An Efficient User Clustering in IRS-Assisted NOMA Systems

Kanchana Katta
Department of ECE
IIIT Manipur, Imphal, 795002, India

Ramesh Ch Mishra
Department of ECE
IIIT Manipur, Imphal, 795002, India

ABSTRACT

The integration of non-orthogonal multiple access (NOMA) and intelligent reflecting surfaces (IRS) has emerged as a promising approach for enhancing spectral efficiency, user connectivity, and fairness in next generation wireless networks. However, the overall system performance is highly contingent on effective user clustering mechanisms, which directly impact successive interference cancellation (SIC) and optimal power allocation. Conventional clustering algorithms, such as K-Means, DBSCAN, and hierarchical clustering, often encounter scalability issues and yield imbalanced user distributions across clusters, adversely affecting system level metrics. In this study, we introduce a Balanced K-Means clustering framework specifically tailored for IRS-assisted NOMA systems. The proposed method addresses the imbalance and overfitting challenges inherent in traditional clustering by enforcing approximately equal cluster sizes, thereby facilitating more equitable resource allocation and robust SIC performance. The simulation results were simulated over a normalized synthetic dataset representing user spatial features and the proposed Balanced K-Means clustering outperforms conventional methods, achieving higher throughput and better bit error rate for varying user densities and IRS configurations.

Keywords

Non-orthogonal multiple access (NOMA), Intelligent reflecting surface (IRS), User clustering, Balanced K-means, Deep learning

1. INTRODUCTION

The increasing demand for ubiquitous connectivity, massive device access, and high data rates has placed significant strain on the capabilities of conventional radio access schemes. Next generation wireless networks such as 6G aim to address these challenges through paradigm shifting innovations that include advanced multiple access techniques, intelligent radio environments, and machine learning aided optimization. In this context, NOMA has emerged as a promising solution to improve spectral efficiency, support massive connectivity, and enhance user fairness [1]-[3]. Unlike orthogonal multiple access (OMA) techniques such as TDMA, FDMA, and OFDMA that assign orthogonal resources to users, NOMA superimposes user signals in the power domain and relies on signal level processing to separate them at the receiver using SIC [4].

In power-domain NOMA, the base station transmits a linear combination of user signals with different power levels, exploiting users heterogeneous channel conditions. While the high channel gain user applies decoding and signal cancellation, the weak user directly decodes its signal. This allows NOMA to outperform OMA in terms of spectral efficiency and throughput [5]. However, the performance of NOMA is highly sensitive to user clustering and power allocation strategies, which determine the effectiveness of multi-user decoding and resource distribution [6, 7].

To further improve propagation conditions and enable environmental control, IRS have been proposed as a novel complement to traditional infrastructure [8, 9]. IRSs consist of a large array of passive reflecting elements that can reconfigure wireless signals via programmable phase shifts. By integrating IRS with NOMA [10], systems can leverage IRS to adjust the channel response and enhance the signal strength received by weaker users. However, the integration of IRS with NOMA increases the system design complexity particularly in user clustering and beamforming [11]-[15].

Recent studies have investigated various user clustering methods to improve NOMA performance. Traditional clustering algorithms such as K-Means, DBSCAN, and Hierarchical Clustering have been widely used in wireless communication to group users based on distance, channel gain, or other features [16, 17]. However, K-Means suffers from unbalanced clustering, especially in heterogeneous scenarios, leading to poor power domain separation. Balanced K-Means has been proposed as an enhanced variant that enforces nearly equal sized clusters, improving fairness and decoding reliability [18]. It is especially suitable for IRS-assisted NOMA where even distribution across IRS-controlled beams can significantly impact performance.

Additionally, machine learning (ML) and deep learning (DL) models have been increasingly applied to automate clustering and user pairing decisions [19]-[21]. These methods learn complex patterns in real time data and adapt clustering based on user behavior, location, and channel statistics. Nevertheless, traditional clustering remains crucial due to its interpretability and lower computational complexity, especially when implemented on embedded wireless devices.

This work proposes and evaluates a Balanced K-Means clustering algorithm tailored for IRS-assisted NOMA downlink systems. The performance compare against standard K-Means, Hierarchical, and DBSCAN techniques in terms of throughput, user fairness, and decoding reliability. MATLAB based simulation validates the supe-

riority of Balanced K-Means in forming efficient user groups under practical constraints. Unlike deep learning methods that require extensive training, the proposed method remains lightweight and interpretable, making it suitable for 6G use cases with energy and latency constraints.
The rest of the paper is organized as follows: Section 2 presents the proposed IRS-assisted NOMA system along with the integration of Balanced K-Means clustering for user grouping. Section 3 provides comprehensive simulation results, evaluating the proposed framework in terms of system throughput and bit error rate under varying SNR conditions. Finally, Section 4 concludes the paper and outlines possible directions for future research.

2. SYSTEM MODEL

Consider a downlink IRS-assisted NOMA system where a single base station (BS) serves K users, indexed by $\mathcal{K} = \{1, 2, \ldots, K\}$. The communication environment includes an IRS composed of N passive reflecting elements. The users are randomly distributed, and each experiences small scale fading and large scale path loss from both the BS and the IRS. The system includes two main wireless links: The direct BS to user link and the BS to IRS then IRS to user cascaded link. Also, it is assumed that there is no direct link between BS and user. The system model of IRS-assisted NOMA for downlink is shown in Fig. 1. The BS transmits a superimposed

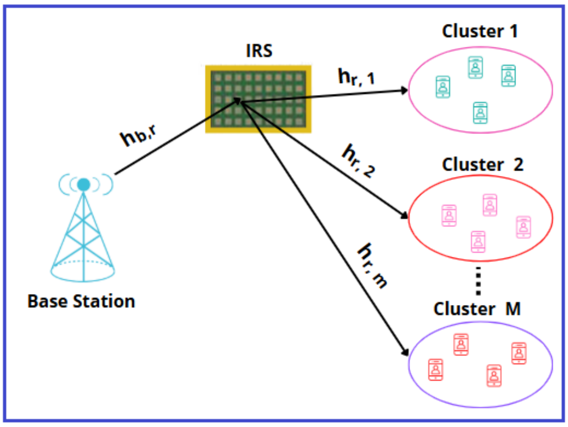

Fig. 1: IRS-assisted NOMA user clustering system model

signal s, which is the power weighted sum of all individual users data symbols:

$$s = \sum_{j=1}^{K} \sqrt{p_j} x_j \tag{1}$$

where x_j is the normalized complex data symbol for user j, often modeled as a unit power symbol, i.e. $E[|x_j|^2] = 1$, p_j is the transmit power allocated to user j, satisfying the total power constraint:

$$\sum_{j=1}^{K} p_j \leq P_{\text{total}} \tag{2}$$

The received signal at user k can be expressed as:

$$y_k = \left(\mathbf{h}_{r,k}^H \boldsymbol{\Theta} \mathbf{h}_{b,r}\right) \sum_{j=1}^{K} \sqrt{p_j} x_j + n_k \tag{3}$$

where $\mathbf{h}_{b,r} \in C^{N\times 1}$ is the channel from the BS to IRS, $\mathbf{h}_{r,k} \in C^{N\times 1}$ is the channel from IRS to user k, $\boldsymbol{\Theta} = \text{diag}(e^{j\theta_1}, \ldots, e^{j\theta_N})$ is the IRS reflection matrix, and $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN noise. Assuming perfect channel state information and effective IRS phase shift design, the total channel gain for user k is:

$$g_k = \left|\mathbf{h}_{r,k}^H \boldsymbol{\Theta} \mathbf{h}_{b,r}\right|^2 \tag{4}$$

The signal-to-interference-plus-noise ratio (SINR) for the k-th user (with SIC decoding) is:

$$\text{SINR}_k = \frac{p_k g_k}{\sum_{j=k+1}^{K} p_j g_k + \sigma^2} \tag{5}$$

The achievable rate of user k is given by:

$$R_k = \log_2 (1 + \text{SINR}_k) \tag{6}$$

The systems total sum rate is:

$$R_{\text{sum}} = \sum_{k=1}^{K} R_k \tag{7}$$

2.1 Balanced K-Means clustering for IRS-Assisted NOMA

In this section, the Balanced K-Means clustering algorithm tailored for IRS-assisted NOMA systems is described. The goal is to partition users into M balanced clusters based on multi dimensional features relevant to NOMA and IRS optimization.

Algorithm 1 Balanced K-Means Clustering for IRS-Assisted NOMA

Input: Feature matrix $X = \{x_1, x_2, ..., x_K\}$ where $x_k = [d_k, g_k, \text{SINR}_k, p_k]$, number of clusters M
Output: Balanced clusters $\mathcal{C} = \{C_1, C_2, ..., C_M\}$

for each user $k \in \mathcal{K}$ **do**
  Compute effective channel gain:
$$g_k = \left|\mathbf{h}_{r,k}^H \boldsymbol{\Theta} \mathbf{h}_{b,r}\right|^2$$
  Form feature vector:
$$x_k = [d_k, g_k, \text{SNR}_k, p_k]$$
end for
Normalize features: $z_k = \frac{x_k - \mu}{\sigma}$
Randomly initialize M centroids: $\mu_1, \mu_2, ..., \mu_M$
Initialize empty clusters $C_1, C_2, ..., C_M$
repeat
  for each user z_k **do**
    Compute distances: $d_{k,m} = \|z_k - \mu_m\|^2$
    Assign z_k to cluster C_m such that $|C_m| < \left\lfloor \frac{K}{M} \right\rfloor$
  end for
  for each cluster C_m **do**
    Update centroid: $\mu_m = \frac{1}{|C_m|} \sum_{z_k \in C_m} z_k$
  end for
until centroids converge
return Balanced clusters $\mathcal{C} = \{C_1, C_2, ..., C_M\}$

3. SIMULATION AND RESULTS

To evaluate the effectiveness of the proposed Balanced K-Means clustering integrated with IRS-assisted NOMA, MATLAB based simulations are conducted using relevant wireless parameters. The simulation scenario includes a single cell downlink NOMA network consisting of a base station, an IRS comprising N passive elements, and K users uniformly distributed in the cell. The clustering performance of different user grouping techniques for IRS-assisted NOMA was evaluated. Fig. 2 through Fig. 5 illustrate the clustering outcomes of each algorithm and the average Silhouette coefficient is employed as a performance metric to quantify the cohesiveness and separability of the formed clusters. Each clustering method affects how users are grouped for NOMA transmission and consequently influences IRS phase alignment and power-domain multiplexing efficiency.

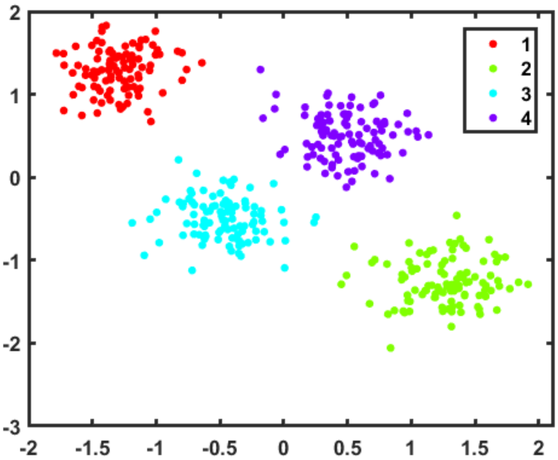

Fig. 2: K-means Clustering

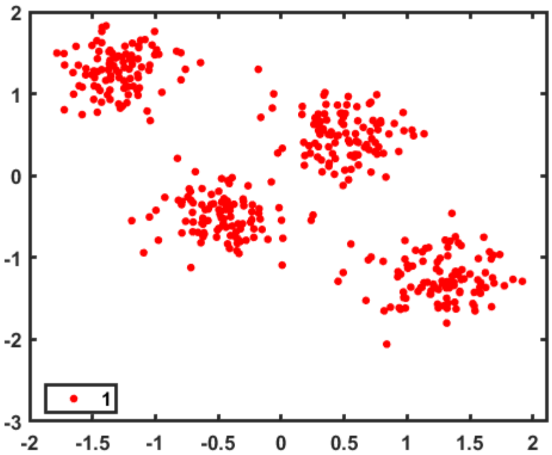

Fig. 3: DBSCAN Clustering

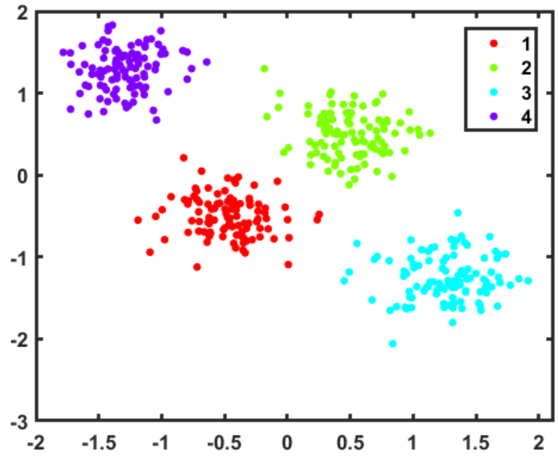

Fig. 4: Hierarchical Agglomerative Clustering

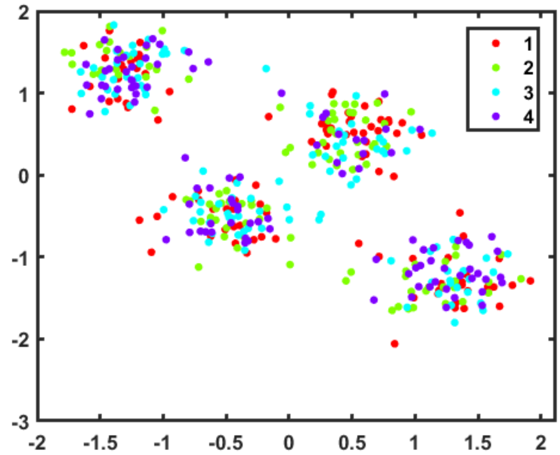

Fig. 5: Balanced K-Means Clustering

The **silhouette score** is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \tag{8}$$

Where $a(i)$ is the average distance of point i to all other points in the same cluster and $b(i)$ is the average distance of point i to all points in the nearest different cluster.
The silhouette score $s(i)$ ranges within:

$$s(i) \in [-1, 1]$$

The clustering performance of the four evaluated techniques is quantitatively measured using the Silhouette Score, which reflects the reflects the compactness and separability of the formed clusters. Among the methods, K-Means achieved the highest Silhouette Score of 0.781, indicating well-separated clusters with tight intra cluster similarity. However, K-Means does not ensure balanced cluster sizes, which is critical in IRS-assisted NOMA for fair resource distribution.
The proposed Balanced K-Means approach attained a competitive score of 0.707, demonstrating its ability to form compact clusters

while simultaneously maintaining uniformity in cluster sizes an essential requirement for optimized IRS beamforming and NOMA power allocation. Hierarchical clustering also showed decent performance with a score of 0.693, but its rigid structure may not adapt well to dynamic user distributions. DBSCAN, with the lowest score of 0.535, was less effective due to its sensitivity to parameter tuning and its handling of border or noise points, making it less suitable for consistent user grouping in the considered wireless system. Overall, Balanced K-Means offers a practical trade off between clustering quality and fairness, justifying its selection as the proposed technique.

We also evaluated performance metrics such as sum rate, and bit error rate for all considered clustering schemes to quantify their impact on system throughput and reliability. The major simulation parameters used for this study, such as transmit power, path loss exponent, noise variance, IRS element count, and NOMA power allocation factors are listed in Table 1.

Table 1. : System and Simulation Parameters

Parameter	Value
Number of Users (K)	400
Number of IRS Elements (N)	64
Path Loss Exponent	4
Noise Power Density	-174 dBm/Hz
Bandwidth	10 MHz
Modulation Scheme	BPSK
Power Allocation	1 watt

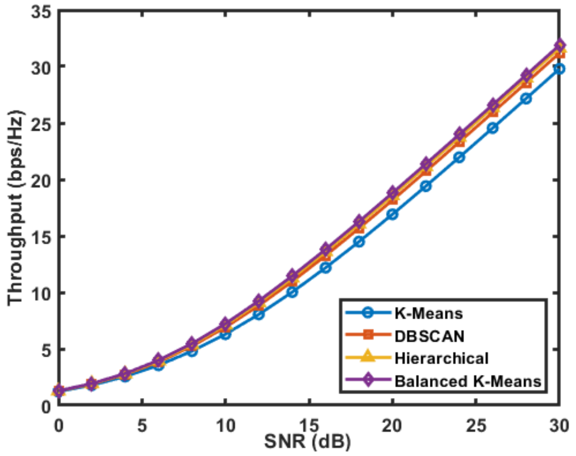

Fig. 6: Sum rate comparison of different clustering schemes

Fig. 6 shows the sum rate comparison of different clustering schemes. This plot depicts how sum rate improves with increasing SNR for each clustering technique. The Balanced K-Means method shows consistently higher throughput, followed by Hierarchical, DBSCAN, and standard K-Means. The result signifies that well distributed clusters in Balanced K-Means enable more efficient spectrum utilization and user multiplexing.

Fig. 7 illustrates the BER performance for four clustering techniques in a IRS-assisted NOMA system across SNR values ranging from 0 to 30 dB. As SNR increases, BER decreases exponentially for all clustering methods, with Balanced K-Means achieving the lowest BER across all SNR levels. This demonstrates that Balanced K-Means clustering provides more distinct and reliable user separation, leading to reduced interference and error rates.

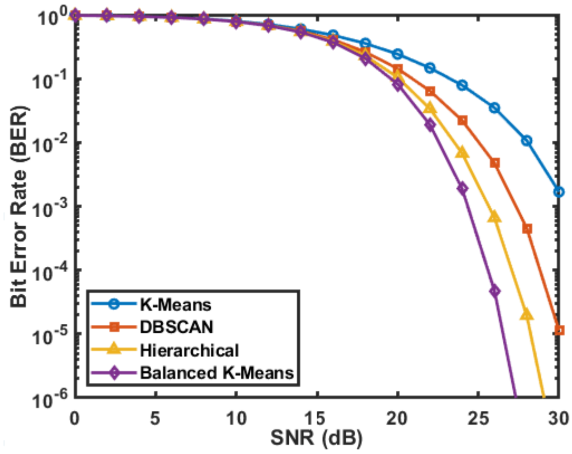

Fig. 7: Ber performance of clustering schemes

4. CONCLUSION

This paper presented a comparative analysis of traditional clustering techniques K-Means, DBSCAN, Hierarchical, and the proposed Balanced K-Means algorithm in the context of an IRS-assisted downlink NOMA system. The simulation results clearly demonstrated that Balanced K-Means significantly outperforms other methods in terms of sum rate, and BER across varying SNR levels, due to its ability to maintain equal user distribution among clusters. The method showed enhanced stability and efficiency in optimizing user pairing and power allocation when integrated with IRS-assisted channels. Overall, the proposed clustering framework effectively supports reliable and high-throughput communication in intelligent wireless environments. Future work can explore adaptive and real time clustering strategies that account for user mobility and dynamic channel conditions. Such methods could leverage machine learning and reinforcement learning paradigms to dynamically optimize clustering decisions, thus further enhancing system throughput, latency, and resilience in next-generation wireless networks.

5. REFERENCES

[1] Z. Ding, et al., “A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends,” IEEE J. Sel. Areas Commun., vol. 35, no. 10, pp. 2181–2195, 2017.

[2] L. Dai, et al., “Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends,” IEEE Commun. Mag., vol. 53, no. 9, pp. 74–81, 2015.

[3] Y.Liu, Z.Qin, M.Elkashlan, Z.Ding, A.Nallanathan, and L.Hanzo, “Non-orthogonal multiple access for 5G and beyond,” *Proceedings of the IEEE*, vol.105, no.12, pp. 2347–2381, 2017.

[4] K. Katta, R. C. Mishra and K. Deka, ”Optimal user pairing using the shortest path algorithm,” 2021 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Hyderabad, India, pp. 1-6, 2021.

[5] S. M. R. Islam, et al., "Power-Domain NOMA in 5G Systems: Potentials and Challenges," IEEE Commun. Surveys Tuts., vol. 19, no. 2, pp. 721–742, 2017.

[6] Y. Zeng, et al., "Graph-Based User Clustering for NOMA Systems," IEEE Trans. Veh. Technol., vol. 70, no. 4, pp. 3753–3764, 2021.

[7] S. Timotheou and I. Krikidis, "Fairness for Non-Orthogonal Multiple Access in 5G Systems," IEEE Signal Process. Lett., vol. 22, no. 10, pp. 1647–1651, 2015.

[8] Q. Wu and R. Zhang, "Towards Smart and Reconfigurable Environment: Intelligent Reflecting Surface Aided Wireless Network," IEEE Commun. Mag., vol. 58, no. 1, pp. 106–112, 2019.

[9] C. Huang, et al., "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4157–4170, 2020.

[10] S. Sharma, K. Deka, D. Dixit, and R. Ahmad, "Intelligent reflecting surface-assisted downlink nonorthogonal multiple access systems," *International Journal of Communication Systems*, vol. 35, no. 3, pp. 1–12, 2022.

[11] G. Lee and Y. Sung, "Optimization of Reflecting Elements for IRS-Aided NOMA Clustering," IEEE Wireless Commun. Lett., vol. 9, no. 10, pp. 1650–1654, 2020.

[12] X. Guo, et al., "Joint IRS Phase Shift and User Clustering for NOMA Systems," IEEE Trans. Commun., vol. 70, no. 3, pp. 1983–1996, 2022.

[13] T. Nguyen, et al., "Deep Learning-Based Resource Allocation in NOMA Systems," IEEE Trans. Veh. Technol., vol. 69, no. 7, pp. 7946–7950, 2020.

[14] Y. Shi, et al., "Deep Learning for IRS-Aided NOMA with Channel Estimation," IEEE Trans. Commun., vol. 69, no. 12, pp. 8256–8268, 2021.

[15] J. Kim and H. Park, "Fairness-Aware Clustering and Power Control for IRS-Assisted NOMA," IEEE Access, vol. 10, pp. 31764–31775, 2022.

[16] M. Hasan, et al., "User Clustering for Downlink NOMA in 5G Wireless Systems," IEEE Access, vol. 6, pp. 74051–74065, 2018.

[17] J. Cui, Z. Ding, et al., "Unsupervised Machine Learning-Based User Clustering in mmWave-NOMA," IEEE Trans. Wireless Commun., vol. 17, no. 11, pp. 7425–7440, 2018.

[18] H. Wang, et al., "Balanced K-Means Clustering With Applications in Wireless Resource Allocation," IEEE Access, vol. 8, pp. 37781–37791, 2020.

[19] Y. Sun, et al., "Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Challenges, and Future Directions," IEEE Commun. Mag., vol. 58, no. 1, pp. 94–100, 2019.

[20] Y. Zhang, et al., "Deep Q-learning for Downlink NOMA User Clustering," IEEE Wireless Commun. Lett., vol. 11, no. 1, pp. 45–49, 2022.

[21] M. Zhao, et al., "Federated Learning for NOMA User Grouping in 6G Networks," IEEE Internet Things J., vol. 9, no. 15, pp. 13544–13556, 2022.