Dynamic Resource Allocation in High-Speed Railway Fog Radio Access Networks with Delay Constraint

Rui Wang[1,2,3(✉)], Jun Wu[1], and Jun Yu[1]

[1] School of Electronics and Information Engineering, Tongji University, Shanghai, China
{ruiwang,wujun,1610986}@tongji.edu.cn

[2] State Key Laboratory of Integrated Services Networks, Xidian University, Xian, China

[3] Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai, China

Abstract. By applying caching resource at the remote radio heads (RRHs), the fog radio access network (Fog-RAN) has been considered as an promising wireless architecture in the future network to reduce the transmission delay and release the heavy burden of backhaul link for huge data delivery. In this paper, we propose to use the Fog-RAN to assist the data transmission in the high-speed railway scenario. In specific, we investigate the dynamic resource allocation in high-speed railway Fog-RAN systems by considering the delay constraint. The instantaneous power allocation at the RRHs and the instantaneous content delivery rate over the backhaul links are jointly optimized with an aim to minimize the total power consumed at the RRHs and over the backhaul links. An alternating optimization (AO) approach is used to find solutions of the instantaneous power and instantaneous content delivery rate in two separate subproblems. The closed-form solutions are derived in two subproblems under certain special conditions. Simulation results demonstrate that the proposed dynamic resource allocation is significantly superior to the constant resource allocation scheme.

Keywords: Fog radio access network · High-speed railway · Mobility · Power allocation

1 Introduction

Cloud radio access network (C-RAN) has been considered a promising wireless access network architecture to deal with the confronted explosive amount of traffic in the current cellular network to carry out the goal of the fifth generation (5G) wireless communication [1]. Thanks to the centralized wireless resource control ability at the base band unit (BBU) pool, the C-RAN can allocate the wireless resource of the network in a more efficient way. In C-RAN system,

H. Gao et al. (Eds.): ChinaCom 2019, LNICST 312, pp. 15–32, 2020.
https://doi.org/10.1007/978-3-030-41114-5_2

the uplink/downlink data transmission between BBU and the rate radio heads (RRHs) uses the backhaul links. However, the capacity of the backhaul links is still insufficient. As a result, when transmission data becomes large, the ginormous content demand creates heavy burden on backhaul links. To overcome this limitation, the fog radio access network (Fog-RAN) is proposed [2]. Compared to the C-RAN, certain caching resources are employed at RRHs. By caching some popular contents at RRHs, the requested data by users can be directly served by edge RRHs. This greatly alleviates the heavy traffic load pressure over the backhaul links and a significant portion of delay could be declined.

Although Fog-RAN is more efficient than C-RAN to achieve low-latency transmission and alleviate huge traffic burden on backhaul links, the configuration of network resources, including caching, power, computation etc., need further optimization to achieve better performance. A good number of studies have already been reported in this direction. For example, in literature [3], the authors investigated the problem of maximizing the delivery rate of Fog-RAN through content prefetching and improved precoding scheme. In literature [4], the authors investigated a proactive probabilistic caching optimization in wireless Fog-RAN by maximizing the successful transmission probability (STP). The authors in [5] studied the subchannel assignment and power control in mmWave-based fog radio access networks. In literature [6], the authors studied sparse beamforming design in a multicast Fog-RAN system.

All the aforementioned contributions show that great efforts have been paid to overcome the challenges of Fog-RAN architecture. However, the investigations are still insufficient as more different application scenarios may emerge in the upcoming 5G network domain. One important application is to employ the Fog-RAN architecture to assist the huge data transmission in high speed railway (HSR) scenario [7]. Recently, HSR is developing rapidly all over the world, especially in China. How to provide a reliable, high data rate and low latency HSR wireless communication has been identified as one of most important technologies needing to break though in the development. Moreover, HSR wireless communication has been categorized as a typical scenario in future 5G systems [8]. Undoubtedly, the Fog-RAN can be treated as an promising solution.

Different from traditional wireless communication problem, HSR wireless communication has to take the high-speed mobility condition into account, which greatly challenges the corresponding key technologies studies. High-speed mobility has already been considered in traditional wireless communication studies. For example, the authors in [9] studied a pilot aided joint channel and frequency offset estimation. In [10], the authors investigated the effect of distributed antenna techniques in HSR communications. In [11], a quality-of-service (QoS) based achievable rate region was characterized and a QoS distinguished power allocation algorithm was proposed to achieve the largest achievable rate region. In [12], the authors studied how to optimize the power to match user-data arrival process and time-varying channel service process with a delay constraint. The work in [13] studied the location-fair beamforming design by considering the Doppler shift. However, to the best of our knowledge, there are limited works studying the Fog-RAN in the HSR scenarios, which motivates the study of this work.

In this paper, we propose to use the Fog-RAN architecture to assist the data transmission in the HSR wireless communications to achieve low latency service. In the considered Fog-RAN system, the train is served by multiple RRHs and each RRH node has a local cache, which stores certain popular contents. When the contents requested by passengers has been cached at a given RRH, this RRH can directly serve the train; otherwise, the RRH needs to fetch the content from BBU pool via backhaul link. We investigate the dynamic resource allocation at Fog-RAN with an aim to minimize total power cost, including the power cost at RRHs and over backhaul links. The instantaneous power allocation at the RRHs and the instantaneous content delivery rate over backhaul links are simultaneously optimized under transmission delay constraint. By adopting smoothed l_0-norm approximation and other techniques, we propose an alternating optimization (AO) approach to find solutions of the instantaneous power and instantaneous content delivery rate in two separate subproblems. To reduce the computational complexity, we derive the closed-form solutions in two subproblem under certain special conditions. A constant resource allocation scheme is also provided as a benchmark to assess the performance of dynamic resource allocation scheme. Simulation results verify our analysis and demonstrate that the proposed dynamic resource allocation is significantly superior to the constant resource allocation scheme.

2 System Model

2.1 Channel Model

In a Fog-RAN served high speed railway wireless communication systems as illustrated in Fig. 1, BBU performs the resource allocation and the RRH selection to archive a high efficiency transmission. We consider the downlink transmission of a Fog-RAN system where a high speed train is served by uniformly deployed base stations (i.e., RRHs) along one side of the railway with equal intervals d. Assume that the distance between each RRH and railway is d_0, and the height of antenna equipped at each RRH is h_0. A high-speed train is traveling along the line railway with a constant velocity v_0. At the system time $t = 0$, the train passes the original point 0, and during time interval $t \in (0, T]$, the train is served by the same set of by N RRHs. Let $\mathcal{N} = \{1, ..., N\}$ denote the set of RRHs. The coordinate of the n-th RRH is denoted as (l_n, d_0). Then we can obtain the transmission distance between the n-th RRH antenna and the access point (AP) at train at time t as $d_n(t) = \sqrt{(v_0 t - l_n)^2 + d_0^2 + h_0^2}$ with $t \in (0, T]$. After the time $t = T$, the BBU will coordinate the handoff process and the train will be served by another set of N RRHs. Since the transmission process along the time is periodic, we only need to investigate the transmission problem during $t \in (0, T]$. Here we assume that the users in the train connect to the RRHs through the help of AP equipped on the roof the train to avoid severe penetration loss and large amounts of handoff operations. And the connection between the users and

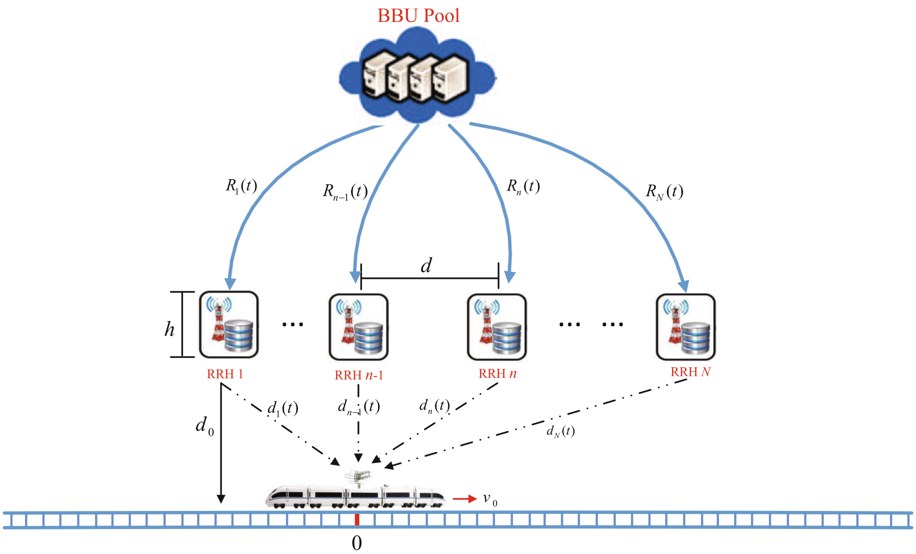

Fig. 1. Fog-RAN based high-speed railway communication system.

the AP is based on some other traditional reliable access networks, e.g. Wi-Fi etc. Therefore, in this study, we only focus on the transmission process between RRHs and AP.

It is supposed that the required data by the passengers in the train are grouped to F different contents. Contents have the same size of Q. We assume that the local storage size of RRH n is F_n and $F_n < QF$, which means that a RRH cannot store all the contents. Therefore, we define a cache placement matrix $C \in \mathbb{B}^{N \times F}$, where $c_{n,f} = 1$ means that the content f is cached in RRH n and $c_{n,f} = 0$ means the opposite. Note that $\forall n, \sum_{f=1}^{F} c_{n,f} Q \leq F_n$. At the beginning of each transmission time interval, the train submits a content request. According to the caching status of all RRHs, the BBU performs the dynamic resource allocation. Denote $x(t)$ the transmit signal from RRHs and $y_n(t)$ the received signal at the AP sent from RRH n. $x(t)$ can be considered as a stochastic process with zero mean and unit variance. Then, the baseband-equivalent instantaneous-time signal transmission between RRH n and AP can be represented as

$$y_n(t) = \sqrt{P_n(t)} h_n(t) x(t) + n_n(t) \tag{1}$$

where $P_n(t)$ is instantaneous transmit power at RRH n, $h_n(t)$ represents the instantaneous channel state information, and $n_n(t)$ denotes the additive complex cycle symmetric Gaussian noise at AP following a distribution of $CN(0, \sigma^2)$. It is noted that we here assume that the signal transmissions from RRHs are over orthogonal bandwidth, then at the receiver, maximal ratio combiner can be used to combine the signals. The corresponding instantaneous information capacity at time t can be expressed as

$$C(t) = B \log_2 \left(1 + \sum_{n=1}^{N} \frac{P_n(t)|h_n(t)|^2}{\sigma^2} \right) \tag{2}$$

where B is the bandwidth allocated for each channel between a RRH and the AP.

Consider that in HSR scenario, the train always runs in plain areas with less scatters. In this case, line-of-sight (LOS) component dominates the channel gain. Therefore, in this work a simple propagation attenuation model, i.e., $h_n(t) = \sqrt{\frac{G}{d_n^{\alpha}(t)}}$, is employed.

2.2 Problem Formulation

Our objective aims to minimize the total network power cost including total RRH power consumption and backhaul power consumption, while satisfying the delay constraint and individual RRH power constraint. Assume that the train request content F_f. If content F_f has been cached at RRH n, RRH n can directly transmit it to the AP without costing backhaul. Otherwise, content F_f needs to be fetched from the BBU via backhaul links. Assuming that instantaneous content delivery rate at time t over backhaul link connecting BBU and RRH n is $R_n(t)$, the total power cost for backhaul content delivery can be represented by

$$\text{Cost}_b = \int_0^T \sum_{n=1}^{N} \beta || \int_0^T P_n(t)dt||_0 (1 - c_{n,f}) R_n(t) dt. \tag{3}$$

In (3), we use term $| \int_0^T P_n(t)dt|_0$ to indicate the active RRH and term $1 - c_{n,f}$ to indicate the RRHs which do not cache the content F_f. It is noted that only the active RRHs which do not cache the content needs to cost backhaul links. Parameter β in (3) represents the ratio relationship between the rate $R_n(t)$ and the power cost.

On the other hand, the total power consumed by the RRHs over time period $(0, T]$ is represented by

$$\text{Cost}_p = \int_0^T \sum_{n \in \mathcal{N}} P_n(t) dt. \tag{4}$$

As a result, the network total power cost can be modeled as

$$\text{Cost} = \text{Cost}_b + \text{Cost}_p. \tag{5}$$

Regarding the delay consideration, as Fog-RAN based HSR system consists of two hops. One hop refers to the backhaul data transmission and the other refers to the content delivery over the wireless channel between the RRHs and the AP. Denote the set of active RRHs which do not cache content F_f as $\Theta = \{n| \| \int_0^T P_n(t)dt \|_0 (1 - c_{n,f}) \neq 0\}$. The instantaneous transmission delay can be represented as

$$\tau_f(t) = \frac{1}{\min\{C(t), \min_{m\in\Theta} R_m(t)\}}. \tag{6}$$

In summary, the overall resource allocation problem is formulated as follows:

$$\min_{P_n(t), R_n(t)} \quad \text{Cost} \tag{7a}$$

$$s.t \quad \frac{1}{T}\int_0^T P_n(t)dt \le P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{7b}$$

$$\tau_f(t) \le \tau_{\max} \tag{7c}$$

$$\int_0^T C(t)dt \ge Q, \quad \int_0^T R_n(t)dt \ge Q, n \in \Theta \tag{7d}$$

$$R_n(t) \le b_n, \quad P_n(t) \ge 0 \tag{7e}$$

where (7b) indicates the average power constraint of each RRH, (7c) represents the instantaneous transmission delay constraint, (7d) means that the content needs to sent out though the network during the time period of T, constraint (7e) indicates that for each backhaul link, we have maximum instantaneous transmission rate, and for each RRH, instantaneous power should not be smaller than zero. Our final objective is to minimize the overall network cost by optimizing the instantaneous power at each RRH and the instantaneous transmission rate over each backhaul link.

3 Dynamic Resource Optimization for HSR with Delay Constraint

In this section, we try to solve problem (7) by using proper optimization techniques. It is noted that different from traditional resource allocation problem, the considered dynamic optimization problem involves integration, which makes the optimization more challenging. Also, as we instantaneously optimize the power at each RRH and the content delivery rates over backhaul links, optimization (7) is a non-convex problem. To find an efficient solution, we apply the alternating optimization to decouple $P_n(t)$ and $R_n(t)$ in optimization (7). Our contribution lies in that in each subproblem, we can approximately find the optimal solution.

3.1 Solving Power $P_n(t)$

With given $R_n(t)$, we optimize the power allocation by solving the following problem

$$\min_{P_n(t)} \text{Cost} \tag{8a}$$

$$s.t \ \frac{1}{T}\int_0^T P_n(t)dt \leq P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{8b}$$

$$\frac{1}{C(t)} \leq \tau_{\max} \tag{8c}$$

$$\int_0^T C(t)dt \geq Q \tag{8d}$$

$$R_n(t) \leq b_n, \quad P_n(t) \geq 0 \tag{8e}$$

In (8), we observe that all constraints are convex with respect to $P_n(t)$. However, since Cost_b in the objective function includes a nonconvex l_0-norm function, the overall problem (8) is nonconvex. To deal with this problem, we approximate the discontinuous l_0-norm with a continuous smooth log-function, i.e., $||x||_0 \approx \frac{\log\left(\frac{x}{\theta}+1\right)}{\log\left(\frac{1}{\theta}+1\right)}$. Here the introduced parameter θ can control the smoothness of the approximation. In general, a larger value of θ leads to a smoother function but a worse approximation and vice versa. With this approximation, the term Cost_b in the objection function of (8) can be approximated as

$$\text{Cost}_b \approx c\beta \sum_{n=1}^{N}(1-c_{n,f}) \log\left(\frac{\int_0^T P_n(t)dt + \theta}{\theta}\right)\int_0^T R_n(t)dt. \tag{9}$$

where $c = \frac{1}{\log\left(\frac{1}{\theta}+1\right)}$. Then the objective function in (8) changes to

$$\text{Cost}_{\text{appro1}} \approx \int_0^T \sum_{n\in\mathcal{N}} P_n(t)dt + \sum_{n\in\mathcal{N}} b_n \log\left(\frac{\int_0^T P_n(t)dt + \theta}{\theta}\right). \tag{10}$$

where $b_n = c\int_0^T \beta(1-c_{n,f})R_n(t)dt$.

In (10), it is found that the objective function is still nonconvex as it is the sum of convex function and a concave function, which cannot be solved directly. We next use the majorization-minimization (MM) algorithm to further approximate it. The key idea is to minimize an upper bound. As the logarithmic function is a concave one, it is upper bounded by its first-order Taylor expansion. In MM algorithm, an optimal solution of problem (10) can be obtained by minimizing the upper-bounded function of objective in an iterative manner.

$$\begin{aligned}\text{Cost}_{\text{appro2}} \approx \int_0^T \sum_{n\in\mathcal{N}} P_n(t)dt + \sum_{n\in\mathcal{N}} b_n \Bigg[\log\left(\frac{\theta + \int_0^T P_n^0(t)dt}{\theta}\right) \\ + \frac{\int_0^T P_n(t)dt - \int_0^T P_n^0(t)dt}{\theta + \int_0^T P_n^0(t)dt}\Bigg]\end{aligned} \tag{11}$$

where $\int_0^T P_n^0(t)dt$ is basis point of the Taylor expansion of $\log\left(\frac{\int_0^T P_n(t)dt+\theta}{\theta}\right)$.

Let $k_n = 1 + b_n \frac{1}{\theta+\int_0^T P_n^0(t)dt}$, we transfer (8) to the following MM problem:

$$\min_{P_n(t)} \int_0^T \left(\sum_{n\in\mathcal{N}} k_n P_n(t)\right) dt \tag{12a}$$

$$s.t\ \frac{1}{T}\int_0^T P_n(t)dt \le P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{12b}$$

$$C(t) \ge \frac{1}{\tau_{\max}} \tag{12c}$$

$$\int_0^T C(t)dt \ge Q \tag{12d}$$

$$P_n(t) \ge 0 \tag{12e}$$

To solve (12), we first give the following theorem.

Theorem 1. *If $\frac{T}{\tau_{\max}} \ge Q$, we have $C(t) = \frac{1}{\tau_{\max}}$ at the optimal solution and optimal solution of (12) can be represented as*

$$\begin{aligned} P_n(t) &= \beta_n \tilde{a}_n(t), \quad n = \{2, 3, \cdots, N\} \\ P_1(t) &= \tilde{a}_0(t) - \sum_{n=2}^{N} \tilde{a}_n(t) P_n(t) \end{aligned} \tag{13}$$

where $\tilde{a}_n(t)$ is defined in (16), and optimal β_n with ordered form is given by

$$\beta_{[n]} = \begin{cases} \frac{TP_{[n],\text{avg}}}{A_{[n]}} \text{ if } & n \le m-1 \\ \frac{b}{k_{[n]}A_{[n]}} & n = m \\ 0 & \text{Otherwise} \end{cases} \tag{14}$$

where $\beta_{[n]}$ is defined in (22), m and b are defined in (23).

Proof. It is found that if $\frac{T}{\tau_{\max}} \ge Q$, constraint (12d) is redundant as it must be satisfied if constraint (12c) is satisfied. In this case, constraint (12c) is active. Otherwise, we can always scale $P_n(t)$ using a positive and less than 1 value to activate constraint (12c) and simultaneously decreases the value of the objective function. With above analysis, problem (12) can be equivalently rewritten as:

$$\min_{P_n(t)} \int_0^T \left(\sum_{n\in\mathcal{N}} k_n P_n(t)\right) dt \tag{15a}$$

$$s.t\ \int_0^T P_n(t)dt \le TP_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{15b}$$

$$B\log_2\left(1 + \sum_{n\in\mathcal{N}} \frac{GP_n(t)}{d_n(t)^\alpha \sigma^2}\right) = \frac{1}{\tau_{\max}} \tag{15c}$$

$$P_n(t) \ge 0 \tag{15d}$$

To proceed, we re-express constraint (15c) as $\sum_{n=1}^{N} a_n(t)P_n(t) = 2^{\frac{1}{B\tau_{\max}}}$ where $a_n(t) = \frac{G}{d_n(t)^\alpha \sigma^2}$. With this relationship between $P_n(t)$, we have

$$P_1(t) = \tilde{a}_0(t) - \sum_{n=2}^{N} \tilde{a}_n(t)P_n(t) \tag{16}$$

where $\tilde{a}_0(t) = \frac{1}{a_1(t)}(2^{\frac{1}{B\tau_{\max}}} - 1)$ and $\tilde{a}_n(t) = \frac{a_n(t)}{a_1(t)}$. Substituting (16) into (15), we have

$$\min_{P_n(t)} \int_0^T \left(\sum_{n=2}^{N} k_n P_n(t)\right) dt + \int_0^T \left(k_1\tilde{a}_0(t) - \sum_{n=2}^{N} k_1 \tilde{a}_n(t)P_n(t)\right) dt \tag{17a}$$

$$s.t \int_0^T P_n(t)dt \leq TP_{n,\text{avg}} \quad \forall n \in \{2, 3, \cdots, N\} \tag{17b}$$

$$\int_0^T \left(\tilde{a}_0(t) - \sum_{n=2}^{N} \tilde{a}_n(t)P_n(t)\right) dt \leq TP_{1,\text{avg}} \tag{17c}$$

$$P_n(t) \geq 0 \tag{17d}$$

From problem (17), we see that at the optimal solution, for given average power, the term $\int_0^T \tilde{a}_n(t)P_n(t)dt$ should be as large as possible. To obtain the optimal solution, using Cauchy-Schwarz inequality, we have

$$\int_0^T \tilde{a}_n(t)P_n(t)dt \leq \sqrt{\int_0^T \tilde{a}_n^2(t)dt \int_0^T P_n^2(t)dt} \tag{18}$$

where the equality succeeds when $P_n(t) = \beta_n \tilde{a}_n(t)$ with β_n being an variable to control the average consumed power. With this observation, finding the optimal solution of problem (17) reduces to finding optimal variables of β_n via solving

$$\min_{\beta_n} \sum_{n=2}^{N} (k_n A_n - k_1 B_n)\beta_n \tag{19a}$$

$$s.t \;\; 0 \leq \beta_n \leq \frac{TP_{n,\text{avg}}}{A_n} \quad \forall n \in \{2, 3, \cdots, N\} \tag{19b}$$

$$\int_0^T \tilde{a}_0(t)dt - TP_{1,\text{avg}} \leq \sum_{n=2}^{N} \beta_n B_n \leq \int_0^T \tilde{a}_0(t)dt \tag{19c}$$

where $A_n = \int_0^T \tilde{a}_n(t)dt$ and $B_n = \int_0^T \tilde{a}_n^2(t)dt$. It is observed that problem (19) is a linear programming which can efficiently solved by interior point algorithm. We next derive the optimal analytical solution. It is noted that if ignoring constraint (19c), the optimal solution can be represented as

$$\beta_n = \begin{cases} \frac{TP_{n,\text{avg}}}{A_n} & \text{if } \; \text{Sign}(k_n A_n - k_1 B_n) < 0 \\ 0 & \text{Otherwise} \end{cases} \tag{20}$$

If the solution given in (20) satisfies constraint (19c), it is the optimal solution of (19). Otherwise, we should increase the values of β_n with corresponding $\mathrm{Sign}(k_n A_n - k_1 B_n) > 0$ if $\sum_{n=2}^{N} \beta_n B_n < \int_0^T \tilde{a}_0(t)dt - TP_{1,\mathrm{avg}}$ to make $\sum_{n=2}^{N} \beta_n B_n = \int_0^T \tilde{a}_0(t)dt - TP_{1,\mathrm{avg}}$; or we should decrease the values of β_n with corresponding $\mathrm{Sign}(k_n A_n - k_1 B_n) < 0$ if $\sum_{n=2}^{N} \beta_n B_n > \int_0^T \tilde{a}_0(t)dt$ to make $\sum_{n=2}^{N} \beta_n B_n = \int_0^T \tilde{a}_0(t)dt$. Then optimal β_n in (19) can be found by solving

$$\min_{\beta_n} \sum_{n=2}^{N} k_n A_n \beta_n \tag{21a}$$

$$s.t \;\; 0 \leq \beta_n \leq \frac{TP_{n,\mathrm{avg}}}{A_n} \quad \forall n \in \{2, 3, \cdots, N\} \tag{21b}$$

$$\sum_{n=2}^{N} \beta_n B_n = a \tag{21c}$$

where $a = \int_0^T \tilde{a}_0(t)dt - TP_{1,\mathrm{avg}}$ or $\int_0^T \tilde{a}_0(t)dt$ depending on the previous analysis. By replacing the variable β_n by a new variable $\beta_n' = k_n A_n \beta_n$, we transfer problem (21) to the one given by (19) can be found by solving

$$\min_{\beta_n} \sum_{n=2}^{N} \beta_n' \tag{22a}$$

$$s.t \;\; 0 \leq \beta_n' \leq \frac{k_n A_n TP_{n,\mathrm{avg}}}{A_n} \quad \forall n \in \{2, 3, \cdots, N\} \tag{22b}$$

$$\sum_{n=2}^{N} \beta_n' \frac{B_n}{k_n A_n} = a \tag{22c}$$

To find the optimal solution of (22), we reorder $\left\{\frac{B_n}{k_n A_n}\right\}$ as $\frac{B_{[2]}}{k_{[2]}A_{[2]}} \geq \frac{B_{[3]}}{k_{[3]}A_{[3]}} \geq \cdots \geq \frac{B_{[N]}}{k_{[N]}A_{[N]}}$. The optimal solution of $\beta_{[n]}'$ can be represented as

$$\beta_{[n]}' = \begin{cases} \frac{k_{[n]}A_{[n]}TP_{[n],\mathrm{avg}}}{A_{[n]}} & \text{if} \;\; n \leq m-1 \\ b & \;\; n = m \\ 0 & \;\; \text{Otherwise} \end{cases} \tag{23}$$

where m is the smallest integer ensuring $\sum_{[n]=2}^{m} \frac{B_{[n]}TP_{[n],\mathrm{avg}}}{A_{[n]}} > a$ and $b = a - \sum_{[n]=2}^{m-1} \frac{B_{[n]}TP_{[n],\mathrm{avg}}}{A_{[n]}}$. This completes the proof of Theorem 1.

For s special case where $N = 1$, Lemma 1 reduces to the following lemma.

Lemma 1. *If* $\frac{T}{\tau_{\max}} \geq Q$ *and* $N = 1$*, we have* $C(t) = \frac{1}{\tau_{\max}}$ *at the optimal solution and optimal solution of* (12) *can be represented as*

$$P_1(t) = \left(2^{\frac{1}{B\tau_{\max}}} - 1\right) \frac{d_1^{\alpha}(t)\sigma^2}{G}. \tag{24}$$

Proof. When $N = 1$, problem (12) can be written as

$$\min_{P_1(t)} \int_0^T \Big(k_1 P_1(t) \Big) dt \tag{25a}$$

$$s.t \ \frac{1}{T} \int_0^T P_1(t) dt \leq P_{1,\text{avg}} \quad \forall n \in \mathcal{N} \tag{25b}$$

$$C(t) = \frac{1}{\tau_{\max}} \tag{25c}$$

$$P_1(t) \geq 0 \tag{25d}$$

If the problem is feasible, the solution is determined by constraint (25c), which completes the proof of Lemma 1.

If condition $\frac{T}{\tau_{\max}} \geq Q$ is not satisfied, the conclusions presented in Theorem 1 and Lemma 1 are not applicable. To find the solution, we decompose the power $P_n(t) = P_{n,1}(t) + P_{n,2}(t)$, where $P_{n,1}(t)$ is used to satisfy

$$\min_{P_{n,1}(t)} \int_0^T \Big(\sum_{n \in \mathcal{N}} k_n P_{n,1}(t) \Big) dt \tag{26a}$$

$$s.t \ \frac{1}{T} \int_0^T P_{n,1}(t) dt \leq P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{26b}$$

$$B \log_2 \left(1 + \sum_{n=1}^{N} \frac{G P_{n,1}(t)}{d_n^{\alpha}(t) \sigma^2} \right) = \frac{1}{\tau_{\max}} \tag{26c}$$

$$P_{n,1}(t) \geq 0 \tag{26d}$$

It is noted that conclusions presented in Theorem 1 and Lemma 1 can be used to find $P_{n,1}(t)$. After determining optimal $P_{n,1}(t)$, denoted by $P^*_{n,1}(t)$, $P_{n,2}(t)$ can found by solving

$$\min_{P_{n,2}(t)} \int_0^T \Big(\sum_{n \in \mathcal{N}} k_n P_{n,2}(t) \Big) dt \tag{27a}$$

$$s.t \ \frac{1}{T} \int_0^T P_{n,2}(t) dt \leq b_n(t) \quad \forall n \in \mathcal{N} \tag{27b}$$

$$\int_0^T B \log_2 \left(c_n(t) + \sum_{n=1}^{N} \frac{G P_{n,2}(t)}{d_n^{\alpha}(t) \sigma^2} \right) dt \geq Q \tag{27c}$$

$$P_{n,2}(t) \geq 0 \tag{27d}$$

where $b_n(t) = P_{n,\text{avg}} - \frac{1}{T} \int_0^T P^*_{n,1}(t) dt$ and $c_n(t) = 1 + \sum_{n=1}^{N} \frac{G P^*_{n,1}(t)}{d_n^{\alpha}(t) \sigma^2}$. As $P_{n,2}(t)$ is non-negative, constraint (12c) must be satisfied when solving $P_{n,2}(t)$ and thus it can be ignored.

It is easy to observe that problem (27) is a convex problem. To get the optimal solution, we next develop a algorithm based on Karush-Kuhn-Tucher (KKT) conditions. To proceed, we first present the Lagrangian function given as

$$\begin{aligned} L = \int_0^T \Bigg(\sum_{n\in\mathcal{N}} \Big(k_n P_{n,2}(t) + \mu_{1,n}(P_{n,2}(t) - Tb_n(t))\Big) \\ - \mu_2\Big(B \log_2(c_n(t) + \sum_{n=1}^{N} \frac{GP_{n,2}(t)}{d_n^{\alpha}(t)\sigma^2}) - Q\Big)\Bigg) dt \end{aligned} \tag{28}$$

where $\mu_{1,n}$ and μ_2 are non-negative multiplier related to constraints (27b) and (27c), respectively. To minimize the Lagrangian function, it is necessary to differentiate the Lagrangian function with respect to $P_{n,2}(t)$ and set the derivative to zero for each time t, that is

$$\frac{\partial L}{\partial P_{n,2}(t)} = k_n + \mu_{1,n} - \frac{\mu_2 B}{\log 2} \frac{\frac{G}{d_n(t)^{\alpha}\sigma^2}}{c_n(t) + \sum_{n\in\mathcal{N}} \frac{GP_{n,2}(t)}{d_n(t)^{\alpha}\sigma^2} - Q} = 0 \tag{29}$$

By combining with the constraint (27d), we can obtain the solution given by

$$\begin{aligned} P_{n,2}(t) = \Bigg[\Bigg(\frac{\mu_2 GB}{\log 2 d_n(t)^{\alpha}\sigma^2(k_n + \mu_1)} + Q - c_n(t) - \sum_{m\neq n}^{N} \frac{GP_{m,2}(t)}{d_m(t)^{\alpha}\sigma^2}\Bigg) \\ \times \frac{d_n(t)^{\alpha}\sigma^2}{G}, 0\Bigg]^{+}. \end{aligned} \tag{30}$$

(30) shows that $P_{n,2}(t)$ with different n are coupled with each other, the final solution of $P_{n,2}(t)$ can be obtained by iterative update them until convergence. During the iteration, Lagrangian multipliers $\mu_{1,n}$ and μ_2 can be obtained via subgradient technique.

3.2 Solving Content Delivery Rate $R_n(t)$

For given $P_n(t)$, we next optimize the content delivery rate $R_n(t)$ over backhaul link by solving

$$\min_{R_n(t)} \int_0^T \sum_{n\in\mathcal{N}} E(t)R_n(t)dt \tag{31a}$$

$$s.t \quad \frac{1}{\min_{n\in\Theta} R_n(t)} \leq \tau_{\max} \tag{31b}$$

$$\int_0^T R_n(t)dt \geq Q, n \in \Theta \tag{31c}$$

$$0 \leq R_n(t) \leq b_n \tag{31d}$$

where $E(t) = \beta|| \int_0^T P_n(t)dt||_0(1 - c_{n,f})$. The optimal solution is given in the following lemma.

Lemma 2. *For problem (31), if* $\frac{T}{\tau_{\max}}$*, the optimal solution is*

$$\begin{cases} R_n(t) = \frac{1}{\tau_{\max}}, n \in \Theta \\ R_n(t) = 0, \quad\quad n \notin \Theta \end{cases} \tag{32}$$

Otherwise, the optimal solution is

$$\begin{cases} R_n(t) = \gamma\frac{1}{\tau_{\max}}, n \in \Theta \\ R_n(t) = 0, \quad\quad\ n \notin \Theta \end{cases} \tag{33}$$

where γ *is chosen to activate constraint (31c).*

Proof. To find optimal $R_n(t)$, we rewrite problem (31) as

$$\min_{R_n(t)} \int_0^T \sum_{n\in\mathcal{N}} E(t)R_n(t)dt \tag{34a}$$

$$s.t \ \ R_n(t) \geq \frac{1}{\tau_{\max}}, n \in \Theta \tag{34b}$$

$$\int_0^T R_n(t)dt \geq Q, n \in \Theta \tag{34c}$$

$$0 \leq R_n(t) \leq b_n \tag{34d}$$

If condition $\frac{T}{\tau_{\max}}$ is met, constraint in (34c) is redundant. The optimal solution is to activate constraint (34b) with the optimal solution given in (32). Otherwise, the optimal $R_n(t)$ should activate constraint (34c) with the optimal solution given in (33).

4 Invariant Resource Optimization for HSR with Delay Constraint

As another simple power allocation scheme, we consider a constant power optimization design where power does not vary with the channel. In this case, the overall optimization problem is modified as

$$\min_{P_n, R_n(t)} T\sum_{n=1}^{N} P_n + \int_0^T \sum_{n=1}^{N} \beta||P_n||_0(1-c_{n,f})R_n(t)dt \tag{35a}$$

$$s.t \ \ 0 \leq P_n \leq P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{35b}$$

$$\frac{1}{\min\{C(t), \min_{m\in\Theta} R_m(t)\}} \leq \tau_{\max} \tag{35c}$$

$$\int_0^T C(t)dt \geq Q, \ \ \int_0^T R_n(t)dt \geq Q, n \in \Theta \tag{35d}$$

$$R_n(t) \leq b_n \tag{35e}$$

where $\Theta = \{n| \ \ ||P_n||_0(1-c_{n,f}) \neq 0\}$ and $C(t) = B\log_2\left(1 + \sum_{n=1}^{N} \frac{GP_n}{d_n(t)^\alpha\sigma^2}\right)$. The alternating optimization approach is also used here to jointly solve P_n and

$R_n(t)$ in an iterative way. For given P_n, the optimization of $R_n(t)$ is the same with the dynamic case. In what follows, we mainly focus on the optimization of P_n by solving

$$\min_{P_n} \sum_{n=1}^{N} k'_n P_n \tag{36a}$$

$$s.t \;\; 0 \leq P_n \leq P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{36b}$$

$$\frac{1}{C(t)} \leq \tau_{\max} \tag{36c}$$

$$\int_0^T C(t)dt \geq Q \tag{36d}$$

where $k'_n = T + \frac{1}{\log(1/\theta+1)} \frac{\beta(1-c_{n,f}) \int_0^T R_n(t)dt}{\theta+P_n^0}$ with P_n^0 being a is basis point of the Taylor expansion.

To solve (36), we consider two specific cases. If $\frac{T}{\tau_{\max}} \geq Q$, we have $C(t) = \frac{1}{\tau_{\max}}$. Then constraint (36d) is redundant. P_n can be found by solving

$$\min_{P_n} \sum_{n=1}^{N} k'_n P_n \tag{37a}$$

$$s.t \;\; 0 \leq P_n \leq P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{37b}$$

$$\sum_{n=1}^{N} \frac{GP_n}{d_n(t)^\alpha \sigma^2} \geq 2^{\frac{1}{\tau_{\max} B}} - 1 \tag{37c}$$

It is noted in (37), constraint (37c) should be satisfied for arbitrary t, which makes that problem (37) contains infinite constraints. To be feasible, we next sample the time period to generate certain discrete time points. If the time interval between two neighboring discrete points is small enough, the obtained solution can be approximately considered as an solution of (37). Denote the discrete time points $\{t_1, t_2, \cdots, t_M\}$, the power P_n can be efficient obtained by solving the following linear programming problem

$$\min_{P_{n,1}} \sum_{n=1}^{N} k'_n P_n \tag{38a}$$

$$s.t \;\; 0 \leq P_n \leq P_{n,\text{avg}} \quad \forall n \in \mathcal{N} \tag{38b}$$

$$\sum_{n=1}^{N} \frac{GP_n}{d_n(t_i)^\alpha \sigma^2} \geq 2^{\frac{1}{\tau_{\max} B}} - 1, \forall i \tag{38c}$$

If $\frac{T}{\tau_{\max}} < Q$, similar to the dynamic case, we represent the power as $P_n = P_{n,1} + P_{n,2}$ where $P_{n,1}$ is used to activate the constraint (38c) in (38). Then $P_{n,2}$ can be obtained by solving

$$\min_{P_{n,2}} \sum_{n=1}^{N} k'_n P_n \tag{39a}$$

$$s.t\ \ 0 \leq P_{n,2} \leq P_{n,\text{avg}} - P_{n,1} \quad \forall n \in \mathcal{N} \tag{39b}$$

$$\int_0^T B \log_2 \left(1 + \sum_{n=1}^{N} \frac{GP_{n,1}}{d_n(t)^\alpha \sigma^2} + \sum_{n=1}^{N} \frac{GP_{n,2}}{d_n(t)^\alpha \sigma^2} \right) dt \geq Q \tag{39c}$$

Problem (39) can be solved similarly to (36) by using Lagrangian method.

5 Numerical Results

In this section, we present some numerical results to illustrate the superiority of the proposed dynamic resource allocation. For the network shown in Fig. 1, it is assumed that we have 2 RRHs and the coordinates of them are $(-200, 100)$ and $(800, 100)$. The height of the RRH is 100 m. We assume that the periodic transmission time is 5 s and only consider the content service during $t \in (0, 5)$.

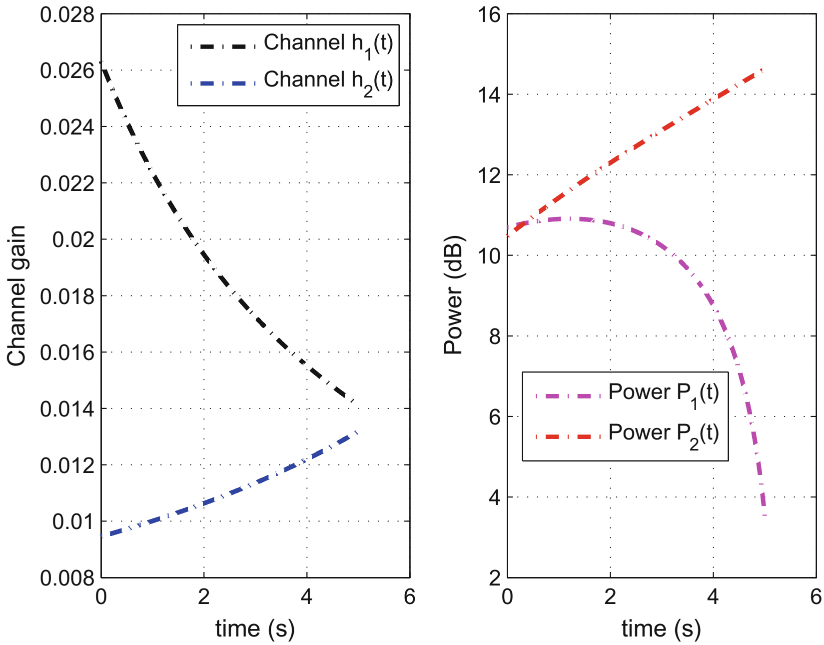

Fig. 2. Power varying with dynamic channel gains with $G = 2$ and average SNR $\frac{P_{1,\text{avg}}}{\sigma^2} = \frac{P_{2,\text{avg}}}{\sigma^2} = 10\,\text{dB}$ at $v_0 = 200\,\text{km/h}$.

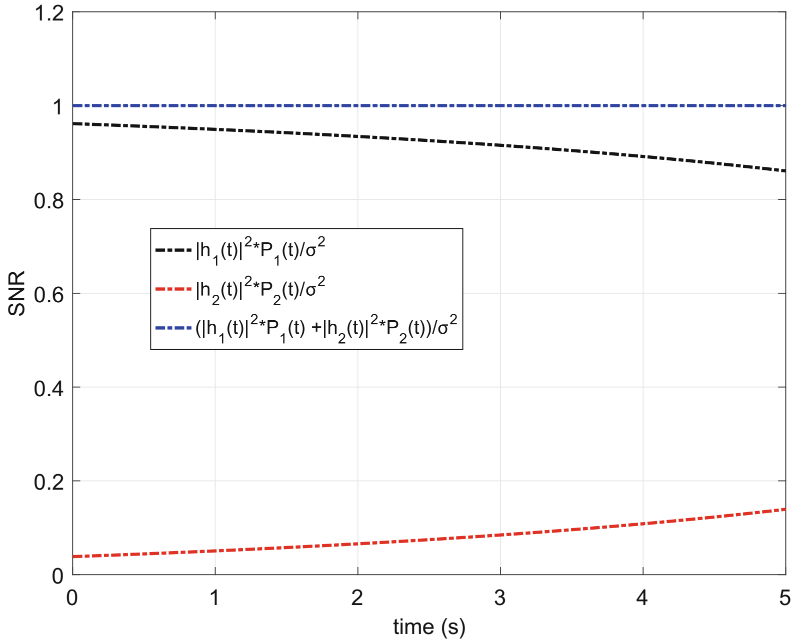

Fig. 3. SNR varying with respect to the time with $G = 10$, $\tau_{\max} = 0.005$ and average SNR $\frac{P_{1,\mathrm{avg}}}{\sigma^2} = \frac{P_{2,\mathrm{avg}}}{\sigma^2} = 10\,\mathrm{dB}$ at $v_0 = 200\,\mathrm{km/h}$.

The bandwidth of the frequency is 200 Hz. We compare the performance of dynamic power allocation with the constant power allocation scheme in terms of total cost defined in (5) with the ratio parameter $\beta = 2.8$. Regarding the cached contents at BBU, we assume that the number of contents is $F = 30$ and the size of all contents is normalized to 1. All contents are independently requested by the passengers in the train with equal probability $\frac{1}{F}$, which implies all the cached contents have the same popularity. For simplicity, we assume that the local storage size at both RRHs is 5. In specific, it is assumed that contents $\{1, 2, 3, 4, 5\}$ are cached at RRH 1 and contents $\{5, 6, 7, 8, 9\}$ are cached at RRH 2.

In Fig. 2, we demonstrate the dynamic power allocation with the time-varying channel gains $h_1(t)$ and $h_2(t)$ at $v_0 = 200\,\mathrm{km/h}$ with $G = 2$ and average signal to noise ratio (SNR) $\frac{P_{1,\mathrm{avg}}}{\sigma^2} = \frac{P_{2,\mathrm{avg}}}{\sigma^2} = 10\,\mathrm{dB}$. It is observed that as the time goes, the channel gain of $h_1(t)$ decreases while the channel gain of $h_2(t)$ increases due to the fact that the train gradually departs from RRH 1 and approaches RRH 2. To satisfy the delay constraint, the power is dynamically allocated over the time period $(0, T)$. The curves in Fig. 2 show that the values of the power $P_1(t)$ and $P_2(t)$ change in the opposite direction of the channels $h_1(t)$ and $h_2(t)$, respectively. That is, the RRH allocates more power to the time instant when the channel gain is low, which accords with the intuition.

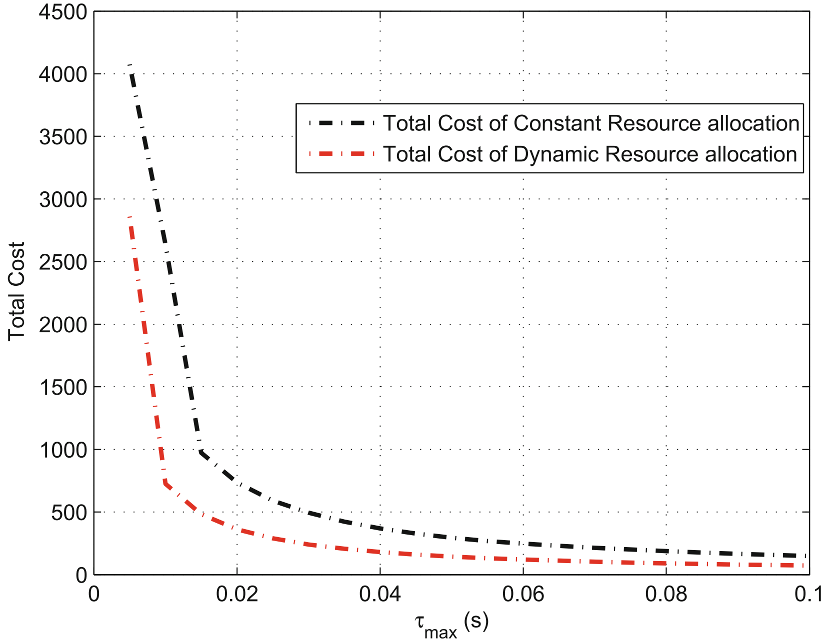

Fig. 4. Total cost comparison between the dynamic power allocation and constant power allocation with $G = 10$ and average SNR $\frac{P_{1,\text{avg}}}{\sigma^2} = \frac{P_{2,\text{avg}}}{\sigma^2} = 10\,\text{dB}$ at $v_0 = 200\,\text{km/h}$.

In Fig. 3, we illustrate the change of the instantaneous SNR over the time by considering a special case where the requested content is cached at both RRHs. With a slight abuse of the notation, we denote $\frac{|h_1(t)|^2 p_1(t)}{\sigma^2}$ and $\frac{|h_2(t)|^2 p_2(t)}{\sigma^2}$ as the instantaneous SNR at the time t over the channel links from RRH 1 to the train and from RRH 2 to the train, respectively. The final instantaneous SNR we obtain at the train after using the maximal ratio combiner is $\frac{|h_1(t)|^2 p_1(t)}{\sigma^2} + \frac{|h_2(t)|^2 p_2(t)}{\sigma^2}$. It is observed that although the values of SNR $\frac{|h_1(t)|^2 p_1(t)}{\sigma^2}$ and $\frac{|h_2(t)|^2 p_2(t)}{\sigma^2}$ vary over the time, their summation equals to a constant value, which is consistent with our analysis in Theorem 1, that is, to minimize the power cost, the constraint (12c) is active if $\frac{T}{\tau_{\max}} \geq Q$.

In Fig. 4, we compare the performance of the proposed dynamic resource allocation and the constant resource allocation with the change of $\tau_{\max}$. The curves show that the proposed dynamic resource allocation is significantly superior to the constant one. Constant resource allocation degrades the performance as it cannot fit the time-varying characteristic of the channel. It is also observed that as the increase of $\tau_{\max}$, the total cost decreases as less power is needed to meet the delay requirement.

6 Conclusion

This paper investigated the dynamic resources allocation in the HSR Fog-RAN system. By optimizing the instantaneous power allocation at the RRHs and the instantaneous content delivery rate over the backhaul links, we minimize the total power consumed at RRHs and backhaul links. We saw that the cached resource at RRHs can help reduce the power cost. Moreover, as dynamic resource allocation considers the time-varying characteristic of the channel, it can significantly outperform the constant resource allocation scheme.

Acknowledgement. This work was supported in part by the National Science Foundation China under Grant 61771345 and Grant 61831018, in part by the fund of the State Key Laboratory of Integrated Services Networks, Xidian University, under Project ISN19-01, and in part by Guangdong Province Key Research and Development Program Major Science and Technology Projects under Grant 2018B010115002.

References

1. Yan, D., Wang, R., Liu, E., Hou, Q.: ADMM-based robust beamforming design for downlink cloud radio access networks. IEEE Access **6**, 27912–27922 (2018)
2. Peng, M., Zhang, K.: Recent advances in fog radio access networks: performance analysis and radio resource allocation. IEEE Access **4**, 5003–5009 (2016)
3. Liu, J., Sheng, M., Quek, T.Q.S., Li, J.: D2D enhanced co-ordinated multipoint in cloud radio access networks. IEEE Trans. Wirel. Commun. **15**(6), 4248–4262 (2016)
4. Wang, R., Li, R., Wang, P., Liu, E.: Analysis and optimization of caching in fog radio access networks. IEEE Trans. Veh. Technol. **68**(8), 8279–8283 (2019)
5. Zhang, H., Zhu, L., Long, K., Li, X.: Energy efficient resource allocation in millimeter-wave-based fog radio access networks. In: 2nd URSI Atlantic Radio Science Meeting (AT-RASC), Meloneras, pp. 1–4 (2018)
6. Tao, M., Chen, E., Zhou, H., You, W.: Content-centric sparse multicast beamforming for cache-enabled cloud RAN. IEEE Trans. Wirel. Commun. **15**(9), 6118–6131 (2016)
7. Ai, B., et al.: Future railway services-oriented mobile communications network. IEEE Commun. Mag. **53**(10), 78–85 (2015)
8. Wu, J., Fan, P.: A survey on high mobility wirless communications: challenges, oppportunities and solutions. IEEE Access **4**, 450–476 (2016)
9. Muneer, P., Sameer, S.M.: Joint ML estimation of CFO and channel, and a low complexity turbo equalization technique for high mobility OFDMA uplinks. IEEE Trans. Wirel. Commun. **14**(7), 3642–3654 (2015)
10. Wang, J., Zhu, H., Gomes, N.J.: Distributed antenna systems for mobile communications in high speed trains. IEEE J. Sel. Areas Commun. **30**(4), 675–683 (2012)
11. Li, T., Xiong, K., Fan, P., Letaief, K.B.: Service-oriented power allocation for high-speed railway wireless communications. IEEE Access **5**, 8343–8356 (2017)
12. Zhang, C., Fan, P., Xiong, K., Fan, P.: Optimal power allocation with delay constraint for signal transmission from a moving train to base stations in high-speed railway scenarios. IEEE Trans. Veh. Technol. **64**(12), 5775–5788 (2015)
13. Liu, X., Qiao, D.: Location-fair beamforming for high speed railway communication systems. IEEE Access **6**, 28632–28642 (2018)