User Clustering based on Correlation in 5G using Semidefinite Programming

Chit Wutyee Zaw*, Yan Kyaw Tun†, Choong Seon Hong‡
Department of Computer Science and Engineering, Kyung Hee University,
Yongin, 446-701 Korea
Email: {*cwyzaw, †ykyawtun7, ‡cshong}@khu.ac.kr

***Abstract*—For 5G cellular networks, non-orthogonal multiple access (NOMA) has been recognized as a promising technique. NOMA allows multiple users to access the same channel by adapting successive interference cancellation (SIC) and multiplexing in power domain. This causes the clustering problem to become critical as higher channel gain users need to cancel out the signals of lower channel gain users while the latter has to receive less interference. In this paper, we formulate the user clustering as a correlation clustering problem. We transform the problem by relaxing the clustering variables, and solve using semidefinite programming (SDP). Moreover, a simple iterative power allocation algorithm is nominated. In the simulation results, it can be seen that the correlation clustering using semidefinite programming outperforms the random clustering of users.**

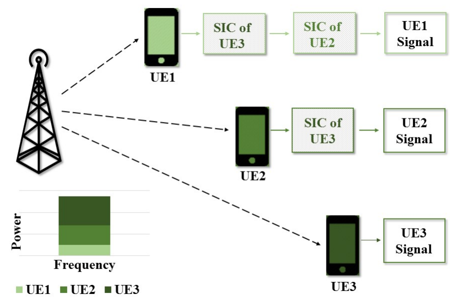

Fig. 1. NOMA Basics

I. Introduction

As an increasing number of Internet of Things (IoT) devices is connected to networks, it is crucial to support the massive connectivity. Moreover, one of the goals of 5G is to provide the high data rate by increasing the spectral efficiency. Non-orthogonal multiple access (NOMA) has been a candidate solution to these goals by granting the multiple users to access the same spectrum resource at the same time. According to [1], the main features of NOMA are improvement in spectrum efficiency, massive connectivity, low transmission latency and signaling cost. Regardless of its advantages, there are numerous factors to be considered from the practical point of view. These factors are examined and taken into accounts to provide the system-level performance in [2].

In NOMA, base station sends superimposed messages to all users in the same group by using Multi-User Superposition Technique (MUST). With the help of successive interference cancellation (SIC), users can cancel out the signals of other users whose channel gains are lower than theirs while taking the signals of higher channel gain users as interference which is shown in Fig. 1. Power is allocated inversely proportional to the channel gains so as to successfully decode the signals.

There is a trade-off in clustering of users in NOMA. When the number of clusters is large which means fewer users in each cluster, users in a cluster have to access less bandwidth but the allocated power could be large. This is inverse for small number of clusters scenario. Because of this trade-off, it is not possible to know the fixed number of cluster in advance. Since channel gains of the users varied largely, every clusters could not have the same number of users.

A. Related Works

Authors in [3] proposed user clustering and optimal power allocation based the result clustering. Depending on the number of clusters, their clustering algorithm produces same number of users for every clusters. [4] proposed four user pairing algorithms and channel state sorting-pairing algorithm (CSSPA) is one of them. CSSPA divides the users into two groups based on their channel gains and pairs two users from different groups. Authors in [5] solved the joint channel and power allocation by using a two-stage dynamic programming in which power is discretized into multiple levels.

Many papers focus only on optimal power allocation such as [6], [7]. The former studied the optimality of power allocation by guaranteeing the Quality of Service (QoS) of users. Majority of NOMA papers target on a single-cell system. However, [7] proposed the distributed power allocation in two interfering cells scenario.

Authors in [8] solved the correlation clustering problem by semidefinite programming. Later, the optimal results are rounded using two rounding approach. One approach is a modified Goemans-Williamson rounding for MAX-CUT problem [9] by choosing two hyperplanes. [10] also formulated the femto-cell clustering as correlation clustering and solved by semidefinite programming. But, the authors used the randomized rounding to approximate the optimal values.

B. Our Contributions

In this paper, we first formulate a joint user clustering and power allocation for NOMA in 5G systems. Later, the whole problem is divided into two independent sub-problems, user clustering and power allocation. User clustering problem is NP-hard, so we transformed it into semidefinite programming by relaxing the clustering variables. Then, the optimal values are approximated by adapting Goemans-Williamson rounding in which different number of hyperplanes are selected. Moreover, we proposed a simple iterative power allocation algorithm which guarantees the individual QoS requirement of users.

II. System Model and Problem Formulation

A single macro-cell downlink system with N users is considered in this paper. These N users are grouped into clusters so as to maximize the sum rate of the system. The system uses a hybrid OFDMA-NOMA approach which means that the clusters are assigned the orthogonal resources so that there is no interference between them. But, users in the same cluster access the same resources. Thus, each cluster applies the NOMA technique and OFDMA is applied among different clusters.

The total available bandwidth is divided and assigned orthogonally to channels which are considered as clusters in this paper. The assigned bandwidth to cluster j is ω_j. In this paper, we assume that the total available power of the macro base station is divided equally among the clusters so that the maximum available power on each channel is P_{total}. Each user has its own minimum data rate requirement which is γ_i. The joint problem of user clustering and power allocation is formulated in (1).

$$
\begin{aligned}
\max_{p_i, x_{i,j}} \quad & \sum_{j=1}^{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}_j|} x_{i,j}\omega_j \log_2\left(1+\frac{p_i h_i}{\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0}\right) \\
\text{subject to} \quad & \sum_{i=1}^{|\mathcal{C}|} x_{i,j} p_i \le P_{total},\ \forall j \in \mathcal{C} \\
& \sum_{j=1}^{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}_j|} \omega_j \log_2\left(1+\frac{p_i h_i}{\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0}\right) \ge \gamma_i, \\
& \qquad \forall i \in \mathcal{N} \\
& \sum_{j=1}^{|\mathcal{C}_j|} x_{i,j} \le 1,\ \forall i \in \mathcal{N} \\
& p_i \ge 0,\ x_{i,j} \in \{0,1\}
\end{aligned}
\tag{1}
$$

The objective is to maximize the sum rate of all users in the system. The first constraint is to make sure that power allocated to users in a cluster is not exceeded the power budget on a channel which is assumed as a cluster in this paper. The second constraint guarantees the QoS requirement of each user. The last constraint is that a user must only belong to one cluster. p_i is the power allocated to user i and $x_{i,j}$ is clustering variable which means user i is grouped in cluster j. The problem is broken down into two independent sub-problems, user clustering and power allocation, to solve efficiently.

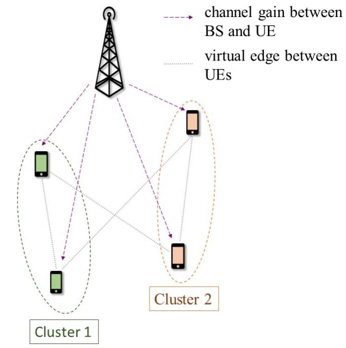

Fig. 2. User Clustering using Correlation Clustering

III. User Clustering using SDP based Correlation Clustering

The sub-problem, user clustering, is presented in this section. In NOMA, users with different channel gains need to be grouped in a cluster to reduce the intra-cluster interference. For instance, if two users with similar channel gains are grouped together in the same cluster, their power allocation will also be similar to achieve the maximum data rate. This leads to the high inter-interference and difficulty for higher channel gain user to decode the signal of lower channel gain user.

As there is trade-off between number of users in the clusters and the number of clusters, it is not possible to define number of clusters in advance. In addition, channel gains of users vary a lot, every cluster would not have the same number of users. For these reasons, we use correlation clustering for user grouping. In correlation clustering, a fully connected graph is considered where there are two weight w^+i,j and $w^-_{i,j}$ between node i and node j. $w^+_{i,j}$ represents the weight that i and j should be in the same cluster and $w^-_{i,j}$ represents the weight that i and j should not be in the same cluster. So, in this paper, we define $w^+_{i,j}$ as the difference of channel gains between user i and user j as shown in Fig. 2.

The correlation clustering problem is formulated as follows.

$$
\begin{aligned}
\max_{x_{i,j}} \quad & \sum_{i}^{N}\sum_{j}^{N} w^+_{i,j}x_{i,j} + w^-_{i,j}(1-x_{i,j}) \\
\text{subject to} \quad & x_{i,i} = 1,\ \forall i \in \mathcal{N} \\
& x_{i,j} = x_{j,i}\ \forall i,j \in \mathcal{N} \\
& x_{i,j} \in \{0,1\}
\end{aligned}
\tag{2}
$$

where $w_{i,j} = |h_i - h_j|$, $w^+_{i,j} = \frac{w_{i,j}-w_{min}}{w_{max}-w_{min}}$, $w^-_{i,j} = 1 - w^+_{i,j}$. $w^+_{i,j}$ is normalized between 0 and 1 where w_{min} is $\min_{i,j\in\mathcal{N}}\{w_{i,j}\}$ and w_{max} is $\max_{i,j\in\mathcal{N}}\{w_{i,j}\}$.

The problem is NP-hard with boolean variables $x_{i,j}$. In order to solve this problem efficiently, the boolean variables are relaxed to the values which are between 0 and 1. The objective function is linear and X is positive semidefinite matrix where $z^T X z \ge 0$ for any nonzero vector z with dimension N [11]. Thus, the relaxed problem is SDP problem which can be solved efficiently by using cvxpy [12].

The optimal value X^* is between 0 and 1 which needs to be rounded back to 0 or 1. We modified Goemans-Williamson rounding for MAX-CUT [9] by choosing multiple hyperplanes. In [9], a random hyperplane r through the origin, in which mean is 0 and variance is 1, is selected to partition nodes into two groups where node i is in S if $x_i.r \geq 0$, otherwise, it is in $\bar{S}$. We modified this approach by choosing k hyperplanes, $r_1, r_2, \cdots, r_k$, which produces at most 2^k clusters.

$$\begin{aligned} \mathcal{C}_1 &= \{i \mid x_i.r_1 \geq 0, x_i.r_2 \geq 0, \cdots, x_i.r_k \geq 0\} \\ \mathcal{C}_2 &= \{i \mid x_i.r_1 \geq 0, x_i.r_2 \geq 0, \cdots, x_i.r_k < 0\} \\ &\vdots \\ \mathcal{C}_{2^k} &= \{i \mid x_i.r_1 < 0, x_i.r_2 < 0, \cdots, x_i.r_k < 0\} \end{aligned} \tag{3}$$

Algorithm 1 User Clustering Algorithm

Input: Channel gains of all users $h_1, h_2, \cdots, h_N$ and number of hyperplanes, k

Output: Clusters of users, $\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_{2^k}$

Solve the relaxed problem by using cvxpy
Calculate random vectors $r_1, r_2, \cdots, r_k$ which follow the normal distribution
for each user $i \in \mathcal{N}$ **do**
    Select a cluster where user i belongs by equation (3)
end for

IV. An Iterative Power Allocation

In this section, an algorithm for an iterative power allocation sub-problem is proposed which can ensure the QoS requirements of users. The power allocation sub-problem for each cluster $\mathcal{C}_j$ is as follows.

$$\begin{aligned} \max_{p_i} \quad & \sum_{i=1}^{|\mathcal{C}_j|} \omega_j \log_2 \left(1 + \frac{p_i h_i}{\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0}\right) \\ \text{subject to} \quad & \sum_{i=1}^{|\mathcal{C}_j|} p_i \leq P_{total} \\ & \omega_j \log_2 \left(1 + \frac{p_i h_i}{\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0}\right) \geq \gamma_i, \\ & \qquad \forall i \in \mathcal{C}_j \\ & p_i \geq 0 \end{aligned} \tag{4}$$

The user i has the interference from the users whose channel gains are higher than its. Thus, for the highest channel gain user in the group would not have interference by using SIC. To maximize the sum rate of all users, power is allocated to the user by making the received data rate equals to the QoS requirement of that user.

$$\omega_j \log_2 \left(1 + \frac{p_i h_i}{\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0}\right) = \gamma_i \tag{5}$$

TABLE I
Simulation Parameters

Number of users	200
Transmit power of Base Station	46 dbm
The total available bandwidth	20 MHz
Power density thermal noise	-174 dBm/Hz
Data rate requirement of users	[1.0, 1.5] Mbps

$$p_i = \frac{2^{\gamma_i/\omega_j} - 1}{h_i} \left(\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0 \right) \tag{6}$$

Then, power is allocated to each user in the cluster according to the ascending order of channel gains while checking the total power budget. The total available power is discretized into different power levels where each step is $\triangle p$.

Algorithm 2 Iterative Power Allocation Algorithm

Input: $\mathcal{C}_j$, ω_j, p_{total}, $h_i, \cdots, h_N$

Output: power allocated to each users in the cluster $\mathcal{C}_j$

$p_i^0 \leftarrow 0$, $\forall i \in \mathcal{C}_j$
$i \leftarrow 0$
$t \leftarrow 0$
repeat
    if $i == |\mathcal{C}_j|$ **then**
        $i \leftarrow 1$
        $t \leftarrow t + 1$
        $p_0^t \leftarrow p_0^{t-1} + \triangle t$
        $p_i^t \leftarrow 0, \forall i \in \{1, 2, \cdots, |\mathcal{C}_j| - 1\}$
    end if
    $p_i^t \leftarrow \frac{2^{\gamma_i/\omega_j} - 1}{h_i} \left(\sum_{s=1}^{i-1} p_s h_i + \omega_j n_0 \right)$
    $i \leftarrow i + 1$
until $\sum_{i \in \mathcal{C}_j} p_i^t < p_{total}$

V. Simulation Results

Simulation parameters are presented in Table. I. The long distance path loss model with *Rayleigh* distribution is used. The total available bandwidth and transmit power are divided equally among clusters. We compare our clustering with channel gain based user pairing approach proposed by [4].

In fig. 3, when the number of users is 50, the channel gain based pairing approach surpasses the proposed clustering with 5 and 6 hyperplanes. When the number of users becomes 200, the proposed clustering with all k values outperforms the channel gain based pairing. The smallest k value which results in 8 clusters at maximum gives the highest received data rate but puts a lot of complexity on SIC. If SIC complexity is considered, the large number of hyperplanes can be chosen.

The number of clusters is compared among different values of k in fig. 4. From this result, we can see that the proposed clustering is trying to pack the more users in the clusters so as to achieve the maximum data rate. As the number of users rises, the number of clusters also rises. This happens when it

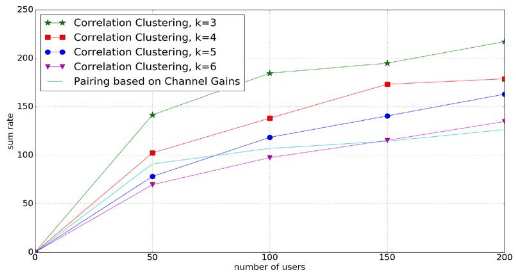

Fig. 3. Comparison between different k and Channel Gains based Pairing

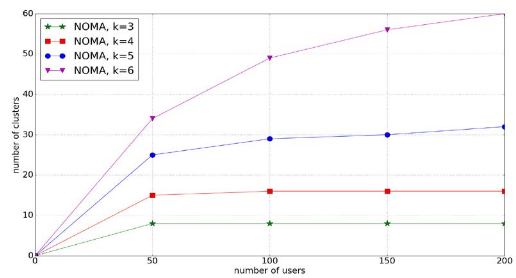

Fig. 4. Number of Clusters Comparison between different k values

tries to minimize the intra-cluster interference by increasing the number of clusters until it reaches to its maximum.

Different from fig. 3 and 4 in which clustering is performed four times for 50, 100, 150 and 200 users, fig. 5 performs clustering for 200 users at once and the individual rate is compared. NOMA with proposed clustering produces much higher received rate than OFDMA. It proves NOMA can give the higher spectrum efficiency by clustering users and allowing a cluster of users to access the same resource. An interesting fact is that from an individual user's point of view, it is difficult to know which values of k gives the highest rate up to the point where the number of users is 150. After that, the proposed clustering where k is 3 achieves the highest data rate.

VI. Conclusion

In this paper, the joint problem of user clustering and power allocation is formulated which is further divided into two independent sub-problems, user clustering and power allocation. User clustering problem is formulated as correlation clustering and solved by using semidefinite programming where the boolean variables are relaxed. Then, the optimal values are rounded by adapting Goemans-Williansom rounding with different numbers of hyperplanes. In addition, an iterative power allocation is nominated which guarantees the minimum data rate requirement of users. In simulation results, we analyzed the different numbers of clusters. We compared our proposed clustering with user paring based on channel gains approach and our clustering outperforms the latter as the number of users increases.

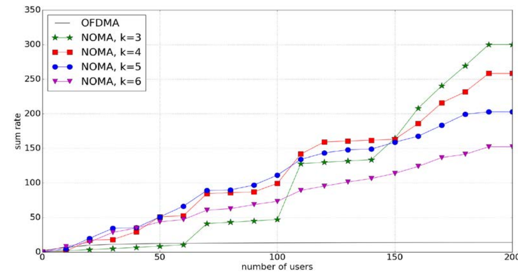

Fig. 5. Comparison with OFDMA

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (NRF-2017R1A2A2A05000995). *Dr. CS Hong is the corresponding author.

References

[1] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5g: solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.

[2] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (noma) for future radio access," in *Intelligent Signal Processing and Communications Systems (ISPACS), 2013 International Symposium on*. IEEE, 2013, pp. 770–774.

[3] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (noma) systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.

[4] H. Zhang, D.-K. Zhang, W.-X. Meng, and C. Li, "User pairing algorithm with sic in non-orthogonal multiple access system," in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[5] L. Lei, D. Yuan, C. K. Ho, and S. Sun, "Power and channel allocation for non-orthogonal multiple access in 5g systems: Tractability and computation," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8580–8594, 2016.

[6] Z. Yang, W. Xu, C. Pan, Y. Pan, and M. Chen, "On the optimality of power allocation for noma downlinks with individual qos constraints," *IEEE Communications Letters*, 2017.

[7] Y. Fu, Y. Chen, and C. W. Sung, "Distributed downlink power control for the non-orthogonal multiple access system with two interfering cells," in *Communications (ICC), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[8] C. Swamy, "Correlation clustering: maximizing agreements via semidefinite programming," in *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2004, pp. 526–527.

[9] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the ACM (JACM)*, vol. 42, no. 6, pp. 1115–1145, 1995.

[10] A. Abdelnasser, E. Hossain, and D. I. Kim, "Clustering and resource allocation for dense femtocells in a two-tier cellular ofdma network," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1628–1641, 2014.

[11] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM review*, vol. 38, no. 1, pp. 49–95, 1996.

[12] S. Diamond, E. Chu, and S. Boyd, "Cvxpy: A python-embedded modeling language for convex optimization, version 0.2," 2014.