Gen-AI for Radio Resource Management in Multi-orbit 5G-NTN Systems

Flor Ortiz*, Eva Lagunas*, Mikko Majamaa†‡, Abuzar B. M. Adam*, Lauri Sormunen†‡, Timo Hämäläinen‡ and Symeon Chatzinotas*

*The Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg

†Magister Solutions, Jyväskylä, Finland

‡Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland

correspondence: flor.ortiz@uni.lu

***Abstract*—Non-terrestrial networks (NTNs) are poised to play a pivotal role in 5G-Advanced technology, delivering global connectivity and complementing terrestrial infrastructure when needed. Among NTNs, nongeostationary orbit (NGSO) systems, such as low Earth orbit (LEO) constellations, are gaining significant traction due to increasing investments in large-scale deployments. As NTNs evolve towards multi-orbit and dynamic configurations, efficient radio resource allocation emerges as a critical challenge. This paper addresses the resource management problem in 5G NTNs, targeting a fair user demand satisfaction while incorporating multiorbit connectivity, where users can simultaneously connect to multiple satellites belonging to different orbits. In an attempt to overcome the computational complexity of pure optimization-based solutions, we propose two Generative AI (Gen-AI) solutions: a transformer-based model and a generative adversarial network (GAN)-based model. Results demonstrate that the transformer-based model outperforms the alternative optimization-based method as well as the GAN-based one, showcasing its potential for scalable and efficient radio resource management in multi-orbit 5G NTN systems.**

***Index Terms*—5G, non-terrestrial networks, RRM, satellite network simulator**

I. Introduction

The rapid evolution of telecommunications networks has expanded their scope beyond traditional terrestrial infrastructures to include non-terrestrial networks (NTNs), driven by the growing need for seamless, global connectivity. As access to high-speed communication becomes essential, particularly in remote and underserved regions, non-geostationary orbit (NGSO) satellite systems have emerged as a transformative technology. These systems, characterized by their lower latency compared to geostationary orbit (GSO) satellites, are increasingly viewed as a critical solution for bridging connectivity gaps.

This shift toward integrating satellite networks into mainstream telecommunications is evident in the rise of initiatives such as the European Union's Infrastructure for Resilience, Interconnectivity and Security by Satellite (IRIS^2) NGSO constellation [1] and similar endeavors by both governmental and private entities. Standardization efforts by organizations like 3rd Generation Partnership Project (3GPP) have further accelerated this integration by establishing protocols that enable seamless collaboration between satellite and terrestrial systems [2], [3]. While NTNs were initially designed to complement terrestrial networks in 5G, the vision for 5G-Advanced involves a more cohesive architecture where terrestrial and non-terrestrial systems function as a unified network [4].

The increasing deployment of NGSO constellations and advances in direct-to-device communication capabilities underscore the strategic importance of leveraging diverse multi-orbital configurations and applications to meet evolving user needs. These networks offer distinct advantages, including global coverage and high throughput, yet they introduce new challenges, particularly in the domain of radio resource management (RRM). Efficient allocation of radio resources is essential to ensure reliable and fair user performance, minimize latency, and accommodate the dynamic nature of multi-orbit constellations.

Existing research on RRM for satellite networks has explored a range of approaches. For instance, [5] examines resource management in NGSO systems operating in non-cooperative environments, while [6] employs machine learning techniques to expedite RRM in single-GSO setting. In integrated low Earth orbit (LEO)-terrestrial networks, the RRM problem has been tackled using optimization techniques, as shown in [7]. Genetic algorithms for multi-layer satellite networks [8] and resource allocation strategies for disaster-resilient communication [9] further illustrate the breadth of this field. However, many of these studies focus on specific scenarios, leaving gaps in addressing the complexities of multi-orbit, multi-application systems.

In parallel, multi-connectivity (MC) has emerged as a complementary approach to improve RRM. By enabling simultaneous connections to multiple satellites, MC facilitates load balancing [10], enhances service continuity [11], [12], and boosts energy efficiency [13]. It is particularly valuable for supporting high-throughput applications, as highlighted in [14].

This work has been partially funded by the Luxembourg National Research Fund (FNR) under the project SmartSpace (C21/IS/16193290) and the Marie Speyer Excellence Grant (BrainSat). The work of Mikko Majamaa has been funded by Wihuri Foundation and Riitta and Jorma J. Takanen Foundation. For the purpose of open access, and in fulfillment of the obligations arising from the grant agreement, the author has applied a Creative Commons Attribution 4.0 International (CC BY 4.0) license to any Author Accepted Manuscript version arising from this submission.

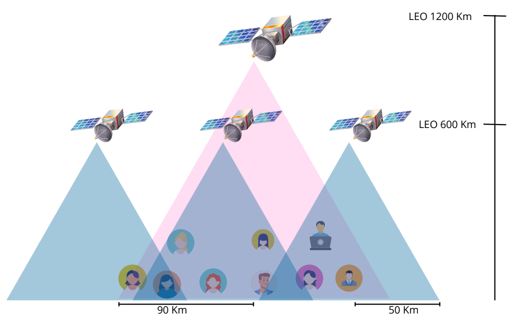

Fig. 1: Illustration of the considered system.

This paper addresses these challenges by investigating RRM in NGSO systems with multi-orbit and multi-application configurations. Our approach accounts for diverse quality of service (QoS) requirements, fairness in resource allocation, and the potential for spectrum coexistence across inter- and intra-orbit systems. To tackle the complexity of resource allocation in such dynamic environments, we design custom Generative AI (Gen-AI) models, specifically transformer-based architectures and generative adversarial networks (GANs). These models are utilized to predict and optimize resource distribution efficiently. At the same time, an optimization method is introduced as a benchmark for comparison, and the speed of inference enables the models to work in real-time.

The remainder of this paper is structured as follows. Section II introduces the system model. Section III details the resource allocation problem and the proposed solutions. Simulation results and analysis are presented in Section V, and Section VI concludes the paper with key findings and future directions.

II. System Model

We consider a multi-constellation LEO satellite network comprising two distinct constellations operating at different orbital altitudes. The first constellation, referred to as LEO600, orbits at an altitude of 600 km and projects beams with a radius of 50 km onto the Earth surface. The second constellation, designated as LEO1200, operates at 1200 km with beam radius of 90 km [2]. The beams are deployed using a quasi-Earth fixed (QEF) configuration, with the set of all beams denoted by $\mathcal{B} = \{1, 2, \ldots, B\}$. Subsets of beams belonging to LEO600 and LEO1200 are denoted by $\mathcal{B}_{600}$ and $\mathcal{B}_{1200}$, respectively. To simplify the analysis, the study confines its coverage area to Luxembourg. The constellations operate on separate frequencies, ensuring negligible inter-orbit interference.

A set of R carriers is assigned to both LEO600 and LEO1200 constellations, with each beam allocated $\frac{R}{3}$ carriers, following the frequency reuse factor (FRF) 3 scheme. The system model is illustrated in Fig. 1. Each satellite in a constellation directs its beam to specific target areas, determined by the satellite's elevation angle.

The network supports U users equipped with omnidirectional antennas (e.g., handheld devices). Each user operates two applications: a low-latency application (e.g., phone calls) and a high-throughput application (e.g., file transfers). These applications are represented by the sets $\mathcal{A}_1$ (low-latency) and $\mathcal{A}_0$ (high-throughput), respectively, with a total of $A = 2U$ applications in the system. Through MC, users can connect their applications to different satellites simultaneously. Low-latency applications are restricted to LEO600 satellites to meet stringent latency requirements, while high-throughput applications can connect to either LEO600 or LEO1200.

The channel model follows the specifications of 3GPP Technical Report (TR) 38.811 [2], assuming the S-band (2-4 GHz) for communication. Signal propagation is affected by various attenuation components, and the cumulative path loss (in dB) is expressed as:

$$\mathrm{PL} = \mathrm{PL}_b + \mathrm{PL}_g + \mathrm{PL}_s, \tag{1}$$

where PL_b is the basic path loss, PL_g accounts for atmospheric gas attenuation, and PL_s represents scintillation-induced attenuation. The basic path loss PL_b is calculated as:

$$\mathrm{PL}_b = \mathrm{FSPL}(d, f_c) + \mathrm{SF} + \mathrm{CL}(\epsilon, f_c), \tag{2}$$

where $\mathrm{FSPL}(d, f_c)$ is the free space path loss, dependent on the slant range d and carrier frequency f_c. Shadow fading loss SF follows a normal distribution, $\mathcal{N}(0, \sigma_{\mathrm{SF}}^2)$, while clutter loss $\mathrm{CL}(\epsilon, f_c)$ depends on the elevation angle ϵ and represents attenuation caused by nearby structures. When the user is in a line-of-sight (LOS) condition, clutter attenuation is minimal.

Atmospheric gas attenuation is influenced by elevation angle, frequency, altitude, and humidity. At frequencies below 10 GHz and elevation angles above 10°, this attenuation is negligible. Similarly, rain and cloud attenuation is insignificant for frequencies under 6 GHz.

The signal-to-noise ratio (SNR) for a user is calculated as:

$$\mathrm{SNR} = \mathrm{EIRP} + \frac{G}{T} - k - \mathrm{PL} - \mathrm{BW}, \tag{3}$$

where EIRP is the effective isotropic radiated power, $\frac{G}{T}$ represents the gain-to-noise temperature ratio of the receiver, k is the Boltzmann constant, and BW is the bandwidth. The parameter $\frac{G}{T}$ is derived from the antenna gain, noise figure, and temperature factors. For detailed NTN link budget analysis, readers are referred to [15].

III. Benchmarking and Training Data Generation

The resource allocation problem considered in this work involves assigning time-frequency resources to applications in both time and frequency domains. Within the 6G-NTN nomenclature, the smallest resource unit assignable by a gNodeB is a Resource Block (RB). Here, we abstract the RB assignment by focusing on carrier frequencies and *fill-rates*. Each beam is allocated a set of carrier frequencies based on the frequency reuse scheme, and the concept of *fill-rate* (ranging between 0 and 1) is introduced to indicate the proportion of a carrier bandwidth that is utilized by each beam and application.

The objective is to maximize the minimum satisfaction level across all applications, defined as the ratio of the supplied throughput s_a to the demanded throughput d_a. The supplied throughput for application a, s_a, is formulated as:

$$s_a = \mathrm{BW} \cdot \sum_{b=1}^{B} \frac{1}{R} \sum_{r=1}^{R} f_{bra} \cdot \mathrm{SE}_{bu}, \tag{4}$$

where BW represents the beam bandwidth, R is the total number of carriers per beam, f_{bra} is the fill-rate of carrier r for beam b assigned to application a, and SE_{bu} is the spectral efficiency of beam b toward user u. Spectral efficiency is calculated using a function $f(\cdot)$ that maps the signal-to-noise ratio (SNR) to spectral efficiency values, tailored to meet specific block error rate (BLER) targets and modulation and coding schemes.

The optimization problem for resource allocation is formally defined in (5a). The objective is to maximize the minimum satisfaction across all applications, where satisfaction is quantified as the ratio of supplied throughput s_a to the demanded throughput d_a. This ensures a fair allocation of resources by prioritizing applications with the least satisfaction levels.

$$\max_{x_{ba}, f_{bra}} \quad \min \frac{s_a}{d_a} \tag{5a}$$

$$\text{s.t.} \quad \sum_{a=1}^{U} f_{bra} \leq 1, \forall b, \forall r \tag{5b}$$

$$x_{ba} = 0, \forall a \in \mathcal{A}_1, \forall b \in \mathcal{B}_{1200} \tag{5c}$$

$$\sum_{b=1}^{B} x_{ba} \leq 1, \forall a \tag{5d}$$

$$x_{ba} = \{0, 1\} \tag{5e}$$

$$f_{bra} \leq x_{ba}, \forall b, \forall a \tag{5f}$$

$$0 \leq f_{bra} \leq 1, \forall b, \forall r, \forall a \tag{5g}$$

$$\epsilon_{\min} \cdot x_{ba} \leq \epsilon_{bu}, \forall b, \forall a \tag{5h}$$

$$\mathrm{SE}_{\min} \cdot x_{ba} \leq \mathrm{SE}_{bu}, \forall b, \forall a \tag{5i}$$

$$f_{bra} = 0, \forall b, \forall r, \forall a, \quad \text{where} \quad r \notin \mathrm{CarrierRange}(b) \tag{5j}$$

The variable $x_{ba} \in \{0, 1\}$ indicates whether application a is assigned to beam b. This binary nature ensures that each application is either fully assigned to a beam or not assigned at all. The variable f_{bra} represents the proportion of carrier r of beam b allocated to application a. It is constrained to values between 0 and 1. Constraint (5b) ensures that the total fill rate across all applications does not exceed the capacity of each carrier. Constraint (5c) enforces that delay-critical applications ($\mathcal{A}_1$) cannot be assigned to LEO1200 beams due to latency requirements. Constraint (5d) ensures each application is assigned to at most one beam. Constraints (5h) and (5i) guarantee that an application can only connect to a beam if minimum elevation angle ($\epsilon_{\min}$) and spectral efficiency ($\mathrm{SE}_{\min}$) requirements are met. Constraint (5j) ensures that each beam operates within its assigned frequency range based on the frequency reuse scheme. The coupling of f_{bra} with x_{ba} in (5f) ensures that fill rates are only allocated for beams assigned to the application.

The max-min optimization problem (5a) is inherently nonlinear due to the minimization within the objective function. To address this, the problem is reformulated by introducing a slack variable ϕ:

$$\max \quad \phi \tag{6a}$$

$$\text{s.t.} \quad s_a \geq \phi \cdot d_a, \forall a \tag{6b}$$

$$(5b), (5c), (5d), (5e), (5f), (5g). \tag{6c}$$

Here, ϕ represents the minimum satisfaction level across all applications, and maximizing ϕ simplifies the objective into a linear form.

Spectral efficiency SE_{bu} is set to zero for beams where either elevation angle or minimum SNR requirements are not met. This reduces the search space and computational overhead. Based on the frequency reuse scheme, carriers are mapped to their designated frequency ranges for each beam. Bandwidth (BW) and carrier count (R) are precomputed per beam.

The problem belongs to the class of mixed-integer linear programming (mixed-integer linear programming (MILP)) problems due to the presence of binary decision variables (x_{ba}). MILP problems are NP-hard, meaning they cannot be solved in polynomial time. However, modern solvers such as Gurobi [16] coupled with CVXPY [17], [18] can provide near-optimal solutions efficiently for moderate problem sizes. The preprocessing steps and problem reformulation significantly reduce the complexity, allowing for feasible computation even with high system dynamics.

In real-world applications, the computational demands of solving MILP problems for high-frequency resource allocation may be prohibitive. To address this, we generate training data using the optimization solver, which is then used to train Gen-AI models, such as transformers and GANs, enabling real-time decision-making without the need for on-the-fly optimization. These models effectively approximate the optimal solutions while maintaining scalability for large-scale systems.

To generate training data for Gen-AI models, we utilize the optimization solver to generate a dataset of optimal resource allocation scenarios under diverse conditions. This data feeds into the AI models, enabling them to learn and predict resource distribution efficiently, bypassing the need for real-time, on-the-fly optimization in dynamic 6G environments. These models simulate potential solutions and accelerate decision-making processes in complex multi-orbit NTN systems, aiming to match the performance of the optimization-based approach while significantly reducing computational overhead.

This approach ensures that AI models are trained with high-quality, realistic scenarios, improving their ability to generalize and perform in real-world settings.

IV. Generative AI Models for Resource Allocation

In this work, we utilize two cutting-edge Gen-AI approaches—Conditional GANs and Convolutional Transform-

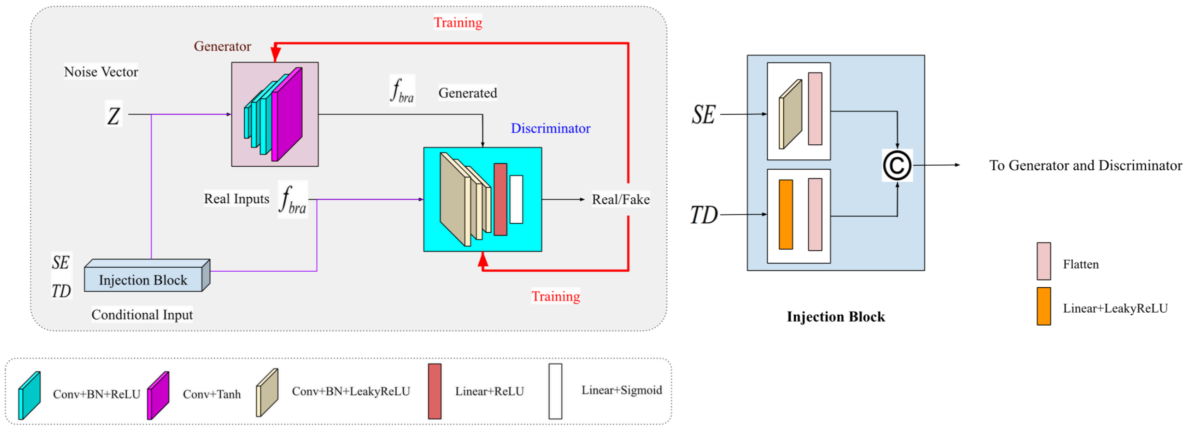

Fig. 2: GAN Architecture

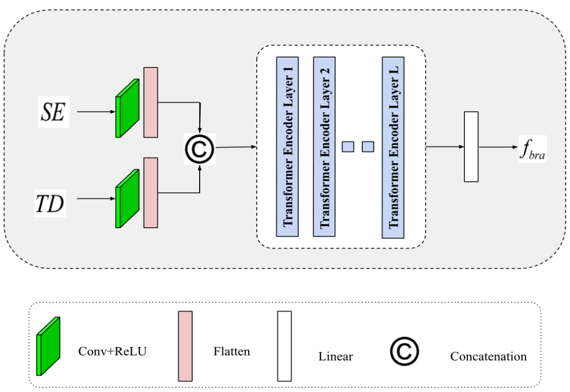

Fig. 3: Convolutional Transformer Architecture

ers—to solve the complex resource allocation problem in (6). These models are designed to learn the intricate patterns of resource distribution in multi-orbit systems, generating solutions adhere to the constraints imposed by user demands and system requirements.

The Conditional GAN (cGAN), as illustrated in Fig. 2, extends the traditional GAN framework by conditioning both the generator and discriminator on additional inputs, specifically the spectral efficiency (*SE*) and throughput demand (*TD*) matrices. The generator in the cGAN takes a noise vector $\mathbf{Z}$ and the conditional inputs to produce fill-rate matrices f_{bra}. The inclusion of conditional inputs ensures that the generated matrices align with the constraints imposed by the satellite system and application-specific QoS requirements. The generator leverages convolutional layers with batch normalization and ReLU activations to stabilize the training process and enhance convergence. The output layer employs a Tanh activation function, producing normalized fill-rate values between 0 and 1. The discriminator, on the other hand, is responsible for distinguishing between real and generated samples. It also receives the conditional inputs to verify the alignment between the generated fill-rates and the given constraints. The discriminator architecture consists of convolutional layers with LeakyReLU activations and a final sigmoid activation for binary classification, determining whether a sample is real or fake. During training, the generator and discriminator compete in an adversarial framework, where the generator aims to produce realistic outputs, while the discriminator learns to distinguish between real and generated data. This adversarial process iteratively improves the quality of the generated fill-rate matrices, leading to outputs that closely resemble real data.

The Convolutional Transformer architecture, depicted in Fig. 3, is designed to model long-range dependencies and capture global patterns in resource allocation. Transformers, with their self-attention mechanisms, excel at identifying relationships across input dimensions, making them highly suitable for generating structured matrices. In this approach, the spectral efficiency (*SE*) and throughput demand (*TD*) matrices are first processed through convolutional layers to capture local spatial dependencies. The resulting feature maps are flattened and linearly transformed into representations that serve as inputs to the transformer encoder. The transformer encoder consists of multiple layers, each employing multi-head self-attention mechanisms and feedforward neural networks to model complex interactions between input features. The self-attention mechanism allows the model to focus on relevant parts of the input, ensuring that the output adheres to the constraints and requirements of the resource allocation problem. The final output of the transformer encoder is passed through a linear layer to generate the fill-rate matrix f_{bra}, which satisfies the constraints imposed by the input conditions.

Both models are designed to generate solutions that maximize the satisfaction of application-specific QoS requirements while adhering to system constraints. The cGAN leverages its adversarial training framework to learn the underlying data distribution, making it highly effective at generating diverse and realistic outputs. In contrast, the transformer-based model focuses on modeling global relationships within the data, ensuring consistency and accuracy in the generated solutions. These complementary capabilities enable a robust evaluation of generative AI techniques for resource allocation in multi-orbit

6G-NTN systems, providing insights into their applicability and effectiveness in solving complex optimization problems.

V. Numerical Results

A. Simulator

The simulations in this study are performed using the Constellation – Dynamic Resource Allocation Management (C-DREaM) simulator [19], [20], which has been developed as part of European Space Agency (ESA)-funded research projects, specifically the C-DREaM and Direct Access Satellite Constellation Emulator (DASCE) projects [21]. The C-DREaM simulator is a comprehensive modular satellite system simulator, designed to model and analyze the performance of satellite communications networks comprehensively.

In each simulation run, the simulator initializes a realistic satellite communication scenario that includes defining users, satellites, and their respective beam coverages. The simulator then computes the SNRs for each link, performs resource allocation, and subsequently calculates the signal-to-interference-plus-noise ratios (SINRs). These metrics are utilized to estimate the realized throughput and generate other relevant statistics, providing insights into the system performance under various conditions.

B. Assumptions and Scenario Setup

This simulation considers the system model described in Section II, with the satellite constellation parameters extensively detailed in Table I. The orbital parameters are adapted from recent constellation data found in literature [22] and slightly modified to fit the scenario requirements.

A total of 6 carriers are utilized for both the LEO600 and LEO1200 configurations, distributed so that each beam is allocated 2 carriers, adhering to the FRF 3 scheme. The constellation comprises 7 beams for LEO600 and a single beam for LEO1200, summing to a total of 8 beams ($B = 8$) in the simulation.

TABLE I: Constellation parameters.

Parameter	LEO600	LEO1200
Beams	7	1
Altitude [km]	600	1200
Tx power [dBW/MHz]	34	40
Orbital planes	36	36
Satellites per plane	36	20
Inclination [degrees]	70	87.9
Beam diameter [km]	50	90
System BW [MHz]	30	30
BW per beam [MHz]	10	10
Center frequency [GHz]	2.0	2.5

These parameters set the stage for a series of simulations designed to evaluate the efficacy of our resource allocation strategies under semi-static conditions. The allocation frequency of 10 seconds ensures dynamic response without overwhelming the system, suitable for the assumed scenario dynamics.

TABLE II: Simulation parameters.

Parameter	Value
Simulation time [s]	250
Users (U)	20
Antenna type	Omnidirectional
User mobility	Static
Traffic type	Constant bitrate
Demand per app [Mbps]	0.5 - 1.5
Channel condition	LOS
Minimum SNR [dB]	-10
MCS index	3 [Table 5.2.2.1-4] [23]
BLER target	0.0
RNG runs	10

C. Simulation Results

This section presents the results of our simulation studies, comparing the performance of the optimization-based algorithm, transformer-based algorithm, and GAN-based algorithm in terms of user satisfaction and computational complexity.

Figure 4 illustrates the Cumulative Distribution Function (CDF) of user satisfaction across different algorithms. The results indicate that all three algorithms achieve high levels of satisfaction, nearing 100% for most users. The transformer-based algorithm and the GAN-based algorithm show marginally lower satisfaction levels compared to the optimization-based algorithm, achieving higher satisfaction for most users with less variation. Specifically, the transformer-based algorithm outperforms the GAN in terms of how quickly it achieves user satisfaction, as is evident from the steeper curve in the initial part of the plot.

Futhermore, figure 4 demonstrates the energy consumption and time required for each algorithm, normalized to the highest values observed across the simulations. The optimization-based algorithm shows the highest energy and time requirements, serving as a benchmark for comparison. Both the transformer and GAN-based algorithms substantially reduce energy consumption, with the transformer algorithm reducing it by approximately 91% and the GAN by 83%. Regarding time complexity, the transformer algorithm again shows a significant reduction, achieving about a 96% reduction, whereas the GAN algorithm reduces the time required by 90%. These results highlight the efficiency of Gen-AI models in handling resource allocation tasks, offering substantial gains in both energy and time efficiency over traditional optimization approaches.

These findings underscore the potential of Gen-AI approaches, particularly the transformer-based model, in enhancing the efficiency and responsiveness of satellite network operations while maintaining high user satisfaction.

VI. Conclusion

This paper presented a comparative analysis of three distinct algorithms—optimization-based, transformer-based, and GAN-based—for resource allocation in satellite communication systems. The numerical results demonstrate that while traditional optimization methods are reliable, they require substantial computational resources. In contrast, the transformer-based and GAN-based algorithms not only achieve comparable levels of

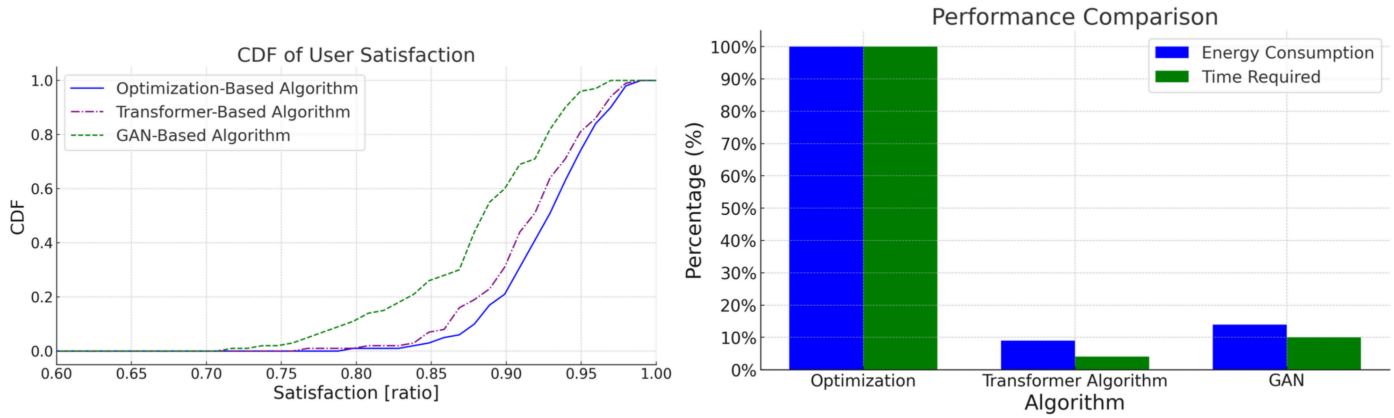

Fig. 4: Performance and complexity comparisons

user satisfaction but also significantly reduce computational complexity, including both energy consumption and time required for computations.

The transformer-based algorithm, in particular, has shown exceptional performance in terms of both high user satisfaction and reduced computational demand. This highlights the efficacy of Gen-AI techniques in managing complex resource allocation tasks in dynamic, multi-orbit satellite networks. These results suggest that AI-driven approaches can be effectively integrated into next-generation satellite communication systems to enhance operational efficiency and user experience.

Future work could explore the scalability of the proposed AI models in larger and more complex satellite constellations. Examining the impact of increased system size and user density on the performance of Gen-AI algorithms will be crucial as satellite networks evolve.

References

[1] "IRIS^2: The new EU secure satellite constellation," accessed on: March 15, 2024. [Online]. Available: https://defence-industry-space.ec.europa.eu/eu-space-policy/iris2_en.

[2] "TR 38.811: Study on new radio (NR) to support non-terrestrial networks," 3GPP, V15.4.0, Sept. 2020.

[3] "TR 38.821: Solutions for NR to support non-terrestrial networks (NTN)," 3GPP, V16.0.0, Jan. 2020.

[4] 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.

[5] E. Re, A. Murrell, and D. Roques, "Radio resource management for large constellations in a spectrum sharing environment," *International Journal of Satellite Communications and Networking*, vol. 39, no. 1, pp. 78–91, 2021.

[6] T. S. Abdu, S. Kisseleff, L. Lei, E. Lagunas, J. Grotz, and S. Chatzinotas, "A deep learning based acceleration of complex satellite resource management problem," in *2022 30th European Signal Processing Conference (EUSIPCO)*, 2022, pp. 1092–1096.

[7] Y. Yuan, L. Lei, T. X. Vu, S. Fowler, and S. Chatzinotas, "Efficient resource scheduling and optimization for over-loaded LEO-terrestrial networks," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 1052–1057.

[8] X. Deng, L. Zhu, Q. Tan, Y. Yang, and Y. Zhang, "Multi-layer satellite network resource management based on genetic algorithm," in *2021 International Symposium on Networks, Computers and Communications (ISNCC)*, 2021, pp. 1–6.

[9] Y. Wang, B. Hu, and S. Chen, "A fair-throughput optimization scheme in emergency satellite communications," in *2021 IEEE/CIC International Conference on Communications in China (ICCC)*, 2021, pp. 618–622.

[10] S. M. Shahid, Y. T. Seyoum, S. H. Won, and S. Kwon, "Load balancing for 5G integrated satellite-terrestrial networks," *IEEE Access*, vol. 8, pp. 132 144–132 156, 2020.

[11] M. López, S. B. Damsgaard, I. Rodríguez, and P. Mogensen, "Connecting rural areas: an empirical assessment of 5G terrestrial-LEO satellite multi-connectivity," in *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1–5.

[12] Y. Sadovaya, O. Vikhrova, S. Andreev, and H. Yanikomeroglu, "Enhancing service continuity in non-terrestrial networks via multi-connectivity offloading," *IEEE Communications Letters*, pp. 1–1, 2024.

[13] M. N. Dazhi, H. Al-Hraishawi, M. R. B. Shankar, S. Chatzinotas, and B. Ottersten, "Energy-efficient service-aware multi-connectivity scheduler for uplink multi-layer non-terrestrial networks," *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 3, pp. 1326–1341, 2023.

[14] M. Majamaa, H. Martikainen, L. Sormunen, and J. Puttonen, "Multi-connectivity in 5G and beyond non-terrestrial networks," *Journal of Communications Software and Systems*, vol. 18, no. 4, pp. 350–358, 12 2022. [Online]. Available: https://doi.org/https://doi.org/10.24138/jcomss-2022-0155

[15] A. Guidotti, A. Vanelli-Coralli, A. Mengali, and S. Cioni, "Non-terrestrial networks: Link budget analysis," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–7.

[16] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," accessed on: March 25, 2024. [Online]. Available: https://www.gurobi.com.

[17] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.

[18] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.

[19] A. Masri, J. Kurjenniemi, J. Puttonen, T. Varjus, and M. Majamaa, "Capacity level simulator for NGSO satellite constellations," *26th Ka and Broadband Communications Conference (KaConf), Arlington, USA, September 27-30, 2021*, 2021.

[20] A. Masri, J. Puttonen, J.-F. Boutillon, J. Capagnon, P. Pontisso, and J. Sprauel, "System level evaluation of LEO 5G non-terrestrial network for handheld user terminals," in *2022 11th Advanced Satellite Multimedia Systems Conference and the 17th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, 2022, pp. 1–6.

[21] "C-DReAM simulator," accessed on: March 25, 2024. [Online]. Available: https://satellite-ns3.com/c-dream-simulator/.

[22] N. Pachler, I. del Portillo, E. F. Crawley, and B. G. Cameron, "An updated comparison of four low earth orbit satellite constellation systems to provide global broadband," in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021, pp. 1–7.

[23] "TS 38.214: NR; Physical layer procedures for data," 3GPP, V18.2.0, March 2024.