

Research Article

Joint Design of Beamforming and Edge Caching in Fog Radio Access Networks

Wenjing Lv,¹ Rui Wang ,^{1,2} Jun Wu ,^{3,4} Zhijun Fang,⁵ and Songlin Cheng ⁶

¹College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China

²Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai 201804, China

³School of Computer Science, Fudan University, Shanghai, China

⁴Peng Cheng Laboratory, Shenzhen, China

⁵College of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

⁶Department of Arts and Sciences, Shanghai Dianji University, Shanghai 201306, China

Correspondence should be addressed to Rui Wang; ruiwang@tongji.edu.cn

Received 16 August 2021; Accepted 30 August 2021; Published 18 September 2021

Academic Editor: Honghao Gao

Copyright © 2021 Wenjing Lv et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we study a novel transmission framework based on statistical channel state information (SCSI) by incorporating edge caching and beamforming in a fog radio access network (F-RAN) architecture. By optimizing the statistical beamforming and edge caching, we formulate a comprehensive nonconvex optimization problem to minimize the backhaul cost subject to the BS transmission power, limited caching capacity, and quality-of-service (QoS) constraints. By approximating the problem using the l_0 -norm, Taylor series expansion, and other processing techniques, we provide a tailored second-order cone programming (SOCP) algorithm for the unicast transmission scenario and a successive linear approximation (SLA) algorithm for the joint unicast and multicast transmission scenario. This is the first attempt at the joint design of statistical beamforming and edge caching based on SCSI under the F-RAN architecture.

1. Introduction

5G is being commercially deployed, while the mobile communication networks still face some challenges. Spectrum resources are scarce, network services are emerging, and the costs of operation are surging. These dilemmas have stimulated the industry and academia to explore novel network technology and wireless technology to support diversified scenarios [1].

From the network technology perspective, exploring a new intelligent wireless access network architecture is sustainable for solving the huge capacity demand of wireless networks. The fog radio access network (F-RAN) architecture is a promising network architecture in next-generation mobile communications [2]. In the F-RAN, the edge nodes, which have functions of storage and CRRM, are collectively evolved to fog access points (F-APs). It is critical that the upgraded F-APs [3] equipped with

functions of storage and cooperative radio resource management (CRRM) can be used to store part of the contents sent by the cloud server [4], which shortens the distance between the content data and the users and relieves the pressure of large data transmission on the backhaul link in the F-RAN [5]. In a word, the F-RAN can achieve better network performance gains by deploying more functions at the F-APs [6]. Users can finally receive the required services in the F-RAN architecture at a faster speed as well as by consuming less power. The technologies deployed in the F-RAN architecture involve edge caching allocation, resource scheduling, and advanced technologies such as massive multiple-input multiple-output (MIMO) and caching strategy. Current corresponding research topics based on the F-RAN architecture are generally reflected on two timescales: delivery-level design from a short timescale or caching-level design from a long timescale.

From the wireless technology perspective, beamforming technology based on multiple antennas can improve the energy efficiency and spectrum efficiency. The existing beamforming problems are generally NP-hard. The corresponding algorithms suffer from poor scalability and high computational complexity in large-scale wireless systems [7]. Even in a basic single-cell single-group multicast transmission scenario, the multicast beamforming design problem is NP-hard. Sidiropoulos et al. [8] obtained a high-quality approximate solution based on Gaussian randomization and semidefinite relaxation (SDR). Wang et al. [9] proposed a global beamforming algorithm named second-order cone programming (SOCP) in multiuser two-way relay systems, while the application scenarios of using this algorithm are limited. Later, Lu and Liu in [10] proposed a global algorithm named the branch-and-bound (BB) algorithm for a multicast beamforming scenario. However, the complexity of these beamforming algorithms usually increases when the network scale increases.

Recently, the team of Tao in [11] studied a content-centric physical layer beamforming framework under the cloud radio access network (C-RAN) architecture. It is aimed at overcoming dilemmas of repeated content transmission in connection-oriented traditional communication architectures. Tao et al. [11] considered only a multicast beamforming design under the C-RAN architecture. From a different research perspective, Chen et al. [12] studied a similar problem but proposed a joint beamforming and BS clustering framework for nonorthogonal unicast and multicast scenarios. They adopted a two-layered-division-multiplexing structure to concurrently achieve unicast and multicast services with different beamformers. Although [11, 12] involve the design of both cache-level and delivery-level strategies, neither works consider the joint optimization of caching and beamforming. The main difficulty is that the optimizations over the caching allocation and traditional beamforming strategy generally occur on different timescales. Traditional beamforming adapts to short-term instantaneous channel information, while caching is generally optimized at the cache allocation phase, which is in the context of a long timescale [13]. Thus, it remains challenging to simultaneously optimize the mixed-timescale beamforming and caching strategy in the F-RAN architecture.

In this article, we present a compromise scheme based on statistical channel state information (SCSI) to address this mixed-timescale deployment problem in the F-RAN architecture. The major contributions are concluded as follows:

- (i) *Joint Framework Based on SCSI.* We first propose a joint framework based on SCSI under the F-RAN architecture. Moreover, our research includes the unicast scenario and the joint nonorthogonal unicast and multicast transmission scenario.
- (ii) *Problem Formulation.* A multiobjective optimization problem is further formulated. Then, we reformulate the nonconvex problems by introducing the l_0 -norm, the approximation of the smooth function, and some other mathematical processing techniques.

- (iii) *Tailored Algorithm.* To solve these problems, we provide a tailored SOCP algorithm for unicast transmission scenarios and a successive linear approximation (SLA) algorithm for the joint unicast and multicast transmission scenario.

- (iv) *Simulation.* Simulation experiments illustrate the effectiveness of the algorithms and show that the joint beamforming-caching is better than the single scheme.

2. System Model and Problem Formulation

2.1. Network Architecture. We study a content-centric transmission framework based on SCSI in the F-RAN. Consider the downlink transmission with K single-antenna users and N multiple-antenna BSs. Each BS has L antennas and a limited-capacity cache with storage Q_n for $n \in \mathcal{N} \triangleq \{1, \dots, N\}$. There are two paths for F-APs to acquire the requested information: the central processor (CP) through the backhaul links or its local cache. BSs can dynamically provide mixed multicast and unicast services. Each user can subscribe to group-specific multicast services and a dedicated unicast request. The users who request the same file can be considered as a group. Each group is served by a BS cluster. The system model is shown in Figure 1. Different transmission mechanisms are shown in Figure 2.

We denote Φ_m to represent the users of group m , $m \in \mathcal{M} \triangleq \{1, 2, \dots, M\}$. In addition, $s_{m,n} = 1$ denotes that the n th BS serves the m th group, and $s_{m,n} = 0$ denotes that the n th BS does not serve the m th group. Similarly, $s_{k,n} = 1$ denotes that the n th BS serves for the k th user; otherwise, $s_{k,n} = 0$. $C_{(k,z),n} = 1$ denotes that the k th user-requested file z has been cached in advance in the n th BS; otherwise, $C_{(k,z),n} = 0$. Correspondingly, $\hat{C}_{(m,z),n} = 1$ denotes that the m th-group-requested file z has been cached in advance in the n th BS; otherwise, $\hat{C}_{(m,z),n} = 0$. The BS cluster serving group m is marked by Ψ_m , $m \in \mathcal{M}$. The channel vectors from all BSs to user k can be written into a wide vector $\mathbf{h}_k = [\mathbf{h}_{k,1}^H, \mathbf{h}_{k,2}^H, \dots, \mathbf{h}_{k,N}^H]^H \in \mathbb{C}^{NL \times 1}$. Similarly, $\mathbf{v}_k \in \mathbb{C}^{NL \times 1}$, $k \in \mathcal{K} \triangleq \{1, 2, \dots, K\}$, and $\mathbf{w}_m \in \mathbb{C}^{NL \times 1}$, $m \in \{0\} \cup \mathcal{M}$, are the network-wide beamforming vectors for the unicast message and multicast message, respectively. Note that $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_K]$ and $\gamma = [\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_M]$ are the SINR vectors, where each element is the minimum required by user k and group m . Correspondingly, the transmission rate is set as $R_k = B \log(1 + \lambda_k)$ or $R_m = B \log(1 + \gamma_m)$, where B denotes the total available bandwidth. The database of Z contents is denoted as $\mathcal{X} \triangleq \{1, 2, \dots, Z\}$, and each file's size is normalized to 1. The popular files can be cached in BSs in advance according to the popularity of the z th content file, which follows a Zipf distribution: $p_z = z^{-r} / \sum_{j=1}^Z j^{-r}$.

2.2. Cost Model. The system cost consists of two parts: backhaul cost and transmission power consumption. The backhaul consumption is generally proportional to its transmission capacity and related to the cache, user's request, and matching service between RRs and the user. It is

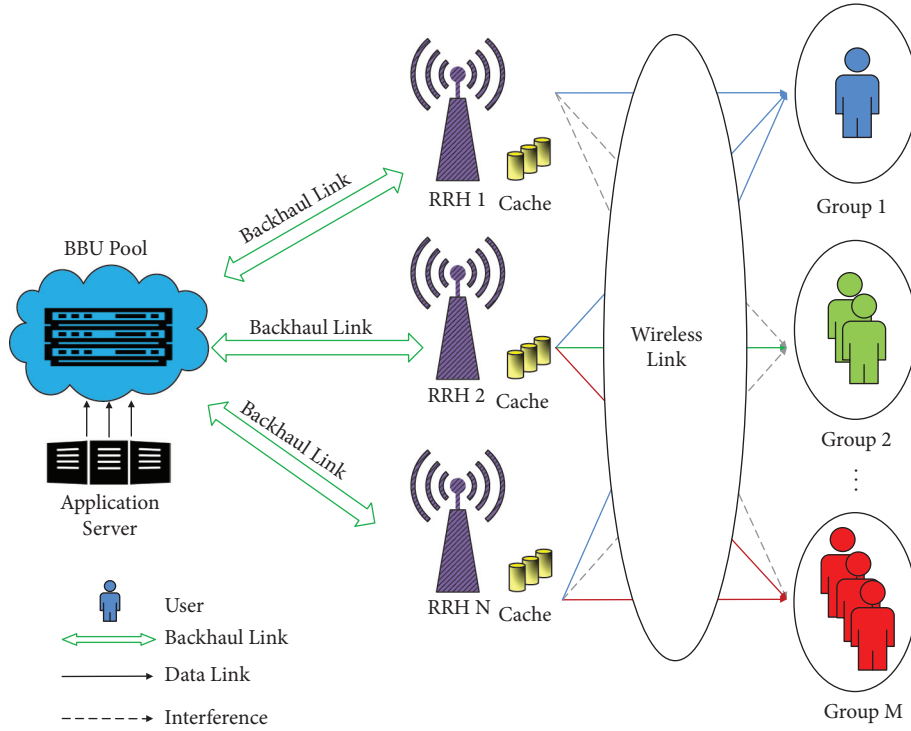


FIGURE 1: The system model based on the F-RAN architecture.

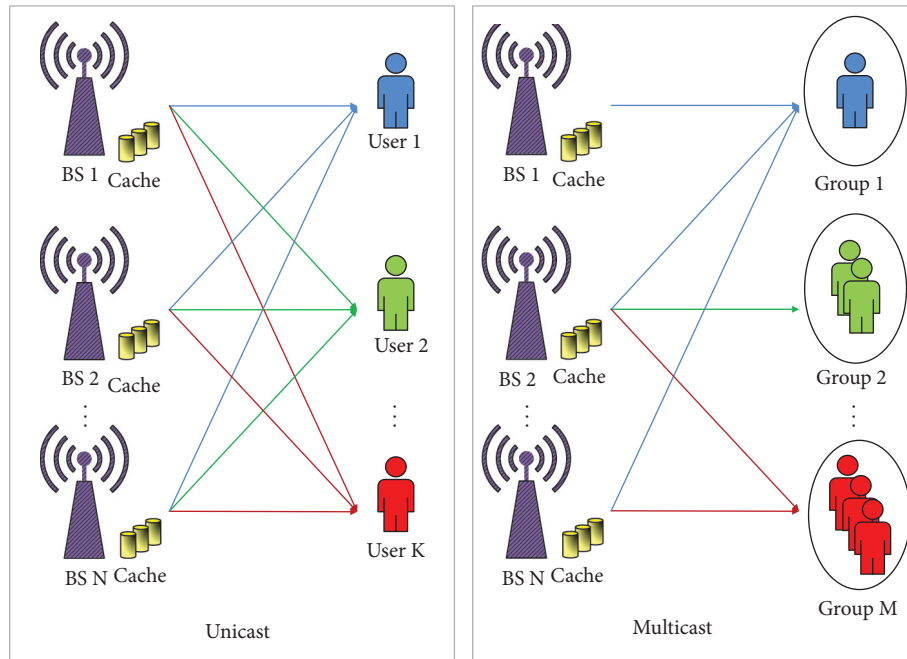


FIGURE 2: Different transmission modes.

worth remarking that if the file has been cached in advance in the BS, it can access the file directly and will not produce the backhaul link cost. Otherwise, it will produce extra backhaul cost. Thus, we use these variables to formulate the objective function on the backhaul cost. Specifically, for the unicast scenario, the system backhaul cost can be modeled as

$$P_B = \sum_{k=1}^K \sum_{n=1}^N s_{k,n} (1 - C_{(k,z),n}) R_k. \quad (1)$$

For the hybrid unicast-multicast scenario, the system backhaul cost consists of T_{B_1} and T_{B_2} , where

$$T_{B_1} = \frac{1}{K} \sum_{z=1}^Z \frac{z^{-\alpha}}{\sum_{j=1}^Z j^{-\alpha}} \sum_{k=1}^K \sum_{n=1}^N s_{k,n} (1 - C_{(k,z),n}) R_k, \quad (2)$$

$$T_{B_2} = \frac{1}{M} \sum_{z=1}^Z \frac{z^{-\alpha}}{\sum_{j=1}^Z j^{-\alpha}} \sum_{m=1}^M \sum_{n=1}^N \hat{s}_{m,n} (1 - \hat{C}_{(m,z),n}) R_m.$$

Additionally, the total BS transmission power can be expressed as a function of beamforming \mathbf{w} or \mathbf{v} . In the unicast scenario, it can be written as

$$P_{F_1} = \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{v}_{k,n}\|_2^2. \quad (3)$$

In the hybrid unicast-multicast scenario, it can be written as

$$P_{F_2} = \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{v}_{k,n}\|_2^2 + \sum_{m=1}^M \sum_{n=1}^N \|\mathbf{w}_{m,n}\|_2^2. \quad (4)$$

The maximum of all BSs' available power is limited to P_f .

2.3. Problem Formulation. The joint design of beamforming and edge caching is deployed based on statistical channel state information. We focus on a cost goal aimed at minimizing the backhaul cost limited to the transmission power, the SINR constraints, and the limited caching capacity. It is different for the transmission modes between the unicast scenario and the joint unicast-multicast scenario. In the joint unicast-multicast scenario, there are both unicast and multicast modes as shown in Figure 2.

- (1) In the unicast scenario, we first study a special case; suppose that there is only one user in each group which is named as a unicast scenario. The received message of the k th user can be written as follows:

$$y_k = h_k^H v_k x_k + \sum_{j=1, j \neq k}^K h_k^H v_j x_j + n_k, \quad \forall k \in \mathcal{K}, \quad (5)$$

where $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ represents the white Gaussian noise, $k \in \mathcal{K}$. Accordingly, user's SINRs can be defined as follows:

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{v}_j|^2 + \sigma_k^2}, \quad \forall k \in \mathcal{K}. \quad (6)$$

We make an average operation on the channel state information. Then, we get the SINR expression of the k th user:

$$\mathbb{E}(\text{SINR}_k) \geq \lambda_k, \quad (7)$$

and it can be written approximately as follows:

$$\frac{\mathbf{v}_k^H \mathbb{E}(\mathbf{h}_k \mathbf{h}_k^H) \mathbf{v}_k}{\sum_{j=1, j \neq k}^K \mathbf{v}_j^H \mathbb{E}(\mathbf{h}_k \mathbf{h}_k^H) \mathbf{v}_j + \sigma_k^2} \geq \lambda_k. \quad (8)$$

The system backhaul cost is written as

$$P_B = \sum_{k=1}^K \sum_{n=1}^N s_{k,n} (1 - C_{(k,z),n}) R_k. \quad (9)$$

The total transmission power consumption is

$$P_F = \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{v}_{k,n}\|_2^2. \quad (10)$$

In the unicast scenario, the optimized problem based on statistical channel state information can be formulated as follows:

$$P^U: \quad \min_{\{\mathbf{v}_{k,n}, s_{k,n}, C_{(k,z),n}\}} \frac{1}{K} \sum_{z=1}^Z \frac{z^{-\alpha}}{\sum_{j=1}^Z j^{-\alpha}} P_B, \quad (11a)$$

$$s.t. \quad \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{v}_{k,n}\|_2^2 \leq P_f, \quad (11b)$$

$$\frac{\mathbf{v}_k^H \mathbb{E}(\mathbf{h}_k \mathbf{h}_k^H) \mathbf{v}_k}{\sum_{j=1, j \neq k}^K \mathbf{v}_j^H \mathbb{E}(\mathbf{h}_k \mathbf{h}_k^H) \mathbf{v}_j + \sigma_k^2} \geq \lambda_k, \quad \forall k \in \mathcal{K}, \quad (11c)$$

$$\sum_{(k,z)} C_{(k,z),n} \leq Q_n, \quad \forall n \in \mathcal{N}. \quad (11d)$$

- (2) Hybrid unicast and multicast scenario: users requesting the same file can be regarded as a small group and served using multicast beamforming as shown in Figure 2. We next study the joint non-orthogonal unicast and multicast transmission scenario. Suppose that the multicast message is first decoded, and then the unicast message is decoded by subtracting the multicast message in advance. We can successively decode the mixed multicast and unicast signal based on statistical channel state information. The signals from other users and groups are treated as interference. The SINRs of the unicast and the multicast are defined as follows:

$$\text{SINR}_k^U = \frac{|\mathbf{h}_k^H \mathbf{v}_k|^2}{\sum_{i=1, i \neq m}^M |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sum_{j=1, j \neq k}^K |\mathbf{h}_k^H \mathbf{v}_j|^2 + \sigma_k^2}, \quad \forall k \in \mathcal{K},$$

$$\text{SINR}_k^M = \frac{|\mathbf{h}_k^H \mathbf{w}_m|^2}{\sum_{i=1, i \neq m}^M |\mathbf{h}_k^H \mathbf{w}_i|^2 + \sum_{j=1}^K |\mathbf{h}_k^H \mathbf{v}_j|^2 + \sigma_k^2}, \quad \forall k \in \Phi_m. \quad (12)$$

The system backhaul cost consists of T_{B_1} and T_{B_2} , where T_{B_1} is

$$T_{B_1} = \frac{1}{K} \sum_{z=1}^Z \frac{z^{-\alpha}}{\sum_{j=1}^Z j^{-\alpha}} \sum_{k=1}^K \sum_{n=1}^N s_{k,n} (1 - C_{(k,z),n}) R_k \quad (13)$$

and T_{B_2} is

$$T_{B_2} = \frac{1}{M} \sum_{z=1}^Z \frac{z^{-\alpha}}{\sum_{j=1}^Z j^{-\alpha}} \sum_{m=1}^M \sum_{n=1}^N \hat{s}_{m,n} \left(1 - \hat{C}_{(m,z),n}\right) R_m. \quad (14)$$

The optimized problem based on statistical channel state information is formulated as follows:

$$P^{UM}: \quad (15a)$$

$$\min_{\{\mathbf{v}_{k,n}, \mathbf{w}_{m,n}, s_{k,n}, \hat{s}_{m,n}, C_{(k,z),n}, \hat{C}_{(m,z),n}\}} T_{B_1} + T_{B_2}, \quad (15a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{v}_{k,n}\|_2^2 + \sum_{m=1}^M \sum_{n=1}^N \|\mathbf{w}_{m,n}\|_2^2 \leq P_f, \quad (15b)$$

$$\mathbb{E}(\text{SINR}_k^U) \geq \lambda_k, \quad \forall k \in \mathcal{K}, \quad (15c)$$

$$\mathbb{E}(\text{SINR}_k^M) \geq \gamma_m, \quad \forall k \in \Phi_m, \quad (15d)$$

$$\sum_{(k,z)} C_{(k,z),n} + \sum_{(m,z)} \hat{C}_{(m,z),n} \leq Q_n, \quad \forall n \in \mathcal{N}. \quad (15e)$$

Due to the combinatorial nature of different variables from multiple levels, to solve the above problems P^U , P^{UM} is still a challenge.

3. Proposed Algorithms for Different Scenarios

To address the aforementioned challenges, we explore a tailored second-order cone programming (SOCP) algorithm for the unicast transmission scenario and a successive linear approximation (SLA) algorithm for the mixed unicast and multicast transmission scenario. In this section, we first reformulate problems as an edge caching and statistical beamforming optimization problem by introducing l_0 -norm approximation, Taylor approximation, and some other techniques. Then, we provide the detailed algorithms in the next section.

For one thing, we reformulate a sparse and tractable beamforming problem equivalently. Denote

$$\hat{s}_{m,n} = \begin{cases} 0 & \text{if } \hat{C}_{(m,z),n} = 0, \\ 0 \text{ or } 1 & \text{if } \hat{C}_{(m,z),n} = 1. \end{cases} \quad (16)$$

$\hat{s}_{m,n}$ can be substituted by the l_0 -norm:

$$\hat{s}_{m,n} = \|\|\mathbf{w}_{m,n}\|_2\|_0. \quad (17a)$$

Similarly, we have

$$s_{k,n} = \|\|\mathbf{v}_{k,n}\|_2\|_0. \quad (17b)$$

By substituting (17a) and (17b) into (11a) and (15a), the corresponding problems can be transformed into two equivalent sparse problems with sparsity from the l_0 -norm.

For another thing, the discontinuous l_0 -norm in the objective functions can be taken place by a continuous function. Specifically, we choose the next frequently used smooth logarithmic concave function:

$$f_\theta(x) = \frac{\log(x/\theta + 1)}{\log(1/\theta + 1)}. \quad (18)$$

Here, θ is a parameter that characterizes the smoothness. Its first-order Taylor expansions $\tilde{f}_\theta(x)$ can be written as follows:

$$\tilde{f}_\theta(x) = \frac{\log(x_0/\theta + 1)}{\log(1/\theta + 1)} + \frac{1}{x_0 + \theta} \cdot \frac{1}{\log(1/\theta + 1)} \cdot (x - x_0). \quad (19)$$

In addition, we denote $\mathbb{E}(\mathbf{h}_k \mathbf{h}_k^H) = \Sigma_k$; if the condition $\text{rank}(\Sigma_k) = 1$ is satisfied, we have the decomposition formula $\Sigma_k = \mathbf{H}_k \mathbf{H}_k^H$. Based on the above processing, the formulated optimization problems can be, respectively, solved by using different optimization algorithms. The details of the tailored algorithms are described as follows.

3.1. SOCP-Based Optimal Algorithm for the Unicast Scenario.

Considering nonconvex SINR (11c), without loss of optimality, there always exists a phase shift version making the formula $\mathbf{H}_k^H \mathbf{v}_k$ to be real as well as positive; that is, multiplying a phase shift transformation $e^{j\theta}$ will not affect the value of the objective function and still keep the constraints satisfied, and we can get $\mathbf{H}_k^H \mathbf{v}_k \geq \sqrt{\gamma_k} \gamma_k + 1 \cdot \sqrt{\sum_{j=1}^K |\mathbf{H}_k^H \mathbf{v}_j|^2 + \sigma_k^2} \mathfrak{I}\{\mathbf{H}_k^H \mathbf{v}_k\} = 0, \quad \forall k \in \{1, 2, \dots, K\}$. This formula can be further written as a standard second-order cone (SOC) constraint form: $\|\mathbf{A}\mathbf{x} + \mathbf{b}\| \leq \mathbf{c}^T \mathbf{x} + d$. In addition, we denote $C_{B_2} = \tilde{f}_\theta(\|\mathbf{I}_n(\mathbf{E}_k \mathbf{X})\|_2^2) (1 - C_{(k,z),n}) R_k$. Then, the problem P^U can be written as follows:

$$P_U: \quad (20a)$$

$$\min_{\{\mathbf{x}, C_{(k,z),n}\}} \frac{1}{K} \sum_{z=1}^Z \frac{z^{-\alpha}}{\sum_{j=1}^Z j^{-\alpha}} \sum_{k=1}^K \sum_{n=1}^N C_{B_2}, \quad (20a)$$

$$\text{s.t.} \quad \sum_{k=1}^K \sum_{n=1}^N \|\mathbf{I}_n(\mathbf{E}_k \mathbf{X})\|_2^2 \leq P_f, \quad (20b)$$

$$\|\mathbf{A}_k \mathbf{X} + \mathbf{b}_k\|_2 \leq \mathbf{C}_k^T \mathbf{X}, \quad \forall k, \quad (20c)$$

$$\sum_{(k,z)} C_{(k,z),n} \leq Q_n, \quad \forall n, \quad (20d)$$

where $\mathbf{E}_k = [0_1^{NL \times NL}, 0_2^{NL \times NL}, \dots, 1_k^{NL \times NL}, \dots, 0_K^{NL \times NL}]$,

$$\begin{aligned}
\mathbf{H}_k &= [\mathbf{H}_{k,1}^H, \mathbf{H}_{k,2}^H, \dots, \mathbf{H}_{k,N}^H]^H \in \mathbb{C}^{NL \times 1}, \quad \mathbf{H}_{k,n} \in \mathbb{C}^{L \times 1}, \\
\mathbf{I}_n &= [0_1^{L \times L}, 0_2^{L \times L}, \dots, 0_n^{L \times L}, \dots, 0_N^{L \times L}], \\
\mathbf{C}_k^T &= \sqrt{\frac{1 + \gamma_k}{\gamma_k}} [0, \dots, \mathbf{H}_k^H, \dots, 0], \\
\mathbf{b}_k &= [0, 0, 0, \dots, \sigma_k]^H \in \mathbb{C}^{(M+1) \times 1}, \\
\mathbf{A}_k &= \begin{bmatrix} \mathbf{H}_k^H, 0, 0, \dots, 0, 0 \\ 0, \mathbf{H}_k^H, 0, \dots, 0, 0 \\ 0, 0, 0, \dots, 0, \mathbf{H}_k^H \\ 0, 0, 0, \dots, 0, 0, 0 \end{bmatrix}.
\end{aligned} \tag{21}$$

Based on the above series of transformations over the nonconvex SINR constraint, approximations, and other techniques, it is shown that the original problem can be equivalently transformed into a standard SOCP paradigm. We can finally use the mature optimization algorithm to efficiently solve this SOCP via the available software package solver.

3.2. Generalized Algorithm for the Nonorthogonal Unicast and Multicast Scenario. SINR constraints (15c) and (15d) are nonconvex. Firstly, we perform an expected value calculation over the nonconvex SINR constraints. For nonconvex unicast SINR constraint (15c), similar to the aforementioned processing techniques over (11c), we can transform (15c) into SOC constraints as follows:

$$\begin{aligned}
\mathbf{H}_k^H \mathbf{v}_k &\geq \lambda_k \sqrt{\sum_{i=1, i \neq k}^M |\mathbf{H}_k^H \mathbf{w}_i|^2 + \sum_{j=1, j \neq k}^K |\mathbf{H}_k^H \mathbf{v}_j|^2 + \sigma_k^2}, \\
\Im(\mathbf{H}_k^H \mathbf{v}_k) &= 0, \quad \forall \lambda_k.
\end{aligned} \tag{22}$$

Different from the unicast SINRs, for given γ_m in the multicast scenario, since there is more than one user sharing a multicast beamforming \mathbf{w}_m and the channel matrices $\{\mathbf{H}_k\}$ are linearly independent, only one of users' SINRs can be rewritten as a SOC constraint; let us assume that it is the K th user; we have

$$\begin{aligned}
\mathbf{H}_K^H \mathbf{w}_m &\geq \gamma_m \sqrt{\sum_{i=1, i \neq m}^M |\mathbf{H}_K^H \mathbf{w}_i|^2 + \sum_{j=1, j \neq K}^K |\mathbf{H}_K^H \mathbf{v}_j|^2 + \sigma_K^2}, \\
\Im(\mathbf{H}_K^H \mathbf{w}_m) &= 0, \quad K \in \Phi_m.
\end{aligned} \tag{23}$$

Others of the multicast SINR constraints are still nonconvex. By introducing the auxiliary variables and letting the

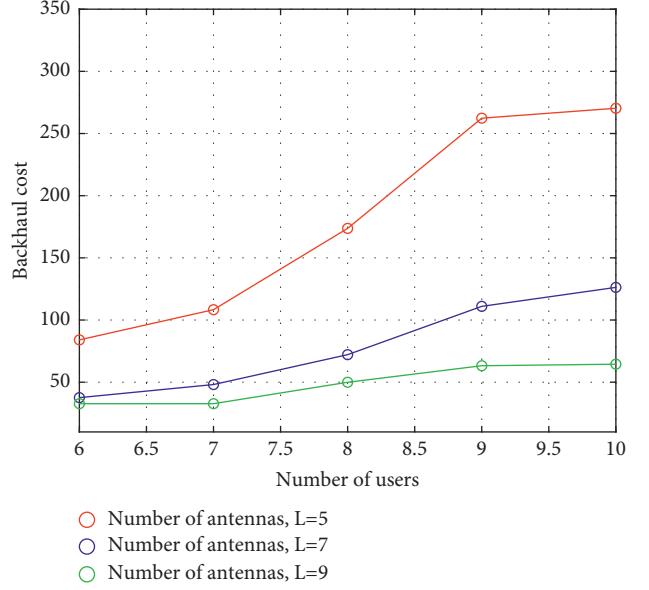


FIGURE 3: Backhaul cost against the number of users in the proposed SOCP algorithm.

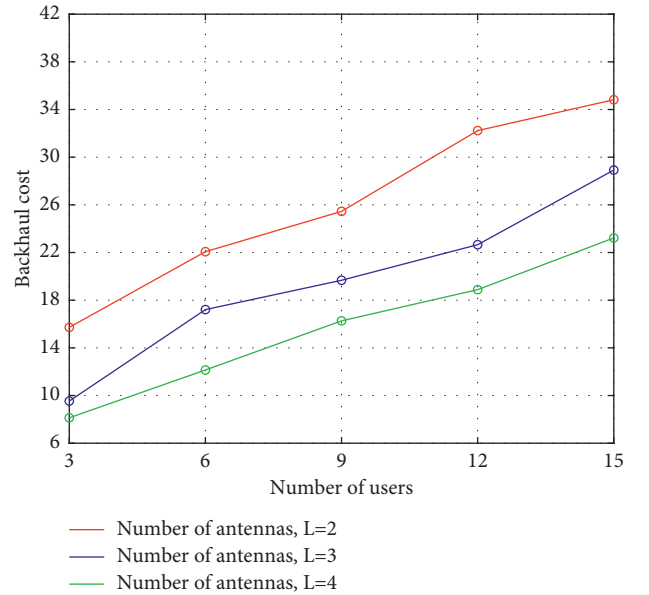


FIGURE 4: Backhaul cost against the number of users (the number of groups $M=2$) in the proposed SLA algorithm.

t th interaction point to be $\{g_{k,m}^t \in \mathbb{R}^2\}$, $\forall k \in \Phi_m / \{K\}$, the nonconvex multicast constraint can be transformed into

$$\begin{aligned}
&\|g_{k,m}^t\|^2 + 2(g_{k,m}^t)^T (g_{k,m} - g_{k,m}^t) \\
&\geq \gamma_m \sqrt{\sum_{i=1, i \neq m}^M |\mathbf{H}_k^H \mathbf{w}_i|^2 + \sum_{j=1}^K |\mathbf{H}_k^H \mathbf{v}_j|^2 + \sigma_k^2}.
\end{aligned} \tag{24}$$

Based on the above series of approximation transformations on nonconvex SINR constraints (15c) and (15d), by

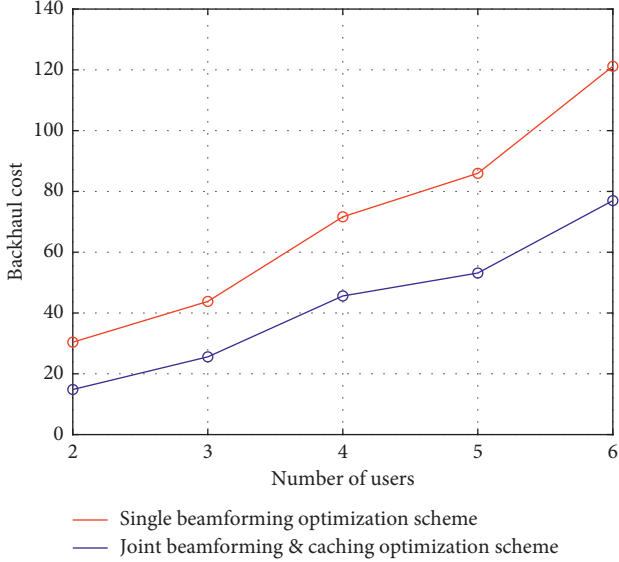


FIGURE 5: Performance comparison between the single beamforming scheme and joint beamforming-caching scheme for the SOCP algorithm.

substituting the formulas l_0 -norm and smooth logarithmic function into the objective function of P^{UM} , we can finally reduce P^{UM} to the following problem as follows:

$$P_{UM}: \quad \min_{\{v_{k,n}, w_{m,n}, C_{(k,z),n}, \hat{C}_{(m,z),n}\}} T_{B_1} + T_{B_2} \quad (24a)$$

$$s.t. (15b), (15e), (22), (23), \text{ and } (24). \quad (24b)$$

Using the SLA method, problem P_{UM} can finally be solved. The main step of the SLA method is to solve a series of convex subproblems as conducted in (24a) and (24b).

4. Simulation Results

4.1. Performance Comparison under Different Setups. Figure 3 demonstrates the backhaul cost against the number of users under different setups of BS antennas. It shows that the backhaul consumption increases with the number of users increasing. Moreover, the more the BS antennas in the network are, the less the backhaul cost the system consumes.

Figure 4 demonstrates the backhaul cost against the number of users under different numbers of BS antennas with three setups. As is shown, the more the antennas of the BS, the less the consumption of the backhaul link. Moreover, the backhaul consumption becomes higher with the number of users increasing. In addition, we can observe that the gap becomes smaller with the number of antennas increasing, which further illustrates the feasibility and effectiveness of the proposed framework.

Figures 3 and 4 separately demonstrate the backhaul cost against the number of users under different setups of BS antennas. Here, the backhaul consumption increases when the number of users increases. Moreover, when there are

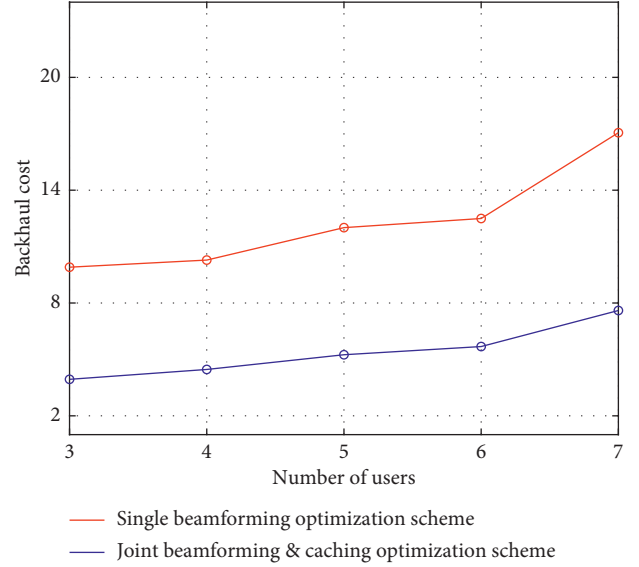


FIGURE 6: Performance comparison between the single beamforming scheme and joint beamforming-caching scheme for the SLA algorithm.

more BS antennas in the network, the system consumes a lower backhaul cost. The explanation for this phenomenon is that when the BS is equipped with more antennas, there is more antenna cooperative transmission. As a result, under the limited number of BSs and caching capacity, the transmission pressure on the backhaul link decreases, which reduces the backhaul consumption.

4.2. Performance Comparison between the Single Transmission Scheme and the Joint Beamforming-Caching Scheme. Figures 5 and 6 demonstrate the performance comparison between the single beamforming framework and the joint beamforming-caching framework for the SOCP algorithm in the unicast scenario and SLA algorithm in the joint unicast-multicast scenario, respectively. We observe that the joint beamforming-caching scheme performs better than the single beamforming scheme, which illustrates that the introduction of local caching can effectively reduce the backhaul consumption. Using the F-APs' caching function, part of the contents can be cached in advance, which shortens the actual distance between the user and content data and effectively alleviates the transmission pressure of the backhaul link.

5. Conclusion

In this article, we study a novel framework which is based on SCSI by incorporating beamforming and edge caching in a content-centric F-RAN architecture. A novel design on joint beamforming, edge caching, and dynamic F-AP clustering strategy with respect to SCSI is first explored. We formulate a mixed nonconvex problem aimed at minimizing backhaul consumption subject to the limited caching capacity, the BS power constraints, and SINR requirements. Accordingly, we provide a tailored second-order cone programming (SOCP)

algorithm for the unicast transmission scenario and a successive linear approximation (SLA) algorithm for the joint unicast and multicast transmission scenario. It is the first attempt about the joint design on beamforming and edge caching based on SCSi under the F-RAN architecture. Simulation results demonstrate the advantages of our framework based on SCSi compared with conventional schemes. In addition, it also makes sense to continue the study for the multicast scenario in the F-RAN architecture based on SCSi.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] H. Gao, X. Qin, R. J. D. Barroso, W. Hussain, Y. Xu, and Y. Yin, "Collaborative learning-based industrial IoT API recommendation for software-defined devices: the implicit knowledge discovery perspective," *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–11, 2020.
- [2] D. You, T. V. Doan, R. Torre et al., "Fog computing as an enabler for immersive media: service scenarios and research opportunities," *IEEE Access*, vol. 7, no. 7, pp. 65797–65810, 2019.
- [3] A. Sengupta, R. Tandon, and O. Simeone, "Fog-aided wireless networks for content delivery: fundamental latency tradeoffs," *IEEE Transactions on Information Theory*, vol. 63, no. 10, pp. 6650–6678, 2017.
- [4] Y. Huang, H. Xu, H. Gao, X. Ma, and W. Hussain, "SSUR: an approach to optimizing virtual machine allocation strategy based on user requirements for cloud data center," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 2, pp. 670–681, 2021.
- [5] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: issues and challenges," *IEEE Network*, vol. 30, no. 4, pp. 46–53, 2016.
- [6] M. Peng and K. Zhang, "Recent advances in fog radio access networks: performance analysis and radio resource allocation," *IEEE Access*, vol. 4, pp. 5003–5009, 2016.
- [7] J. Xiao, H. Xu, H. Gao, M. Bian, and Y. Li, "A weakly supervised semantic segmentation network by aggregating seed cues: the multi-object proposal generation perspective," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 1s, pp. 1–19, 2021.
- [8] N. D. Sidiropoulos, T. N. Davidson, and Z. Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, 2006.
- [9] R. Wang, M. Tao, and Y. Huang, "Linear precoding designs for amplify-and-forward multiuser two-way relay systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4457–4469, 2012.
- [10] C. Lu and Y. F. Liu, "An efficient global algorithm for single-group multicast beamforming," *IEEE Transactions on Signal Processing*, vol. 65, no. 14, pp. 3761–3774, 2017.
- [11] M. Tao, E. Chen, H. Zhou, and W. Yu, "Content-centric sparse multicast beamforming for cache-enabled cloud RAN," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6118–6131, 2016.
- [12] E. Chen, M. Tao, and Y. F. Liu, "Joint base station clustering and beamforming for non-orthogonal multicast and unicast transmission with backhaul constraints," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6265–6279, 2018.
- [13] H. Gao, K. Xu, M. Cao, J. Xiao, Q. Xu, and Y. Yin, "The deep features and attention mechanism-based method to dish healthcare under social IoT systems: an empirical study with a hand-deep local-global net," *IEEE Transactions on Computational Social Systems*, pp. 1–12, 2021.