

RESEARCH

Open Access

Utility-based resource allocation in high-speed railway wireless networks

Shengfeng Xu, Gang Zhu, Chao Shen^{*}, Yan Lei and Zhangdui Zhong

Abstract

In this paper, we investigate the utility-based resource allocation problem at a base station in high-speed railway (HSR) wireless networks, jointly taking into account the power allocation along the time and the packet allocation among services. The problem to maximize the total utility under the average power constraint is formulated as a mixed-integer non-linear programming (MINLP) problem. Through the integer constraint relaxation, the MINLP problem can be simplified into a convex optimization problem. The detailed analysis reveals that the relaxed problem can be equivalently decomposed into power allocation problem along the time and packet allocation problem among services, which can reduce the problem size. When the optimality of the relaxed problem is achieved, the power allocation along the time and the packet allocation along the time for each service are both proportionally fair. Since the integer relaxation causes a non-integer solution not implementable in practice, a greedy algorithm is proposed to obtain a near-optimal integer solution of the MINLP problem. Finally, the performance of the proposed algorithm is analyzed by simulations under realistic conditions for HSR wireless networks.

Keywords: Utility-based resource allocation; Power allocation; High-speed railway wireless networks

1 Introduction

In recent years, high-speed trains are being deployed rapidly all over the world, serving as a fast, convenient and green public transportation system. The passengers on the train have an increasingly high demand on multimedia services. Meanwhile, more and more data related with the railway controlling information needs to be transmitted between the train and the ground in order to guarantee that the train is moving safely. However, the dominant wireless communication system, GSM for railway (GSM-R), is specifically designed for train control rather than passenger communications and can only support a maximum data rate of 200 kbps [1]. Therefore, it is crucial to investigate the new network architectures and transmission technologies to satisfy the needs of high-speed railway (HSR) wireless communication.

On one hand, some broadband wireless communication systems for HSR have been developed, and a review of network architectures has been presented in [2]. The relay-assisted HSR network architecture proposed in [1]

and [3] becomes a promising architecture for future HSR wireless communication [4] and was considered to be a better choice than direct transmission in case of large penetration loss [5]. On the other hand, there have been some recent works to improve the transmission performance in HSR wireless networks. The radio-over-fiber (ROF) technology for HSR wireless networks was proposed in [6], which can reduce the number of handoffs and increase throughput effectively. Multi-input multi-output (MIMO) technology was introduced in order to improve the throughput performance of the HSR wireless networks [7]. However, further investigation is needed to improve the system performance when considering the multimedia service transmission in HSR wireless networks.

The broadband wireless communications for HSR should have the function of transmitting multimedia services and enable a variety of applications such as entertainment and safety. There are many types of HSR services with different rate requirements and priorities [8]. In particular, the HSR services are classified into four main categories [9], i.e., pure passenger internet, passenger comfort services, security-related applications, and cost-saving applications. The provisioning of wireless data

^{*}Correspondence: shenchao@bjtu.edu.cn
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China

services in a moving train is a technologically demanding challenge due to the fast-varying wireless channel and heterogeneous service requirements. Since the network resources are limited in HSR wireless networks, in order to take full advantage of them, the resource allocation to the heterogeneous services should be considered. In addition, the data transmission rate is highly determined by the transmit power and the distance between the base station and the train, which makes it feasible to implement power allocation along the time in HSR wireless networks [10]. Therefore, when transmitting multiple services from the base station to the moving train, it is necessary to consider not only the power allocation along the time but also the resource allocation among the services. To the best of our knowledge, the resource allocation problem in HSR wireless networks which can jointly consider resource allocation among services and power allocation along time is still an open problem.

In this paper, we investigate the utility-based resource allocation problem at a base station in HSR wireless networks, jointly taking into account the power allocation along time and the packet allocation among services. This optimization problem is formulated as a mixed-integer non-linear programming (MINLP) problem with the objective to maximize the total utility at a base station under the average power constraint. The MINLP problem is in general non-deterministic polynomial-time (NP)-hard [11], so the integer constraint relaxation is adopted to obtain some engineering insights to solve it effectively. Since the size of the relaxed problem is also large, to further reduce the computational complexity, we develop theoretical insights from the nature of optimal solutions and carry out the problem transformation. It is shown that the relaxed problem can be equivalently decomposed into two subproblems, i.e., power allocation problem along time and packet allocation problem among services. The former can be solved by bisection search method with low complexity and then the closed-form solution of the latter can be computed by the obtained power allocation. Moreover, the optimal solution of the relaxed problem can achieve not only the proportional fair power allocation along time but also the proportional fair packet allocation along the time for each service. Since the integer relaxation causes a non-integer solution not implementable in practice, a greedy algorithm is proposed to obtain a near-optimal integer solution of the MINLP problem. Finally, we present the analysis of the proposed algorithm performance by simulations under realistic conditions for HSR wireless networks.

The rest of the paper is organized as follows. In Section 2, we review previous works in the literature which are the most relevant to our work. Section 3 describes the system model. The problem formulation is provided in Section 4. The problem transformation and

a bisection search method are presented in Section 5. In Section 6, a greedy algorithm is proposed to obtain the integer solution. Some numerical results and discussions are shown in Section 7. Finally, conclusions are drawn in Section 8.

2 Related work

Resource allocation, as a critical part of radio resource management, plays an important role in improving the throughput and ensuring service fairness in HSR wireless networks. Among the works related to resource allocation in HSR wireless networks, [12] proposed a scheduling and resource allocation mechanism in HSR networks with a cell array architecture, which can maximize the service rate based on periodical signal quality changes. Considering the intermittent network connectivity in a cellular/infostation integrated HSR network, [13] and [14] investigated the resource allocation problem for transmitting multiple services from the ground to the train. Different algorithms were proposed since they involved different optimization objectives, while the transmit power was assumed to be constant along time. In a relay-assisted HSR network, [15] studied delay-aware fair resource allocation problem with heterogeneous packet arrivals and delay requirements for the HSR services. However, the above works did not make full use of three unique features of HSR wireless networks [16], i.e., the deterministic moving direction, relatively steady moving speed, and the accurate train location information.

The data transmission rate is highly determined by the transmit power and the distance between the base station and the train; thus, the power allocation along the time has a large influence on transmission performance in HSR wireless networks. Four power allocation schemes were presented in [10] for different design objectives. Among these schemes, to achieve a tradeoff between the power efficiency and the fairness along the time, a proportional fair power allocation scheme was proposed and its ϵ -optimal solution was presented. However, the schemes proposed in [10] do not consider cross-layer design. The forecast channels in the HSR scenario were exploited in [17] to minimize the total transmit power along the time under the packet deadline constraint, while only one type of service was considered. Under the average power constraint in HSR wireless network, [18] investigated a joint admission control and resource allocation problem, which aims at maximizing the system utility while stabilizing all transmission queues.

An effective cross-layer resource allocation is necessary to improve the performance of wireless communication system. The cross-layer resource allocation problem was formulated as a utility-based resource allocation problem in [19], where the utility function is used to define the relationship between resource allocated to application and

the satisfaction brought by this application. With the help of the utility function, we present a utility-based resource allocation framework in HSR wireless networks, which can jointly consider the power allocation along the time and packet allocation among services.

3 System model

In this paper, a two-hop HSR wireless network architecture is considered, as shown in Figure 1. The base station (BS) deployed along the rail line is connected to the backbone network via a wireline link. The relay station (RS) with powerful antennas installed on the top of the train is used for communicating with the BS. The RS is further connected to the access point (AP) which can be accessed by the users inside the train. Thus, there is a two-hop wireless link, consisting of the BS-RS link and the AP-users link, which has several advantages. Firstly, it is the RS not each user to implement the handover procedure, which can achieve better handover performance and reduce the drop-off rate significantly [4]. Secondly, with this two-hop wireless link, signals do not penetrate into the carriage; thus, the large penetration loss can be dramatically reduced [1]. Finally, since the users are nearly stationary with respect to the AP, the AP-users link can provide a stable and high-speed wireless data transmission [13].

We consider the downlink data transmission in this two-hop architecture. The AP-users link inside the carriage can provide a large data transmission rate by using wireless local area network (WLAN) technologies, while the BS-RS link suffers from the fast-varying wireless channel and may become the bottleneck in this architecture. Therefore, the transmission in the BS-RS link will be

mainly considered in this paper with the assumption for the downlink data always being successfully received when delivered to RS.

3.1 Time-distance mapping

We consider a train that travels at a constant speed v through a single cell with radius R . The total time the train spends is $T_{\text{all}} = 2R/v$, which is divided into slots of equal duration T_s . Then, the total number of slots is given by $T = T_{\text{all}}/T_s$, where we assume that T_{all} can be exactly divided by T_s and T is even. Without loss of generality, we assume that the train goes into and out of the cell coverage at slot 0 and slot T , respectively. The traveled distance until slot t is given by $s(t) = vtT_s$. A time-distance mapping function $d(t)$ is defined as the distance between the BS and RS at slot t [20], i.e., $d(t) : [0, T] \rightarrow [d_0, d_{\text{max}}]$, where $d_{\text{max}} = \sqrt{R^2 + d_0^2}$ and d_0 is the distance between the BS and the rail line as shown in Figure 1. The mapping function $d(t)$ can be expressed by

$$d(t) = \sqrt{|s(t) - R|^2 + d_0^2}, \quad t \in [0, T], \quad (1)$$

where we assume that the distance $d(t)$ does not change within slot t since T_s is very small. There are two inherent properties about the mapping function $d(t)$: (i) the train is closest to the BS at slot $\frac{T}{2}$, i.e., $d(\frac{T}{2}) = d_0$. (ii) For any slot $t \in [0, T]$, we have $d(t) = d(T - t)$ due to the distance symmetry.

3.2 BS-RS link capacity

In HSR wireless networks, the channel condition in BS-RS link is fast varying due to the large-scale fading and

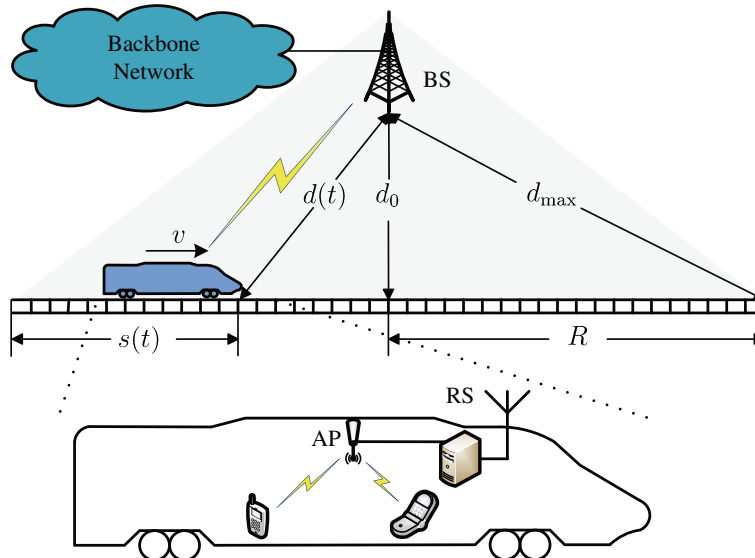


Figure 1 System model.

small-scale fading [21]. Confirmed by engineering measurements [22,23], the line-of-sight path in the BS-RS link is available at most time, and the effect of large-scale fading is more obvious than that of small-scale fading. Therefore, we ignore the small-scale fading and assume that the channel condition variation results only from the fast-varying distance between the BS and RS. As shown in [10], power allocation along the travel time plays a key role in improving the performance of HSR wireless networks. We denote $P(t)$ as the transmit power of the BS at slot t , which is limited by the average value P_{av} . The assumption of average power constraint is the same as that used in [10]. Given $P(t)$ and $d(t)$, the received signal-to-noise ratio (SNR) by RS at slot t can be expressed by

$$\text{SNR} = \frac{P(t)}{WN_0d^\alpha(t)} = \frac{P(t)}{N(t)}, \quad (2)$$

where $N(t) = WN_0d^\alpha(t)$, W is the system bandwidth, N_0 is the noise power spectral density, and α is the path loss exponent. The instantaneous transmission rate in the BS-RS link at slot t is

$$R(t) = W \log_2 \left(1 + \frac{P(t)}{N(t)} \right) \text{ bits/s}. \quad (3)$$

For the communications in the BS-RS link, we consider the frame structure proposed in [24] which is specifically designed for high-speed trains with a speed up to 360 km/h. Suppose that a packet is the fundamental unit of transmission with equal size L bits and the overhead of packetization can be ignored; hence, the link capacity $C(t)$ at slot t can be denoted as the maximum number of packets [18], which can be expressed by

$$C(t) = \lfloor \tilde{C}(t) \rfloor = \lfloor R(t)T_s/L \rfloor, \quad (4)$$

where $\tilde{C}(t) = R(t)T_s/L$ and $\lfloor x \rfloor = \max\{n \in \mathbb{Z} | n \leq x\}$.

3.3 Utility-based resource allocation

Assume that there are K types of elastic services with infinite packets to be transmitted from the BS to RS, and the service set is denoted by $\mathcal{K} \triangleq \{1, \dots, K\}$. The elastic services, e.g., file transfer, electronic mail, and web surfing, can tolerate long delays. To allocate the network resources based on the type of elastic services, utility-based resource allocation can be employed [19]. Figure 2 gives an example of the relationship between utility and data rate for three different types of services. For any service, the utility grows as the allocated data rate increases. On one hand, equal data rate allocation does not provide equal utility, which is interpreted as equal service satisfaction. On the other hand, to achieve equal utility, the different data rates should be allocated to the services according to their types, which results in utilizing the network resources more efficiently [25]. Thus, we consider utility-based resource allocation instead of rate-based resource allocation in this paper.

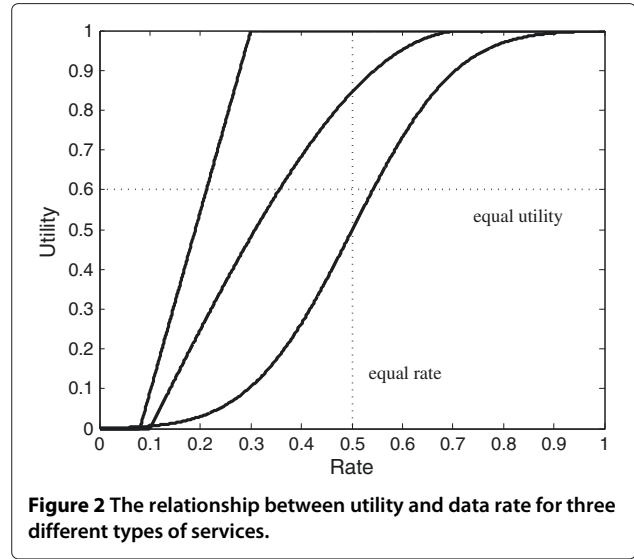


Figure 2 The relationship between utility and data rate for three different types of services.

Suppose, in general, that service k maintains an increasing and concave function $U_k(v_k)$ as its utility function, which indicates a service's degree of satisfaction on the allocated v_k packets. Instead of maximizing network throughput performance, our objective is to maximize the overall network utility, which is the summation of all services' utility functions. From [26] and [27], the utility function for service k can be defined as

$$U_k(v_k) = \begin{cases} \omega_k \frac{v_k^{1-\alpha}}{1-\alpha}, & \alpha \geq 0, \alpha \neq 1, \\ \omega_k \ln(v_k), & \alpha = 1, \end{cases} \quad (5a)$$

$$(5b)$$

where α is a parameter dictating the shape of the utility function, and ω_k represents the weight of service k . When the weights of all services are the same, e.g., $\omega_k = 1$, the optimization objective can be specialized into the following four cases according to different values of α :

- (i) throughput maximization, $U_k(v_k) = v_k$ when $\alpha = 0$;
- (ii) proportional fairness, $U_k(v_k) = \ln(v_k)$ when $\alpha = 1$;
- (iii) $(1, \alpha)$ -proportional fairness, $U_k(v_k) = \frac{v_k^{1-\alpha}}{1-\alpha}$ when $\alpha \in (0, 1) \cup (1, \infty)$;
- (iv) max-min fairness, $U_k(v_k) = \lim_{\alpha \rightarrow \infty} \frac{v_k^{1-\alpha}}{1-\alpha}$ when $\alpha \rightarrow \infty$.

Similar to [28], we choose $U_k(v_k) = \omega_k \ln(v_k)$ to obtain weighted proportionally fair resource allocation, where ω_k is assumed to be an integer for $k \in \mathcal{K}$ in this paper.

4 Problem formulation

In this section, we develop a mathematical formulation of the optimal resource allocation problem in HSR wireless networks. Let $\mathbf{v}(t) = [v_1(t), \dots, v_K(t)]^T$ represent the packet allocation vector at slot t , where $v_k(t)$ denotes the

number of packets allocated to service k at slot t . The optimization problem consists in maximizing the BS utility to find the optimal power allocation along the time and the optimal packet allocation among the services, and two necessary constraints are added: (i) The BS has an average power constraint along the time. (ii) The total number of allocated packets is no more than the link capacity at any slot. Thus, the utility-based resource allocation optimization problem is formulated as

$$(P1) \text{ maximize } \sum_{t=0}^T \sum_{k \in \mathcal{K}} \omega_k \ln(v_k(t)) \quad (6a)$$

$$\text{subject to } \frac{1}{T+1} \sum_{t=0}^T P(t) \leq P_{av}, \quad (6b)$$

$$0 \leq \sum_{k \in \mathcal{K}} v_k(t) \leq C(t), \quad \forall t \in [0, T], \quad (6c)$$

$$\text{variables } v_k(t) \in \mathbb{N}, P(t) \geq 0, \forall k \in \mathcal{K}, t \in [0, T]. \quad (6d)$$

Problem P1 is a MINLP problem, including $T+1$ continuous variables $P(t)$ and $K(T+1)$ integer variables $v_k(t)$, which is in general NP-hard [11]. The main difficulty of analyzing problem P1 comes from the integer nature of $v_k(t)$. To significantly improve the computational efficiency and obtain some engineering insights for solving the MINLP problem, we adopt integer constraint relaxation for problem P1, where $\mu_k(t) \in \mathbb{Q}^+$ substitutes the constraint $v_k(t) \in \mathbb{N}$ and $\tilde{C}(t)$ substitutes $C(t)$.

As a result, problem P1 is simplified into a relaxed problem P2 as follows:

$$(P2) \text{ maximize } \sum_{t=0}^T \sum_{k \in \mathcal{K}} \omega_k \ln(\mu_k(t)) \quad (7a)$$

$$\text{subject to } \frac{1}{T+1} \sum_{t=0}^T P(t) \leq P_{av}, \quad (7b)$$

$$0 \leq \sum_{k \in \mathcal{K}} \mu_k(t) \leq \tilde{C}(t), \quad \forall t \in [0, T], \quad (7c)$$

$$\text{variables } \mu_k(t) \geq 0, P(t) \geq 0, \forall k \in \mathcal{K}, t \in [0, T]. \quad (7d)$$

Notice that the optimal solution of problem P2 provides an upper bound to that of problem P1 since the constraints in problem P2 are looser than those in problem P1. There are totally $(K+1)(T+1)$ continuous variables in problem P2, where T is typically of the order of 10^4 to 10^5 . Standard convex optimization tools such as CVX [29] can be employed to solve P2; however, the computational complexity is very high due to the large size of the problem [30]. In order to obtain a low-complexity and effective algorithm for problem P2, we carry out the problem transformation in Section 5. Before we present the

solution for problem P2, we consider the problem decomposition to determine some characteristics which will be useful in understanding the structure of problem P2 better.

By decoupling of the optimization variables in constraint (7c), problem P2 can be decomposed into two subproblems: (i) power allocation along the time (PAT): how to implement power allocation along the time under the average power constraint at the BS. (ii) Packet allocation among services (PAS): how to allocate resources to multiple services at each slot by the given power allocation. Next, we will discuss these two subproblems separately.

4.1 PAT Problem

In this subsection, we investigate the power allocation problem along the time under the average power constraint at the BS. Since the channel state in BS-RS link is time-varying, to achieve different optimization objectives, four power allocation schemes have been proposed in [10].

4.1.1 Constant power allocation

The most straightforward scheme is the constant power allocation, where BS maintains a constant transmit power at all times, i.e., $P(t) = P_{av}$. Thus,

$$\tilde{C}(t) = \frac{T_s W}{L} \log_2 \left(1 + \frac{P_{av}}{N(t)} \right), \quad \forall t \in [0, T]. \quad (8)$$

4.1.2 Channel inversion power allocation

It tries to maintain a constant link capacity $\tilde{C}(t)$ at the BS all the time. Therefore, based on (3) and (4), the ratio of $P(t)$ to $N(t)$ is a constant for all slots. Without loss of generality, we suppose that $P(t) = k_0 N(t)$. And then by solving $\sum_{t=0}^T P(t) = (T+1)P_{av}$, we have $k_0 = \frac{(T+1)P_{av}}{\sum_{t=0}^T N(t)}$ and

$$\tilde{C}(t) = \frac{T_s W}{L} \log_2 (1 + k_0), \quad \forall t \in [0, T]. \quad (9)$$

4.1.3 Water-filling power allocation

To maximize the total link capacity at the BS, we formulate the following optimization problem:

$$\text{maximize } \sum_{t=0}^T \tilde{C}(t) \quad (10a)$$

$$\text{subject to } \frac{1}{T+1} \sum_{t=0}^T P(t) \leq P_{av}, \quad (10b)$$

$$\text{variables } P(t) \geq 0, t \in [0, T], \quad (10c)$$

whose solution can be obtained by a water-filling scheme [10].

4.1.4 Proportional fair power allocation

To achieve a tradeoff between the total link capacity and the fairness along the time, a proportional fair power allocation optimization problem is formulated as

$$\text{maximize } \sum_{t=0}^T \ln \tilde{C}(t) \quad (11a)$$

$$\text{subject to } \frac{1}{T+1} \sum_{t=0}^T P(t) \leq P_{av}, \quad (11b)$$

$$\text{variables } P(t) \geq 0, t \in [0, T], \quad (11c)$$

whose ϵ -optimal solution can be obtained by the proposed algorithm in [10].

4.2 PAS problem

In this subsection, the packet allocation problem among the services is studied under the link capacity constraint by fixing the power allocation at all slots, which can be obtained according to the power allocation schemes in the above subsection. This setup is less complicated compared to our more general model, and its solution can provide us with some insights for solving problem P2.

Given the fixed power allocation $P(t)$ at any slot t , the link capacity $\tilde{C}(t)$ can be computed by (4). Problem P2 can be divided into $T+1$ packet allocation problems, and the problem at any slot $t \in [0, T]$ can be given by

$$\text{maximize } \sum_{k \in \mathcal{K}} \omega_k \ln(\mu_k(t)) \quad (12a)$$

$$\text{subject to } 0 \leq \sum_{k \in \mathcal{K}} \mu_k(t) \leq \tilde{C}(t), \quad (12b)$$

$$\text{variables } \mu_k(t) \geq 0, \forall k \in \mathcal{K}, \quad (12c)$$

which is a convex optimization problem. By applying the Karush-Kuhn-Tucker (KKT) conditions, we obtain the closed-form optimal solution of (12) and an important structural characteristic in the following lemma.

Lemma 1. *For the optimal solution vector $\mu^*(t) = [\mu_1^*(t), \mu_2^*(t), \dots, \mu_K^*(t)]^T$ at any slot $t \in [0, T]$, there must be the case that*

$$\frac{\mu_i^*(t)}{\omega_i} = \frac{\mu_j^*(t)}{\omega_j}, \forall i, j \in \mathcal{K}, \quad (13)$$

and the closed-form optimal solution of (12) is

$$\mu_k^*(t) = \frac{\omega_k \tilde{C}(t)}{\sum_k \omega_k}, \forall k \in \mathcal{K}. \quad (14)$$

Proof. The proof of Lemma 1 is provided in Appendix 1. \square

Based on (13) in Lemma 1, there is an inherent relationship among the optimal solutions at any slot. More

resources are allocated to the service with a larger weight. In particular, for any service k , the allocated resource $\mu_k^*(t)$ at slot t is proportional to its weight ω_k . Notice that the service weights are assumed to be integer so as to simplify the analysis in the sequel, since the optimal solutions depend on the relative size rather than on the absolute size of the weights. If the service weights are non-integer, the corresponding integer weights can be obtained after being multiplied by the same factor. For simplicity of expression in the sequel, we introduce a virtual variable at any slot $t \in [0, T]$, which is given by

$$x(t) = \frac{\tilde{C}(t)}{\sum_k \omega_k}. \quad (15)$$

From (15), we can see that $x(t)$ is only determined by the power allocation $P(t)$ at any slot t . By plugging (15) into (14), the optimal solution of (12) can be rewritten by

$$\mu_k^*(t) = \omega_k x(t), \forall k \in \mathcal{K}, t \in [0, T]. \quad (16)$$

Thus, if the optimal power allocation solution in problem P2 can be obtained, then the optimal $x^*(t)$ and the optimal packet allocation solution in problem P2 can be calculated by (15) and (16), respectively.

5 Problem transformation

In this section, to reduce the computational complexity, we consider the problem transformation for problem P2, where the number of the optimization variables dramatically decreases from $(K+1)(T+1)$ to $\frac{T}{2} + 1$. Based on the bisection search method, a greedy algorithm with low complexity is proposed for solving problem P2.

Firstly, since Lemma 1 provides the necessary condition for the optimal solutions of (12) and problem P2, based on (16), the resource allocation variables $\mu_k(t)$ at slot t can be substituted by one single variable $x(t)$. Then, the objective function of problem P2 can be simplified into

$$\begin{aligned} \sum_{t=0}^T \sum_{k \in \mathcal{K}} \omega_k \ln(\mu_k(t)) &= \sum_{t=0}^T \sum_{k \in \mathcal{K}} (\omega_k (\ln(\omega_k) + \ln(x(t)))) \\ &= \alpha + \gamma \sum_{t=0}^T \ln(x(t)), \end{aligned} \quad (17)$$

where $\alpha = (T+1) \sum_k (\omega_k \ln(\omega_k))$ and $\gamma = \sum_k \omega_k$ are both constant. Similarly, for the constraint (7c), $\sum_k \mu_k(t) = \gamma x(t)$. Thus, problem P2 can be transformed into

$$(P3) \text{ maximize } \sum_{t=0}^T \ln(x(t)) \quad (18a)$$

$$\text{subject to } \frac{1}{T+1} \sum_{t=0}^T P(t) \leq P_{av}, \quad (18b)$$

$$\gamma x(t) \leq \tilde{C}(t) = \frac{T_s W}{L} \log_2 \left(1 + \frac{P(t)}{N(t)} \right), \quad \forall t \in [0, T], \quad (18c)$$

$$\text{variables } x(t) \geq 0, P(t) \geq 0, \forall k \in \mathcal{K}, t \in [0, T]. \quad (18d)$$

Lemma 2. Suppose that the optimal solution of problem P3 exists, the optimal solution provides proportionally fair resource allocation along the time for each service.

Proof. The proof of Lemma 2 is provided in Appendix 2. \square

After the problem transformation, the total number of variables decreases from $(K+1)(T+1)$ to $2(T+1)$, and hence, the computational complexity is dramatically reduced when K is large. Based on the investigation on problem P3, the total number of variables can be further reduced to $T+1$ as shown below.

It is easy to show that at the optimality of problem P3, the constraints (18b) and (18c) are both tight; otherwise, one can increase the value of $x(t)$ and $P(t)$, such that the objective function can be further maximized. Thus, we have

$$\frac{1}{T+1} \sum_{t=0}^T P(t) = P_{av}, \quad (19)$$

and

$$\gamma x(t) = \tilde{C}(t) = \frac{T_s W}{L \ln 2} \ln \left(1 + \frac{P(t)}{N(t)} \right), \quad \forall t \in [0, T]. \quad (20)$$

Based on (20), there exists a one-to-one correspondence established between $x(t)$ and $P(t)$, which is expressed by

$$x(t) = \eta \ln \left(1 + \frac{P(t)}{N(t)} \right), \quad (21)$$

where $\eta = \frac{T_s W}{\gamma L \ln 2}$.

Thus, plugging (19) and (21) into problem P3 yields

$$(P4) \text{ maximize } \sum_{t=0}^T \ln \left(\eta \ln \left(1 + \frac{P(t)}{N(t)} \right) \right) \quad (22a)$$

$$\text{subject to } \sum_{t=0}^T P(t) = (T+1)P_{av}, \quad (22b)$$

$$\text{variables } P(t) \geq 0, \forall t \in [0, T]. \quad (22c)$$

Lemma 3. The optimal solution of problem P4 is the same as that of the proportional fair power allocation (PFPA) problem.

Proof. The proof of Lemma 3 is provided in Appendix 3. \square

Based on Lemma 3, the optimal solution of problem P4 provides proportionally fair power allocation along the time. Furthermore, we can observe that problem P2 can be equivalently decomposed into two subproblems: problem P4 and PAS problem, which are corresponding to power allocation problem along the time and packet allocation problem among the services, respectively. The explanation of the equivalence problem is given as follows. On the one hand, problem P2 can be decomposed into $T+1$ PAS problems for the given power allocation along the time. Based on the Lemma 1, if the optimal power allocation $P^*(t)$ is given, the optimal packet allocation $\mu_k^*(t)$ of problem P2 can be obtained by solving the PAS problems. On the other hand, based on (15) and (16), the virtual variable $x(t)$ builds a bridge between $\mu_k(t)$ and $P(t)$. By variable substitution, problem P2 can be equivalently transformed into problem P3, which is further simplified into problem P4 based on the necessary optimality conditions. These equivalent transforms can guarantee that the optimal power allocation $P^*(t)$ of problem P4 is the same as that of problem P2 at any slot t . Based on the above two points, we can conclude that problem P2 can be equivalently decomposed into problem P4 and PAS problem. Thus, in order to obtain the optimal solutions of problem P2, we can solve problem P4 at first, and then by using the power allocation results, the packet allocation solution can be obtained by using (16) and (21).

To solve problem P4 effectively, the following lemma allows us to further reduce the computational complexity based on the distance symmetry at the base station, which has been mentioned in Section 3.1.

Lemma 4. In the optimal solution vector $\mathbf{P}^* = [P^*(0), \dots, P^*(T)]$, there exists a symmetry on the optimal solution, i.e., $P^*(t) = P^*(T-t)$, $\forall t \in [0, T]$.

Proof. The proof of Lemma 4 is provided in Appendix 4. \square

As a consequence of Lemma 4, problem P4 can be simplified into the power allocation problem from slot 0 to slot $\frac{T}{2}$, which is labeled as P5.

$$(P5) \text{ maximize } \sum_{t=0}^{T/2} g(P(t)) \quad (23a)$$

$$\text{subject to } \sum_{t=0}^{T/2} P(t) = \left(\frac{T}{2} + 1\right) P_{\text{av}}, \quad (23b)$$

$$\text{variables } P(t) \geq 0, \forall t \in [0, T/2], \quad (23c)$$

where $g(P(t)) = \ln\left(\eta \ln\left(1 + \frac{P(t)}{N(t)}\right)\right)$ and the total number of variables decrease nearly half from $T + 1$ to $\frac{T}{2} + 1$.

Problem P5 is a convex optimization problem, which can be solved by CVX [29]. In addition, since problem P5 has a similar structure to the PFPA problem, the proposed algorithm in [10] can be used to find the ϵ -optimal solution of problem P5. However, the Lambert W function was introduced in the proposed algorithm, resulting in the high computing time. In this paper, the bisection search method is employed to reduce the computing time of searching the optimal solution.

Specifically, using the standard optimization technique, the corresponding Lagrangian function is obtained as

$$\begin{aligned} L(\{P(t)\}, \lambda) &= \sum_{t=0}^{T/2} g(P(t)) - \lambda \left(\sum_{t=0}^{T/2} P(t) - \left(\frac{T}{2} + 1\right) P_{\text{av}} \right) \\ &= \sum_{t=0}^{T/2} \left[\ln\left(\eta \ln\left(1 + \frac{P(t)}{N(t)}\right)\right) - \lambda P(t) + \lambda P_{\text{av}} \right]. \end{aligned} \quad (24)$$

Based on the KKT conditions, we have

$$\frac{\partial L(\{P(t)\}, \lambda)}{\partial P(t)} = \frac{1}{\eta \ln\left(1 + \frac{P(t)}{N(t)}\right)} \frac{\eta}{P(t) + N(t)} - \lambda = 0, \quad (25)$$

which can be rewritten by

$$\ln\left(1 + \frac{P(t)}{N(t)}\right) (P(t) + N(t)) = \frac{1}{\lambda}. \quad (26)$$

Let $f(P(t)) = \ln\left(1 + \frac{P(t)}{N(t)}\right) (P(t) + N(t))$, which is a monotonically increasing function of $P(t)$ at any slot t . Let $\beta = \frac{1}{\lambda}$, then (26) is equal to $f(P(t)) = \beta$. Due to the monotonicity of $f(P(t))$, the bisection search method can be used to find $P(t)$, satisfying $f(P(t)) = \beta$ for a given β at each slot t . In addition, for any slot t , $P(t) = f^{-1}(\beta)$ is also a monotonically increasing function of β . Thus, to satisfy the average power constraint (23b), the bisection search method can also be used to find the optimal β^* .

The specific steps of the bisection search method is provided in Algorithm 1. The search regions of $P(t)$ and β should be initialized based on their maximums and minimums. At first, it is easy to verify that the maximum and minimum of $P(t)$ at each slot t can be set as $P_{\text{max}} = \left(\frac{T}{2} + 1\right) P_{\text{av}}$ and $P_{\text{min}} = 0$, respectively. And then the

maximum and minimum of β can be obtained by the following lemma:

Lemma 5. *Based on the equality $f(P(t)) = \beta$, the maximum of β can be obtained when $P(t) = P_{\text{max}}$ and $t = 0$ in function $f(P(t))$, i.e., $\beta_{\text{max}} = f(P_{\text{max}})|_{t=0}$ and the minimum of β can be set as $\beta_{\text{min}} = 0$.*

Proof. The proof of Lemma 5 is provided in Appendix 5. \square

The Algorithm 1 consists of two loops to find the optimal power allocation. The outer loop is used for the bisection search of β , and the inner loop is used to solve $f(P(t)) = \beta$ for a given β . In addition, the convergence of Algorithm 1 is ensured by the bisection search, where $\epsilon_{\Delta P}$ and $\epsilon_{\Delta \beta}$ are small constants to control the convergence accuracy.

Algorithm 1 Bisection search method

```

1: Initialize  $\beta_{\text{min}}, \beta_{\text{max}}, P_{\text{min}}, P_{\text{max}}, \epsilon_{\Delta \beta} = 10^{-6},$ 
    $\epsilon_{\Delta P} = 10^{-6}$ .
2: while  $(\beta_{\text{max}} - \beta_{\text{min}}) \geq \epsilon_{\Delta \beta}$  do
3:    $\beta = (\beta_{\text{max}} + \beta_{\text{min}})/2$ ;
4:   for  $t = 0$  to  $\frac{T}{2}$  do
5:     while  $P_{\text{max}} - P_{\text{min}} \geq \epsilon_{\Delta P}$  do
6:        $P(t) \leftarrow (P_{\text{max}} + P_{\text{min}})/2$ ;
7:       Calculate  $f(P(t))$ ;
8:       if  $f(P(t)) \geq \beta$  then
9:          $P_{\text{max}} \leftarrow P(t)$ ;
10:      else
11:         $P_{\text{min}} \leftarrow P(t)$ ;
12:      end if
13:    end while
14:  end for
15:  if  $\sum_{t=0}^{\frac{T}{2}} P(t) \geq \left(\frac{T}{2} + 1\right) P_{\text{av}}$  then
16:     $\beta_{\text{max}} \leftarrow \beta$ ;
17:  else
18:     $\beta_{\text{min}} \leftarrow \beta$ ;
19:  end if
20: end while

```

6 The greedy algorithm

In the above section, we obtain the power allocation results using Algorithm 1, and then the solution of virtual variable $x^*(t)$ for any slot t can be computed by using (21). Since the solution $x^*(t)$ is continuous, based on (16), it can not ensure that the packet allocation solution $\mu_k^*(t)$ is an integer for any service k and slot t . As a result, the solution $\mu_k^*(t)$ is not valid for practical purposes since the number of allocated packets must take integer value. As an alternative, an integer solution $v_k^*(t)$ in problem P1 can

be obtained if the non-integer solution $\mu_k^*(t)$ is rounded to the nearest integer, but there is no guarantee about satisfying the problem constraints in problem P1.

According to the above analysis, if $x^*(t)$ is an integer, then the integer solution of problem P1 can be obtained by (16). Based on this idea, we propose a greedy algorithm to find the integer solution of the virtual variable $x(t)$ for any slot t , which is denoted as $y^*(t)$, and $y(t)$ is an integer variable corresponding to $x(t)$. Since $x^*(t)$ and $y^*(t)$ largely coincide, the non-integer solution $x^*(t)$ can be used to obtain the integer solution $y^*(t)$. Then, substituting $y(t)$ for $x(t)$ in (21), the power allocation solution can be computed by

$$P(t) = \left(\exp \left(\frac{y(t)}{\eta} \right) - 1 \right) N(t). \quad (27)$$

Since (27) establishes a one-to-one mapping between $y(t)$ and $P(t)$, the average power constraint should be considered when finding the integer solution $y^*(t)$. In particular, the proposed greedy algorithm

- (i) provides a valid integer solution $y^*(t)$, derived from the non-integer solution $x^*(t)$ at any slot t ;
- (ii) fulfills the average power constraint;
- (iii) achieves an objective value as close as possible to the maximum objective value in problem P5; and
- (iv) has low complexity.

Notice that simply rounding $x^*(t)$ does not ensure to always satisfy (ii) and (iii).

The pseudocode of the greedy algorithm is provided in Algorithm 2 and its main steps are sketched as follows. In step 2, each non-integer solution $x^*(t)$ is rounded to its floor integer $y(t)$ by the floor integer function $\lfloor \cdot \rfloor$, and then the corresponding power allocation $P(t)$ is calculated based on (27) in step 3. This may cause that the total power is underutilized and the maximum objective value is not achieved. Thus, the remaining power will be allocated along the time to increase the objective function value in the following steps. In step 4, $y(t)$ is assumed to be added one for any slot t , and then the increased power $\Delta P(t)$ and increased objective function value $\Delta g(t)$ are calculated in step 5 and step 6, respectively. The allocation process from step 8 to step 16 is repeated to add one to the selected $y(t)$ at each process until the set \mathcal{T}_A is empty. \mathcal{T}_A in step 9 represents the set of active slots at which $y(t)$ can possibly be added one under the average power constraint, given by

$$\mathcal{T}_A = \left\{ t | t \in [0, \frac{T}{2}], \Delta P(t) + \sum_{t=0}^{T/2} P(t) \leq \left(\frac{T}{2} + 1 \right) P_{av} \right\}. \quad (28)$$

In step 10, the slot t' in set \mathcal{T}_A which can achieve the maximal ratio of $\Delta g(t)$ to $\Delta P(t)$ is selected, which implies that the increase of objective function value per power is maximal at slot t' . Then, only $y(t')$ can be added one and the corresponding power consumption $P(t')$ can be assigned in step 11 and step 12, respectively. Next, $\Delta P(t')$ and $\Delta g(t')$ can be updated from step 13 to step 15. Finally, the integer solution $y^*(t)$ and the power allocation solution $P^*(t)$ can be obtained in step 18.

According to Algorithm 2, $y^*(t)$ and $P^*(t)$ at slot $t \in [0, \frac{T}{2}]$ have been obtained. Based on Lemma 4 and (27), $y^*(t)$ and $P^*(t)$ at slot $t \in [0, T]$ can be calculated. And then substituting $y^*(t)$ for $x^*(t)$ in (16), we can obtain the packet allocation solution for any service $k \in \mathcal{K}$ at slot $t \in [0, T]$. Furthermore, the greedy algorithm with low complexity leads to a near optimal rather than an optimal solution of problem P5, which implies that the obtained integer packet allocation solution and power allocation solution of problem P1 are both near optimal.

Algorithm 2 The greedy algorithm

Require: $x^*(t)$ for slot $t \in [0, \frac{T}{2}]$;

```

1: for  $t = 0$  to  $\frac{T}{2}$  do
2:    $y(t) \leftarrow \lfloor x^*(t) \rfloor$ ;
3:   Calculate  $P(t)$  using  $y(t)$  by (27);
4:   Calculate  $\tilde{P}(t)$  using  $\tilde{y}(t) = y(t) + 1$  by (27);
5:    $\Delta P(t) = \tilde{P}(t) - P(t)$ ;
6:    $\Delta g(t) = g(\tilde{P}(t)) - g(P(t))$ ;
7: end for
8: repeat
9:    $\mathcal{T}_A$  calculation;
10:   $t' = \arg \max_{t \in \mathcal{T}_A} \left( \frac{\Delta g(t)}{\Delta P(t)} \right)$ ;
11:   $y(t') \leftarrow \tilde{y}(t')$ ;
12:   $P(t') \leftarrow \tilde{P}(t')$ ;
13:  Calculate  $\tilde{P}(t')$  using  $\tilde{y}(t') = y(t') + 1$  by (27);
14:   $\Delta P(t') = \tilde{P}(t') - P(t')$ ;
15:   $\Delta g(t') = g(\tilde{P}(t')) - g(P(t'))$ ;
16: until  $\mathcal{T}_A = \emptyset$ 
17: for  $t = 0$  to  $\frac{T}{2}$  do
18:   $y^*(t) \leftarrow y(t), P^*(t) \leftarrow P(t)$ ;
19: end for
```

Ensure: $y^*(t)$ and $P^*(t)$ for slot $t \in [0, \frac{T}{2}]$

7 Numerical results and discussions

In this section, we implement the proposed algorithm using MATLAB (The MathWorks, Inc., Natick, MA, USA) and present simulation results to illustrate its performance. In order to emphasize different service weights, without loss of generality, we set the integer weight value $\omega_k = k$ for any $k \in \mathcal{K}$. In addition, we summarize the

simulation parameters in Table 1. A single simulation runs the algorithm when the train moves from the edge to the center of the BS coverage (25,000 slots).

The power allocation and link capacity along the time for the four power allocation schemes are presented in Figures 3 and 4, respectively. The advantage and disadvantage of these schemes can be observed from these two figures. For the sake of convenience in the engineering implementation, a constant power is allocated along the time in constant power allocation (CPA) scheme while it ignores the variation of channel gain and results in the great unfairness in term of link capacity. In order to provide a stable link capacity and achieve the best fairness along the time, the channel inversion power allocation (CIPA) scheme spends much power to compensate those bad channel states when the train is far from the BS. Similar to the traditional water-filling method, the water-filling power allocation (WFPA) scheme can maximum the total link capacity at the BS, whereas all the services will generally suffer from starvation when the train is near the edge of the BS coverage. In addition, the PFPA scheme can achieve a trade-off between the total link capacity and the link capacity fairness along the time. Finally, from Figure 3, we can observe that the power allocation solutions of PFPA problem and problem P4 are the same, which verifies Lemma 3 by simulations.

Figure 5 presents a comparison of the optimal packet solution $\mu_4^*(t)$ of problem P2 with the solution obtained by the other schemes. It can be observed that the trend of the curves in Figure 5 is similar to that in Figure 4. This can be explained by (14), which shows that the packet allocation solution $\mu_4^*(t)$ is linear with respect to the link capacity $\tilde{C}(t)$. Moreover, we can see that the optimal packet solution $\mu_4^*(t)$ of problem P2 is the same as the solution obtained by the PFPA+PAS scheme, which implies that problem P2 can be equivalently decomposed into PFPA problem and PAS problem. Service 4 is just an example

Table 1 Parameters in simulation

Parameter	Description	Value
P_{av}	Average power constraint	30 W
K	Number of services	6
W	System bandwidth	10 MHz
v	Constant moving speed	100 m/s
L	Packet size	240 bits
R	Cell radius	2.5 km
T_s	Slot duration	1 ms
d_0	Distance between BS and rail	100 m
α	Pathloss exponent	4
N_0	Noise power spectral density	-157 dBm/Hz

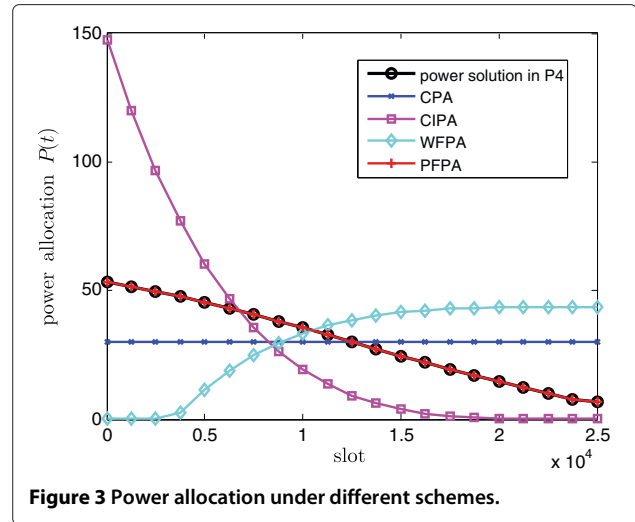


Figure 3 Power allocation under different schemes.

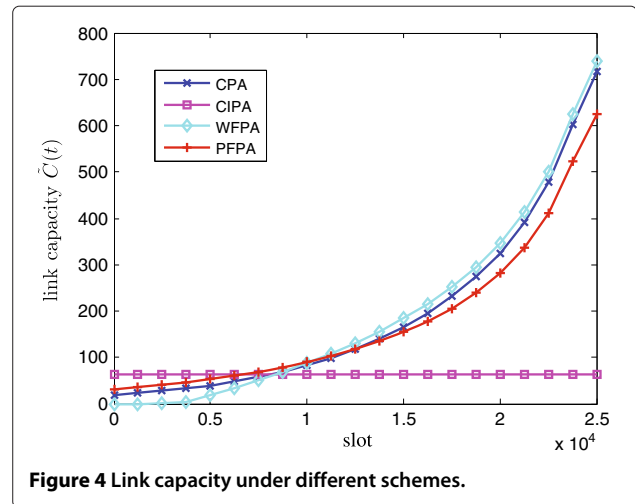


Figure 4 Link capacity under different schemes.

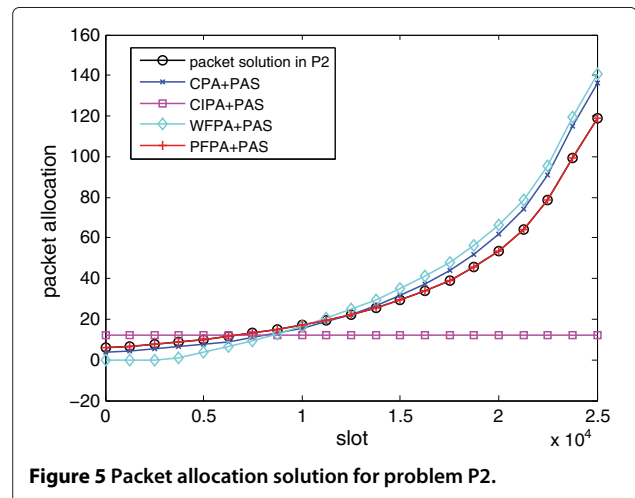


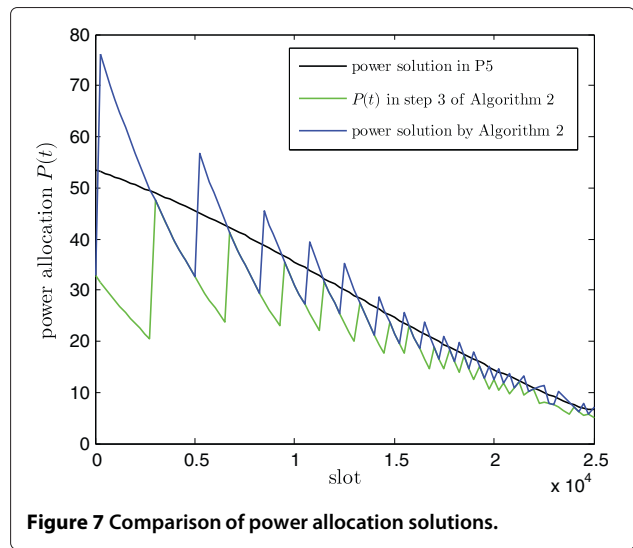
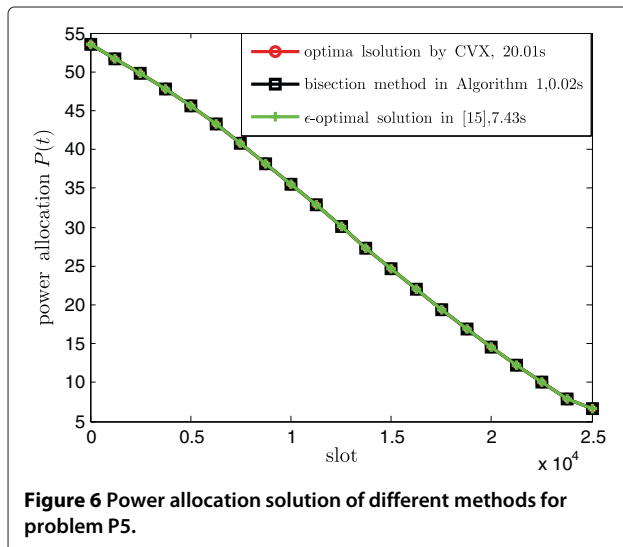
Figure 5 Packet allocation solution for problem P2.

for illustrating the characteristics of problem P2, and the same results can be obtained for the other services.

Figure 6 compares the power allocation solutions of three different methods solving problem P5 and their computational complexities, where the complexity is represented by the computing time in an Intel Core 3.30-GHz computer (Santa Clara, CA, USA). In this figure, the power allocation solutions are plotted in a single simulation and the computing times are obtained by averaging 100 simulations. It can be observed that the power allocation solutions of these methods are almost the same, which implies that the optimal power allocation solution of problem P5 can be obtained by the bisection search method when $\varepsilon_{\Delta P}$ and $\varepsilon_{\Delta \beta}$ are arbitrarily small. In addition, it is worth noting that the computing time of the bisection search method is much lower than those of the other two methods, which illustrates the high effectiveness of Algorithm 1.

Figure 7 compares different power allocation solutions under the situations, where the solution of virtual variable $x(t)$ is non-integer or integer. We can observe that the curve of the power allocation solution in problem P5 is smooth, which corresponds to the non-integer solution of $x(t)$. For obtaining the integer solution of $x(t)$, the non-integer solution $x^*(t)$ is rounded to its floor integer, while the power allocation $P(t)$ can be calculated in step 3 of Algorithm 2, which results in the underutilization of the total power as shown in this figure. As for the power allocation solution obtained by Algorithm 2, we can see that more power is allocated to some slots for better utilization of the total power. This causes that the curve of the power allocation solution obtained by Algorithm 2 is just around that of the power allocation solution in problem P5.

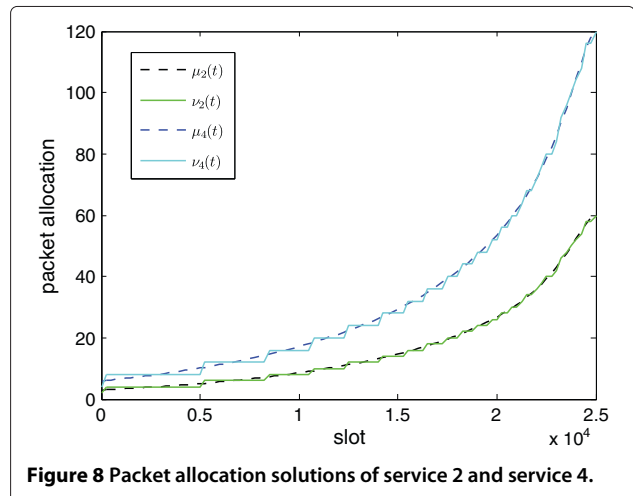
The packet allocation solutions of service 2 and service 4 are described in Figure 8, including the non-integer



packet allocation solution of problem P2 and the integer packet allocation solution obtained by Algorithm 2. From this figure, we can see that more and more packets are allocated to each service when the train moves from the edge to the center of the BS coverage. Similar to the phenomenon presented in Figure 7, the curve of the integer solution is just around that of the non-integer solution for both two services. Similar results can be obtained for the other services. In addition, the number of packets allocated to service 4 equals twice the number of packets allocated to service 2, which can be explained by Lemma 1.

8 Conclusion

In this paper, we investigated the utility-based resource allocation problem at a base station in HSR wireless networks, jointly taking into account the power allocation along time and the packet allocation among services. To



maximize the total utility at a base station under the average power constraint, the resource allocation optimization problem was formulated as an MINLP problem. With the help of integer constraint relaxation, the MINLP problem was simplified into a convex optimization problem. However, the size of the relaxed problem is very large, which motivates us to carry out the problem transformation and find some structural characteristics of the problem. The detailed analysis has been performed, which revealed that the relaxed problem can be equivalently decomposed into power allocation problem along time and packet allocation problem among the services. When the optimality of the relaxed problem is achieved, the power allocation along the time and the packet allocation along the time for each service are both proportionally fair. In addition, the integer relaxation provides a non-integer solution not implementable in practice, which motivates us to propose a greedy algorithm to obtain a near-optimal integer solution of the MINLP problem. Finally, the simulation results coincide with the structural characteristics we have shown in this paper.

Notice that only single-cell resource allocation problem is considered in this paper. The single-cell resource allocation can be regarded as a special case of the resource allocation in the entire trip. The single-cell resource allocation method can be extended to the multi-cell resource allocation problem in HSR wireless networks.

Appendices

Appendix 1

Proof of Lemma 1. Introduce a dual variable $\lambda \geq 0$ for the constraint (12b). The Lagrangian function of the optimization problem (12) is

$$\begin{aligned} L(\{\mu_k\}, \lambda) &= \sum_{k \in \mathcal{K}} \omega_k \ln(\mu_k) - \lambda \left(\sum_{k \in \mathcal{K}} \mu_k - \tilde{C} \right) \\ &= \sum_{k \in \mathcal{K}} (\omega_k \ln(\mu_k) - \lambda \mu_k) + \lambda \tilde{C}. \end{aligned} \quad (29)$$

The dual function $g(\lambda)$ can now be stated as

$$g(\lambda) = \begin{cases} \max L(\mu_k, \lambda) \\ \text{s.t. } \mu_k \geq 0, \forall k \in \mathcal{K}. \end{cases} \quad (30)$$

From the solution of (30), the resource allocation vector μ can be determined by solving K decomposed problems with an explicit solution $\mu_k = \frac{\omega_k}{\lambda}$. The dual problem of (12) can be expressed as

$$\min_{\lambda \geq 0} g(\lambda) = \sum_{k \in \mathcal{K}} \left(\omega_k \ln \left(\frac{\omega_k}{\lambda} \right) - \omega_k \right) + \lambda \tilde{C}, \quad (31)$$

whose optimal solution is $\lambda^* = \frac{\sum_k \omega_k}{\tilde{C}}$. Since the problem (12) is convex and satisfies Slater's condition, strong duality holds between (12) and its dual problem (31). Based on

[30], the solution of (12) can be obtained by the solution of (31) and is given by

$$\mu_k^* = \frac{\omega_k}{\lambda^*} = \frac{\omega_k \tilde{C}}{\sum_k \omega_k}, \forall k \in \mathcal{K}. \quad (32)$$

In addition, we can obtain that

$$\frac{\mu_i^*}{\omega_i} = \frac{\mu_j^*}{\omega_j} = \frac{1}{\lambda^*}, \forall i, j \in \mathcal{K}. \quad (33)$$

This completes the proof of Lemma 1. \square

Appendix 2

Proof of Lemma 2. Based on Definition 3 in [10], an optimal solution vector $\mathbf{x}^* = [x^*(1), \dots, x^*(T)]^T$ is said to be proportionally fair along the time if and only if, for any feasible solution vector $\mathbf{x}' = [x'(1), \dots, x'(T)]^T$, we have

$$\sum_{t=0}^T \frac{x'(t) - x^*(t)}{x^*(t)} \leq 0. \quad (34)$$

Firstly, we will prove that the solution of problem P3 satisfies (34) for any feasible vector \mathbf{x}' . For the ease of exposition, let $U(\mathbf{x})$ denote the objective function (18a) in problem P3. Since the strictly concave increasing property of $U(\mathbf{x})$, the following condition holds at $\mathbf{x} = \mathbf{x}^*$ [31]

$$\sum_{t=0}^T \frac{\partial U(\mathbf{x})}{\partial x(t)} \Big|_{x(t)=x^*(t)} (x'(t) - x^*(t)) = \sum_{t=0}^T \frac{x'(t) - x^*(t)}{x^*(t)} \leq 0. \quad (35)$$

This can be explained by the fact that movement along any direction $(\mathbf{x} - \mathbf{x}^*)$ at the optimal vector \mathbf{x}^* can not improve the objective function. Thus, the optimal solution vector \mathbf{x}^* is proportionally fair.

Secondly, due to $\mu_k(t) = \omega_k x(t)$ in (16), for any service k , we have

$$\begin{aligned} \sum_{t=0}^T \frac{\mu'_k(t) - \mu_k^*(t)}{\mu_k^*(t)} &= \sum_{t=0}^T \frac{\omega_k x'(t) - \omega_k x^*(t)}{\omega_k x^*(t)} \\ &= \sum_{t=0}^T \frac{x'(t) - x^*(t)}{x^*(t)} \leq 0, \end{aligned} \quad (36)$$

where $\mu'_k(t)$ is the feasible solution corresponding to $x'(t)$ and $\mu_k^*(t)$ is the optimal solution corresponding to $x^*(t)$. Thus, the optimal solution of problem P3 provides proportionally fair resource allocation along the time for each service. \square

Appendix 3

Proof of Lemma 3. First, substituting $\tilde{C}(t)$ into PFPA problem yields

$$\text{maximize } \sum_{t=0}^T \ln \left(\xi \ln \left(1 + \frac{P(t)}{N(t)} \right) \right) \quad (37a)$$

$$\text{subject to } \frac{1}{T+1} \sum_{t=0}^T P(t) \leq P_{av}, \quad (37b)$$

$$\text{variables } P(\tau) \geq 0, \tau \in [0, T], \quad (37c)$$

where $\xi = \frac{T_s W}{L \ln 2}$. For the constraint (37b), when the optimal solution is achieved, the equality holds. In this sense, the constraint in problem P4 and PFPA problem is same. To compare these two problems, the only difference is that η is in problem P4 and ξ is in the PFPA problem. Thus, if we can prove that the optimal solutions of these two problems are independent of η and ξ , respectively, then the optimal solution of problem P4 is equivalent to that of the PFPA problem.

Consider the Lagrangian function of PFPA problem

$$\begin{aligned} L(\{P(t)\}, \lambda) &= \sum_{t=0}^T \ln \left(\eta \ln \left(1 + \frac{P(t)}{N(t)} \right) \right) \\ &\quad - \lambda \left(\sum_{t=0}^T P(t) - (T+1)P_{av} \right) \quad (38) \\ &= \sum_{t=0}^T \left[\ln \left(\eta \ln \left(1 + \frac{P(t)}{N(t)} \right) \right) - \lambda P(t) + \lambda P_{av} \right]. \quad (39) \end{aligned}$$

Since the PFPA problem is convex, by applying the KKT conditions, we have

$$\frac{\partial L(\{P(t)\}, \lambda)}{\partial P(t)} = \frac{1}{\ln \left(1 + \frac{P(t)}{N(t)} \right) (P(t) + N(t))} - \lambda = 0, \quad (40)$$

and

$$\lambda \left(\sum_{t=0}^T P(t) - (T+1)P_{av} \right) = 0. \quad (41)$$

Thus, the optimal solution of PFPA problem can be obtained by solving (40) and (41), which is independent of η . Similarly, we can show that the optimal solution of problem P4 is independent of ξ . Therefore, the optimal solution of problem P4 is the same as that of the PFPA problem, which completes the proof. \square

Appendix 4

Proof of Lemma 4. We prove this lemma by contradiction. Without loss of generality, we assume that $P^*(t_1) \neq$

$P^*(T - t_1)$ for certain $t_1 \in [0, T]$. Construct another solution vector \mathbf{P}' by replacing the elements $P^*(t_1), P^*(T - t_1)$ in \mathbf{P}^* with $P'(t_1), P'(T - t_1)$ and keeping all other elements unchanged, where $P'(t_1) = P'(T - t_1) = \frac{P^*(t_1) + P^*(T - t_1)}{2}$. Notice that the following equality holds

$$P'(t_1) + P'(T - t_1) = P^*(t_1) + P^*(T - t_1), \quad (42)$$

which implies the solution vector \mathbf{P}' satisfies the constraint (22b) and \mathbf{P}' is a feasible solution vector.

For any $t \in [0, T]$, since $d(t) = d(T - t)$ and $N(t) = WN_0 d^\alpha(t)$, we have $N(t) = N(T - t)$. Since $\ln(\ln(\cdot))$ is a concave function, based on Jensen's inequality, we can obtain

$$\begin{aligned} &\ln \left(\eta \ln \left(1 + \frac{P^*(t_1)}{N(t_1)} \right) \right) + \ln \left(\eta \ln \left(1 + \frac{P^*(T - t_1)}{N(T - t_1)} \right) \right) \\ &= \ln \left(\eta \ln \left(1 + \frac{P^*(t_1)}{N(t_1)} \right) \right) + \ln \left(\eta \ln \left(1 + \frac{P^*(T - t_1)}{N(t_1)} \right) \right) \\ &\leq \ln \left(\eta \ln \left(1 + \frac{P'(t_1)}{N(t_1)} \right) \right) + \ln \left(\eta \ln \left(1 + \frac{P'(T - t_1)}{N(t_1)} \right) \right) \\ &= \ln \left(\eta \ln \left(1 + \frac{P'(t_1)}{N(t_1)} \right) \right) + \ln \left(\eta \ln \left(1 + \frac{P'(T - t_1)}{N(T - t_1)} \right) \right), \quad (43) \end{aligned}$$

which illustrates the solution vector \mathbf{P}^* is not optimal, which contradicts with the assumption. Therefore, for any $t \in [0, T]$, there must be the case $P^*(t) = P^*(T - t)$ in the optimal solution vector \mathbf{P}^* . \square

Appendix 5

Proof of Lemma 5. According to the equality $f(P(t)) = \beta$, we can find the maximum and minimum of function $f(P(t))$ among all slots instead of finding the maximum and minimum of β directly. Firstly, considering a certain slot t , $N(t)$ is fixed and $P(t) \in [P_{\min}, P_{\max}]$ is variable. Since $f(P(t))$ is a monotonically increasing function of $P(t)$, the maximum and the minimum of function $f(P(t))$ can be achieved when $P(t) = P_{\max}$ and $P(t) = P_{\min}$, respectively.

Secondly, to find the maximum and minimum of function $f(P(t))$ among all slots, it is necessary to consider different values of $N(t)$ along the time. The derivative of $f(P(t))$ with respect to $N(t)$ is given by

$$\frac{\partial f(P(t))}{\partial N(t)} = \ln \left(1 + \frac{P(t)}{N(t)} \right) - \frac{P(t)}{N(t)} \geq 0, \quad (44)$$

where the inequality holds based on $\ln(1 + z) \geq z$ for $z \geq 0$. Thus, given a constant $P(t)$ along the time, for any $t_1, t_2 \in [0, \frac{T}{2}]$ and $t_1 > t_2$, since $N(t_1) < N(t_2)$, we have $f(P(t))|_{t=t_1} < f(P(t))|_{t=t_2}$, which implies that the maximum and the minimum of function $f(P(t))$ can be achieved when $t = 0$ and $t = \frac{T}{2}$, respectively.

From the above analysis, we conclude that the maximum of β can be set as $\beta_{\max} = f(P_{\max})|_{t=0}$, and the minimum of β can be set as $\beta_{\min} = f(P_{\min})|_{t=\frac{T}{2}} = 0$. \square

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work is supported by the Fundamental Research Funds for the Central Universities (Grant No. 2014YJ5026), the Key Projects of State Key Lab of Rail Traffic Control and Safety (No. RCS2012ZZ004), the China Postdoctoral Science Foundation (Grant No. 2013M530519), and the Key Grant Project of Chinese Ministry of Education (No. 313006).

Received: 30 January 2014 Accepted: 9 April 2014

Published: 27 April 2014

References

1. J Wang, H Zhu, NJ Gomes, Distributed antenna systems for mobile communications in high-speed trains. *IEEE J. Selected Areas Commun.* **30**(4), 675–683 (2012)
2. JY Zhang, ZH Tan, ZD Zhong, Y Kong, A multi-mode multi-band and multi-system-based access architecture for high-speed railways, in *IEEE vehicular technology conference (fall)* (Ottawa, ON, Canada, 6–9, Sept 2010, pp. 1–5
3. Fokum D, Frost V, A survey on methods for broadband internet access on trains. *IEEE Commun. Surv. Tutorials.* **12**(2), 171–185 (2010)
4. L Tian, J Li, Y Huang, J Shi, J Zhou, Seamless dual-link handover scheme in broadband wireless communication systems for high-speed rail. *IEEE J. Selected Areas Commun.* **30**(4), 708–718 (2012)
5. J You, Z Zhong, R Xu, G Wang, Transmission schemes for high-speed railway: direct or relay?, in *Paper presented at the IEEE international wireless communications and mobile computing conference* Limassol, Cyprus, 27–31 Aug 2012, pp. 1103–1107
6. B Lannoo, D Colle, M Pickavet, P Demeester, Radio-over-fiber-based solution to provide broadband internet access to train passengers. *IEEE Commun. Mag.* **45**(2), 56–62 (2007)
7. W Luo, X Fang, M Cheng, Y Zhao, Efficient multiple-group multiple-antenna (MGMA) scheme for high-speed railway viaducts. *IEEE Trans. Vehicular Technol.* **62**(6), 2558–2569 (2013)
8. D Pareit, EV de Velde, D Naudts, J Bergs, J Keymeulen, I De Baere, W Van Brussel, C Vangeneugden, P Hauspie, G De Vos, I Moerman, C Blondia, P Demeester, A novel network architecture for train-to-wayside communication with quality of service over heterogeneous wireless networks. *EURASIP J. Wireless Commun. Netw.* **2012**(1), 1–30 (2012)
9. W van Brussel, Bringing ICT services to trains technical and economical challenges. *Proceedings of conference on telecommunications internet and media techno economics*, 1–7
10. YQ Dong, PY Fan, KB Letaief, High-speed railway wireless communications: efficiency v.s. fairness. *IEEE Trans. Vehicular Technol.* **63**(2), 925–930 (2014)
11. J Lee, S Leyffer, *IMA Volume in Mathematics and its Applications: Mixed Integer Nonlinear Programming*. (Springer, New York 2012)
12. OB Karimi, J Liu, C Wang, Seamless wireless connectivity for multimedia services in high speed trains. *IEEE J. Selected Areas Commun.* **30**(4), 729–739 (2012)
13. H Liang, WH Zhuang, Efficient on-demand data service delivery to high-speed trains in cellular/infostation integrated networks. *IEEE J. Selected Areas Commun.* **30**(4), 780–791 (2012)
14. TY Chen, HG Shan, X Wang, Packet scheduling for on-demand data services to high-speed trains over wireless links, in *Proceedings of IEEE Globecom Atlanta*, 9–13 Dec 2013, pp. 4787–4792
15. SF Xu, G Zhu, C Shen, Y Lei, Delay-aware fair scheduling in relay-assisted high-speed railway networks, in *Proceedings of international conference on communications and networking* Guilin, China, 14–16 Aug 2013, pp. 17–21
16. SY Lin, ZD Zhong, L Cai, Y Luo, Finite state Markov modelling for high speed railway wireless communication channel, in *Proceedings of IEEE Globecom Anaheim*, CA, USA, 3–7 Dec 2012, pp. 5421–5426
17. CD Thai, M Berbineau, F Sun, Exploiting the forecast channels for power allocation in high-speed railway communications, in *International conference on ITS telecommunications* Tampere, 5–7 Nov 2013, pp. 402–406
18. SF Xu, G Zhu, C Shen, Y Lei, Z Zhong, Analysis and optimization of resource control in high-speed railway wireless networks. *Math. Probl. Eng.* **2014**(781654), 13 (2014)
19. G Song, Y Li, Utility-based resource allocation and scheduling in OFDM-based wireless broadband networks. *IEEE Commun. Mag.* **43**(12), 127–134 (2005)
20. SF Xu, G Zhu, C Shen, B Ai, A QoS-aware scheduling algorithm for high-speed railway communication system, in *Proceedings of IEEE international conference on communications* Sydney, 10–14 June 2014
21. B Sklar, Rayleigh fading channels in mobile digital communication systems. I. Characterization. *IEEE Commun. Mag.* **35**(7), 90–100 (1997)
22. L Liu, C Tao, J Qiu, H Chen, Y Li, W Dong, Y Yuan, Position-based modeling for wireless channel on high-speed railway under a viaduct at 2.35 GHz. *IEEE J. Selected Areas Commun.* **30**(4), 834–845 (2012)
23. R He, ZD Zhong, B Ai, J Ding, An empirical path loss model and fading analysis for high-speed railway viaduct scenarios. *IEEE Antenn. Wireless Propag. Lett.* **10**, 808–812 (2011)
24. DH Ho, S Valaee, Information raining and optimal link-layer design for mobile hotspots. *IEEE Trans. Mobile Comput.* **4**(3), 271–284 (2005)
25. M Mehrjoo, MK Awad, M Dianati, X Shen, Design of fair weights for heterogeneous traffic scheduling in multichannel wireless networks. *IEEE Trans. Commun.* **58**(10), 2892–2902 (2010)
26. F Kelly, A Maulloo, D Tan, Rate control for communication networks: shadow prices, proportional fairness and stability. *J. Oper. Res. Soc.* **49**(3), 237–252 (1998)
27. Y Long, HY Li, M Pan, YG Fang, TF Wong, A fair QoS-aware resource-allocation scheme for multiradio multichannel networks. *IEEE Trans. Vehicular Technol.* **62**(7), 3349–3358 (2013)
28. A Eryilmaz, R Srikant, Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control, in *Proceedings of IEEE INFOCOM*, Miami, 13–17 March 2005, pp. 1794–1803
29. S MGrant, CVX Boyd, Matlab software for disciplined convex programming, v2.0 beta, Sep 2013. <http://cvxr.com/cvx>. Accessed 1 Dec 2013
30. S Boyd, L Vandenberghe, *Convex Optimization*. (Cambridge University Press, Cambridge, 2009)
31. KD Lee, VCM Leung, Fair allocation of subcarrier and power in an OFDMA wireless mesh network. *IEEE J. Selected Areas Commun.* **24**(11), 2051–2060 (2006)

doi:10.1186/1687-1499-2014-68

Cite this article as: Xu et al.: Utility-based resource allocation in high-speed railway wireless networks. *EURASIP Journal on Wireless Communications and Networking* 2014 **2014**:68.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com