three as the optimal number of clusters because our dataset indicates three different types of cells. Since we use unlabeled data in unsupervised learning, consequently, models must identify patterns or similarities in the data. In our NTN simulator, we designed the intra-satellite handover architecture with one satellite and three cells (the candidate cell, serving cell, and target cell). Therefore, the satellite-based navigation system provides location information to users with Global Positioning System (GPS) receivers. Figure 10 shows the cluster of the 62908 users according to their location.

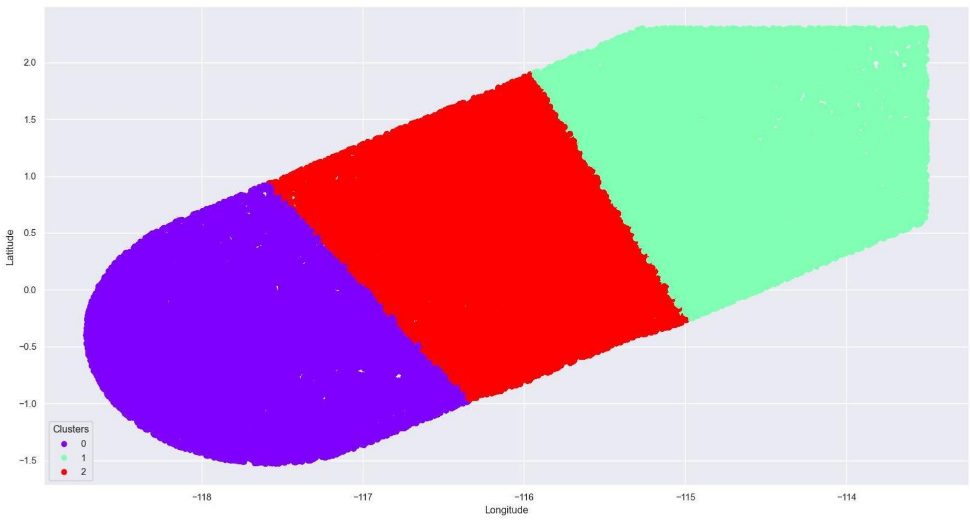


Figure 10. UE clustered based on their location: Cell type, latitude, and longitude.

One of the issues in [11] is the signaling storm brought on by moving all of the UE in a cell to a different cell. As a result, it envisaged that every few seconds, all users in a cell will switch to another cell due to handover. UE location is a crucial factor to consider while executing handover because 3GPP has defined location (UE and Satellite) triggering as one of the conditional handovers; this is why we developed the UE-based clustering approach by using the centroid-based method.

The clusters in the centroid-based technique are created by how close the UE is to the centroid. Here, the centroid of the cluster is calculated with the least possible distance from the UE; therefore, we calculated the centroid of each cell, as shown in Table 2, so that we can determine the distance between all the users in a cell and its center point. In addition, Figure 11 shows the clustering of 49,000 users.

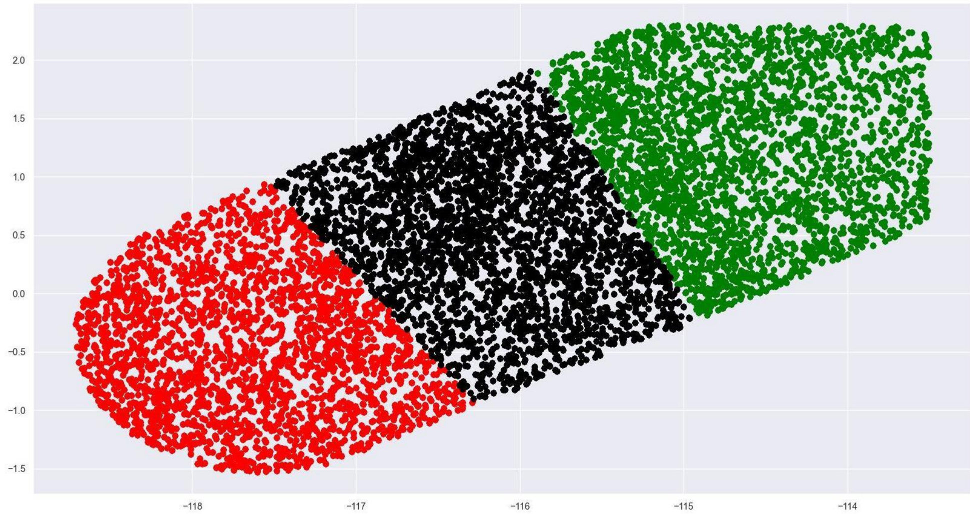


Figure 11. 49,000 UE clustered based on their location: Cell type, latitude, and longitude.

Table 2. Centroid of each cluster.

Cell	Centroid [Latitude, Longitude]
First Cluster [Green]	1.22293499, −114.66066589
Second Cluster [Black]	0.51339138, −116.18096421
Third Cluster [Red]	−0.4651767, −117.6435267

4.2. Performance Result of the Classification Algorithm

Conditional handover is used to support mobility management in NTNs by allowing the system to make more informed decisions about when and how to perform a handover based on network conditions and the requirements of the mobile device. Since we know where each cluster centroid is, we can determine the distance between every UE and the cell center. The handover decisions are designed as a binary classification problem by assuming that UE close to the cell center ($C3_{UE}$ and $C2_{UE}$) is not ready to handover; as a result, they belong to category 0; those who are close to the cell edge ($C1_{UE}$) are ready to handover, and they belong to category 1, as shown in Figure 8. One metric for measuring classification model performance is accuracy. Informally, accuracy is the percentage of predictions that our model correctly predicted. Formally, accuracy has the following definition:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \tag{10}$$

Accuracy can also be determined in terms of positives and negatives for binary classification, as seen below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives. Building an efficient machine learning model requires evaluating a model, and there are various evaluation measures in machine learning. Along with accuracy, the following measures have been used to check our model:

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

Precision: The number of correct positive results divided by the number of positive results predicted by the classifier.

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

$$F1\ Score = 2\left(\frac{Precision\ \times\ Recall}{Precision + Recall}\right) \tag{14}$$

A measure of the quality of the estimator is the Mean Square Error. It is always a positive number that gets smaller as the error gets closer to zero since it is derived from the square of the Euclidean distance.

$$MSE = \frac{1}{N}\sum_{j=1}^{N}(y_j - \hat{y}_i)^2 \tag{15}$$

During our experimentation, after collecting the data, we preprocessed them; we used preprocessed datasets to train the model using various machine learning algorithms. The machine learning model is trained offline, and on-device ML can be used to perform inference models directly on a device since our proposed architecture is for a regenerative onboard satellite payload. Once the model training was finished, we used evaluation metrics to test our model. The results presented in Table 3 describe the accuracy of different machine learning classifiers and the artificial neural network (ANN). Here, for the random forest, the value of the n-estimator was 30, and we found out that the max-depth equal to 9 was a good value for this specific generated dataset. Consequently, the random forest has a higher accuracy than the ANN, but the ANN has a lower MSE than the random forest.

Table 3. Machine Learning Classifiers.

Num	Model	Accuracy	MSE
1	Logistic Regression	88%	0.1150
2	Naive Bayes	71%	0.2835
3	Support Vector Machine	82%	0.1791
4	Random Forest	**99**%	0.0038
5	ANN	**98**%	**0.0017**

Most frequently, we use classification accuracy to evaluate the performance of our model, but it is not sufficient. F1-Score, AUC, recall, and precision were used to evaluate the performance of the model. Besides the logistic regression, naves, and the support vector machine, the ANN has a better recall and precision, as shown in Table 4.

Table 4. F-1 Score, AUC, Recall, and Precision of Machine Learning Classifiers.

Num	Model	F1-Score	AUC	Recall	Precision
1	Logistic Regression	0.8849	0.8849	0.89	0.87
2	Naive Bayes	0.6942	0.7168	0.45	0.97
3	Support Vector Machine	0.8185	0.8210	0.71	0.92
4	Random Forest	**0.9961**	0.9961	**1**	**1**
5	ANN	0.9820	0.9817	**1**	**1**

Table 1 described the features of the UE; using feature creation, one of the techniques in feature engineering, also using the haversine formula, we created a new feature and named it a distance. It is the distance between a UE and its cell center.

The feature importance helps us to understand the solved problem and sometimes leads to model improvements. Figure 12 describes which features are relevant to the model, and the feature distance comes at the first position. In addition, the feature importance section explains how features like distance, latitude, and longitude are significant in demonstrating why the location (UE and Satellite) triggering [35] is one of the most crucial conditional handover triggerings.

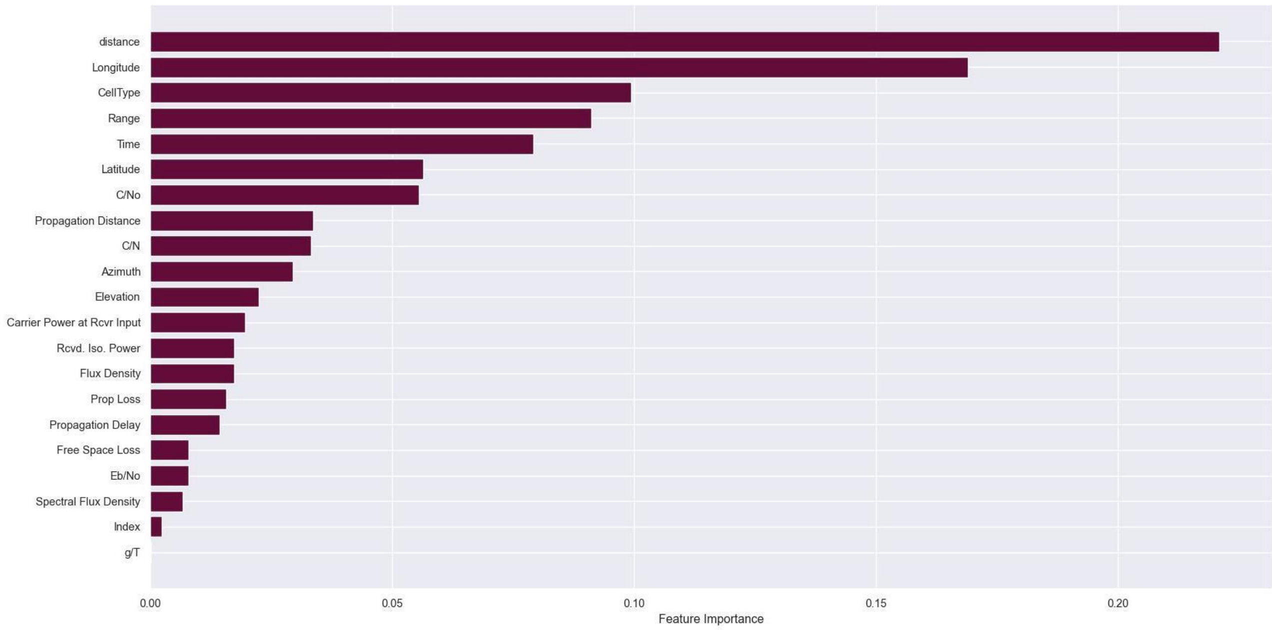


Figure 12. Feature Importance.

Besides the machine learning classifiers, we also built an artificial neural network using the sequential keras model.

Figure 13 shows us the training process of the model. This sequential model API is a way of creating deep learning models where an instance of the sequential class is created and model layers are created and added to it. Therefore, we built a simple classification network that has 21 neurons in the input layer, 2 hidden layers with 4 neurons each, and an output layer with 1 output. Rectified linear activation functions and random-normal as kernel-initializer are used in each hidden layer and a sigmoid activation function is used in the output layer, for binary classification; then, we compiled the model with binary cross-entropy as the loss and adam as the optimizer to minimize the loss and adjust input weights. We then fitted the model to the training and validation data set. Figure 13 shows us that the accuracy depends on the number of epochs, with 75 epochs we obtain 98% accuracy. The recall and precision of the ANN are described in Table 4. In addition, the loss functions are used to determine the amount that a model should try to minimize during training, and as shown in Figure 14, the ANN has a lower loss of 0.0017 than all the models.

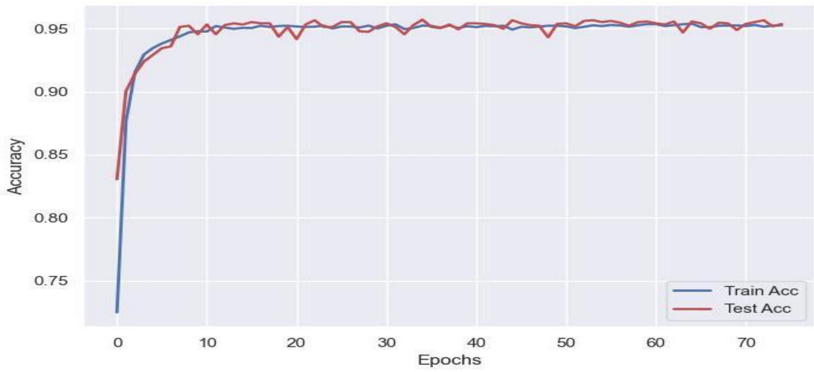


Figure 13. Accuracy of the neural network.

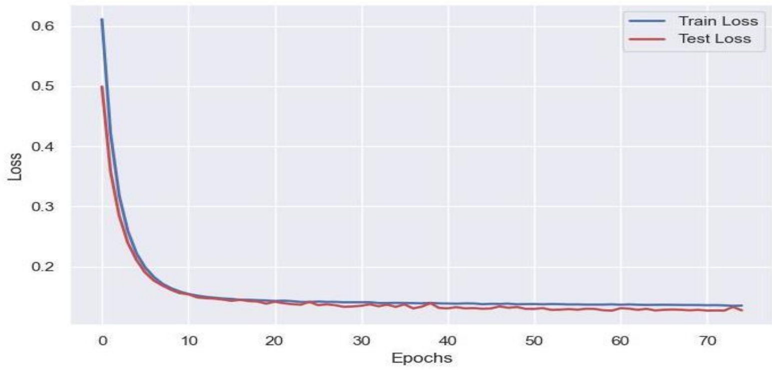


Figure 14. Loss of the neural network.

5. Conclusions

Conditional handover is a technique used in non-terrestrial networks (NTNs) to control the handover of mobile devices between various network nodes while maintaining an uninterrupted communication session. The network technology and design of an NTN are distinct from those of a standard terrestrial network; it might make handover more difficult because a mobile device needs to hand over between various satellites and a ground station in a satellite-based NTN. We created an NTN handover simulator for this study to model and replicate the behavior of user equipment and satellite in non-terrestrial communication networks. Machine learning can improve the performance and efficiency of the handover process in wireless networks by providing more accurate and efficient handover decisions, optimizing the handover process, and mobility management. Therefore, first, we proposed a K-means algorithm with an optimal number of clusters, using the Elbow technique. After that, we built different machine learning classification models that decide or classify whether a UE is ready to hand over or not based on the distance from its cell center. The feature importance describes that features such as distance, latitude, and longitude are relevant to the model, proving why the location (UE and Satellite) triggering is very important among the conditional handover triggering defined by the 3GPP. The 3GPP study on NR support for non-terrestrial networks did not yet provide any dataset; therefore, we generated the dataset using the simulator that models user and satellite communication. Before applying clustering and classification algorithms, we preprocessed the data and created the distance feature using the Haversine formula. The experimental results demonstrate that the distance between a user and its cell center is a crucial factor in Handover decisions in NTN. In addition, the random forest outperformed all models with a higher accuracy of 99% and a superior F1-score of 0.9961.

Author Contributions: Conceptualization, M.K.D.; methodology, M.K.D. and S.J.; software, M.K.D. and S.J.; validation, M.K.D. and S.J.; formal analysis, M.K.D.; investigation, M.K.D.; resources, S.J.; data curation, S.J. and M.K.D.; writing—original draft preparation, M.K.D.; writing—review and

editing, M.K.D. and S.J.; visualization, M.K.D. and S.J.; supervision, I.J.; project administration, I.J.; funding acquisition, I.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2020-0-00107, Development of the technology to automate the recommendations for big data analytic models that define data characteristics and problems).

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author, I.J. The data are not publicly available due to their containing information that could compromise the privacy of research participants.

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

Sample Availability: The data for this project are confidential but may be obtained with Data Use Agreements with the Computer Software Department of Hanyang University. Researchers interested in access to the data may contact Mwamba Kasongo Dahouda at dahouda37@hanyang.ac.kr, also see http://wm.hanyang.ac.kr/xe/contact (accessed on 6 March 2023). It can take some months to negotiate data use agreements and gain access to the data. The author will assist with any reasonable replication attempts for two years following publication.

Abbreviations

The following abbreviations are used in this manuscript:

3GPP	3rd Generation Partnership Project
5G	Fifth-generation technology standard
ACK	Acknowledge
CHO	Conditional Handover
DDPG	Deep Deterministic Policy Gradient
GEO	Geostationary Orbit
gNB	Next Generation Node B
HAPS	High Altitude Platform System
HEO	High Elliptical Orbit
HO	Handover
ISL	Inter-Satellite Links
LEO	Low Earth Orbit
LLA	Latitude, Longitude, and Altitude
MADM	Multi-Attribute Decision Making
MEO	Medium Earth Orbit
ML	Machine Learning
NR	New Radio
NTN	Non-Terrestrial Network
PCI	Physical Cell ID
QoS	Quality of Service
RSRP	Reference Signals Received Power
RSS	Received Signal Strength
SINR	Signal-to-interference-plus-noise ratio
TN	Terrestrial Networks
UAS	Unmanned Aerial System
UE	user equipment
WCSS	within-cluster sum squares

References

1. Yang, P.; Xiao, Y.; Xiao, M.; Li, S. 6G Wireless Communications: Vision and Potential Techniques. *IEEE Netw.* **2019**, *33*, 70–75. [CrossRef]
2. Bhushan, N.; Li, J.; Malladi, D.; Gilmore, R.; Brenner, D.; Damnjanovic, A.; Sukhavasi, R.T.; Patel, C.; Geirhofer, S. Network densification: The dominant theme for wireless evolution into 5G. *IEEE Commun. Mag.* **2014**, *52*, 82–89. [CrossRef]
3. Maksymyuk, T.; Brych, M.; Klymash, M.; Jo, M. Cooperative channels allocation in unlicensed spectrum for D2D assisted 5G cellular network. In Proceedings of the 2017 2nd International Conference on Advanced Information and Communication Technologies (AICT), Lviv, Ukraine, 4–7 July 2017; pp. 197–200. [CrossRef]

4. Hwang, I.; Song, B.; Soliman, S. A Holistic View on Hyper-Dense Heterogeneous and Small Cell Networks. *Commun. Mag. IEEE* **2013**, *51*, 20–27. [CrossRef]
5. Gazda, J.; Šlapak, E.; Bugar, G.; Horvath, D.; Maksymyuk, T.; Jo, M. Unsupervised Learning Algorithm for Intelligent Coverage Planning and Performance Optimization of Multitier Heterogeneous Network. *IEEE Access* **2018**, *6* , 39807–39819. . [CrossRef]
6. Peng, M.; Li, Y.; Zhao, Z.; Wang, C. System architecture and key technologies for 5G heterogeneous cloud radio access networks. *IEEE Netw.* **2015**, *29*, 6–14. [CrossRef]
7. Maksymyuk, T.; Gazda, J.; Yaremko, O.; Nevinskiy, D. Deep Learning Based Massive MIMO Beamforming for 5G Mobile Network. In Proceedings of the 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS), Lviv, Ukraine, 20–21 September 2018; pp. 241–244. [CrossRef]
8. Maksymyuk, T.; Brych, M.; Masyuk, A. Fractal geometry based resource allocation for 5G heterogeneous networks. In Proceedings of the 2015 Second International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T), Kharkiv, Ukraine, 13–15 October 2015; pp. 69–72. [CrossRef]
9. Lin, X.; Rommer, S.; Euler, S.; Yavuz, E.A.; Karlsson, R.S. 5G from Space: An Overview of 3GPP Non-Terrestrial Networks. *IEEE Commun. Stand. Mag.* **2021**, *5*, 147–153. [CrossRef]
10. Han, S.; Xie, T.; I, C.L.; Chai, L.; Liu, Z.; Yuan, Y.; Cui, C. Artificial-Intelligence-Enabled Air Interface for 6G: Solutions, Challenges, and Standardization Impacts. *IEEE Commun. Mag.* **2020**, *58*, 73–79. [CrossRef]
11. 38.821. Solutions for NR to Support Non-Terrestrial Networks (NTN). V16.0.0. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3525 (accessed on 27 February 2023).
12. Lin, X.; Chen, M.; Rydén, H.; Jeong, J.; Lee, H.; Sundberg, M.; Timo, R.; Shokri-Razaghi, H.; Poor, H.V. Fueling the Next Quantum Leap in Cellular Networks: Embracing AI in 5G Evolution towards 6G. *arXiv* **2021**, arXiv:2111.10663.
13. 23.737. Study on Architecture Aspects for Using Satellite Access in 5G. V17.1.0. Available online: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx (accessed on 27 February 2023).
14. Gupta, V.; Bonde, P. Handover Techniques in New Generation Wireless Networks. In *Smart Computing and Informatics*; Satapathy, S.C., Bhateja, V., Das, S., Eds.; Springer: Singapore, 2018; pp. 127–138.
15. Tetarwal, M.L.; Kuntal, A.; Karmakar, P. A Review on Handoff Latency Reducing Techniques in IEEE 802.11 WLAN. Available online: https://research.ijcaonline.org/nwnc/number2/nwnc1424.pdf (accessed on 27 February 2023).
16. El Fachtali, I.; Saadane, R.; Koutbi, M. A survey of handovers decision algorithms for next generation wireless networks. *Network* **2015**, *4*, 159–165. https://doi/10.17148/IJARCCE.2015.4133.
17. Seth, A. Vertical handoff decision algorithms for next generation wireless networks: Some issues. *Int. J. Adv. Res. IT Eng.* **2013**, *2*, 64–79.
18. Chandavarkar, B.; Reddy, G.R.M. Survey paper: Mobility management in heterogeneous wireless networks. *Procedia Eng.* **2012**, *30*, 113–123. [CrossRef]
19. Ravichandra, M.; Gowda, H.K.; Kumar, C.U. A survey on handovers literature for next generation wireless networks. *Int. J. Adv. Res. Comput. Commun. Eng.* **2013**, *2*, 4671–4677.
20. Sinai, N.K.; Namgung, J.I.; shin, S.y.; Park, S.H. Handover Protocol in Ad-hoc Diver Networks using Visual Light Communication. In Proceedings of the 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN), Split, Croatia, 2–5 July 2019; pp. 604–609. [CrossRef]
21. TalebiFard, P.; Leung, V.C. A dynamic context-aware access network selection for handover in heterogeneous network environments. In Proceedings of the 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Shanghai, China, 10–15 April 2011; pp. 385–390.
22. Huang, Z.H.; Hsu, Y.L.; Chang, P.K.; Tsai, M.J. Efficient Handover Algorithm in 5G Networks using Deep Learning. In Proceedings of the GLOBECOM 2020-2020 IEEE Global Communications Conference, Taipei, Taiwan, 7–11 December 2020; pp. 1–6. [CrossRef]
23. Khosravi, S.; Ghadikolaei, H.S.; Petrova, M. Learning-based Load Balancing Handover in Mobile Millimeter Wave Networks. *arXiv* **2020**, arXiv.2011.01420.
24. Bahra, N.; Pierre, S. A Hybrid User Mobility Prediction Approach for Handover Management in Mobile Networks. *Telecom* **2021**, *2*, 199–212. [CrossRef]
25. Yajnanarayana, V. Proactive Mobility Management of UEs using Sequence-to-Sequence Modeling. *arXiv* **2021**, arXiv:2110.07262.
26. Shubyn, B.; Lutsiv, N.; Syrotynskyi, O.; Kolodii, R. Deep Learning based Adaptive Handover Optimization for Ultra-Dense 5G Mobile Networks. In Proceedings of the 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukrain, 25–29 February 2020; pp. 869–872. [CrossRef]
27. Kumar, S.; Vanmathi, S.; Sanjay, B.; Bharathi, S.; Meena, M. Handover forecasting in 5G using machine learning. *Int. J. Eng. Technol.* **2018**, *7*, 76–79. [CrossRef]
28. Paropkari, R.A.; Thantharate, A.; Beard, C.C. Deep-Mobility: A Deep Learning Approach for an Efficient and Reliable 5G Handover. *arXiv* **2021**, arXiv:2101.06558.
29. Ali, Z.; Baldo, N.; Mangues-Bafalluy, J.; Giupponi, L. Machine learning based handover management for improved QoE in LTE. In Proceedings of the NOMS 2016-2016 IEEE/IFIP Network Operations and Management Symposium, Istanbul, Turkey, 25–29 April 2016; pp. 794–798. [CrossRef]

30. Majid, S.; Shah, S.; Marwat, S.N.K.; Hafeez, A.; Ali, H.; Jan, N. Using an Efficient Technique Based on Dynamic Learning Period for Improving Delay in AI-Based Handover. *Mob. Inf. Syst.* **2021**, *2021*, 2775278. [CrossRef]
31. Lee, C.; Cho, H.; Song, S.; Chung, J.M. Prediction-Based Conditional Handover for 5G mm-Wave Networks: A Deep-Learning Approach. *IEEE Veh. Technol. Mag.* **2020**, *15*, 54–62. [CrossRef]
32. Ortiz, F.; Monzon Baeza, V.; Garcés-Socarrás, L.; Vasquez, J.; González Rios, J.L.; Fontanesi, G.; Lagunas, E.; Querol, J.; Chatzinotas, S. Onboard Processing in Satellite Communications Using AI Accelerators. *Aerospace* **2023**, *10*, 101. [CrossRef]
33. Na, S.; Xumin, L.; Yong, G. Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm. In Proceedings of the 2010 Third International Symposium on Intelligent Information Technology and Security Informatics, Jinggangshan, China, 2–4 April 2010; pp. 63–67. [CrossRef]
34. Umargono, E.; Suseno, J.; Gunawan, S. K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula. In Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019), Yogyakarta, Indonesia, 25–26 November 2020. [CrossRef]
35. Stanczak, J.; Karabulut, U.; Awada, A. Conditional Handover in 5G: Principles, Future Use Cases and FR2 Performance. In Proceedings of the 2022 International Wireless Communications and Mobile Computing (IWCMC), Dubrovnik, Croatia, 30 May–3 June 2022. [CrossRef]