

Article

RRH Clustering Using Affinity Propagation Algorithm with Adaptive Thresholding and Greedy Merging in Cloud Radio Access Network

Seju Park ¹, Han-Shin Jo ^{2,*}, Cheol Mun ³ and Jong-Gwan Yook ^{1,*}

¹ Department of Electrical and Electronic Engineering, Yonsei University, Seoul 03722, Korea; dream158@yonsei.ac.kr

² Department of Electronics and Control Engineering, Hanbat National University, Daejeon 34158, Korea

³ Department of Electronic Engineering, Korea National University of Transportation, Chungju 27469, Korea; chmun@ut.ac.kr

* Correspondence: hsjo@hanbat.ac.kr (H.-S.J.); jgyook@yonsei.ac.kr (J.-G.Y.);
Tel.: +82-42-821-1132 (H.-S.J.); +82-2-2123-4618 (J.-G.Y.)

Abstract: Affinity propagation (AP) clustering with low complexity and high performance is suitable for radio remote head (RRH) clustering for real-time joint transmission in the cloud radio access network. The existing AP algorithms for joint transmission have the limitation of high computational complexities owing to re-sweeping preferences (diagonal components of the similarity matrix) to determine the optimal number of clusters as system parameters such as network topology. To overcome this limitation, we propose a new approach in which preferences are fixed, where the threshold changes in response to the variations in system parameters. In AP clustering, each diagonal value of a final converged matrix is mapped to the position (x,y coordinates) of a corresponding RRH to form two-dimensional image. Furthermore, an environment-adaptive threshold value is determined by adopting Otsu's method, which uses the gray-scale histogram of the image to make a statistical decision. Additionally, a simple greedy merging algorithm is proposed to resolve the problem of inter-cluster interference owing to the adjacent RRHs selected as exemplars (cluster centers). For a realistic performance assessment, both grid and uniform network topologies are considered, including exterior interference and various transmitting power levels of an RRH. It is demonstrated that with similar normalized execution times, the proposed algorithm provides better spectral and energy efficiencies than those of the existing algorithms.

Keywords: machine learning; clustering; affinity propagation; C-RAN; exterior interference



Citation: Park, S.; Jo, H.-S.; Mun, C.; Yook, J.-G. RRH Clustering Using Affinity Propagation Algorithm with Adaptive Thresholding and Greedy Merging in Cloud Radio Access Network. *Sensors* **2021**, *21*, 480. <https://doi.org/10.3390/s21020480>

Received: 2 December 2020

Accepted: 9 January 2021

Published: 12 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The 5th generation mobile communication system (5G system) has three types of requirements: Enhanced mobile broadband, massive machine-type communications, and ultra-reliable and low-latency communications (URLLC) [1]. Among these requirements, reliability is an important factor that directly affects the quality of the 5G system service. Therefore, many studies have focused on this aspect of 5G systems [2,3]. The cloud radio access network (C-RAN) has a centralized processor baseband unit (BBU) and a distributed radio remote head (RRH) to provide a better network structure for URLLC [4]. In the C-RAN of the 5G system, the joint transmission technique between RRHs is an essential technique that can increase reliability by eliminating the overhead caused by the inter-cell interference and hand-off [5–7]. There are two types of joint transmission: Non-coherent joint transmission and coherent joint transmission [8–10]. As a representative wireless communication technology that uses coherent joint transmission, the cell-free massive multiple-input and multiple-output (MIMO) is capable of simultaneously transmitting high data rates to multiple devices without any interference [11–13]. In this study, we assume coherent joint transmission. Although joint transmission exhibits strong performance, the

number of RRHs that can be combined on a network is limited. Therefore, it is necessary to appropriately configure a cluster between RRHs based on the channel environment using an appropriate clustering algorithm. Even in traditional wireless sensor networks, clustering algorithms are important for configuring user-oriented multiple sensors. Therefore, many clustering algorithms have already been proposed [14–16].

Traditionally, clustering is a method of discovering the regularity present in a large quantity of unstructured information and appropriately grouping this information based on user-defined rules. Depending on how the information is grouped, the interpretation and results can be completely different. Therefore, the user intending to use clustering must adopt an appropriate clustering method to group the desired information. This paper aims to address the clustering problem, which is a non-deterministic polynomial hard problem, using a clustering algorithm with low time complexity and excellent performance. Clustering algorithms are often used in medicine, chemistry, and biotechnology to categorize information [17–19]. Among the clustering algorithms, the well-known affinity propagation (AP) clustering algorithm is a representative clustering method based on graph theory [20–23]. Graph theory-based clustering is characterized by information exchange among data points. This method has a lower time complexity than that of other clustering methods [20]. The AP clustering algorithm is extremely efficient in terms of energy consumption [24]. Furthermore, low time complexity is a key requirement that enables dynamic clustering to be implemented in a wireless communication system where the channel state between a RRH and user equipment (UE) changes in real time.

The AP clustering algorithm is a machine learning-based clustering technique that obtains a reasonable cluster combination and number of clusters through iteration. That is, the number of clusters does not need to be determined before the algorithm is executed. First, a similarity matrix is designed based on the characteristics of the data points, and each component of this matrix is updated and converged through iteration. Finally, the converged matrix is used for clustering [20]. As the similarity matrix is the only input value, the design of the matrix and setting the initial values almost dominate the clustering performance. The similarity matrix is divided into diagonal and off-diagonal components. Diagonal components are called preferences and are terms that pertain mainly to the number of clusters K . Off-diagonal components are values that pertain mainly to the combinations of data points that constitute the cluster. Although it is not an input value of this algorithm, there is one more parameter to effectively control the number of clusters, which is a threshold set to the default value of zero. Among the diagonal components of the final matrix after the iteration is completed, the components that exceed this threshold are called exemplars. The number of exemplars is equal to the number of clusters.

Recently, studies on AP clustering algorithms have focused on improving the joint transmission performance among RRHs in wireless communication networks. In these studies, the similarity matrix is designed on the basis of the expected or instant value of the channel gain between transmitting and receiving nodes, because the channel gain is a representative index that measures the similarity between wireless communication nodes (data points) [25–27]. Wesemann et al. [25] verified that the application of a similarity metric, defined as the sum of mutual inter-cell interference, can outperform static clustering in uplink joint transmission. Zhang et al. [26] proposed BS clustering using a similarity matrix consisting of the signal-to-interference and noise ratio (SINR) gain of joint transmission over non-joint transmission in a downlink pico-cellular network. Both the offline and online phases of AP clustering are also provided for practical implementation. This approach to clustering achieved better performance than that of the existing clustering algorithms. In a typical small-cell C-RAN, the RRHs are not sufficiently separated out of the reach of radio waves, causing exterior interference [28], which refers to interference emitted from RRHs connected to the adjacent BBUs using the same frequency and resource elements. With this background, Park et al. [27] proposed a new preference for controlling the number of clusters considering not only inter-cluster and intra-cluster interferences but also exterior interference.

Contribution and Organization

As all AP clustering methods proposed in previous studies [25–27] use a threshold fixed to zero, the change in the number of clusters is significant in response to a change in the preference value. Therefore, the clustering performance also varies significantly. Studies have suggested that the number of clusters increases as the value of the preference increases [21]. However, there is no formula that specifically defines the relationship between the two values. Therefore, the preference value must be swept to achieve optimal clustering performance. However, this approach inevitably entails an extremely high computational complexity, as shown in [29].

To overcome this problem, we propose a new approach in which the preference value is fixed and the conventional fixed zero threshold is replaced with a value that changes adaptively in response to various system parameters. To determine the adaptive threshold, the diagonal components of the converged responsibility matrix \mathbf{R} + availability matrix \mathbf{A} ($\mathbf{R} + \mathbf{A}$ matrix) that has finished iteration in the similarity matrix (see Section 3 Algorithm 1 for detail) are imaged in three dimensions as the RRH position. The Otsu method, which determines the threshold that divides the background and foreground based on statistical decision theory using a gray-scale histogram, is applied to the imaged information. The proposed AP clustering technique using the Otsu threshold works effectively in situations where many exemplars are determined based on a fixed preference and adjusts the number of exemplars reasonably [30]. Since an exemplar represents the center of a cluster, the adjacent exemplars indicate that the interference among clusters (inter-cluster interference) can be enlarged. The existing AP clustering method does not have the ability to correct the exemplars adjacent phenomenon. The greedy merging algorithm can increase the performance of the AP clustering system by combining the exemplars into one when they are adjacent to one another and solve this problem. We verified that the proposed AP clustering algorithms have better performances than those of the existing AP clustering algorithms using the Otsu threshold and greedy merging methods, which are simpler than the existing preference sweeping AP clustering algorithms. In Section 3, the process flow of the AP clustering algorithm is discussed in detail.

In this study, the conventional and proposed AP clustering algorithms are compared using a system-level simulation in a small-cell C-RAN with various realistic models and parameters. The simulation considers not only inter-cluster and intra-cluster interferences but also exterior interference, and a channel model using stochastic path loss and shadowing for mmWave of IEEE TR 38.900 and flat fading [31], as well as various levels of RRH transmitting power in exterior areas. Moreover, the performance evaluation is performed for both the grid and uniform random network topology, as depicted in Figure 1.

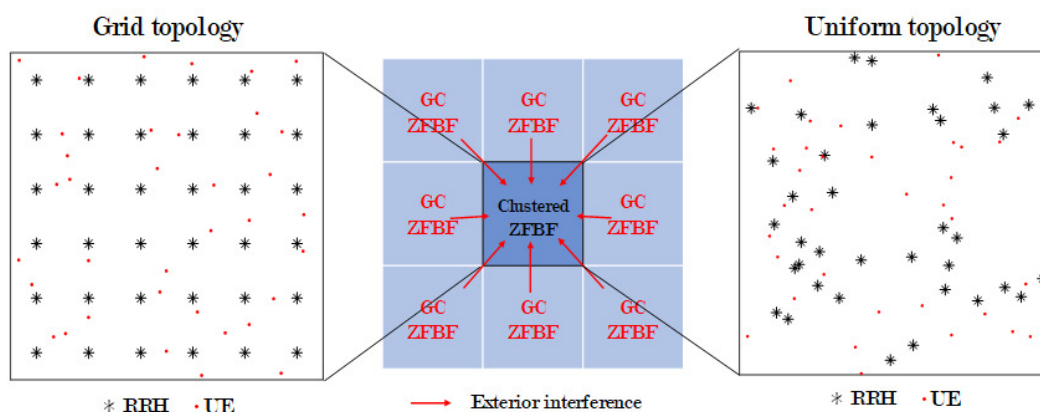


Figure 1. Concept of exterior interference from eight baseband units (BBU) environment with globally coordinated zero-forcing beamforming (ZF-BF) in two topologies.

This paper is organized as follows. Section 2 describes the system model and simulation settings. The conventional and proposed AP algorithms are discussed in Section 3. Performance evaluation and simulation results are presented in Section 4, and concluding remarks summarize the main results in Section 5.

2. System Model

In this section, the spectral efficiency (SE) for performance verification is derived, and the environment for the system-level simulation is defined. Figure 1 depicts the total area of the system-level simulation. This area is divided into nine sections in two network topologies. Each of these nine sections has 36 RRHs and 36 UE connected to one BBU. The scheduling algorithm is generally not assumed. If the scheduling algorithm is assumed, it is possible to simulate a situation in which the number of UE is significantly larger than that of RRHs. The clustered ZFBF area in Figure 1 is depicted in Figure 2b as the desired clustering area. This is the area where the performance of the existing clustering algorithm is compared with that of the algorithm proposed in this paper. This area is evaluated using two network topologies: Grid and uniform. In the first topology, the RRHs are installed in a grid pattern in the form of a square cell, with the distance between adjacent RRHs being equal. In this topology, the UE are randomly located within the square cell boundary of the central RRH. In the second topology, the RRHs and UE are randomly located within the area. In this topology, the RRH with the highest reception power at the UE is assumed to be the RRH serving the UE. The remaining eight areas in Figure 1 are regions for exterior interference. These areas use the same frequency and resource blocks as does the middle area and assume global coordination ZFBF i.e., $K = 1$. The following describes the evaluation of the clustering algorithm in the middle area. Figure 2a depicts the desired signal in a conventional cellular system with inter-cell interference, and Figure 2b depicts the desired cluster region C receiving inter-cluster interference from C' and exterior interference from C'' . The received signal at user i served by the desired cluster C with ZFBF can be described as:

$$y_i = \underbrace{\mathbf{h}_i^C \mathbf{w}_i^C s_i}_{\text{Desired signal}} + \underbrace{\sum_{j=1, j \neq i}^N \mathbf{h}_i^C \mathbf{w}_j^C s_j}_{\text{Intra-cluster interference}} + \underbrace{\sum_{k=1, k \neq i}^M \sum_{l \in C'} \mathbf{h}_{i_k}^{C'} \mathbf{w}_{l_k}^{C'} s_{l_k}}_{\text{Inter-cluster interference}} + \underbrace{\sum_{o=1}^E \sum_{q \in C''} \mathbf{h}_{i_o}^{C''} \mathbf{w}_{q_o}^{C''} s_{q_o}}_{\text{Exterior interference}} + n_i, \quad (1)$$

where s_i is the desired signal from the desired i th RRH to the desired i th UE, and \mathbf{h}_i^C is the channel gain received in the desired cluster C . N is the number of RRHs in desired cluster C . s_j is the signal from the N RRHs, except for the i th RRH in cluster C to the i th UE. Assuming perfect channel state information, the intra-cluster interference converges to zero owing to the orthogonal characteristics of the ZF precoding. As depicted in Figures 1 and 2b, $\mathbf{h}_{i_k}^{C'}$ is the channel gain received from the M RRHs of other cluster C' in the desired area, and s_{l_k} is the signal from the M RRHs of the other cluster C' to the i th UE. $\mathbf{h}_{i_o}^{C''}$ is the channel gain received from the exterior clusters C'' in the eight exterior areas ($E = 8$). n_i denotes the additive white Gaussian noise with zero mean and variance σ_i^2 of the i th UE. The channel coefficient \mathbf{h} used in (1) assumes a flat fading channel. For the path-loss and shadowing model, the urban micro-street canyon model of IEEE TR 38.900 is used [31]. s_i , s_j , and s_{l_k} occupy the same resource block. In (1), interference is categorized into three types: Intra-cluster, inter-cluster, and exterior interferences. The intra-cluster interference is generated inside the precoding matrix of the desired cluster C , inter-cluster interference is generated by the weight matrices from the other clusters C' in the desired area, and exterior interference is applied from clusters C'' of the eight exterior areas. \mathbf{w} is the normalized ZF weight vector for each channel coefficient vector, \mathbf{h} , and is given by:

$$\widetilde{\mathbf{W}} = [\widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2, \dots, \widetilde{\mathbf{w}}_Q] = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1} \quad (2)$$

$$\mathbf{w}_i = \widetilde{\mathbf{w}}_i / \|\widetilde{\mathbf{w}}_i\|,$$

where H is the Hermitian of the matrix, and Q is the number of UE constituting the desired cluster. The SINR of the i th user in the downlink are given by:

$$\text{SINR}_i = \frac{P_i |\mathbf{h}_i^C \mathbf{w}_i^C|^2}{\underbrace{\sum_{k=1, k \neq i}^M \sum_{l \in C'} P_l |\mathbf{h}_{i_k}^{C'} \mathbf{w}_{l_k}^{C'}|^2}_{\text{Inter interference}} + \underbrace{\sum_{o=1}^E \sum_{q \in C''} P_e |\mathbf{h}_{i_o}^{C''} \mathbf{w}_{q_o}^{C''}|^2}_{\text{Exterior interference}} + \sigma_i^2}, \quad (3)$$

where the intracuster interference is eliminated by ZF precoding. σ_i^2 denotes the additive white Gaussian noise with zero mean and variance σ_i^2 . P_e is the transmitting power of the RRH in the eight exterior areas ($E = 8$), and P_i is the transmitting power of the RRHs in the desired clustering areas C and C' . This study evaluates the SE performance by fixing P_i and changing only the P_e of the RRHs in the exterior area C'' . The SE of the i th UE is obtained from the SINR as follows:

$$\text{SE}_i = \log_2(1 + \text{SINR}_i). \quad (4)$$

In this paper, we compared and evaluated the joint transmission performance of the RRH group by applying the clustering algorithm using (4). In addition, the energy efficiency performance was evaluated as shown in (5):

$$\text{EE}_i = \frac{\text{SE}_i}{P_i}. \quad (5)$$

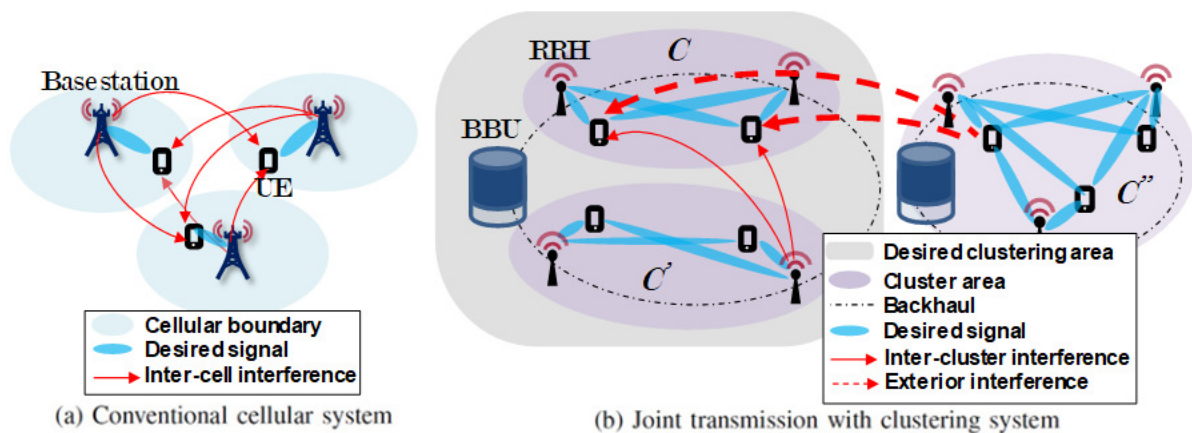


Figure 2. Concept of conventional cellular and joint transmission a with clustering system. (a): conventional cellular system; (b): joint transmission with clustering system.

3. RRH Clustering with AP Algorithm Schemes

This section describes an AP clustering algorithm that determines a joint transmission group among RRHs in a small-cell C-RAN environment. The AP clustering algorithm is a graph theory-based clustering method that uses information exchange among data points [20]. Parametric-based clustering has a disadvantage in that the clustering performance variance is large because the cluster configuration can vary according to the initial value. However, the AP clustering method always creates the same cluster when the initial value is set, but it has the disadvantage of the time complexity being slightly higher than the parametric-based method because of the iteration process. The AP clustering method is particularly suitable for joint transmission systems in wireless communication networks. The AP algorithm has an input value based on initial information among data points called the similarity matrix \mathbf{S} . This matrix contains information relating data point pairs, so it is a square matrix with the size of the number of data points, and is divided into off-diagonal

components and diagonal components called preferences. These two components play different roles in creating a cluster. Off-diagonal components are values related to a set of data points in a cluster. Diagonal components are values related to K . Therefore, both of these two components have a very important effect on the cluster performance. However, the focus of this study is on improving the clustered joint transmission performance by appropriately adjusting K :

$$s(i, p) = \begin{cases} \log\left(\frac{SINR_{CoMP}(i, p)}{SINR_{non}^i}\right), & i \neq p \\ \varepsilon \cdot \left[\log\left(\frac{1}{SINR_{non}^i}\right)\right], & i = p \end{cases} \quad (6)$$

As shown in Algorithm 1, the similarity matrix \mathbf{S} is the input value, and the responsibility matrix \mathbf{R} and the availability matrix \mathbf{A} , which have the same matrix size as \mathbf{S} , are set to 0 matrix. Lines 3 to 7 calculate and update the \mathbf{R} and \mathbf{A} matrices. ψ and ζ in line 8 are oscillating factors, representing the weights for the past and present values of the \mathbf{R} and \mathbf{A} matrix that are iterated by lines 9 to 13, and both values are defined as 0.5. Lines 14 to 20 determine the exemplar. Among the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix after iteration, components larger than threshold 0 are called the exemplar. The cluster set included in the exemplars is determined by line 19. Figure 3 is a graph showing the convergence process of the diagonal components of matrices \mathbf{R} and \mathbf{A} with the iteration progresses for data points 2, 5, 16, and 26 as an example. In the case of Algorithm 1, it can be observed that convergence is usually achieved with 10 iterations. Figure 4 shows the convergence process of the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix, and the two data points 16 and 26 selected as the exemplar larger than threshold 0. (5) is the similarity matrix used in the uplink joint transmission proposed in [26]. In this paper, this similarity matrix is used for down-link joint transmission. The off-diagonal components of this similarity matrix are defined as the value dividing the down-link SINR to the i th UE by the SINR value when non-CoMP pair CoMP between the i th RRH and all other RRHs (p indexing). This is the value representing the gain when performing joint transmission with the i th RRH. Here, the meaning of $SINR_{non}^i$ indicates the ratio of the sum of the interference signal power received from all RRHs except the i th RRH plus noise power and the desired signal power received from the i th UE. The meaning of $SINR_{CoMP}(i, p)$ indicates the ratio of the desired signal power received by the i th UE from the joint transmission of i th RRH and p th RRH and the sum of the interference signal power received from all RRHs except i th RRH and p th RRH plus noise power. Diagonal components can be obtained by multiplying the values remaining after excluding the pair CoMP SINR from the off-diagonal components by ε , a coordinative parameter. In (5), ε adjusts the size of the preference and this value is determined to be 0.3. As ε decreases, the value of the diagonal components decreases and K decreases, and as ε increases, K also increases. We implement a basic AP algorithm using this similarity matrix in the following subsections, and propose a new clustering algorithm by adding an image processing method and another clustering algorithm.

Algorithm 1 AP clustering algorithm

- 1: (A) Initialization: Similarity matrix \mathbf{S} according to (5)
- 2: Set initial availability matrix $\mathbf{A} = [\mathbf{0}]$, responsibility matrix $\mathbf{R} = [\mathbf{0}]$, and # of cluster $K = 0$
- 3: Calculate $\mathbf{R}(i, :)$ and $\mathbf{A}(:, p)$
- 4: $r(i, p) = s(i, p) - \max_{p' \neq p} (a(i, p') + s(i, p'))$
- 5: $a(i, p) = \min (0, r(p, p) + \sum_{i' \neq i, p} \max(0, r(i', p)))$
- 6: $a(p, p) = \sum_{i' \neq p} \max\{0, r(i', p)\}$
- 7: Update $\mathbf{R}(i, :)$ and $\mathbf{A}(:, p)$
- 8: Oscillatory decay: (φ and $\zeta \in [0, 1]$)
- 9: (B) Iteration:
- 10: **while** Convergence **do**
- 11: $\mathbf{R}(\text{iter}) = \varphi \cdot \mathbf{R}(\text{iter}) + (1 - \varphi) \cdot \mathbf{R}(\text{iter} - 1)$
- 12: $\mathbf{A}(\text{iter}) = \zeta \cdot \mathbf{A}(\text{iter}) + (1 - \zeta) \cdot \mathbf{A}(\text{iter} - 1)$
- 13: (C) Exemplar judgment:
- 14: **for** $\text{ind} = 1$ to number of RRH **do**
- 15: **if** $r(\text{ind}, \text{ind}) + a(\text{ind}, \text{ind}) > 0$ **then**
- 16: $K = K + 1$
- 17: $\text{exemplar}(\text{ind}) = \text{argmax}_p (a(\text{ind}, p) + r(\text{ind}, p))$
- 18: (D) Output: exemplar, set of clusters, and K

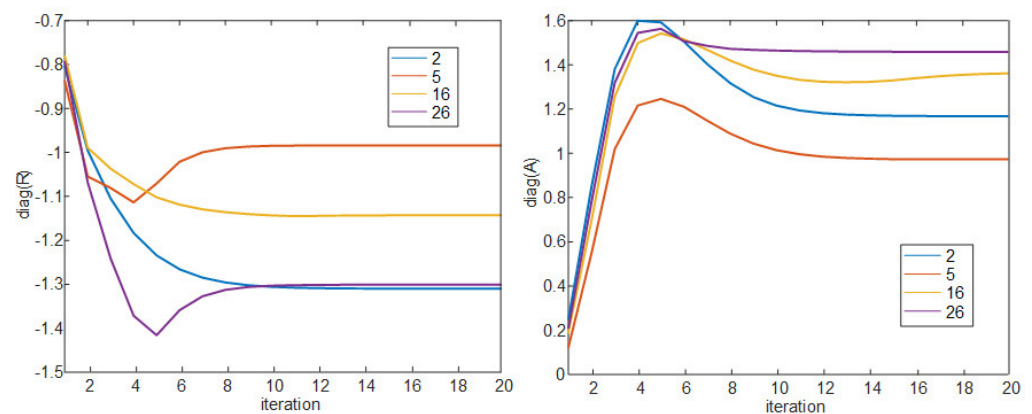


Figure 3. Diagonal component values of \mathbf{R} , \mathbf{A} matrix by iteration in Algorithm 1 (2, 5, 16, and 26 data points).

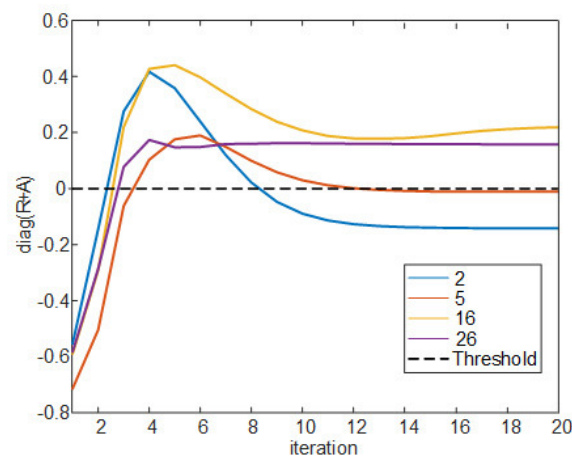


Figure 4. Diagonal component values of $\mathbf{R} + \mathbf{A}$ matrix by iteration in Algorithm 1 (2, 5, 16, and 26 data points) with 0 threshold.

3.1. Proposed Clustering Algorithm 1

The proposed clustering Algorithm 1 is a method of changing the fixed threshold 0 of the conventional AP clustering algorithm to a more suitable value for a given environment. One of the traditional image processing methods, Otsu's method, is used in this clustering algorithm [30]. Otsu's method can be selected as a method for binarization and transforms an image with several levels of gray including white and black into an image with only white and black. In [30], the criterion for dividing white and black is called the Otsu threshold. In this study, as a result of the mesh on the right of Figures 5–8, the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix that determine the cluster in AP clustering are imaged through mapping to the topology. In Algorithm 2, the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix mapping information are used to determine the Otsu threshold in the Otsu's method. The process of determining the Otsu threshold uses the distribution of black and white in the gray-scale histogram, and can be explained by (6)–(9). In the example of Figures 5–8, the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix can be represented as a gray-scale histogram as shown in Figure 7. (6) shows the foreground and background weight value of the gray-scale histogram, and shows black as foreground(f) and white as background(b). T is the bin index, $T_f + T_b = A$, and A is the total number of bins. For both T_f and T_b , all values from 0 to A are obtained for (6)–(9) (see Figure 7 as an example, $A = 17$). (7) shows the foreground and background mean value of the gray-scale histogram, and γ shows the middle value of the bin width (Figure 7 as an example, $-0.65, -0.55, \dots, 0.95$). (8) shows the foreground and background variance value of the gray-scale histogram. (9) indicates within class variance, and eventually finds the index T that minimizes σ_{wcv}^2 , and determines γ_T with that index as the Otsu threshold.

$$W_{T_{f,b}} = \frac{\sum_{\alpha=0}^{T_{f,b}} \beta_{\alpha}}{\text{tr}(\mathbf{R} + \mathbf{A})} \quad (7)$$

$$\mu_{T_{f,b}} = \frac{\sum_{\alpha=0}^{T_{f,b}} (\gamma_{\alpha} \beta_{\alpha})}{\sum_{\alpha=0}^{T_{f,b}} \beta_{\alpha}} \quad (8)$$

$$\sigma_{T_{f,b}}^2 = \frac{\sum_{\alpha=0}^{T_{f,b}} [(\gamma_{\alpha} - \mu_{T_{f,b}})^2 \beta_{\alpha}]}{\sum_{\alpha=0}^{T_{f,b}} \beta_{\alpha}} \quad (9)$$

$$\sigma_{wcv}^2 = W_{T_b} \sigma_{T_b}^2 + W_{T_f} \sigma_{T_f}^2 \quad (10)$$

Algorithm 2 Proposed clustering Algorithm 1

- 1: (C) Exemplar judgment:
- 2: **for** $ind = 1$ to number of RRH **do**
- 3: **if** $r(ind, ind) + a(ind, ind) > \text{Otsu threshold}$ **then**
- 4: $K = K + 1$
- 5: $\text{exemplar}(ind) = \text{argmax}_p (a(ind, p) + r(ind, p))$
- 6: (D) Output: exemplar, set of clusters, and K

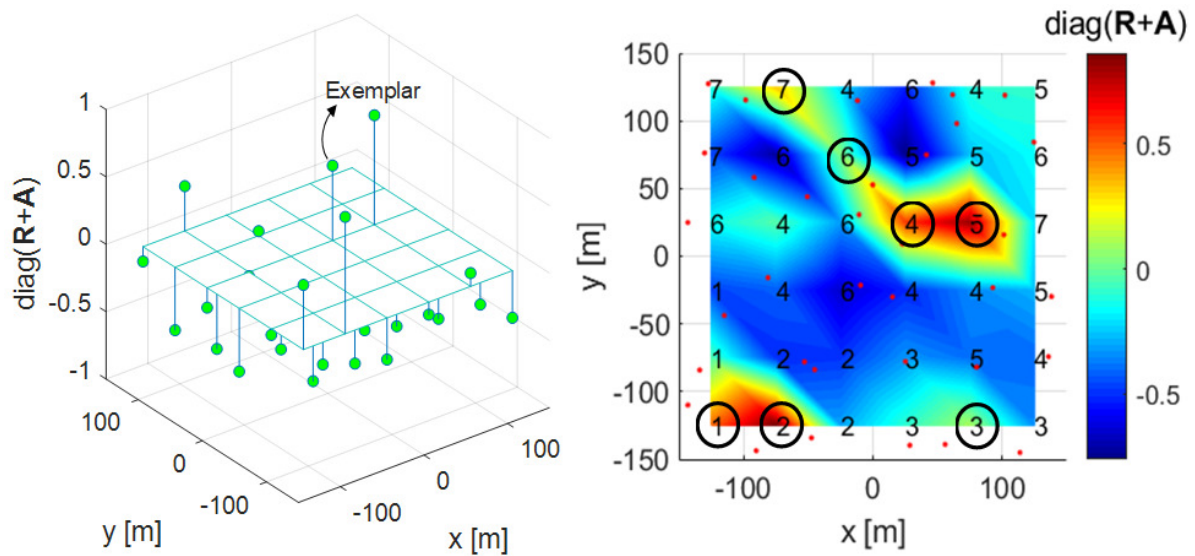


Figure 5. 3-D plot (left) and mesh plot (right) with the diagonal components of the $\mathbf{R} + \mathbf{A}$ in Algorithm 1.

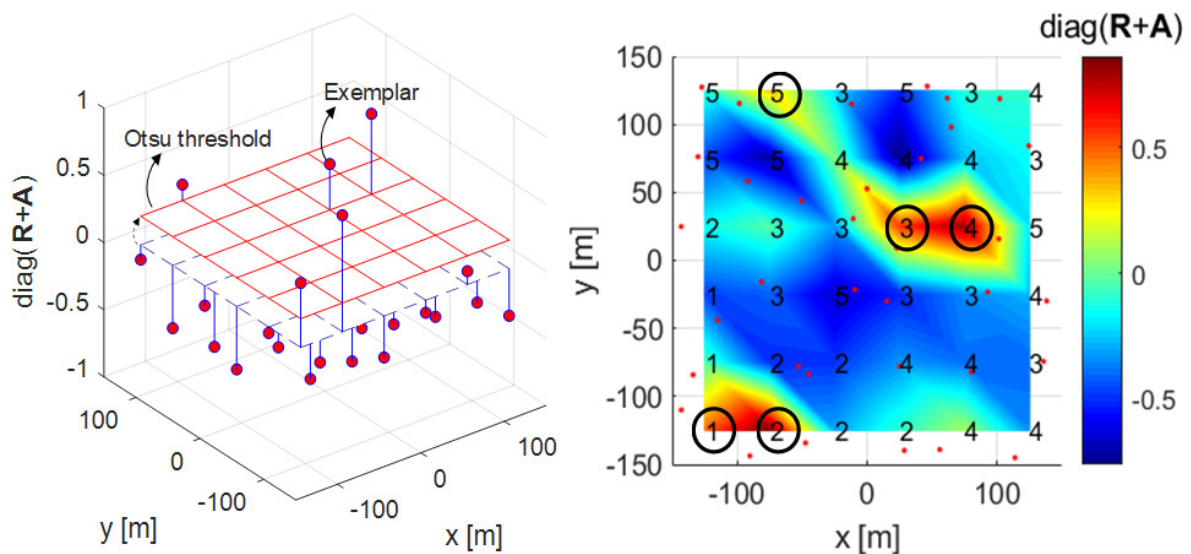


Figure 6. 3-D plot (left) and mesh plot (right) with Otsu threshold applied to the diagonal components of $\mathbf{R} + \mathbf{A}$ in Algorithm 2.

The figure on the right of Figure 5 shows the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix as a mesh plot in the desired clustering area C and C' in Figure 1. The left figure of Figure 5 shows the diagonal components of the $\mathbf{R} + \mathbf{A}$ matrix and threshold 0 at each

RRH position as a 3-D plot. In these two figures, the diagonal component values of the $\mathbf{R} + \mathbf{A}$ matrix having a threshold of 0 or more become exemplars, and in this case, there are 7 exemplars ($K = 7$). Algorithm 2 is the proposed clustering Algorithm 1 to determine the Otsu threshold by constructing a grayscale histogram of the diagonal component values of each RRH's $\mathbf{R} + \mathbf{A}$ matrix. The left figure of Figure 6 shows that the threshold is changed by Otsu's method, and the number of exemplars is eventually changed from 7 to 5. In the figure on the right of Figure 5, exemplars 2 and 7 have lower values than the Otsu threshold, and therefore these data points cannot be selected as exemplars in the figure on the right of Figure 6. Algorithm 2 represents the proposed clustering Algorithm 1, lines 1 to 13 are the same as for Algorithm 1, and there is a difference in how the exemplar is determined after that.

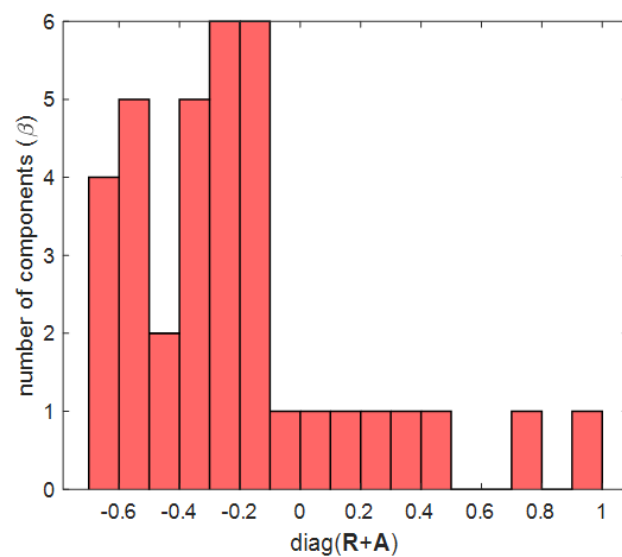


Figure 7. Histogram of diagonal component values of \mathbf{R}, \mathbf{A} matrix for Otsu threshold.

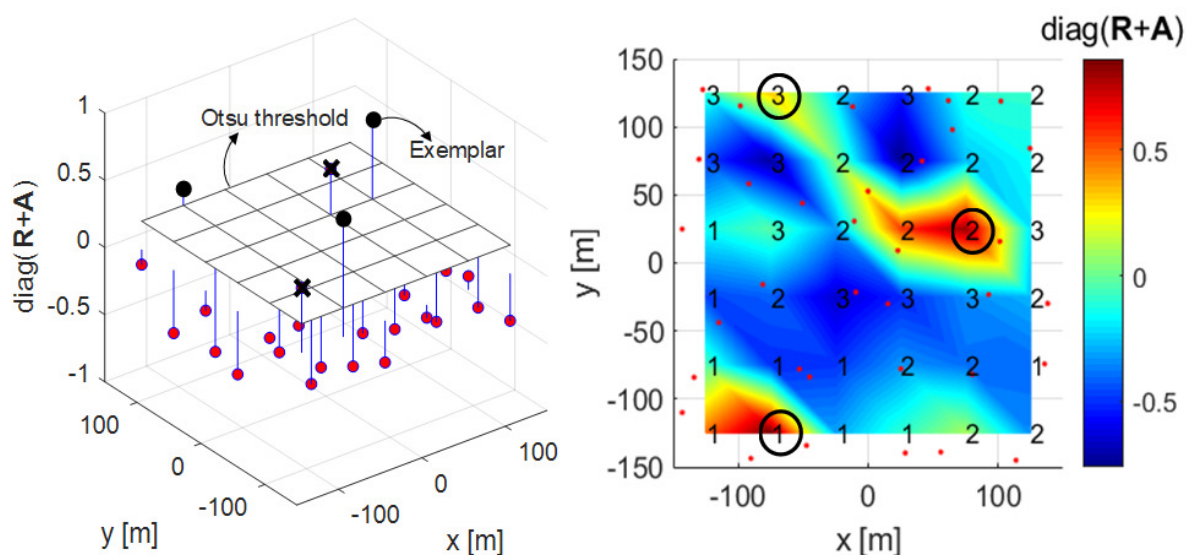


Figure 8. 3-D plot (left) and mesh plot (right) with Otsu threshold and greedy merging algorithm applied to the diagonal components of $\mathbf{R} + \mathbf{A}$ in Algorithm 3.

3.2. Proposed Clustering Algorithm 2

The second proposed clustering algorithm is a clustering technique that uses a greedy merging algorithm in addition to the proposed clustering Algorithm 1 proposed in Section 3.1. In this paper, a greedy merging algorithm creates one representative exemplar

by grouping adjacent exemplars if exemplars are adjacent to each other. An exemplar serves as the center of the cluster, and there will be greater interference when these exemplars are adjacent to each other. Algorithm 3 shows the process flow of the proposed clustering Algorithm 2. Lines 1 to 13 are the same as for Algorithm 1, and lines 14 to 18, the process of applying the Otsu threshold, are the same as for Algorithm and these lines apply the Otsu threshold. Lines 19 to 23 determine the exemplar with the largest diagonal component of $\mathbf{R} + \mathbf{A}$ as the representative exemplar when it finds an adjacent exemplar among the exemplars that exceed the Otsu threshold. In Figure 6, 2, 3, and 4 exemplars are adjacent among the 5 exemplars that have passed the Otsu threshold. As exemplar 3 has the highest value, the rest are eliminated as exemplars, and exemplar 3 becomes the representative.

Algorithm 3 Proposed clustering Algorithm 2

```

1: (C) Exemplar judgment:
2: for  $ind = 1$  to number of RRH do
3:   if  $r(ind, ind) + a(ind, ind) > \text{Otsu threshold}$  then
4:      $K = K + 1$ 
5:   if A exemplar has adjacent exemplar then
6:      $K = K - \text{number of adjacent exemplar}$ 
7:   else Overlapped adjacent exemplar
8:      $K = K$ 
9:    $\text{exemplar}(ind) = \text{argmax}_p a(ind, p) + r(ind, p)$ 
10: (D) Output: exemplar, set of clusters, and  $K$ 

```

3.3. Maximum Performance with AP Algorithm

This method selects the most suitable K from a given similarity matrix and serves as the upper bound of the three AP-based clustering algorithms. The method of selecting the exemplar in Algorithm 4 is to calculate (3) in all cases from when K is 1 to the number of RRHs (=36), and the exemplar is the K having the highest SINR. This clustering method provides the optimal performance for K . Therefore, in the Section 4, this method will be referred to as AP w/max K , and the results of this algorithm will also be derived to address the extent to which the proposed clustering algorithms can exhibit a performance comparable to that of this clustering algorithm.

Algorithm 4 Maximum performance with AP algorithm

```

1: (C) Exemplar judgment:
2: for  $t = 1$  to number of RRH do
3:    $K = K + t$ 
4:    $\text{temporal exemplar}(1 \text{ to } t) = \text{argmax}_p a(t, p) + r(t, p)$ 
5:   Solve the (3) - SINR
6:   Find  $K \leftarrow \max(\text{SINR})$ 
7: (D) Output: exemplar, set of clusters, and  $K$ 

```

4. Simulation Results

In this section, we compare the SE performance against the γ of several clustering methods in grid and uniform topologies with the C-RAN environment designed in Section 2 and

compare the normalized execution time. The following shows the six different clustering methods covered in this section:

1. **AP w/ max K**: AP clustering algorithm with max K exhibits the best performance for Algorithm 1 and is shown in Algorithm 4;
2. **Proposed AP 2**: AP clustering algorithm with ghd Otsu threshold and the greedy merging algorithm (Algorithm 3);
3. **Proposed AP 1**: AP clustering algorithm with Otsu threshold (Algorithm 2);
4. **Conventional AP 1** [26]: Conventional AP clustering Algorithm 1;
5. **Conventional AP 2** [25]: Conventional AP clustering Algorithm 2;
6. **Static CoMP**: Static CoMP scheme assumes coordination of four adjacent cells ($K = 9$).

However, in the uniform topology, the static CoMP method is difficult to define, and therefore it is not added to the result. The non-CoMP result is also added as the lower bound result. Table 1 shows the simulation parameters used in the system-level simulator to derive the results. P_i , which has a direct effect on the γ value, is set to 25.4 dBm, and P_e is set to increase by 3 dBm from 10 dBm to 31 dBm, where $\gamma = E(I_{Non-CoMP})/E(I_{exterior})$ [dB] on the x-axis is defined as the ratio of the average non-CoMP interference power to the average exterior interference power. Thus, the $\gamma = \infty$ values in Figures 9–12 indicate that the exterior interference is 0. The number of RRHs and UEs is set to 36, and each RRH and UE has a single antenna. In the grid topology, the inter-site distance between RRHs is assumed to be 50 m. In the uniform topology, the inter-site distance is not defined, but the RRH has a random location within the same area as the grid topology. The clustering SE performance is calculated by (4) in the situation where P_i and P_e are set, and both average SE and 5-percentile UE SE results are derived from Figures 9–12.

Table 1. Simulation parameter. RRD: Radio remote head.

Parameter	Value
Center frequency (GHz)	28
Bandwidth (MHz)	20
Number of RRHs and UEs	36
MS receive antenna	1
RRH transmit antenna	1
P_i (dBm)	25.4
P_e (dBm)	10,13,16,19,22,25,28,31
RRH inter-site distance (m)	50
Height of RRH (m)	5
Height of MS (m)	1.5

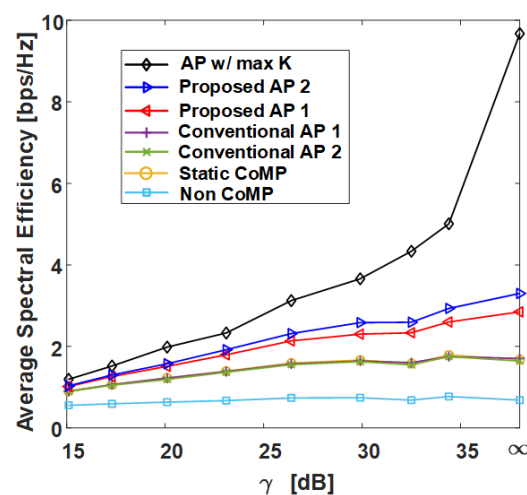


Figure 9. Average spectral efficiency of various clustering schemes with respect to γ in grid topology. AP: Affinity propagation.

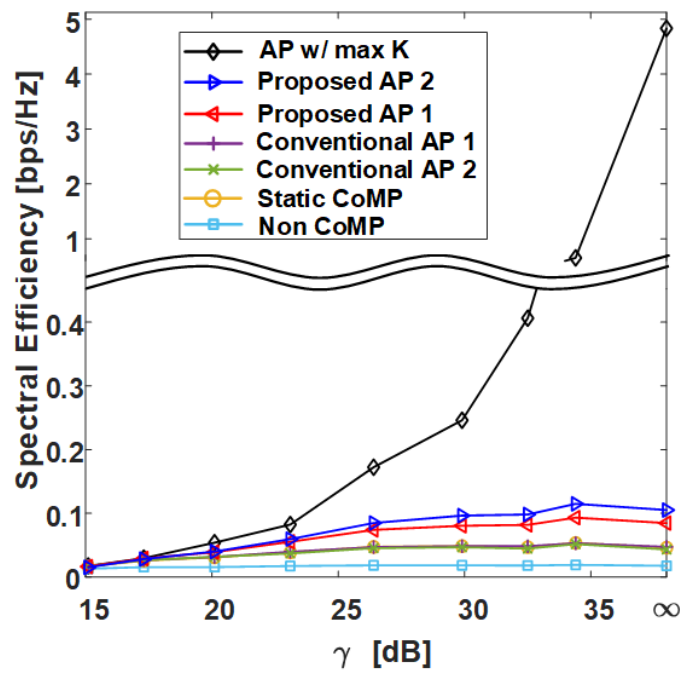


Figure 10. Spectral efficiency at 5% outage level of various clustering schemes with respect to γ in grid topology.

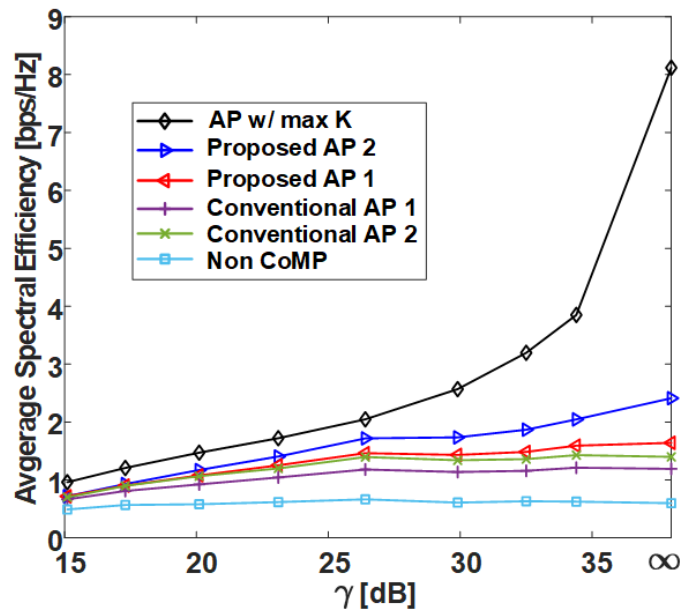


Figure 11. Average spectral efficiency of various clustering schemes with respect to γ in uniform topology.

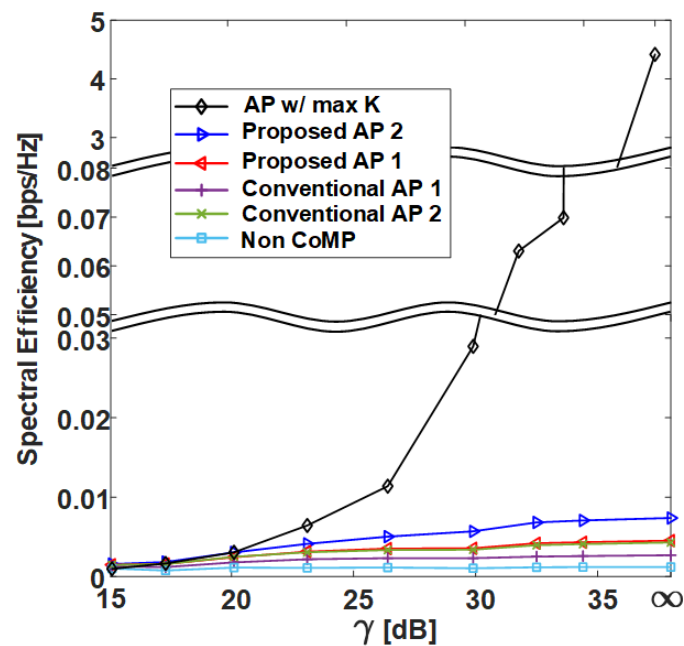


Figure 12. Spectral efficiency at 5% outage level of various clustering schemes with respect to γ in uniform topology.

Figures 9 and 11 show the average SE by γ in grid topology and uniform topology, respectively, and show that the proposed clustering Algorithms 1 and 2 exhibit better performance than the conventional clustering algorithms. The proposed clustering Algorithm 2 shows better performance than the proposed clustering Algorithm 1. Figures 10 and 12 show the results of 5-percentile UE SE in the grid topology and uniform topology, respectively. Figures 10 and 12 show the results of 5-percentile UE SE in grid topology and uniform topology, respectively. The trend is similar to the average SE result, but the SE performance of AP w/ max K is much better than that of the other clustering algorithms. The reason is that the proposed clustering Algorithms 1 and 2 are clustering methods that consider only K without considering the combination of clusters.

Figure 13 is the result of calculating the normalized execution time of AP-based clustering methods by the number of cells (RRHs) in the network. In Figure 13, the conventional AP clustering Algorithms 1 and 2 do not search the preference but use a initially defined similarity matrix. Naturally, the result of AP w/ max K shows the highest normalized execution time because the iteration is performed at least 36 times (the number of RRHs) more than the other algorithms. However, although the proposed clustering algorithms exhibit high performance in Figures 9–12, they have a normalized execution time similar to that of the other existing AP clustering algorithms. Hence, the proposed clustering algorithms exhibit better performance than the conventional clustering algorithms without a significant increase in time complexity. This result is excellent in terms of energy efficiency because it shows a similar time complexity even with a better clustering performance than the conventional clustering algorithm. Figures 14 and 15 show the average energy efficiency according to (5). This results show that the proposed AP clustering algorithm has a higher performance than the existing AP clustering algorithm in terms of energy efficiency.

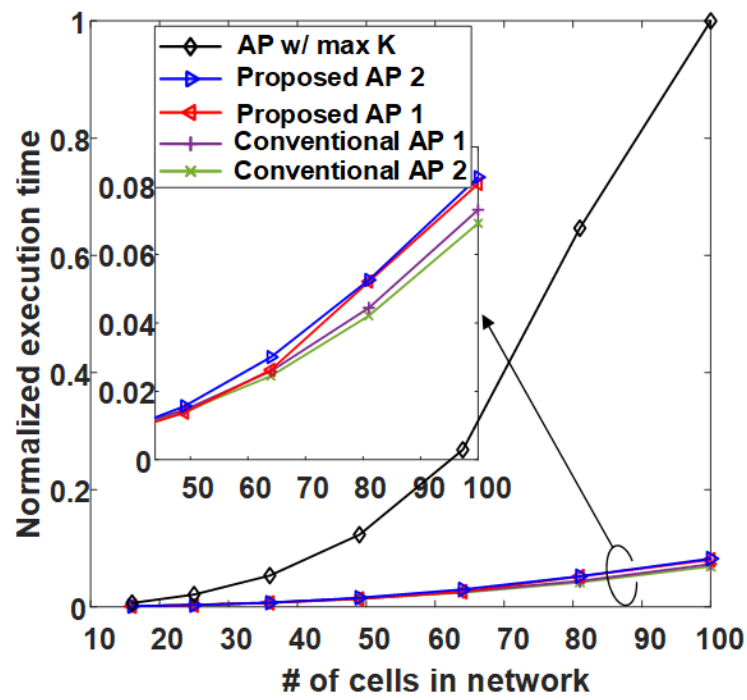


Figure 13. Average normalized execution time comparison of the conventional and the proposed clustering algorithms.

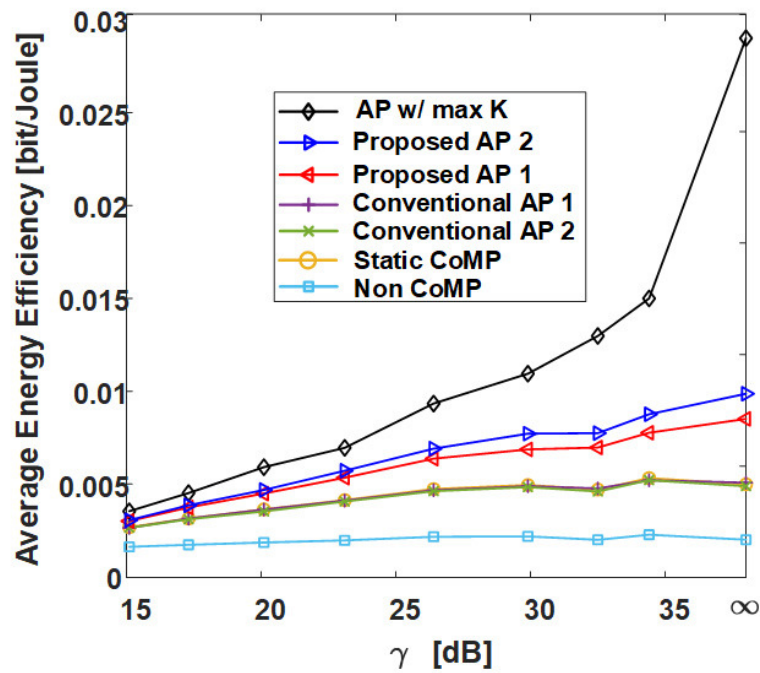


Figure 14. Average energy efficiency of various clustering schemes with respect to γ in grid topology.

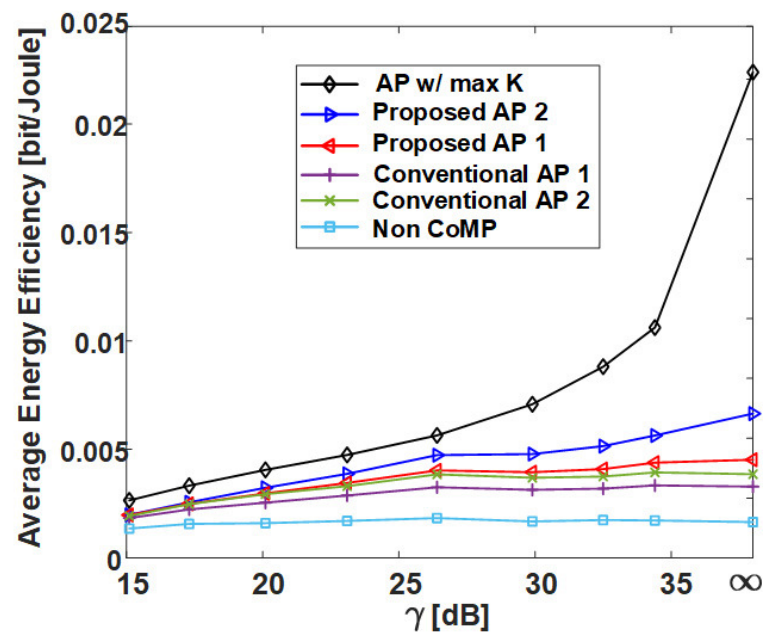


Figure 15. Average energy efficiency of various clustering schemes with respect to γ in uniform topology.

5. Conclusions

We proposed the use of adaptive thresholding and greedy merging for the AP clustering of the RRHs for dynamic joint transmission. The proposed method is a simple imaging method, without preference searching and the accompanying high time complexity. The clustering performance for a varying γ value, which is the ratio of the exterior interference to the non-CoMP interference, was analyzed in grid and uniform network topologies. While the existing AP-based clustering algorithms create an excessive number of clusters for some channel environments, the proposed clustering Algorithms 1 and 2 generated a relatively lower number of clusters, providing up to 37.7% and 102.4% improved SEs, respectively. Furthermore, the increases in the normalized execution time were small, at most 10.4% and 13%, respectively. Moreover, similar trends for SE were indicated for both the grid and uniform topologies. In addition, we could infer the average SE result of the actual distribution of RRHs as a value between the grid and uniform topology results [32]. The proposed algorithms could be extended to consider a more realistic joint transmission, such as imperfect symbol synchronization owing to the time difference of arrival, in future studies. Furthermore, we intend to perform an additional experimental verification of the results of this study.

Author Contributions: Conceptualization: S.P. and H.-S.J.; methodology: C.M.; software: S.P.; validation: H.-S.J. and J.-G.Y.; formal Analysis: S.P., H.-S.J., C.M., and J.-G.Y.; investigation: S.P.; resources: S.P.; data curation: S.P. and J.-G.Y.; writing—original draft preparation: S.P. and C.M.; writing—review and editing: S.P. and H.-S.J.; visualization: S.P.; supervision: H.-S.J., C.M., and J.-G.Y.; project administration: H.-S.J. and J.-G.Y.; funding acquisition: H.-S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Technology Advancement Research Program funded by the Ministry of Land, Infrastructure, and Transport of the Korean government under grant 20CTAP-C151968-02, in part by the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) under grant 2018R1D1A3B07050327.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. IMT Vision. *Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond*; International Telecommunication Union (ITU), Document, Radiocommunication Study Groups: Geneva, Switzerland, 2015.
2. Han, Y.; Elayoubi, S.E.; Galindo-Serrano, A.; Varma, V.S.; Messai, M. Periodic Radio Resource Allocation to Meet Latency and Reliability Requirements in 5G Networks. In Proceedings of the 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), Porto, Portugal, 3–6 June 2018; pp. 1–6.
3. Murakami, T.; Kishi, Y.; Ishibashi, K.; Kasai, K.; Shinbo, H.; Tamai, M.; Tsuda, K.; Nakazawa, M.; Tsukamoto, Y.; Yokoyama, H.; et al. Research project to realize various high-reliability communications in advanced 5G network. In Proceedings of the 2020 IEEE Wireless Communications and Networking Conference (WCNC), Seoul, Korea, 25–28 May 2020; pp. 1–8.
4. Wu, J.; Zhang, Z.; Hong, Y.; Wen, Y. Cloud radio access network (C-RAN): A primer. *IEEE Netw.* **2015**, *29*, 35–41. [[CrossRef](#)]
5. Nigam, G.; Minero, P.; Haenggi, M. Coordinated multipoint joint transmission in heterogeneous networks. *IEEE Trans. Commun.* **2014**, *62*, 4134–4146. [[CrossRef](#)]
6. Tanbourgi, R.; Singh, S.; Andrews, J.G.; Jondral, F.K. Analysis of non-coherent joint-transmission cooperation in heterogeneous cellular networks. In Proceedings of the 2014 IEEE International Conference on Communications (ICC), Sydney, Australia, 10–14 June 2014; pp. 5160–5165.
7. Peng, M.; Li, Y.; Jiang, J.; Li, J.; Wang, C. Heterogeneous cloud radio access networks: A new perspective for enhancing spectral and energy efficiencies. *IEEE Wirel. Commun.* **2014**, *21*, 126–135. [[CrossRef](#)]
8. Yu, X.; Cui, Q.; Haenggi, M. Coherent joint transmission in downlink heterogeneous cellular networks. *IEEE Wirel. Commun. Lett.* **2017**, *7*, 274–277. [[CrossRef](#)]
9. Tanbourgi, R.; Singh, S.; Andrews, J.G.; Jondral, F.K. A tractable model for noncoherent joint-transmission base station cooperation. *IEEE Trans. Wirel. Commun.* **2014**, *13*, 4959–4973. [[CrossRef](#)]
10. Zhang, H.; Safaei, F.; Tran, L.C. Joint Transmission Power Control and Relay Cooperation for WBAN Systems. *Sensors* **2018**, *18*, 4283. [[CrossRef](#)]
11. Ngo, H.Q.; Ashikhmin, A.; Yang, H.; Larsson, E.G.; Marzetta, T.L. Cell-free massive MIMO versus small cells. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1834–1850. [[CrossRef](#)]
12. Buzzi, S.; D’Andrea, C. Cell-free massive MIMO: User-centric approach. *IEEE Wirel. Commun. Lett.* **2017**, *6*, 706–709. [[CrossRef](#)]
13. Nayebi, E.; Ashikhmin, A.; Marzetta, T.L.; Yang, H. Cell-free massive MIMO systems. In Proceedings of the 2015 49th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 8–11 November 2015; pp. 695–699.
14. Abbasi, A.A.; Younis, M. A survey on clustering algorithms for wireless sensor networks. *Comput. Commun.* **2007**, *30*, 2826–2841. [[CrossRef](#)]
15. Wohwe Sambo, D.; Yenke, B.O.; Förster, A.; Dayang, P. Optimized clustering algorithms for large wireless sensor networks: A review. *Sensors* **2019**, *19*, 322. [[CrossRef](#)]
16. Song, L.; Song, Q.; Ye, J.; Chen, Y. A Hierarchical Topology Control Algorithm for WSN, Considering Node Residual Energy and Lightning Cluster Head Burden Based on Affinity Propagation. *Sensors* **2019**, *19*, 2925. [[CrossRef](#)] [[PubMed](#)]
17. Ronan, T.; Qi, Z.; Naegle, K.M. Avoiding common pitfalls when clustering biological data. *Sci. Signal.* **2016**, *9*, re6. [[CrossRef](#)] [[PubMed](#)]
18. Leber, B.; Maier, B.; Fuchs, F.; Chi, J.; Riffel, P.; Anderhub, S.; Wagner, L.; Ho, A.D.; Salisbury, J.L.; Boutros, M.; et al. Proteins required for centrosome clustering in cancer cells. *Sci. Transl. Med.* **2010**, *2*, 33ra38. [[CrossRef](#)] [[PubMed](#)]
19. Agarwal, S.; Ganesh, S. Perinuclear mitochondrial clustering, increased ROS levels, and HIF1 are required for the activation of HSF1 by heat stress. *J. Cell Sci.* **2020**, *133*, jcs245589. [[CrossRef](#)]
20. Frey, B.J.; Dueck, D. Clustering by passing messages between data points. *Science* **2007**, *315*, 972–976. [[CrossRef](#)]
21. Dueck, D. *Affinity Propagation: Clustering Data by Passing Messages*; University of Toronto: Toronto, ON, Canada, 2009.
22. Wang, J.; Gao, Y.; Wang, K.; Sangaiah, A.K.; Lim, S.J. An affinity propagation-based self-adaptive clustering method for wireless sensor networks. *Sensors* **2019**, *19*, 2579. [[CrossRef](#)]
23. Subedi, S.; Pyun, J.Y. Lightweight workload fingerprinting localization using affinity propagation clustering and Gaussian process regression. *Sensors* **2018**, *18*, 4267. [[CrossRef](#)]
24. Morosi, S.; Mucchi, L.; Marabissi, D.; Dolfi, M.; Marini, K. On the trade-off between Secrecy and Energy-Efficiency in Multi-Layer Cellular Networks. In Proceedings of the 2019 IEEE 5th International forum on Research and Technology for Society and Industry (RTSI), Florence, Italy, 9–12 September 2019; pp. 132–137.
25. Wesemann, S.; Fettweis, G. Decentralized formation of uplink CoMP clusters based on affinity propagation. In Proceedings of the 2012 International Symposium on Wireless Communication Systems (ISWCS), Paris, France, 28–31 August 2012; pp. 850–854.
26. Zhang, H.; Liu, H.; Jiang, C.; Chu, X.; Nallanathan, A.; Wen, X. A practical semidynamic clustering scheme using affinity propagation in cooperative picocells. *IEEE Trans. Veh. Technol.* **2014**, *64*, 4372–4377. [[CrossRef](#)]
27. Park, S.; Jo, H.S.; Mun, C.; Yook, J.G. Radio Remote Head Clustering with Affinity Propagation Algorithm in C-RAN. In Proceedings of the 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), Honolulu, HI, USA, 22–25 September 2019; pp. 1–2.
28. Rahman, M.M.U.; Ghauch, H.; Imtiaz, S.; Gross, J. RRH clustering and transmit precoding for interference-limited 5G CRAN downlink. In Proceedings of the 2015 IEEE Globecom Workshops (GC Wkshps), San Diego, CA, USA, 6–10 December 2015; pp. 1–7.

29. Moiane, A.F.; Machado, Á.M.L. Evaluation of the clustering performance of affinity propagation algorithm considering the influence of preference parameter and damping factor. *Bol. Ciências Geodésicas* **2018**, *24*, 426–441. [[CrossRef](#)]
30. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
31. 3GPP. *Study on Channel Model for Frequency Spectrum Above 6 GHz*; Technical Report (TR) 38.900, 3rd Generation Partnership Project (3GPP), Version 14.2.0.; 3GPP: Sophia Antipolis, France, 2017.
32. Jo, H.S.; Sang, Y.J.; Xia, P.; Andrews, J.G. Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 3484–3495. [[CrossRef](#)]