Fair and Efficient Transmission of Multi-Omics Data via NOMA: A Hybrid Machine Learning Approach using Random Forest and KNN

Seda KIRTAY[1*][0000-0003-2415-9131], Veysel Gökhan BOCEKCI[2][0000-0003-4559-7173], Kazım YILDIZ[3][0000-0001-6999-1410] and Murat KOCAK[4][0000-0001-6510-3666]

[1*,2] Marmara University, Faculty of Technology, Electrical and Electronics Engineering, Department of Telecommunication, Istanbul 34840, Turkiye
[3] Marmara University, Faculty of Technology, Computer Engineering, Department of Computer Sciences, Istanbul 34840, Turkiye
[4] Baskent University, Faculty of Medicine, Department of Medical Informatics, Ankara, Turkiye.

seda.kirtay@marmara.edu.tr

Abstract. Multi-omics data are produced by merging many biological data types, including proteomics, metabolomics, and genomes. These data help us gain a more thorough understanding of people's biological states and are essential for the creation of individualized healthcare treatment plans and diagnostic techniques. However, a major obstacle to the security and accuracy of health data is the safe and effective transfer of multi-omics data. For the efficient transfer of multi-omics data in wire-free health systems, a method based on non-orthogonal multiple access (NOMA) has been suggested in this paper. NOMA is a technique that uses the same frequency band for several users to share transmission power in an equitable manner. Three distinct users' distances from the base station were used to determine the channel gains for the study, and power allocation was done in accordance with these gains. The inverse channel gain approach has been used to optimize power allocation, while machine learning techniques (K-Nearest Neighbors and Random Forest) have been used to predict power allocation. In addition, the Successive Interference Cancellation (SIC) algorithm has been used to reduce user interference. The results reveal that the proposed NOMA-based approach is very reliable and efficient for multi-omics data transmission, with digital success rates surpassing 95%. Furthermore, with a short processing time (2.18 seconds) and low memory utilization (314.24 MB → 304.49 MB), the suggested system has shown considerable computational efficiency. This study describes an important strategy for increasing the integration of multi-omics data into health information systems.

Keywords: Fair Power Allocation, Machine Learning, Multi-omics Data, NOMA, Non-Orthogonal Multiple Access, Telemedicine

1 Introduction

The incorporation of multi-omics data has been a crucial development in medical research and healthcare in recent years. To provide a more comprehensive understanding of health and disease, multi-omics includes the thorough investigation of multiple biological levels, such as proteomics, metabolomics, and genomes. Researchers can find intricate connections between genetic, protein expression, and metabolic pathways by integrating multiple data types, which will enable more individualized and accurate treatment plans [1]. In this regard, the effective transmission and interpretation of multi-omics data, especially in remote healthcare settings, requires the use of cutting-edge communication technology [2].

With the growing reliance on telemedicine and health information systems, efficient and dependable communication mechanisms are critical for ensuring the secure and real-time transfer of sensitive health data, such as multi-omics data [3]. NOMA is a cutting-edge technology that has received a lot of attention in the field of communication systems. It allows you to share resources efficiently in environments with several users [4]. Communication performance is maximized by separating user signals using techniques like power allocation and SIC (Successful Interference Cancellation). NOMA offers both low latency and high efficiency with equitable power allocation, particularly in situations with a high user density and fluctuating channel circumstances [5,6]. NOMA provides high throughput while reducing latency, making it excellent for data-intensive applications such as telemedicine.

In the context of medical informatics and healthcare communication, this research investigates the relationship between multi-omics data transmission and NOMA-based communication systems. Using NOMA, we suggest a unique power allocation strategy that is fair to users with different channel circumstances. We provide a model that predicts the best power allocation for efficient communication by combining machine learning algorithms like Random Forest and K-Nearest Neighbors. This ensures good data integrity and little interference during multi-omics data transmission.

The purpose of this study is to improve the efficiency of multi-omics data allocation in telemedicine systems, ultimately leading to better healthcare outcomes through improved data-driven decision-making. The suggested strategy has the potential to transform healthcare communications by providing more consistent, accurate, and rapid access to vital health information across many clinical settings.

2 Methodology

2.1 Data Generation

Synthetic multi-omics datasets were created for three users, each representing genomic (10 binary features), proteomic (8 continuous features, 0-100 range), and metabolomic (6 continuous features, 0-50 range) data. These features were organized into hierarchical matrices identified by user.

2.2 Power Allocation via Hybrid AI (Random Forest and K-Nearest Neighbors)

Users (patients with multi-omics data) were stationed at 50, 150, and 400 kilometers from the base station. Corresponding channel gains were calculated as inversely proportional to distance. These gains were used to determine initial power levels, which were then normalized to assure total transmission power conservation. To predict power allocations, a hybrid machine learning model was trained using user distances and channel gains, combining the Random Forest

Regressor and the K-Nearest Neighbors Regressor. Predictions from both models were averaged and re-normalized to ensure compliance with NOMA restrictions.

2.3 System Performance under SIC

Transmission performance was simulated with SNR (Signal-to-Noise Ratio)'s ranging from -20 dB to 150 dB. For every SNR level:

- Optimal SNR per user was calculated after SIC decoding. The Bit Error Rate (BER) was calculated using QPSK modulation assumptions. The throughput was computed using $\log_2(1+SNR)$ and capped at 1 bps to simulate channel saturation.
- Accuracy has been expressed as (1-BER).
- Graphs were created for BER, throughput, and accuracy versus SNR.

3 Result and Discussion

The simulation results show how successful the suggested NOMA-based transmission model is for multi-omics data in wireless healthcare systems. The usage of hybrid machine learning models (Random Forest and K-Nearest Neighbors) allowed for accurate prediction of equitable power allocation across users under various channel circumstances. Notably, the BER stayed below 5% for all users across a wide SNR range, and data delivery accuracy exceeded 95%, demonstrating the model's durability. Digital success rates above 95% show that the suggested solution performs well in terms of dependability and data transfer efficiency. This ratio was computed using the high accuracy value (1 - BER) and the low BER attained under high SNR circumstances. The findings obtained show that the system has nearly excellent fault tolerance and that it is possible to successfully protect the integrity of various omic data transferred in multi-user telemedicine situations. Because of its flexibility and resilience to changing channel circumstances, the hybrid artificial intelligence model that was created in this context allows for continuous and secure data exchange. These findings confirm that implementing AI-driven power allocation algorithms can greatly improve the reliability and performance of telemedicine communications.

- The distant user (400 km) received proportionally more power allocation while keeping a low BER.
- The average accuracy across users improved as SNR increased, demonstrating the hybrid model's predictive stability.
- Graphs demonstrated the proposed system's robustness under various channel circumstances.
- The hybrid solution outperformed static allocation and single-model approaches in terms of fairness and communication dependability.

4 Conclusion

This paper proposes an intelligent power allocation paradigm for power domain NOMA-based wireless devices that transmit complicated multi-omics data into health information systems. The hybrid AI technique assures fairness, security, and high transmission quality even for users at great distances. This approach is highly suited for inclusion into telemedicine infrastructures that require reliable, real-time health data transfer.

Using NOMA, we presented a new hybrid AI-assisted power allocation model in this work for the safe, equitable, and effective transfer of multi-omic data in multi-user telemedicine systems.

Robust and adaptable power forecasts were made possible by the combination of Random Forest and KNN regressors with distance-based channel gain estimations, which achieved a success rate of over 95% under high SNR circumstances.

With a processing time of just 2.18 seconds and a low memory consumption (first 314.24 MB, then 304.49 MB), the suggested hybrid AI-supported NOMA system has proven its high efficiency and computational lightness.

These results have important ramifications for the creation of intelligent and scalable healthcare communication systems. The suggested approach enhances the dependability of customized remote health monitoring by guaranteeing high data integrity and reducing bit error rates. Additionally, this method enables real-time, physiologically informed telehealth decision-making by bridging the domains of artificial intelligence, communication systems, and medical informatics.

In order to further enhance the usability of this strategy in contemporary healthcare ecosystems, future research may involve clinical validation using actual patient data and real-time deployment utilizing embedded devices.

5 References

[1]. Babu, M., & Snyder, M. (2023). Multi-omics profiling for health. Molecular & Cellular Proteomics, 22(6).

[2]. Thacharodi, A., Singh, P., Meenatchi, R., Tawfeeq Ahmed, Z. H., Kumar, R. R., V, N., ... & Hassan, S. (2024). Revolutionizing healthcare and medicine: The impact of modern technologies for a healthier future—A comprehensive review. Health Care Science, 3(5), 329-349.

[3]. Vishweshwara, A., & Ramya, R. (2025). Transforming Telemedicine: Reducing Latency Through Edge Computing and 5G—A Review. *Biomedical Materials & Devices*, 1-14.

[4]. Kirtay, S., Yildiz, K., & Bocekci, V. G. (2024). Artificial Intelligence-based Fair Allocation in NOMA Technique: A Review. International Journal of Sensors Wireless Communications and Control, 14(3), 161-174.

[5]. Rehman, S. U., Ahmad, J., Manzar, A., & Moinuddin, M. (2024). Beamforming techniques for mimo-noma for 5g and beyond 5g: Research gaps and future directions. *Circuits, Systems, and Signal Processing*, *43*(3), 1518-1548.

[6]. Kirtay, S., Bocekci, V. G., & Yildiz, K. (2025). Genetic Algorithm Based Approach for Optimization of Fair Power Allocation in Pd-NOMA Systems. *Tehnički vjesnik*, *32*(3), 948-957.