Toward Multi-Connectivity in Beyond 5G Non-Terrestrial Networks: Challenges and Possible Solutions

Mikko Majamaa

***Abstract*—Non-terrestrial networks (NTNs) will complement terrestrial networks (TNs) in 5G and beyond, which can be attributed to recent deployment and standardization activities. Maximizing the efficiency of NTN communications is critical to unlock its full potential and reap its numerous benefits. One method to make communications more efficient is by the usage of multi-connectivity (MC), which allows a user to connect to multiple base stations simultaneously. It is standardized and widely used for TNs, but for MC to be used in the NTN environment, several challenges must be overcome. In this article, challenges related to MC in NTNs are discussed, and solutions to the identified challenges are proposed.**

INTRODUCTION

Mobile communications and access to the Internet play a critical role in a prosperous society. At the individual level, this can be explained by the increased opportunities that access to the Internet brings, such as educational and employment opportunities. At the societal level, it can be explained by factors such as business transformation and economic modernization. Yet, according to the statistics published by the International Telecommunications Union (ITU), there were 2.7 billion people in the world who had no access to the Internet in 2022 [1]. To this end, non-terrestrial networks (NTNs) can be used to provide connectivity to un(der)served areas.

Recent standardization and deployment activities have made the deployment of 5G and beyond through NTNs a practical possibility [2]. The standardization efforts enhance cooperation between terrestrial and satellite operators. In the future, the different networks can be seen as complementary rather than competing, while converging into one indistinguishable network from the user's point of view. In addition, non-geosynchronous orbit (NGSO) satellites, in particular, have been the subject of intense research and deployment interest in recent years, due to advances in launch vehicles that allow many small satellites to fit into a single launch, and the mass production of satellites, which lowers production costs.

NTNs offer a variety of ways to improve the performance of wireless communications. Extending coverage to unserved and underserved areas means services such as backhaul connections to remote networks, direct-to-handheld services, maritime coverage, and remote IoT connectivity. In addition to extending coverage, NTNs can provide load balancing to overloaded terrestrial networks (TNs). In cases of natural disasters or emergencies, NTNs can provide critical access when TNs are out of coverage.

To ensure the efficiency of NTN communications, research is needed to explore methods on how to achieve this. Multi-connectivity (MC) is a promising method that allows a user equipment (UE) to connect to multiple base stations (BSs) simultaneously. For example, MC in NTNs can be used in load balancing [3], and to enhance service continuity [4] and energy efficiency [5]. Further, MC for throughput enhancement may be critical to enable applications with high throughput requirements to cell edge users, such as live video streaming, where a single link cannot achieve the required performance [6]. MC between satellite and cellular networks in rural areas can help reduce latency, increase availability, resiliency, and reliability, and thus help meet the low latency requirements of applications such as environmental and livestock monitoring [7].

MC can be implemented at different levels such as physical (PHY) layer, medium access control (MAC) layer, packet data convergence protocol (PDCP) layer, or core network (CN) level. PDCP layer MC solutions are attractive because they can quickly adapt to changing radio conditions. Multi-radio dual connectivity (MR-DC) is one such solution and is the main focus of the article. In MR-DC, a UE is connected to two BSs one providing 5G access and the other 4G/5G access. MC in NTNs can include only NTN nodes or a mix of TN and NTN nodes.

MR-DC has been standardized for TNs, yet its standardization for NTNs has not been actualized to date. Additionally, although there is a substantial body of literature addressing MC in TNs [8], analogous comprehensive research exploring MC in NTNs remains comparatively underdeveloped. This article aims to help fill this research gap by providing a comprehensive introduction and discussion on MC in NTNs, including related challenges and possible solutions. To achieve this, the article examines 5G and beyond specifications, existing literature, and insights. Notably, this article is the first published tutorial/survey article related to MC in NTNs, to the best of the author's knowledge. Although the examples in this article mainly center around NGSO-based NTNs, given their inherently more complex nature due to their high velocity relative to users, the discussion can be applied to other types of NTNs.

This work has been partially funded by the European Union Horizon-2020 Project DYNASAT (Dynamic Spectrum Sharing and Bandwidth-Efficient Techniques for High-Throughput MIMO Satellite Systems) under Grant Agreement 101004145. The views expressed are those of the authors and do not necessarily represent the project. The Commission is not liable for any use that may be made of any of the information contained therein.

The author is with Magister Solutions, and University of Jyväskylä, Finland.

The article's structure is as follows: first, 5G and beyond networks are introduced, followed by a discussion on the way toward MC in NTNs. Then, considerations about MC in NTNs are analyzed, including challenges and potential solutions. Finally, the article is concluded.

5G and Beyond Networks

The 3rd Generation Partnership Project (3GPP) is a leading standards body for mobile communications standards. The purpose of standardization is to ensure high-quality systems and that the various network operators and equipment manufacturers can work together seamlessly. 3GPP Rel-15 marked the beginning of the 5G era in terms of 3GPP standardization. Subsequent releases 16 and 17 continued by enhancing the features. Releases 15 through 17 focused on 5G, while the ongoing Rel-18 marks the beginning of 5G-advanced (5G-A) and beyond 5G (B5G). For example, Rel-18 addresses network energy savings, mobility, and coverage enhancements, artificial intelligence (AI)/machine learning (ML) for next-generation radio access network, and dynamic spectrum sharing enhancements [9]. In addition, future 6G use cases include holographic telepresence applications, tactile Internet, teleoperated driving, nanonetworks, and a surge in the number of IoT devices [10].

History was made in March 2022 when 3GPP included NTNs in its specifications in its Release 17 (Rel-17), although it should be noted that Rel-15 and Rel-16 included necessary preparatory work in the form of technical reports (TRs). NTN work in 3GPP covers low Earth orbit (LEO), medium Earth orbit (MEO), and geosynchronous orbit (GSO) satellite communications, as well as airborne vehicles such as high altitude platform stations (HAPSs). Rel-17 provides specifications for new radio (NR)-based satellite access in S- and L-bands, where NR is the air interface of 5G. The focus is on transparent payload architecture and frequency division duplexing serving handheld terminals. Rel-18 will further enhance NR operations in NTNs, for example, by improving coverage to Ka-bands serving very small aperture terminals (VSATs), as well as addressing mobility and service continuity.

Toward Multi-Connectivity in Non-Terrestrial Networks

This section discusses the path toward MC in NTNs in the context of 5G technology. The goal of using MC is to improve end-user service performance. Various forms of MC have already been standardized for TNs [11].

MC in 5G can be used on different protocol layers. Multi-transmission and reception point (TRP) (technical specification (TS) 38.300) is a PHY layer MC solution. Carrier aggregation (CA) (TS 38.331) can be used to achieve MC on the MAC layer. PDCP layer MC solutions include new radio dual connectivity (NR-DC) (TS 37.340), where the UE is connected to two BSs providing 5G access, MR-DC (TS 37.340), where the UE is connected to two BSs providing 5G and 4G/5G access, and dual active protocol stack (DAPS) (TS 38.300), where the UE is connected to two 5G BSs to ensure a seamless handover. MC can also be achieved at the CN in terms of access traffic steering, switching, and splitting (ATSSS) (TS 24.193), where simultaneous 3GPP and non-3GPP access is provided.

Each MC solution has some limitations. The PHY layer solutions have strict latency and synchronization requirements. The MAC layer solutions introduce increased scheduler complexity. The PDCP layer solutions require additional hardware and software capabilities. The CN-based MC can be problematic because the CN may not be able to quickly adapt to dynamic radio link conditions.

A CN-based MC for NTNs, that is, upper-layer traffic steer, switch, and split over dual 3GPP access (TR 22.841), will be included in the upcoming Rel-19. This type of MC is similar to ATSSS. In addition, MR-DC supporting NTN was also planned for Rel-19, but was scoped out. Although scoped out of Rel-19, it is a promising solution (and a candidate for future releases) because of its ability to quickly adapt to dynamic radio link conditions. In this article, the focus is primarily on the PDCP layer MC, namely, MR-DC, and the term MC is used interchangeably to mean MR-DC in the article unless otherwise noted. The analysis of MC implementations on different layers is left for future work.

MR-DC is a generalization of evolved universal terrestrial radio acces (E-UTRA) dual connectivity (DC), that is, 4G DC. In MR-DC, one of the nodes serves as a master node (MN) and the other as a secondary node (SN). One of the nodes provides 5G access and the other either 5G or 4G access. The nodes are connected via the Xn interface for control signaling as well as to steer traffic from the MN to the SN. In principle, more than two connections to a user could be formed but hardware and software requirements can be a limiting factor. Further, according to TR 38.821, the MR-DC specification for TNs (TS 37.340) may need to be adapted to support NTN MC, for example, to accommodate extended latency, variable latency (e.g. Xn interface traversing multiple satellites on different orbital planes), and differentiated delay between different nodes involved.

The SN addition for a UE is triggered by the MN (i.e., the current serving node of the UE) that sends SN addition request to a candidate SN. After the SN addition procedure is completed, the MN can send data to the SN through the Xn interface. The SN then forwards the data down the protocol stack to allocate the necessary resources and finally send it over the air to the UE. This does not significantly change the BS architecture. On the UE side, in the downlink direction, the UE must be able to receive data from the two connections, which means that the UE must have two receiving antenna elements. At the protocol stack level, this means that the UE must receive the two different transmissions and then combine them at the PDCP layer which has an impact on the UE's architecture. When uplink MC is considered, the UE must be able to transmit to the MN and to the SN simultaneously.

MC in NTNs is illustrated in Fig. 1. Some of the main functions of the layers in the protocol stack shown in the figure are briefly summarized as follows [12]:

- Service data adaptation protocol (SDAP), PDCP, radio link control (RLC). Responsible for assembling and reordering packets, security, automatic repeat request (ARQ), and integrity protection.

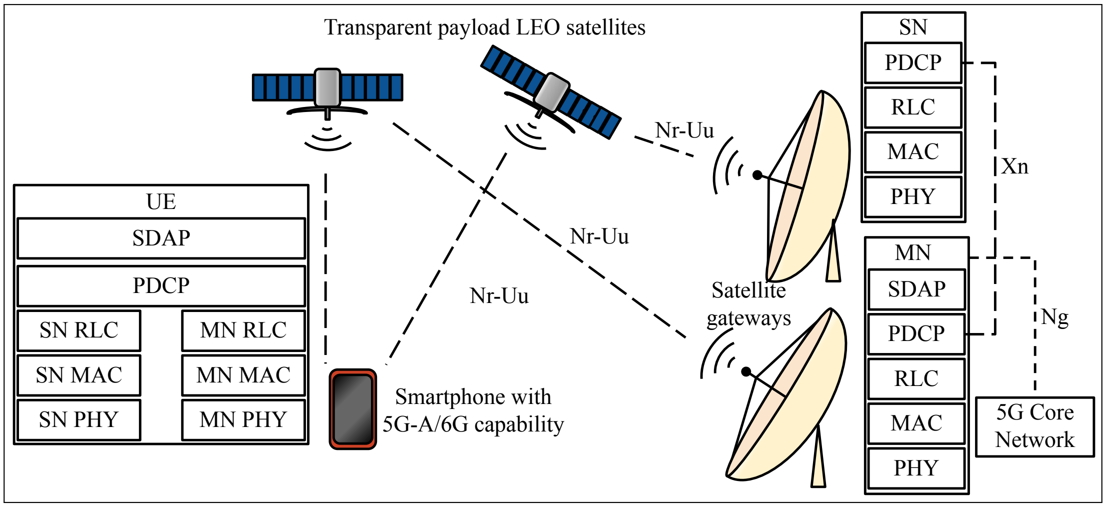

Fig. 1. MC illustrated in the NTN environment. The figure shows the related protocol stacks, where changes to the UE's architecture can be seen: the UE must be able to receive transmissions from multiple sources, namely the MN and SN in the downlink direction, and similarly, to transmit to these different nodes when uplink MC is considered. A transparent payload architecture is considered, so the service and feeder links use the Nr-Uu interface. The MN and SN are connected via the Xn interface for the exchange of control plane and user plane data. The MN is connected to the CN via the Ng interface. Adapted from [6].

- MAC. Defines how to access the medium. Responsible for logical channel prioritization, random access (RA), and retransmissions through hybrid automatic repeat request (HARQ).
- PHY. Physical layer operations, such as (de)modulation, (inverse) fast Fourier transform, and orthogonal frequency division multiplexing operations, that is, the responsibility for transmitting the raw data over the air.

Considerations for Multi-Connectivity in Non-Terrestrial Networks

MC in TNs is well-studied and documented but few works have explored MC in NTN, and the unique challenges of this environment need to be considered. The problems and solutions for MC in NTN are discussed below and summarized in Table I. Please note that the solutions listed in the table with specific references are derived from existing scientific work. Solutions without such references represent original contributions proposed by the author.

Different Types of Nodes Involved in Multi-Connectivity

The two main satellite payload architectures are transparent and regenerative. Satellites with transparent payloads act as RF repeaters, while the next-generation node B (gNB) (i.e., the BS providing 5G access) is on the ground. The gNB sends signals through a gateway and the satellite performs frequency conversion and power amplification, and relays the signals. This architecture is considered in 3GPP Rel-17 and Rel-18. In contrast, regenerative payload satellites may have (some of) the BS, that is, gNB, functionalities on board.

MC in NTNs can include only NTN nodes or a combination of TN and NTN nodes. In the former case, the nodes could be transparent or regenerative payload NTN nodes or a mix of both. Table II outlines various MC scenarios involving NTN nodes. Scenarios that involve regenerative payload satellites assume the presence of PDCP on board the satellite, otherwise, these cases would be reduced to cases where transparent payload (PDCP on the ground) was considered.

Careful consideration must be given when selecting the combination of nodes for MC in NTNs with respect to the possible need for: 1) increased buffering; 2) formation of inter-satellite links (ISLs); and 3) increased path length for the data transmission. For example, regenerative payload NTN as a MN and transparent payload NTN as a SN may not be feasible because the user's downlink data directed from the MN to the SN must first travel to the satellite (MN), then to the node on the ground, and then through the transparent payload satellite to the user.

Delay Differences

NR PDCP supports two types of delivery. The first one is out-of-order delivery, where packets can be delivered to upper layers as they are received without buffering. The delay differences with MC in NTNs do not cause increased buffering requirements on UE's PDCP when this mode is used. When considering using MC in NTNs, this mode should be used whenever possible. However, this requires the application to support out-of-order packet reception. File transfer or non-live video streaming are examples of such applications. The second delivery mode is in-order delivery, in which PDCP packets are buffered at the recipient's PDCP buffer, reordered, and delivered in order. This mode should be used for applications such as voice or live video streaming.

For TNs, TS 38.331 defines the length of the PDCP reordering window, which can vary from 0 ms (out-of-order delivery) to 3 s. Additionally, the window can be set to infinity, which corresponds to forced in-order delivery. These window lengths can cover even the largest delay differences when using MC in NTNs between heterogeneous nodes, that is, these values can be reused from the TN specifications. Since the larger

TABLE I
PROBLEMS AND SOLUTIONS FOR MC IN NTNS.

Attribute	Problem	Solutions
Different node types	MC in NTNs may include a heterogeneous set of nodes involved (transparent/regenerative payload satellites and terrestrial nodes).	• Choosing the possible combination so that it is feasible in terms of possibly required: 1) increased buffering; 2) formation of ISLs; and 3) increased path length for the data transmission.
Delay differences	Delay may vary between the different paths.	• Usage of out-of-order delivery mode whenever possible, or limitation on the usage of MC in NTNs to applications that support such mode • Increase the buffer size of UE • Usage of high timer value to discard outdated data • Joint coordination of the transmissions through time compensation • Path optimization
Mobility	NTN nodes move at a high speed, resulting in frequent SN handovers, modifications, and releases.	• Predictive schemes for MC-related operations using: 1) UEs' GNSS capabilities; and/or 2) satellite ephemeris data
SN addition	The standard methods for SN addition (based on signal strength) may not be sufficient due to the NTN environment. Also, SN addition is time consuming.	• SN addition based on: 1) location; or 2) elevation angle • Optimized SN addition without RA
Traffic steering	Traffic steering between the nodes can be difficult due to the long distances.	• Traffic steering schemes that require the minimal amount of control signaling • By network configuration, the disablement of MC involving nodes that are too far apart
Resource allocation	MN/SN resources used must be coordinated.	• Providing the secondary connection users resources only if there are left from the primary connection users • Piggybacking the user's QoS requirements in the SN addition message and taking them into account in resource scheduling on the SN side • Treating users equally regardless of the connection types • Joint optimization of the resource allocation between the MN and SN via Xn signaling • Using beam hopping for increased flexibility in resource allocation • In the case of PD, dynamic detection of the need for PD, for example, based on HARQ feedback
Power consumption	Power consumption, particularly by UE in NTNs, poses a significant challenge.	• Usage of the best link at any given time in the uplink direction • Utilization of efficient scheduling strategies [13] which may include AI/ML

distance between MN and SN in the NTN MC case results in larger transmission delay differences than in the TN MC case, a relatively higher reordering window length should typically be used. This may also require increasing the PDCP buffer size of the UE.

An alternative option to increasing the UE's PDCP buffer size is to coordinate MN/SN transmissions. This can be accomplished by time alignment or by path optimization, which is the process of choosing the best path for communication based on factors such as congestion levels and propagation delays. However, both methods require signaling between nodes.

Mobility

The beam deployment (shown in Fig. 2) affects the density for SN handovers, modifications, and releases. In a quasi-earth fixed beam deployment, the beam pattern is fixed to a target on the ground. In an earth-moving beam deployment, the beams of the satellites are directed in a fixed direction, causing the beams to move on the ground as the satellite moves. In a quasi-earth fixed beam deployment, the SN handovers, modifications, and releases may be less frequent than in an earth-moving beam deployment because the UE typically stays in the service area longer in a quasi-earth fixed beam deployment.

However, with either beam deployment SN handovers, modifications and releases can be highly time consuming, frequent, and caused by satellite movement. In contrast, in TNs, UE

TABLE II
DIFFERENT MC SCENARIOS INVOLVING NTN NODES DETAILED.

Master node	Secondary node	Considerations
Transparent payload NTN node	Transparent payload NTN node	[1]
Transparent payload NTN node	Regenerative payload NTN node	[2]
Transparent payload NTN node	Terrestrial node	[1]
Regenerative payload NTN node	Transparent payload NTN node	This setting could be infeasible because the data directed from the MN to the SN needs to travel first to the satellite (MN), then to the BS on the ground, and then to the user through the transparent payload satellite. [2]
Regenerative payload NTN node	Regenerative payload NTN node	Xn is between the BSs onboard the NTN nodes, that is, ISL needs to be formed.
Regenerative payload NTN node	Terrestrial BS	[2]
Terrestrial BS	Transparent payload NTN node	The terrestrial BS may need to consider the extra delay caused by the NTN node by buffering some of the data to mitigate the delay spread in some applications where packets must be received sequentially by the user. [1]
Terrestrial node	Regenerative payload NTN node	[2]

[1] Both BSs reside on the ground, that is, the Xn interface can be fiber. [2] Xn must be between the BS on the ground and the satellite.

mobility is the primary cause of such events. Indeed, in TNs the base stations are static.

Predictive schemes for SN handovers/modifications/releases can be used in NTNs to anticipate the movement of users/satellites and proactively initiate the necessary MC-related operations. For user movement, this can be achieved by exploiting the UEs' global navigation satellite system (GNSS) capabilities assumed in Rel-17. For satellite motion, ephemeris data can be used for prediction. This can be done either online by computing the positions of the satellites, or offline by precomputing the positions and using a timer that gives information about the positions of the satellites.

Secondary Node Addition

The denser deployment of TN base stations allows cells to be formed more precisely in the desired locations, thus improving the quality of the links. Due to this setup, cells can be separated with high accuracy and therefore do not interfere with each other even with full frequency reuse. With full frequency reuse, UEs can seamlessly collect reference signal strength measurements from adjacent cells without measuring different frequencies.

In the NTN environment, frequency reuse is typically required due to the large diameter of the beams and the relatively small differences in signal strength for cell edge and nadir UEs. Frequency reuse is used to reduce inter-cell interference (see Fig. 2). Frequency reuse implies that UEs are required to apply measurement gaps to measure the frequencies of adjacent beams. During the measurement gaps, the UE may have to stop receiving and/or transmitting data. Therefore, the use of measurement gaps should be carefully designed so that they do not unduly affect the user experience due to paused data reception/transmission. Further, to enable MC for a user, the user must be under the coverage of multiple beams. This can be achieved through constellation design and/or directional antennas.

However, even when measurement gaps are carefully designed, the problem of relatively small differences in signal strength between users in a beam remains. Thus, unlike MC in TNs, where reference signal strength is typically used as a criterion for SN addition, other criteria for SN addition should be considered when using MC in NTNs. These criteria might include location or elevation angle.

In addition, the SN addition process itself could be optimized. The process in TNs is specified in TS 37.340. The SN addition is initiated by the MN, which sends a SN addition request to the candidate SN. The SN then responds with an acknowledgment or rejection. If the request is acknowledged, the MN sends an RRC reconfiguration required message to the UE. The UE performs the required configurations and responds to the MN. The MN then indicates to the SN that the SN reconfiguration is complete. Non-contention-based RA is used by the UE to connect to the SN. RA may occur after the UE receives the RRC reconfiguration message. The RA response contains timing advance and scheduling grant for the UE to use for the uplink transmission configuration. In the SN addition, the RA process takes at least twice the round-trip time, which is a significant time in the NTN environment compared to the TN environment. Methods to perform handovers without RA have already been proposed [14]. Similarly, methods to perform SN addition without RA could be investigated.

Traffic Steering

Traffic steering refers to steering the user's traffic from MN to SN. In TN MC, it can be done smoothly in a coordinated manner because the MN and SN are usually located relatively close to each other, allowing for a fast exchange of information. This allows for flexible traffic control strategies that require a lot of control signaling. In NTNs, the BSs involved in MC could even be located on different continents. This imposes limits on the traffic control strategies chosen. Indeed, the same amount of control signaling may not be feasible between the nodes as in the TN environment.

This problem can be overcome by using traffic steering schemes that include a minimum amount of control signaling.

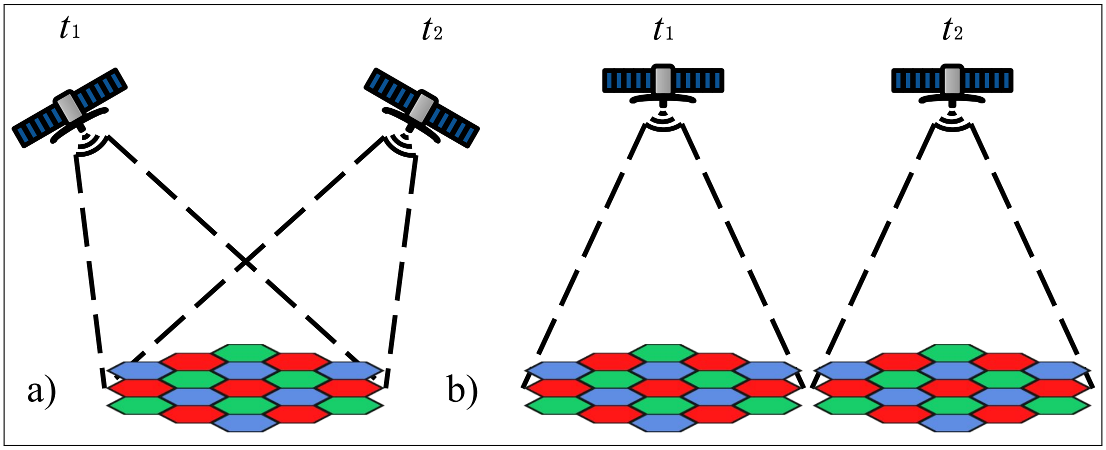

Fig. 2. a) Quasi-earth fixed and b) earth-moving beam deployments. Frequency reuse 3 is considered where adjacent beams operate at different frequencies corresponding to the different colors to mitigate inter-cell interference. The figure shows the difference in the beam deployments as the satellite moves in space from its initial position between times t_1 and t_2.

One option is to use a static split, where the MN steers a fixed proportion, say 40% (the exact amount depends on the use case), of the data for the SN to send to the UE. By using flow control messages, this split could be made semi-static, making it more flexible, if a bit more complex. Now, the split would otherwise be static, but the SN could send flow control messages to the MN if the split is too high/low for the SN, and a new split ratio would be decided in cooperation between the MN and SN.

An alternative option to mitigate the traffic steering problem is to use strategic network configuration by disabling MC between BSs that are not spatially close or sufficiently adjacent in the network infrastructure. This approach allows network administrators to exercise deliberate control by ensuring that MC is prohibited between BSs whose proximity threshold is not met.

Resource Allocation

Since secondary connections to a BS may affect the existing users, a resource allocation scheme based on connection type, that is, whether the UE is a primary connection or a secondary connection UE, could be used. Here, a primary UE refers to a UE that has the BS as its MN, and a secondary UE refers to a UE that has the BS as its SN. Prioritizing UEs with primary connections in resource allocation can be used to avoid impacting existing users. This means that leftover capacity from the primary connection UEs will be distributed to the secondary connection UEs, rather than affecting the UEs already being served by the BS. Another solution would be to treat all connections equally and not distinguish between primary and secondary UEs connections. However, the secondary connection users typically have worse channel conditions than the primary connection users, which means that the secondary connection users require more resources.

An example of a more complex resource allocation scheme includes one where the quality of service (QoS) requirements of the UEs can be taken into account. One way to accomplish this would be to piggyback this information into the SN addition request message. Further, if the SN and MN can exchange information about their resource scheduling strategies, the resource allocation at each node can be jointly optimized via Xn signaling. In addition, beam hopping can be used to switch beams between different coverage areas. This allows for flexible provisioning of resources according to demand. MC could be used in conjunction with beam hopping, and by intelligently designing the beam hopping patterns, more flexibility in resource allocation could be introduced.

MC can also be used to enhance reliability through packet duplication (PD) [15]. In PD, a packet is duplicated and transmitted over different links to increase the probability that the transmission over at least one of the links is successful. While PD enhances reliability it also severely increases resource consumption. Instead of using blind PD, in which all of the packets are duplicated for a given user, dynamic schemes for detecting the need for PD should be utilized. One of the suggested schemes includes PD activation based on HARQ feedback [6].

Power Consumption

Power consumption is a significant issue, especially on the UE side in NTNs, especially when MC is used in the uplink direction with simultaneous transmission to MN and SN. However, MC can be used in the uplink by selecting the best link in terms of SINR at any given time. This could alleviate the power consumption problem on the UE side, since fewer resources are needed when the quality of the link is better. In this approach, MC is used to enable a backup link, which also contributes to service continuity if a link fails. In addition, a likely Rel-19 feature called High Power UE (HPUE), which allows a UE to transmit at higher power, helps to enable simultaneous transmission to MN and SN.

The use of efficient scheduling strategies [13] can lead to efficient resource utilization and an increase in spectral efficiency. Indeed, the use of MC can lead to a greater number of possible paths through which the data can be routed. This results in a complex path selection process that requires intelligent scheduling strategies. AI/ML can be used to solve such complex problems, for example, by learning online from the environment or by learning offline from historical data. It should be noted that AI/ML is a tool that can also be used to help solve the problems listed in the previous subsections.

Conclusion

The convergence of NTNs and TNs into a unified network that is indistinguishable to the user is underway. To efficiently utilize the available resources, the use of disruptive technologies is required. One of these is MC in NTNs, where a user can be connected to multiple BSs, which can include NTN and TN nodes.

Challenges that need to be overcome to achieve MC in NTNs have been discussed in this article and solutions have been proposed. The proposed solutions help pave the way for future research that needs to consider the technical details of the solutions. By overcoming the identified challenges, the use of MC in NTNs can bring significant benefits to various applications. These include, but are not limited to, load balancing, throughput enhancement for cell edge users to enable high throughput applications such as live video streaming, and livestock and environmental monitoring.

While this article primarily examines challenges related to the application of PDCP layer MC in NTNs, it is important to note that some of the proposed solutions may also be relevant to MC implementations on other layers. In future work, there is a need for a comprehensive analysis of MC in NTNs applied on different layers, but also the potential cross-layer applicability of the solutions proposed in this article. Further, the consideration of MC in multi-layer NTN scenarios should also be included in future research.

Acknowledgment

The author thanks Prof. Timo Hämäläinen and Dr. Jani Puttonen, his Ph.D. supervisors, for their valuable feedback on the article.

References

[1] "Internet surge slows, leaving 2.7 billion people offline in 2022," accessed on: Nov. 25, 2023. [Online]. Available: https://www.itu.int/en/mediacentre/Pages/PR-2022-09-16-Internet-surge-slows.aspx.

[2] A. Guidotti *et al.*, "The path to 5G-Advanced and 6G Non-Terrestrial Network systems," in *Adv. Sat. Multimedia Syst. Conf. and 17th Signal Processing for Space Commun. Workshop*, 2022, pp. 1–8.

[3] S. M. Shahi *et al.*, "Load Balancing for 5G Integrated Satellite-Terrestrial Networks," *IEEE Access*, vol. 8, pp. 132 144–132 156, 2020.

[4] M. López *et al.*, "Connecting Rural Areas: an Empirical Assessment of 5G Terrestrial-LEO Satellite Multi-Connectivity," in *2023 IEEE 97th Veh. Technol. Conf.*, 2023, pp. 1–5.

[5] M. Dazhi *et al.*, "Energy-Efficient Service-Aware Multi-Connectivity Scheduler for Uplink Multi-Layer Non-Terrestrial Networks," *IEEE Trans. Green Commun.*, vol. 7, no. 3, pp. 1326–1341, 2023.

[6] M. Majamaa *et al.*, "Satellite-Assisted Multi-Connectivity in Beyond 5G," in *2023 IEEE 24th Int. Symp. on a World of Wireless, Mobile and Multimedia Netw.*, 2023, pp. 413–418.

[7] M. López *et al.*, "An Empirical Analysis of Multi-Connectivity between 5G Terrestrial and LEO Satellite Networks," in *2022 IEEE Globecom Workshops*, 2022, pp. 1115–1120.

[8] T. Sylla *et al.*, "Multi-Connectivity for 5G Networks and Beyond: A Survey," *Sensors*, vol. 22, no. 19, 2022.

[9] X. Lin, "An Overview of 5G Advanced Evolution in 3GPP Release 18," *IEEE Commun. Mag.*, vol. 6, no. 3, pp. 77–83, 2022.

[10] M. Giordani *et al.*, "Toward 6G Networks: Use Cases and Technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, 2020.

[11] C. Pupiales *et al.*, "Multi-Connectivity in Mobile Networks: Challenges and Benefits," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 116–122, 2021.

[12] H.-L. Määttänen *et al.*, "Radio interface protocols and radio resource management procedures for 5G new radio non-terrestrial networks," *Int. J. Satell. Commun.*, Dec. 2022.

[13] M. Dazhi *et al.*, "Terminal-Aware Multi-Connectivity Scheduler for Uplink Multi-Layer Non-Terrestrial Networks," in *2022 IEEE Globecom Workshops*, 2022, pp. 1133–1139.

[14] J. Stańczak, "Mobility enhancements to reduce service interruption time for LTE and 5G," in *2016 IEEE Int. Conf. Commun. Syst. Netw.*, 2016, pp. 1–5.

[15] M.-T. Suer *et al.*, "Multi-Connectivity as an Enabler for Reliable Low Latency Communications—An Overview," *IEEE Commun. Surv. Tutor.*, vol. 22, no. 1, pp. 156–169, 2020.

Biographies

Mikko Majamaa [STM'23] (mikko.majamaa@magister.fi) works as a senior researcher at Magister Solutions. He is also a Ph.D. candidate at the University of Jyväskylä, Finland. His research interests include multi-connectivity, radio resource management, and the use of machine learning in 5G and beyond non-terrestrial networks.