User Clustering Techniques for Massive MIMO-NOMA Enabled mmWave/THz Communications in 6G

Md. Shahjalal, *Student Member, IEEE,* Md. Habibur Rahman, *Student Member, IEEE,* Md. Osman Ali, ByungDeok Chung*, and Yeong Min Jang, *Member, IEEE*
Department of Electronics Engineering, Kookmin University, Seoul 02707, Korea
*ENS. Co. Ltd, Ansan 15655, Korea
Email: mdshahjalal26@ieee.org; rahman.habibur@ieee.org; osman@kookmin.ac.kr; yjang@kookmin.ac.kr; *bdchung@ens-km.co.kr

***Abstract*—Recently, Cooperative massive multiple-input multiple-output and non-orthogonal multiple access (mMIMO-NOMA) has been considered as a promising solution that can significantly improve the system capacity and the spectral efficiency of the sixth-generation (6G) high frequency spectrum such as Millimeter Wave and Terahertz networks. In this paper, we consider a mMIMO-NOMA enabled base station that can support a number of single antenna users in different clusters. Cooperative use of NOMA can support the users in a cluster by sharing the same frequency and time resources. However, in 6G the networks will be congested with ultra-massive interconnected users and that arises challenges in clustering the users efficiently. Therefore. we briefly summarize the studies about user clustering solutions in mMIMO-NOMA systems and divided them into two categories; resource aware user clustering (RAUC) and learning assisted user clustering (LAUC) approaches. A comparison among those techniques has been tabulated considering the computational complexities. The result depicts that the RAUC demonstrates a polynomial complexity function while that for the LAUC is comparatively low.**

***Index Terms*—User clustering, cooperative mMIMO-NOMA, energy efficiency, computational complexity, massive interconnectivity, 6G.**

I. Introduction

As a successor of fifth-generation, sixth-generation (6G) of wireless networks is being designing to provide ultra-low-latency, ultra-high-throughput, ultra-massive connectivity, highly reliable energy and spectrum efficiency, etc. [1]. In recent years, the increase in data traffic has skyrocketed owing to the various new services provided by smart devices and machines [2]. Therefore, millimeter-wave (mmWave) and terahertz (THz) domains are being extensively researched to support the anticipated ultra-high demand [3]. mmWave, core of 5G networks, offers a bandwidth (24-300 GHz) to meet the increased demand [4]. On the other hand, THz band, having a bandwidth of 3-4 times compared with the present wireless technology, 10 Gbps transmission rate has been already achieved. This massive capacity is advantageous in terms of improved directionality, security, and anti-interference capability [5]. As the path loss is extremely high at this high frequencies, multi-antenna technique is adopted to remediate this problem [3]. Aside from these, massive multiple-input multiple-output (mMIMO) non-orthogonal multiple access (NOMA) is also considered as a significant solution for data rate augmentation that can be applied in mmWave and THz communication systems [3]. In NOMA, multiple users can be served with the same spectrum simultaneously by superimposing the desired signals in power domain where each user is served independently in time and frequency domain in OMA [6], [7]. As a result, on the receiver side, inter user interference is generated that is eliminated utilizing successive interference cancellation to decode the desired signal. NOMA outperforms the OMA in terms of throughput, fairness, and spectral efficiency [4].

Most of the existing mMIMO-NOMA studies in mmWave/THz networks focuses on the performance analysis and hardly focuses on the importance of user clustering. Moreover, massive user connectivity with better spectral efficiency requirements in 6G creates an obligation towards structured user grouping in NOMA enabled networks. In addition, though sufficient studies have been conducted on user clustering in low-frequency band networks, the area of mmWave/THz networks are yet to be investigated extensively. However, there are few studies that considered user pairing when clustering the users in MIMO-NOMA system for limited active users [8]–[10]. As a more recent work, [11] divides users into two groups: cellular and device-to-device (D2D) and applies channel correlation based cluster matching algorithm. However, the system produces polynomial computational complexities in user clustering. Learning assisted clustering techniques shows less complexity however has a drawbacks of effective initialization of cluster heads.

Therefore, in this paper, a brief study has been conducted on the proposed clustering techniques in mMIMO-NOMA-based mmWave/THz communication architectures. We summarized the techniques and divided into two categories; joint resource aware user clustering (RAUC) and learning assisted user clustering (LAUC). The RAUC techniques include channel correlation metrics, power and energy efficiency (EE), and spa-

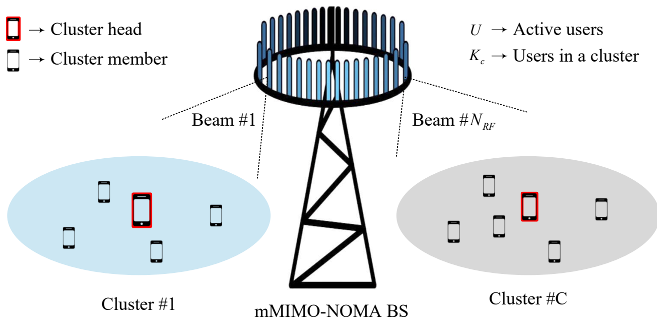

Fig. 1. User clustering in mMIMO-NOMA downlink network.

tial positions. Different supervised and unsupervised machine learning (ML) approaches such as k-means algorithm, fuzzy c-means algorithms, and classification-based ML algorithms are categorized into the LAUC techniques. Finally, we compared those techniques considering their computational complexities. The remainder of this paper is arranged as follows. Section II illustrates the mMIMO-NOMA system and channel modeling of the system. Section III presents the clustering algorithms and comparative analysis among them. Finally, Section IV concludes the paper.

II. mMIMO-NOMA System Model

In this paper, we consider a mMIMO-NOMA enabled base station equipped with N transmit MIMO antennas and N_{RF} RF chain connections that supports high frequency bands (mmWave/THz) for the downlink communications. The base station can support U users equipped with single antenna. NOMA is used to incorporate more than one user in a group at each transmission beam. In general, the number of clusters is made equal to the number of RF chain to maximize the multiplexing gain. Different clustering techniques that can be applied to group the users are discussed in Section III. As shown in Fig. 1, a clustering algorithm is applied to the mMIMO-NOMA system that partitions U users into C clusters with K_c user points per cluster, where $c \in C$. According to the proposed system, each RF chain of the mMIMO setup is cooperating with the power domain NOMA concept that can produce power separated beamforming for the every clusters.

The received signal of the ith user at the cth cluster can be given as

$$y_c^{(i)}(t) = \overline{h}_c^{T(i)} A_n \sum_{c=1}^{C} D_c X_c(t) + \eta_c^{(i)}. \quad (1)$$

The $N \times 1$ channel vector is represented by $h_c^{(i)} = d_c^{-\upsilon/2^{(i)}} \xi_c^{(i)}$, where we consider Rayleigh multipath fading channel $\xi_c^{(i)}$ between the ith user in the cth cluster and the antenna $n \in N$. The signal amplitude decays with increasing distance between the ith user in the cth cluster and the antenna n according to $d_c^{-\upsilon/2}$, where υ denotes the path loss exponent. In (1), A_n represents the analog precoding matrix of size $N \times N_{RF}$ for a fully-connected architecture, D_c is the digital precoding vector of size $N_{RF} \times 1$, and η_c is a complex Gaussian thermal noise represented as $\mathcal{CN}(0, \sigma_b^2)$. The $X_c(t)$ represents the superimposed signal for the all K_c users in cth cluster that can be written as

$$X_c(t) = \sum_{i=1}^{K_c} \sqrt{\rho_c^{(i)} P_c} x_c^{(i)}(t), \quad (2)$$

where $x_c^{(i)}$ represents the signal of the ith user in the cth cluster at time t. P_c is the set of transmit power and $\rho_c^{(i)}$ denotes the power coefficient of the ith user in the cth cluster, where both of them are applied with the conditions $\sum_{c=1}^{C} P_c \leq P_t, \quad \forall n \in N$, and $\sum_{i=1}^{K_c} \rho_c^{(i)} = 1, \quad \forall n \in N, \forall c \in C$. Here, P_t is the total transmit power. The received signal in (1) can be divided into several parts considering desired signal, the intra-cluster interference, and inter-cluster interference as

$$y_c^{(i)}(t) = d_c^{-\upsilon/2^{(i)}} \xi_c^{(i)} A_n D_c \sqrt{\rho_c^{(i)} P_c} x_c^{(i)}(t) + I_{intra-cluster} + I_{inter-cluster} + \eta_c^{(i)}, \quad (3)$$

where $I_{intra-cluster}$ and $I_{inter-cluster}$ denote the intra-cluster interference and inter-cluster interference, respectively.

The achievable rate of the ith user in cth cluster can be found as follow

$$R_c^{(i)} = \frac{\xi}{N} \log_2 \left(1 + \frac{\rho_c^{(i)} P_c \left\| h_c^{H(i)} A_n D_c \right\|_2^2}{\lambda_c^{(i)}} \right), \quad (4)$$

where ξ is the bandwidth and $\lambda_c^{(i)}$ is the interference and noise term of the signal-to-interference-plus-noise-ratio (SINR) $SINR_c^{(i)}$. Finally, the achievable sum rate can be written as

$$R_c^{sum} = \sum_{j=1}^{C} \sum_{i=1}^{K_c} R_c^{(i)}. \quad (5)$$

III. Clustering Approaches in mMIMO-NOMA System

A brief survey has been carried out on the clustering algorithms for mMIMO-NOMA-enabled networks and divided into the following two categories.

A. Joint Resource Aware User Clustering (RAUC)

Channel correlation coefficient-based clustering: Interference aware approaches can be considered as one of the resource constrained clustering technique. In [11], an interference aware graph based clustering approach was proposed for two different types of users: cellular and D2D. A channel graph was constructed by the BS for cellular users by measuring the channel correlation between two users. The users are assigned in clusters exhibit low channel correlation to eliminate inter-beam interference and within clusters have high channel correlations to elevate the system robustness. However, highly correlated channels can create intra-channel interference. On the other hand, the D2D users are partitioned

Algorithm 1 K-means user clustering in mMIMO-NOMA

Input: U and C
Output: K_c and C, $\forall\ c \in C$
Initialize cluster head set $C_v = \emptyset$ and $c = 1$;
while $c \leq C$ **do**
  Randomly select a cluster head v_c from U ;
  Update $C_v = \{v_c,\ \forall c \in C\}$;
  $c = c + 1$;
end while
repeat
  For each user $m \in U/C$, calculate the minimum Euclidean distance from the v_c ;
  Users will fit to the closest clusters;
  Update the cluster head v_c by taking the average of all the users m ;
until The cluster members do not change

into different clusters by considering interference graph. Here, the highly interfered users are placed in separate clusters to reduce the intra-cluster and inter-cluster interference.

When completing both the channel and interference graph construction for cellular and D2D users, respectively, clusters matching is applied to perform best match between each D2D pair cluster to its best cellular user cluster. This matching technique enables better spectrum efficiency for this kind of user scenario under a single NOMA cluster. The proposed scheme has a polynomial complexity comprising the complexities for cellular user clustering, D2D pair clustering, and matching operation.

Clustering based on energy efficient user admission: EE is a major concern for mMIMO-NOMA systems in 6G. Multiple user clustering employing efficient power allocation can maximize the EE. As a study in this area, a multi-user admission concept in multi-cluster was developed with predefined QoS requirement considering the EE in [12]. This concept was formulated as

$$\max_{\rho_c^{(i)}} \sum_{m=1}^{C} \sum_{i=1}^{K_c} \Omega_c^{(i)} \tag{6a}$$

$$s.t.\ \ R_c^{(i)} \geq min(R_c^{(i)})\Omega_c^{(i)}, \tag{6b}$$

$$\sum_{m=1}^{C} \sum_{i=1}^{K_c} \rho_c^{(i)} \leq 1, \tag{6c}$$

$$\Omega_c^{(i)} \in \{0, 1\}, \tag{6d}$$

where $\Omega_c^{(i)}$ is a binary decision variable. The user admission process can be briefly described as follows: users of highest channel gain from each cluster $c \in C$ are selected and computed the resource requirement; the user with minimum of the requirement is admitted until the total resources exceeds.

Algorithm 2 FCM clustering algorithm in mMIMO-NOMA

Input: Channel QoS metrics $Q = [q_{11}, ..., q_{K_cF}]$
Output: S and v_c
Initialize $r \geq 1$ and $C \geq 2$;
Partitioning of S with Q;
Compute v_c and μ_{ic};
if $J_r(S, Q, L) \leq \epsilon$ **then**
  take final output;
else
  update v_c and μ_{ic}
end if

Joint user clustering and power allocation: Another resource aware clustering approach can be performed by efficient power allocation. A joint user clustering and power allocation algorithm was proposed in [13], where the main target was to elevate the sum-rate by selecting two best users (defined as cell center (CC) and cell edge (CE) user at $\hat{h_1}$ and $\hat{h_2}$, respectively) in a cluster. This work followed a fixed transmit power allocation scheme and a minimum distance metric was derived to differentiate the CC and CE user. It was represented as

$$MDF = \left(\frac{1 - 2P_{CC}}{P_{CC}^2 SINR_c}\right)^{-\frac{1}{\gamma}}, \tag{7}$$

where P_{CC} indicates the power allocation coefficient for the CC user and $SINR_c$ is the transmit SINR. Therefore, the position of the CC and CE user will be determined when $\hat{h_1} \leq MDF$ and $\hat{h_2} > MDF$, respectively.

A different cluster formation and power control technique was presented in [14] for an uplink MIMO-NOMA scenario. A high-power cluster and a low-power cluster was designed to partition active mobile users by exploiting the composite channel gains rather than individual one.

Clustering based on spatial positions: Spatial position of users in a mMIMO-NOMA system can also be considered while performing user clustering. This technique shows lower complexity than any other schemes. In [15] a clustering approach was used for multi-user downlink MIMO-NOMA system where multiple clusters were formed according to the users spatial positions. With the help of global position tracking systems this information can be collected. This approach aims to group users having lower spatial separation and a multi-antenna cluster head is selected to serve them. Therefore, this system demonstrates interference awareless approach while clustering and can experience higher performance degradation in achievable sum rate if no additional interference mitigation technique is applied.

B. Learning Assisted User Clustering (LAUC)

K-means: K-means user clustering is a low-complex unsupervised ML algorithm often used in mMIMO-NOMA systems [5], [16]. A specified number of user cluster C of cluster head v_c is targeted to achieve. Firstly, initial cluster heads are

TABLE I
COMPLEXITY COMPARISON OF THE USER CLUSTERING TECHNIQUES IN MMIMO-NOMA SYSTEM.

Ref.	Clustering techniques	Computational Complexity
[11]	Channel correlation coefficient-based clustering	$O(U_{CEU}^3 + (U_{D2D}^2/2 + U_{D2D}/2 + C) + C^3)$
[12]	Clustering based on energy efficient user admission	$O(C^2 K_c)$
[14]	Joint user clustering and power allocation	$O(U^2 N + N^3 \phi)$
[16]	K-means	$O(U v_c I N)$
[16]	K-means++	$O(U v_c N)$
[18]	Fuzzy C-means	$O(U I d C^2)$
[19]	Clustering as a multi-level classification problem	$O(UBM \log_2(E) W_{tr})$

ϕ, I, d, and E denotes division size, no. of iterations, no. of dimensions, and no. of leaf nodes, respectively.

randomly selected for the predetermined clusters. Secondly, every user is assigned to its nearest cluster head. Then the head position is changed by the associated cluster members by taking average for all the users. The whole process is described in Algorithm 1.

K-means++: K-means algorithm is very sensitive to the selection of initial cluster heads. It may result in inefficient user grouping if there is a fluctuation with initial cluster heads or exists more than one cluster heads into the same cluster. Therefore, an advancement is proposed in the initialization of that head selection process in a mMIMO-NOMA system and named as k-means++ algorithm [16]. It can be summarized as follows: •

- Step 1: The first starting cluster head is selected randomly from U.
- Step 2: $\forall\, m \in (U-1)$ the distance from the nearest cluster head is calculated.
- Step 3: The next head is selected from a user in $\forall\, m \in U/C$ such that it has a maximum distance from the nearest cluster head.
- Step 4: Repeats the step 2 and 3 until C cluster heads are assigned.
- Step 5: Remaining procedures are the same as k-means algorithm described in Algorithm 1.

Fuzzy C-means: Fuzzy C-means (FCM) is an unsupervised clustering algorithm developed for feature analysis from where the algorithm classifies users into several clusters. The algorithm first initializes the cluster numbers and fuzzy exponent. After that, a membership function is assigned to each user to determine a fractional relation with a cluster. The cluster head is being computed repeatedly by updating the membership coefficient of each data points with a view to minimizing the objective function at a certain threshold [17]. The detail FCM clustering algorithm is provided in Algorithm 2.

The objective of the FCM algorithm is to identify the fuzzy matrix and a set of mean of the ith points in the cth cluster i.e. the cluster head. Thereby, we can define the objective function as

$$\min_{(S,\,Q,\,L)} \left(J_r(S,\,Q,\,L) = \sum_{i=1}^{K_c} \sum_{c=1}^{C} \mu_{ic}^{r} \| q_{if} - v_c \|^2 \right). \quad (8)$$

where L is a positive definite matrix, $r \geq 1$ denotes the fuzzy exponent, v_c represents the mean vector of the ith points in the cth cluster, and S is generally known as the fuzzy matrix of element μ_{ic} termed as the membership function of the cth cluster satisfying $0 \leq \mu_{ic} \leq 1$. Q is a data matrix of F number of quality-of-service (QoS) features for the K_c users satisfying $q_{if} \in \mathbb{N}$, $i \in K_c$, and $f \in F$. As one of our previous work in [18], FCM algorithm was proposed in mMIMO-NOMA system and the fuzzy partitioning coefficient was evaluated on a received signal strength indicator data sets.

Clustering as a multi-level classification problem: In [19], the clustering for MIMO-NOMA system was formulated as a multi-level classification problem. The authors adopted gradient boosted decision tree along with a classifier chain as base learners. Users are partitioned into two disjoint clusters C_1 and C_2 such that $C = C_1 + C_2$. Then users are shorted in such a way that spatial correlation of consecutive users is high and they are assigned to different clusters to maximize the sum-rate. These two users out of U active users can be found by checking the vector channel response alignment. The optimization problem was formulated as

$$\max_{(l_m, \ldots l_U)} \sum_{m \in U} \log_2(1 + \bar{l}_m SINR_m^{(1)} + l_m SINR_m^{(1)}), \quad (9)$$

where l_m is a binary indicator variable and $\bar{l}_m$ indicates the opposite of it. Specifically, this problem can be termed as multi-label binary classification task and the computational complexities for the training and test cases are defined as $O(UBM \log_2(E) W_{tr})$ and $O(UM \log_2(E) W_{tst})$, where B and W are the bundle size and training or test data size, respectively.

Finally, we compared the given clustering algorithms for mMIMO-NOMA system using its individual computational complexities shown in Table 1. The system performance depends on the clustering efficiency and that relies on the way of user classifications and requirements.

IV. CONCLUSION

In this paper, we considered a downlink mMIMO-NOMA-based mmWave/THz system where all the single-antenna users are grouped into multiple clusters to increase the spectral-efficiency. Therefore, a brief study has been conducted on the proposed clustering techniques in the mMIMO-NOMA systems. We summarized the techniques into two sections: joint RAUC and LAUC. The RAUC techniques takes considerations of channel correlation metrics, power and energy efficiency, and spatial positions based schemes. Different supervised and unsupervised ML approaches such as k-means algorithm, fuzzy c-means algorithms, and classification-based ML algorithms are categorized into the LAUC. The computational complexities of those schemes are also compared. The

RAUC techniques shows higher order polynomial function of complexities whereas that of the LAUC algorithms is comparatively low.

ACKNOWLEDGMENT

This work was partly supported by the Technology development Program of MSS [S3098815] and the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-0-01396) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

REFERENCES

[1] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, ”6G Wireless Communication Systems: Applications, Requirements, Technologies, Challenges, and Research Directions,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957-975, 2020.

[2] W. A. Al-Hussaibi and F. H. Ali, “Efficient User Clustering, Receive Antenna Selection, and Power Allocation Algorithms for Massive MIMO-NOMA Systems,” *IEEE Access*, vol. 7, pp. 31865–31882, 2019.

[3] X. Hu, C. Zhong, X. Chen, W. Xu, and Z. Zhang, “Cluster Grouping and Power Control for Angle-Domain mmWave MIMO NOMA Systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 1167–1180, 2019.

[4] S. Solaiman, L. Nassef, and E. Fadel, “User Clustering and Optimized Power Allocation for D2D Communications at mmWave Underlaying MIMO-NOMA Cellular Networks,” *IEEE Access*, vol. 9, pp. 57726–57742, 2021.

[5] H. Zhang, H. Zhang, W. Liu, K. Long, J. Dong, and V. C. M. Leung, ”Energy Efficient User Clustering, Hybrid Precoding and Power Optimization in Terahertz MIMO-NOMA Systems,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2074-2085, Sep. 2020.

[6] S. Chinnadurai et al., “User Clustering and Robust Beamforming Design in Multicell MIMO-NOMA System for 5G Communications,” *International Journal of Electronics and Communications*, vol. 78, pp. 181–191, 2017.

[7] M. Shahjalal, M. K. Hasan, M. M. Islam, M. Z. Chowdhury, and Y. M. Jang, “A Two-Stage Power Allocation-Based NOMA Architecture for Optical Camera Communication,” *IEEE Systems Journal*, pp. 1–10, 2020. doi: 10.1109/JSYST.2020.3015766.

[8] Q. Sun, S. Han, C. L. I, and Z. Pan, “On the Ergodic Capacity of MIMO NOMA Systems,” *IEEE Wireless Communications Letters*, vol. 4, no. 4, pp. 405–408, 2015.

[9] Z. Ding, R. Schober, and H. V. Poor, “A General MIMO Framework for NOMA Downlink and Uplink Transmission Based on Signal Alignment,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4438–4454, 2016.

[10] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, “On the Sum Rate of MIMO-NOMA and MIMO-OMA Systems,” *IEEE Wireless Communications Letters*, vol. PP, no. 99, p. 1, Jun. 2017.

[11] S. Solaiman, L. Nassef, and E. Fadel, ”User Clustering and Optimized Power Allocation for D2D Communications at mmWave Underlaying MIMO-NOMA Cellular Networks,” *IEEE Access*, vol. 9, pp. 57726-57742, 2021.

[12] M. Zeng, A. Yadav, O. A. Dobre, and H. V. Poor, ”Energy-Efficient Power Allocation for MIMO-NOMA With Multiple Users in a Cluster,” *IEEE Access*, vol. 6, pp. 5170-5181, 2018.

[13] S. Chinnadurai, P. Selvaprabhu, Y. Jeong, A. L. Sarker, H. Hai, W. Duan, and M. H. Lee, ”User clustering and robust beamforming design in multicell MIMO-NOMA system for 5G communications,” *International Journal of Electronics and Communications*, vol. 78, pp. 181-191, 2017.

[14] W. A. Al-Hussaibi, ”Optimal Cluster Formation and Power Control for High Connectivity Wireless MIMO-NOMA Applications,” *Electronics Letters*, vol. 55, no. 20, pp. 1110-1112, 2019.

[15] T. M. Hoang, B. C. Nguyen, X. N. Tran, and L. T. Dung, ”Outage Probability and Ergodic Capacity of User Clustering and Beamforming MIMO-NOMA Relay System With Imperfect CSI Over Nakagami-m Fading Channels,” *IEEE Systems Journal*, vol. 15, no. 2, pp. 2398-2409, 2021.

[16] Q. N. Le, V.-D. Nguyen, N.-P. Nguyen, S. Chatzinotas, O. A. Dobre, and R. Zhao, ”Learning-Assisted User Clustering in Cell-Free Massive MIMO-NOMA Networks,” *arXiv:2011.07549*.

[17] J. C. Bezdek, R. Ehrlich, and W. Full, ”FCM: The Fuzzy C-Means Clustering Algorithm,” *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191-203, 1984.

[18] M. Shahjalal, M. F. Ahmed, M. M. Alam, M. H. Rahman, and Y. M. Jang, ”Fuzzy C-Means Clustering-Based mMIMO-NOMA Downlink Communication for 6G Ultra-Massive Interconnectivity,” in *Proc. International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Jeju, Korea, 2021, pp. 421-424.

[19] C. Ben Issaid, C. Antón-Haro, X. Mestre, and M. -S. Alouini, ”User Clustering for MIMO NOMA via Classifier Chains and Gradient-Boosting Decision Trees,” *IEEE Access*, vol. 8, pp. 211411-211421, 2020.