Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Cross-layer resource allocation in wireless multi-hop networks with outdated channel state information*

Wei FENG[1], Sui-li FENG[†‡1], Yue-hua DING[1], Xin HUANG[1,2]
([1]*School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China*)
([2]*School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou 510006, China*)
†E-mail: Fengsl@scut.edu.cn

Received Oct. 29, 2013; Revision accepted Jan. 10, 2014; Crosschecked Apr. 11, 2014

Abstract: The cross-layer resource allocation problem in wireless multi-hop networks (WMHNs) has been extensively studied in the past few years. Most of these studies assume that every node has the perfect channel state information (CSI) of other nodes. In practical settings, however, the networks are generally dynamic and CSI usually becomes outdated when it is used, due to the time-variant channel and feedback delay. To deal with this issue, we study the cross-layer resource allocation problem in dynamic WMHNs with outdated CSI under channel conditions where there is correlation between the outdated CSI and current CSI. Two major contributions are made in this work: (1) a closed-form expression of conditional average capacity is derived under the signal-to-interference-plus-noise ratio (SINR) model; (2) a joint optimization problem of congestion control, power control, and channel allocation in the context of outdated CSI is formulated and solved in both centralized and distributed manners. Simulation results show that the network utility can be improved significantly using our proposed algorithm.

Key words: Wireless multi-hop networks, Outdated channel state information, Cross-layer resource allocation, Distributed algorithm
doi:10.1631/jzus.C1300315 **Document code:** A **CLC number:** TP393

1 Introduction

We consider the resource allocation problem in multi-radio and multi-channel wireless multi-hop networks such as ad hoc, mesh, and sensor networks. To achieve optimal resource allocation, joint cross-layer optimization draws substantial attention. In general, cross-layer optimization methods break through the conventional layered architecture and allow direct communication between protocols at nonadjacent layers or share variables between layers to obtain performance gains (Akyildiz and Wang, 2008).

Various cross-layer design methods for multi-radio and multi-channel wireless multi-hop networks have been proposed recently. Nguyen *et al.* (2012) proposed a cross-layer optimization framework for cognitive radio networks (CRNs) by jointly taking the physical layer and transport layer into account. Chen *et al.* (2013) designed a cross-layer framework to jointly schedule the sleep time and control the transmission power of mobile stations in the uplink of IEEE 802.16 networks. Qu *et al.* (2010), Augusto *et al.* (2011), Huang *et al.* (2011), and Shi *et al.* (2011) have also introduced different cross-layer methods to improve the energy efficiency of the networks. These cross-layer designs depend on two assumptions: the network is static; nodes have the real-time and perfect channel state information knowledge for resource allocation. However, network is dynamic, and the channel state information (CSI) obtained at nodes is imperfect due to the varying channel and transmitting delay, which make the above two assumptions

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (No. 61340035)

impractical.

'Imperfect CSI' is usually referred to as 'limited CSI' or 'outdated CSI'. Limited CSI refers to incomplete CSI induced by limited CSI feedback or the channel estimation error and can be improved by the use of strong channel coding. Outdated CSI signifies that there is a delay from CSI estimation to CSI utilization. CSI outdatedness is induced by the following two reasons: first, the CSI is estimated and may be fed back to the 1-hop transmitter or even the n-hop neighbors in a wireless multi-hop network; second, the CSI cannot be exchanged frequently due to network overhead. The unavailability of perfect CSI necessitates the design of resource allocation techniques that ensure communication with the required quality of service (QoS) despite the imperfect CSI. In fact, some researchers (Aggarwal *et al.*, 2011; Biyanwilage *et al.*, 2011; Li XC *et al.*, 2011; Ahmad and Assaad, 2013) have considered the impact of CSI outdatedness in the cross-layer design for the orthogonal frequency-division multiple-access (OFDMA) downlink. In the studies of this type, the base station (BS) runs the centralized algorithm to allocate power and subcarriers to the downlink according to outdated CSI fed back by the users around the BS. Ahmad and Assaad (2013) studied optimal power and subcarriers allocation for the downlink of a multi-user OFDMA system with imperfect channel knowledge. Li Y *et al.* (2011) proposed a sub-optimal power allocation scheme to address the impact of outdated CSI on single-user multi-channel OFDM systems. Aggarwal *et al.* (2011) jointly considered user scheduling, power allocation, and rate selection in OFDMA downlink systems to maximize the expected sum-utility under a sum-power constraint with imperfect CSI. Biyanwilage *et al.* (2011) investigated the power allocation problem in an OFDM relay link with a non-regenerative relay where the available CSI is an outdated version of the actual instantaneous CSI. There is also literature (Kim *et al.*, 2012; Lim *et al.*, 2012) which considers the impact of outdated CSI in spectrum sharing environments in cognitive radio (CR) systems. Kim *et al.* (2012) investigated the impact of outdated CSI between secondary and primary users in spectrum sharing environments, and derived the ergodic capacity of the secondary users with optimal power allocation under the average received-power constraint. Lim *et al.* (2012) proposed a power allocation scheme based on the mean value of the outdated channel gain between secondary transmitters and primary receivers in a spectrum sharing system. In addition, Xu and Dong (2012) designed an optimized relay precoding strategy for a multi-input multi-output (MIMO) relay downlink system by taking the effects of both the channel quantization error and feedback delay into account. Bin *et al.* (2013) and Cui *et al.* (2013) both analyzed the impact of outdated CSI in a simple amplify-and-forward (AF) relay system consisting of one source, N relays, and one destination.

However, these centralized allocation strategies in a single hop environment may not be directly applicable in a distributed wireless multi-hop network. To the best of our knowledge, limited research on dynamic multi-hop wireless networks has been carried out in terms of either performance analysis or design of dynamic cross-layer resource allocation with outdated CSI. In practical settings, however, the impact of feedback delay is more prominent in a multi-hop environment, since nodes need to send outdated information back to the BS or other nodes more than 1-hop away. Therefore, it is necessary to design a distributed algorithm accounting for outdated CSI, to improve the accuracy of resource allocation, especially in wireless multi-hop networks. In this paper, we investigate cross-layer resource allocation in dynamic wireless multi-hop networks with outdated CSI under channel conditions where there is correlation between the outdated CSI and current CSI; as a consequence, the network utility is significantly enhanced. The main contributions of this paper are listed as follows:

1. A closed-form expression of conditional average capacity under the signal-to-interference-plus-noise ratio (SINR) model is derived over Rayleigh fading channels by considering the probability distribution function of current CSI conditioned on outdated CSI. The SINR model is widely regarded as a better model for interference characterization. However, there are many difficulties in carrying out analysis with this model due to the computational complexity SINR involves, particularly when we study it in a multi-hop environment.

2. To the best of our knowledge, we have considered for the first time the impact of outdated CSI on cross-layer resource allocation in dynamic multi-hop wireless networks, and proposed a framework with outdated CSI to jointly optimize congestion control

in the transport layer, channel allocation in the data link layer, and power control in the physical layer. In the proposed framework, the network is modeled as a generalized network utility maximization (NUM) problem with elastic link data rate and power constraints. By applying the Lagrangian dual decomposition technique, the NUM problem is solved in a distributed manner.

2 System model

2.1 Network model

We consider a dynamic multi-radio and multi-channel wireless multi-hop network where nodes can move randomly. Let N denote the set of wireless nodes, L the set of unidirectional logical links, and (m,n) a logical link with transmitter m and receiver n. The sets of the incoming and outgoing logical links of node n are defined as $L_n^{\text{in}} \subset L$ and $L_n^{\text{out}} \subset L$ respectively, and the sets of in- and out-neighbors of node n are defined as $N_n^{\text{in}} = \{m : (m,n) \in L_n^{\text{in}}\}$ and $N_n^{\text{out}} = \{m : (n,m) \in L_n^{\text{out}}\}$, respectively. I_n denotes the set of available radios for node $n \in N$. L_n^k denotes the set of logical links that use radio k at node n. $\Theta = \{1, 2, \cdots, C\}$ represents the available orthogonal frequency channel set. In addition, it is assumed that the network operates in slotted time with slot set T and the Hyacinth algorithm proposed in Raniwala and Chiueh (2005) is used to form our ripple-effect free logical topology.

2.2 Channel model

With the aforementioned logical link model, the signal received by node n can be written as follows:

$$y_n = h_{mn} x_m + w_n, \tag{1}$$

where $x_m \in \Omega$ is the transmitted symbol from node m with power $P_m = E[|x_m|^2]$, $w_n \in \Omega$ is the additive white Gaussian noise (AWGN) with zero mean and variance σ_w^2 (independent of n), and $h_{mn} \in \Omega$ is the channel response of link (m,n), modeled as Rayleigh fading. For mathematical convenience, we assume a block-fading channel where the channel response remains constant during one time-slot and that different channels (for different links (m,n)) are independent and identically distributed (i.i.d.).

Each link receiver estimates the CSI and sends it to the link transmitter or to its neighbors more than 1-hop away in a multihop network. Due to feedback delay, the receiver has only outdated CSI at its disposal and can allocate only the resources at time t according to the CSI estimated at time $t-\tau$, where τ is the feedback delay. The effects of delay and the time-varying characteristic of channel are analyzed based on the general correlation model, in which the actual CSI accounting for the CSI outdatedness is modeled as (Kim *et al.*, 2012; Lim *et al.*, 2012)

$$h_{mn}(t) = \rho \hat{h}_{mn}(t) + \sqrt{1-\rho^2} e_{mn}(t), \tag{2}$$

where $\hat{h}_{mn}(t)$ denotes the outdated channel response, e_{mn} is a complex Gaussian random variable with zero mean and unit variance, uncorrelated with $\hat{h}_{mn}$, and ρ is the time correlation coefficient defined as

$$\rho = E[h_{mn}(t) \hat{h}_{mn}^*(t)]. \tag{3}$$

For mathematical tractability, we assume ρ is the same for all the nodes. In this study, the Jakes scattering model is adopted; hence, $\rho = J_0(2\pi f_{\text{d}} \tau)$, where f_{d} represents the Doppler frequency, τ is the delay in time units, and $J_0(\cdot)$ denotes the zero-order Bessel function of the first kind. Actually, this model just gives us an easy way to relate the correlation coefficient ρ to the delay τ. The analysis and designs throughout this paper do not depend on the specific Jakes model as long as ρ is given. Note that ρ denotes the degree of CSI uncertainty; hence, it is equal to 0 when $\hat{h}_{mn}$ is completely unreliable for current resource allocation. This worst case is not considered in this paper. Instead, we limit the discussion under general channel conditions where there is correlation between the outdated CSI and current CSI as in most of the literature (Li Y *et al.*, 2011; Kim *et al.*, 2012; Ahmad and Assaad, 2013).

According to the channel model given in Eq. (2), it can be derived that h_{mn} and $\hat{h}_{mn}$ follow a joint complex Gaussian distribution with correlation coefficient ρ; as a consequence, the probability density function (PDF) of h_{mn} conditioned on $\hat{h}_{mn}$ can be easily obtained by applying Bayes' theorem:

$$h_{mn} | \hat{h}_{mn} \sim \text{CN}(\rho \hat{h}_{mn}, 1-\rho^2). \tag{4}$$

Let $\gamma_{mn} = |h_{mn}| / \sigma_w^2$ denote the instantaneous channel-to-noise ratio (CNR) of link (m,n) in a given time slot and $\hat{\gamma}_{mn} = \left|\hat{h}_{mn}\right| / \sigma_w^2$ the corresponding outdated value. Based on the above discussion, it is straightforward that γ_{mn} conditioned on $\hat{\gamma}_{mn}$ follows a non-central chi-square distribution with two degrees of freedom. The PDF takes the following expression (Lim *et al.*, 2012):

$$f_{\gamma_{mn}|\hat{\gamma}_{mn}} = \frac{\sigma_w^2}{1-\rho^2}\exp\left[-\frac{\sigma_w^2}{1-\rho^2}(\gamma_{mn}+\rho^2\hat{\gamma}_{mn})\right] \cdot I_0\left(\frac{2\sigma_w^2}{1-\rho^2}\sqrt{\rho^2\gamma_{mn}\hat{\gamma}_{mn}}\right), \quad (5)$$

where $I_0(\cdot)$ stands for the zero-order modified Bessel function of the first kind.

3 Problem formulation

Our objective is to jointly optimize the data rate, channel and power allocation with outdated CSI to improve network utility in a multi-hop network environment. To solve the problem, the following constraints are considered:

1. Channel allocation constraint. Let $\boldsymbol{X} = [\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}, \cdots, \boldsymbol{x}_{|L|}^{\mathrm{T}}] \in \mathbb{R}^{C\times|L|}$ be the binary channel allocation matrix where $\boldsymbol{x}_l = [x_{l1}, x_{l2}, \cdots, x_{lC}]$. The elements $x(l,c)$ in $\boldsymbol{X}$ are equal to 1 if channel $c \in \Theta$ is allocated to logical link $l \in L$; otherwise, it is equal to 0. So, the channel allocation must satisfy the following two equations:

$$\boldsymbol{x}_l\boldsymbol{1}^{\mathrm{T}} = 1,\ \forall l \in L, \quad (6)$$

$$\boldsymbol{x}_l = \boldsymbol{x}_i, \forall l, i \in L_n^k, n \in N, k \in I_n, \quad (7)$$

where $\boldsymbol{1}$ denotes a $1 \times C$ vector with all elements equal to 1.

Eq. (6) means that only one channel can be assigned to each logical link at one time slot. Eq. (7) represents that the logical links which share a radio should be assigned the same channel.

2. Total power constraint. Let $P_n^{\max}$ denote the total power budget at node n in each time slot and $\boldsymbol{P} = [P_1, P_2, \cdots, P_{|L|}]$ the transmit power vector. Thus, we have

$$0 \le \sum_{m\in N_n^{\mathrm{out}}} P_{nm} \le P_n^{\max}. \quad (8)$$

3. Link capacity constraint. Assume that there are $S = \{1, 2, \cdots, s\}$ flows in the multi-radio and multi-channel wireless multi-hop network and flow s is transmitted through the path $L(s)$. $\boldsymbol{F} = \{f_1, f_2, \cdots, f_s\}$ is the data rate vector. Since the total traffic rate of the flows on each link must be less than its link capacity, we have

$$\sum_{s:l\in L(s)} f_s \le C_l(\boldsymbol{X}, \boldsymbol{P}),\ \forall l \in L. \quad (9)$$

We consider the scheme of concurrent transmitting, according to which all links are active simultaneously. The instantaneous data rate for link l is given as

$$\begin{aligned} C_l(\boldsymbol{X}, \boldsymbol{P}) &= B\log_2(1 + K\cdot \mathrm{SINR}_l(\boldsymbol{X}, \boldsymbol{P})) \\ &= B\log_2\left(1 + K\cdot\frac{P_l\gamma_{ll}}{\sum_{i\neq l}\boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}P_i\gamma_{il}+1}\right), \\ & i\neq l,\ \forall i, l\in L. \end{aligned} \quad (10)$$

Here, constant B is the transmission bandwidth on each channel, SINR_l denotes the SINR of link l, $K = -\phi_1/\log(\phi_2\cdot \mathrm{BER})$, where ϕ_1 and ϕ_2 are constants determined by the modulation and BER is the required bit-error rate, and γ_{il} denotes the instantaneous channel-to-noise ratio (CNR) of the link from the transmitter of link i to the receiver of link l.

As aforementioned, the node cannot obtain the current CSI because of delay; however, it has the PDF of instantaneous CSI conditioned on outdated CSI. So, we can derive the expectation of conditional capacity of link l as follows:

$$\begin{aligned} &\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X}, \boldsymbol{P}) \\ &= E_{\gamma|\hat{\gamma}}\left[B\log_2\left(1+K\frac{P_l\gamma_{ll}}{\sum_{i\neq l}\boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}P_i\gamma_{il}+1}\right)\right], \\ &i\neq l,\quad \forall i, l \in L, \end{aligned} \quad (11)$$

where $\gamma = \{\gamma_{1l}, \gamma_{2l}, \cdots, \gamma_{|L|l}\}$ and $\hat{\gamma} = \{\hat{\gamma}_{1l}, \hat{\gamma}_{2l}, \cdots, \hat{\gamma}_{|L|l}\}$ are the actual CNR set and outdated CNR set respectively, and $E_{\gamma|\hat{\gamma}}$ denotes the expectation with respect to γ conditioned on $\hat{\gamma}$. Hence, we can remodel the link capacity constraint in terms of the expectation of conditional capacity as follows:

$$\sum_{s:l\in L(s)} f_s \le \bar{C}_l(\boldsymbol{X}, \boldsymbol{P}),\quad \forall l \in L. \quad (12)$$

It is obvious that the expectation of conditional capacity under the SINR model is a complex multiple integral, which is difficult to solve. However, by applying Eq. (33) in Ahmad and Assaad (2013) iteratively, we can derive the closed-form expression for it:

$$\begin{aligned} \bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X}, \boldsymbol{P}) &= \frac{B}{\log 2}\left[\tau_l - \tau_{1'}\beta_l + \tau_{2'}\beta_l(\beta_{1'}-1)\right. \\ &- \tau_{3'}\beta_l\ (\beta_{1'}-1)(\beta_{2'}-1) + \cdots + (-1)^m\tau_{m'}\beta_l \\ &\cdot (\beta_{1'}-1)(\beta_{2'}-1)\cdots(\beta_{(m-1)'}-1) + (-1)^{m+1} \\ &\left.\cdot\ \beta_l(\beta_{1'}-1)(\beta_{2'}-1)\cdots(\beta_{m'}-1) + \beta_l\log K\right], \end{aligned} \quad (13)$$

where $M' = \{1', 2', \cdots, m'\}, l \notin M'$ denotes the interference link set of link l, $\boldsymbol{x}_i \boldsymbol{x}_l^{\mathrm{T}} = 1\ \forall i \in M'$, $\tau_i = \alpha_i + \beta_i \psi(a_i) + \beta_i \log_2(P_i \theta_i)$, $\alpha_i = \log_2(1+z_i) - \frac{z_i}{1+z_i}\log_2 z_i$, $\beta_i = \frac{z_i}{1+z_i}$,

$$z_i = \begin{cases} \dfrac{K P_i \hat{\gamma}_{il}}{\sum_{j\in\{i+1,i+2,\dots,m'\}} P_j \hat{\gamma}_{jl} + 1}, & i \neq l, i \in M', \\ \dfrac{K P_i \hat{\gamma}_{ii}}{\sum_{j\in\{1,2,\dots,m'\}} P_j \hat{\gamma}_{ji} + 1}, & i = l, \end{cases}$$

$$a_i = \begin{cases} \dfrac{\left(\frac{\sigma_w^2}{1-\rho^2}\rho^2 \hat{\gamma}_{il} + 1\right)^2}{2\frac{\sigma_w^2}{1-\rho^2}\rho^2 \hat{\gamma}_{il} + 1}, & i \neq l, i \in M', \\ \dfrac{\left(\frac{\sigma_w^2}{1-\rho^2}\rho^2 \hat{\gamma}_{ii} + 1\right)^2}{2\frac{\sigma_w^2}{1-\rho^2}\rho^2 \hat{\gamma}_{ii} + 1}, & i = l, \end{cases}$$

$$\theta_i = \begin{cases} \dfrac{\left(\rho^2 \hat{\gamma}_{il} + \frac{\sigma_w^2}{1-\rho^2}\right)^2}{a_i}, & i \neq l, i \in M', \\ \dfrac{\left(\rho^2 \hat{\gamma}_{ii} + \frac{\sigma_w^2}{1-\rho^2}\right)^2}{a_i}, & i = l, \end{cases}$$

$\forall i \in M' \cup l$. The derivation process for Eq. (13) is provided in Appendix A.

According to the above discussion, the NUM problem can now be stated as

$$\max_{\boldsymbol{F},\boldsymbol{X},\boldsymbol{P}} \sum_{s\in S} U(f_s)$$

subject to

$$\begin{aligned} & \boldsymbol{x}_l \mathbf{1}^{\mathrm{T}} = 1,\ \forall l \in L, \\ & \boldsymbol{x}_l = \boldsymbol{x}_i,\ \forall l, i \in L_n^k, n \in N, k \in I_n, \\ & 0 \leq \sum_{m\in N_n^{\text{out}}} P_{nm} \leq P_n^{\max}, \\ & \sum_{s:l\in L(s)} f_s \leq \bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X}, \boldsymbol{P}),\ \forall l \in L, \end{aligned} \tag{14}$$

where function $U(f_s)$ is a continuously differentiable, increasing and strict concave. We use $U(f_s) = \log f_s$ in Mo and Walrand (2000) to obtain the proportional fairness among the flows.

4 Centralized optimal algorithm

Problem (14) is a non-linear mixed integer programming problem, since it has a binary matrix $\boldsymbol{X}$, real vectors $\boldsymbol{F}$ and $\boldsymbol{P}$, as well as mixed binary-real cubic constraints. Even relaxing the binary constraints, it is still non-convex. In this section, we use binary linearization techniques (Rad and Wong, 2008) and log-transformed convex optimization techniques (Chiang, 2005) to obtain the global optimal solution of the NUM problem (14) in a centralized manner.

According to the binary linearization techniques (Rad and Wong, 2008), we firstly define a $C \times 1$ auxiliary link channel assignment vector

$$\boldsymbol{u}_{il} = \boldsymbol{x}_i \circ \boldsymbol{x}_l, \tag{15}$$

where $\circ$ denotes the Hadamard product. Eq. (15) can be described as the following linear constraints:

$$\boldsymbol{x}_i + \boldsymbol{x}_l - \boldsymbol{u}_{il} \leq \mathbf{1},\quad \boldsymbol{x}_i + \boldsymbol{x}_l - 2\boldsymbol{u}_{il} \geq \mathbf{0}. \tag{16}$$

Then, we define an auxiliary real scalar variable z_{il} to linearize the cubic constraint:

$$z_{il} = \boldsymbol{x}_i \boldsymbol{x}_l^{\mathrm{T}} P_i = \boldsymbol{u}_{il} \mathbf{1}^{\mathrm{T}} P_i. \tag{17}$$

So, we have

$$\begin{aligned} & 0 \leq z_{il} \leq P_i,\ \forall i, l \in L, \\ & \sum_{i\in L_n^{\text{out}}} z_{il} = \sum_{i\in L_n^{\text{out}}} \boldsymbol{x}_i \boldsymbol{x}_l^{\mathrm{T}} P_i \leq \sum_{i\in L_n^{\text{out}}} P_i \leq P_n^{\max}. \end{aligned} \tag{18}$$

Furthermore, using log-transformation, we define $Q_l = \log P_l$, $R_{il} = \log z_{il}$. Problem (14) can be transformed into the following optimization problem:

$$\max_{\boldsymbol{F},\boldsymbol{X},\boldsymbol{P}} \sum_{s\in S} U(f_s)$$

subject to

$$\begin{aligned} & \boldsymbol{x}_l \mathbf{1}^{\mathrm{T}} = 1,\quad \forall l \in L, \\ & \boldsymbol{x}_l = \boldsymbol{x}_i,\quad \forall l, i \in L_n^k, n \in N, k \in I_n, \\ & 0 \leq \sum_{m\in N_n^{\text{out}}} P_{nm} \leq P_n^{\max}, \\ & \boldsymbol{x}_i + \boldsymbol{x}_l - \boldsymbol{u}_{il} \leq \mathbf{1},\quad \boldsymbol{x}_i + \boldsymbol{x}_l - 2\boldsymbol{u}_{il} \geq \mathbf{0}, \\ & 0 \leq z_{il} \leq P_i,\quad \forall i, l \in L, \\ & \sum_{i\in L_n^{\text{out}}} z_{il} \leq \sum_{i\in L_n^{\text{out}}} P_i \leq P_n^{\max},\quad \forall l \in L, n \in N, \\ & \sum_{s:l\in L(s)} f_s \leq \\ & E_{\gamma|\hat{\gamma}}\left\{ B\left[\log_2(K \mathrm{e}^{Q_l} \gamma_{ll}) - \log_2\left(\sum_{i\neq l} \mathrm{e}^{R_{il}} \gamma_{il} + 1 \right) \right] \right\}. \end{aligned} \tag{19}$$

It is known that the utility function $U(f_s)$ is a continuously differentiable, increasing and strict concave. The constraints except the last one are all linear functions. For the last constraint, the first part in the curly braces $\log_2(K\mathrm{e}^{Q_l}\gamma_{ll})$ is obviously linear in Q_l, and the second part $-\log_2\left(\sum_{i\neq l}\mathrm{e}^{R_{il}}\gamma_{il}+1\right)$ is concave in R_{il} because the logarithm of a sum of exponentials of linear functions of Q_l is convex. Therefore, problem (19) is a convex optimization problem. We can use the branch and bound algorithm (Bertsekas, 1995) to find the global optimal solution.

5 Distributed sub-optimal algorithm

The centralized algorithm helps us find the optimal solution of the NUM problem (14). In this section, we propose a sub-optimal distributed design which is more practical. According to the Lagrangian dual decomposition technique (Boyd and Vandenberghe, 2004), we first obtain the Lagrange dual of the original optimization problem (14), and then divide the dual problem into smaller sub-problems which are coordinated by a resource price. Consider the Lagrange dual to primal problem (14):

$$\min_{\boldsymbol{\xi}\geq\mathbf{0}} D(\boldsymbol{\xi}) \tag{20}$$

with partial dual function

$$D(\boldsymbol{\xi})=\max_{\boldsymbol{F},\boldsymbol{X},\boldsymbol{P}}\left[\sum_{s\in S}U(f_s)-\sum_{l\in L}\xi_l\left(\sum_{s:l\in L(s)}f_s-\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X},\boldsymbol{P})\right)\right], \tag{21}$$

where $\xi_l \geq 0$ is the Lagrange multiplier and also called the dual variable. As mentioned, we divide the Lagrange dual (21) into smaller sub-problems:

$$D_1(\boldsymbol{\xi})=\max_{\boldsymbol{F}}\left(\sum_{s\in S}U(f_s)-\sum_{l\in L}\xi_l\sum_{s:l\in L(s)}f_s\right), \tag{22}$$

and

$$D_2(\boldsymbol{\xi})=\max_{\boldsymbol{X},\boldsymbol{P}}\sum_{l\in L}\xi_l\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X},\boldsymbol{P}). \tag{23}$$

The first sub-problem (22) is a classical congestion control problem. It can be solved as follows:

$$f_s^*=(U_s^{'})^{-1}\left(\sum_{l\in L(s)}\xi_l\right). \tag{24}$$

The second sub-problem (23) is a joint channel allocation and power control problem with outdated CSI, which will be discussed later. The interaction between the two sub-problems is coordinated by the dual variable ξ_l. Since each sub-problem depends on ξ_l to decide the amount of resources to be used, ξ_l can be interpreted as the congestion price of link l. Using the sub-gradient method, ξ_l can be updated at each intermediate node as follows:

$$\xi_l(t+1)=\left[\xi_l(t)+\frac{\mu}{\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X},\boldsymbol{P})}\left(\sum_{s:l\in L(s)}f_s(t)-\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X},\boldsymbol{P})\right)\right]^+, \tag{25}$$

where f_s, $\boldsymbol{X}$, and $\boldsymbol{P}$ are the solutions of sub-problems (22) and (23) respectively, $\{\cdot\}^+=\max(0,\cdot)$, μ is a sufficiently small step size, and t is the iteration time slot.

Furthermore, to solve sub-problem (23), according to the primal decomposition method (Chiang *et al.*, 2007), we decompose (23) directly into two smaller sub-problems: (1) channel allocation problem with outdated CSI, and (2) power control problem with outdated CSI.

The channel allocation problem with outdated CSI can be written as

$$\max_{\boldsymbol{X}}\sum_{l\in L}\xi_l E_{\gamma|\hat{\gamma}}\left[B\log_2\left(1+K\frac{P_l^*\gamma_{ll}}{\sum_{i\neq l}\boldsymbol{x}\boldsymbol{x}_l^{\mathrm{T}}P_i^*\gamma_{il}+1}\right)\right]$$
$$\text{subject to}\quad \boldsymbol{x}_l\mathbf{1}^{\mathrm{T}}=1,\ \boldsymbol{x}_l=\boldsymbol{x}_i,\ \ \forall l,i\in L,n\in N, \tag{26}$$

where P_l^* is obtained by solving sub-problem (28). To reduce the computational complexity, sub-problem (26) can be implemented in a distributed manner, which can be described as

$$\max_{\boldsymbol{x}_j,j\in L_n}\sum_{l\in L}\xi_l E_{\gamma|\hat{\gamma}}\left[B\log_2\left(1+K\frac{P_l^*\gamma_{ll}}{\sum_{i\neq l}\boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}P_i^*\gamma_{il}+1}\right)\right]$$
$$\text{subject to}\quad \boldsymbol{x}_l\mathbf{1}^{\mathrm{T}}=1,\ \ \boldsymbol{x}_l=\boldsymbol{x}_i,\ \ \forall l,i\in L,n\in N. \tag{27}$$

During the implementation, each node n is responsible for assigning the optimal channel to link $j\in L_n$ in sequence and periodically exchanges its channel information $\boldsymbol{x}_j, j\in L_n$, as well as its collected data $\xi_j, j\in L_n$ with all other nodes. Let $V_{\max}$ denote the maximum number of links connected to the node in the network. The possibility for $\boldsymbol{x}_j, j\in L_n$ is C, which is the number of available channels. Thus, the computational complexity of sub-problem (27) for each node is at most $C \cdot V_{\max}$. Since the value of $V_{\max}$ generally keeps stable even though the network grows in size, the computational complexity of the local channel allocation sub-problem is a stable value, independent of the network size.

The power control problem with outdated CSI can be formulated as

$$\max_{\boldsymbol{P}} \sum_{l\in L} \xi_l E_{\gamma|\hat{\gamma}}\left[B\log_2\left(1+K\frac{P_l\gamma_{ll}}{\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i\gamma_{il}+1}\right)\right]$$

subject to

$$0\leq \sum_{m\in N_n^{\text{out}}} P_{nm}\leq P_n^{\max},\quad \forall n\in N, \tag{28}$$

where $\boldsymbol{x}_l^*$, $\forall l\in L$ is obtained by solving sub-problem (27). Sub-problem (28) can be implemented in a centralized manner after log-transformation; however, we solve it using a distributed algorithm due to the computational complexity and computing time.

5.1 Distributed power control with outdated CSI

Let $\tilde{P}_l=\log P_l$, $\forall l\in L$. Then, we can define the Lagrangian function in $\tilde{\boldsymbol{P}}$-space as

$$\begin{aligned} L(\tilde{\boldsymbol{P}},\boldsymbol{\zeta}) =& \sum_{n\in N}\sum_{l\in L_n^{\text{in}}}\xi_l\bar{C}_{l,\gamma|\hat{\gamma}}(\tilde{\boldsymbol{P}}) \\ &-\sum_{n\in N}\zeta_n\left(\sum_{l\in L_n^{\text{in}}}\mathrm{e}^{\tilde{P}_l}-P_n^{\max}\right), \end{aligned} \tag{29}$$

where $\zeta_n\geq 0$ is the dual variable. The corresponding dual problem for (28) can be formulated as $\min\limits_{\boldsymbol{\zeta}}\max\limits_{\tilde{\boldsymbol{P}}} L(\tilde{\boldsymbol{P}},\boldsymbol{\zeta})$. Furthermore, according to the strong duality theorem, the solution of the Lagrange dual is the global optimum of original problem (28).

We update dual variable ζ_n through a gradient descent (Bertsekas, 1995; Papandriopoulos and Evans, 2006):

$$\zeta_n(t+1)=\left[\zeta_n(t)+\varepsilon\left(\sum_{l\in L_n^{\text{in}}}P_l(t)-P_n^{\max}\right)\right]^+, \tag{30}$$

where $[\cdot]^+=\max(0,\cdot)$, ε is a sufficiently small step size, and t is the iteration time slot. It is obvious that ζ_n is updated locally by node n. ζ_n can also be interpreted as power price: as the sum of power consumption exceeds the total power budget at node n, node n will increase the power price, and vice versa. Next, for each ζ_n, we need to find the solution of dual problem (29) with respect to $\tilde{\boldsymbol{P}}$. By taking the derivative of $L(\tilde{\boldsymbol{P}},\boldsymbol{\zeta})$ with respect to $\tilde{P}_l$, we obtain

$$\begin{aligned} \nabla_l L(\tilde{\boldsymbol{P}},\boldsymbol{\zeta}) =& \frac{\xi_l B\beta_l}{\log 2} \\ &-\sum_{n\in N}\sum_{i\neq l, l\in L_n^{\text{in}}}\boldsymbol{x}_l^*(\boldsymbol{x}_i^*)^{\mathrm{T}}\frac{\xi_i B\beta_i\beta_{li}}{\log 2}-\zeta_n\mathrm{e}^{\tilde{P}_l}, \end{aligned} \tag{31}$$

where β_l and β_i can be derived from Eq. (A2), and β_{li} can be obtained from Eq. (A5). The derivation process for Eq. (31) is provided in Appendix B.

Then, since $\nabla_l L(P_l,\boldsymbol{\zeta})=(1/P_l)\nabla_l L(\tilde{P}_l,\boldsymbol{\zeta})$, by transforming Eq. (31) back to the $\boldsymbol{P}$ solution space instead of $\tilde{\boldsymbol{P}}$, we have

$$\begin{aligned} \nabla_l L(\boldsymbol{P},\boldsymbol{\zeta}) =& \frac{\xi_l B\beta_l}{P_l\log 2} \\ &-\frac{1}{P_l}\sum_{n\in N}\left(\sum_{i\neq l, l\in L_n^{\text{in}}}\boldsymbol{x}_l^*(\boldsymbol{x}_i^*)^{\mathrm{T}}\frac{\xi_i B\beta_i\beta_{li}}{\log 2}-\zeta_n\right). \end{aligned} \tag{32}$$

We now use the gradient method (Boyd and Vandenberghe, 2004) to solve the dual problem (29) with respect to P_l. With $\boldsymbol{\zeta}$ fixed,

$$\begin{aligned} P_l(t+1) &= P_l(t)+\eta\nabla_l L(\boldsymbol{P},\boldsymbol{\zeta}) \\ &= P_l(t)+\eta\frac{\xi_l B\beta_l}{P_l(t)\log 2}-\eta\zeta_n \\ &\quad -\eta\frac{B}{P_l(t)\log 2}\sum_{n\in N}\sum_{i\neq l, l\in L_n^{\text{in}}}\boldsymbol{x}_l^*(\boldsymbol{x}_i^*)^{\mathrm{T}}\xi_i\beta_i\beta_{li}. \end{aligned} \tag{33}$$

Here, η is a sufficiently small step size. We can rewrite the gradient steps as the following distributed power update algorithm through a combination of measurement and message passing:

$$\begin{aligned} P_l(t+1) &= P_l(t)+\eta\frac{\xi_l B\beta_l}{P_l(t)\log 2}-\eta\zeta_n \\ &\quad -\eta\frac{B}{P_l(t)\log 2}\sum_{n\in N}\sum_{i\neq l, l\in L_n^{\text{in}}}\boldsymbol{x}_l^*(\boldsymbol{x}_i^*)^{\mathrm{T}}m_{li}(t), \end{aligned} \tag{34}$$

where $m_{li}(t)=\xi_i\beta_i\beta_{li}$ is the message passed from the transmitter of link i to the transmitter of link l, η and B are predefined constants, and $\xi_l(t)$, β_l, and ζ_n can be obtained through local measurement. The physical meaning of Eq. (34) is that: the transmitter of link l increases power proportional to the local congestion price and decreases power by the sum of the congestion messages from all other transmitters.

According to the message passing method, we briefly describe the power update algorithm with outdated CSI as follows:

1. Each transmitter of link $i \in L$ calculates a message $m_{li}(t)$ based on locally measurable quantities for interference link $l \in L$, and passes the message to the transmitter of link l.

2. After receiving $m_{il}(t)$, the transmitter of link l updates the power according to Eq. (34).

3. After updating the power, the transmitter updates the power price according to Eq. (30).

4. The transmitter computes $\beta_l(t)$ as follows:

$$\begin{aligned}\beta_l(t) &= \frac{z_0(t)}{1+z_0(t)},\\ z_0(t) &= K\cdot\frac{P_l^*\hat{\gamma}_{ll}}{\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i^*\hat{\gamma}_{il}+1}.\end{aligned} \tag{35}$$

5. Return to step 1.

5.2 Distributed joint congestion control, channel allocation, and power control algorithm with outdated CSI

The proposed distributed joint congestion control, channel allocation, and power control algorithm with outdated CSI is described as follows:

1. We initially generate the network topology using the algorithm in Raniwala and Chiueh (2005) and then set $\boldsymbol{x}_l=[1,0,\cdots,0], \forall l\in L$.

2. During each time slot, the following three operations are carried out simultaneously:

(1) Each transmitter $n \in N$ updates the congestion price $\xi_n = \{\xi_l, \forall l \in L_n^{\text{out}}\}$, power message $m_n = \{m_l, \forall l \in L_n^{\text{out}}\}$, and power price $\zeta_n = \{\zeta_l, \forall l \in L_n^{\text{out}}\}$.

(2) Each transmitter $n \in N$ sends ξ_l back to the source nodes of the flows if link l is on the paths of the flows.

(3) Each transmitter passes m_n to corresponding transmitters by the routing protocol.

3. In the time slots belonging to the set $\Gamma_{D,n}$, each node $n \in N$ carries out the following algorithms at a period of T_D time slot:

(1) If the node belongs to the source node of a flow, the node updates the f_s^* according to Eq. (24).

(2) Each node updates the local channel allocation according to the solution of sub-problem (27) and informs other nodes the new results of local channel allocation.

(3) Each transmitter updates the transmitted power according to Eq. (34).

In the proposed distributed resource allocation algorithm, each node should calculate five parameters: congestion price ξ_n, power message m_n, power price ζ_n, traffic rate f_s^*, and the transmitted power; it also needs to solve the local channel allocation sub-problem. So, the computational complexity of each node is $O(5+CV_{\max})$.

In this study, the primal problem (14) is decomposed into sub-problems (22), (27), and (28). The convergence of the algorithm for sub-problem (22) is obvious. For (27), the convergence is guaranteed by global signal exchange and local exhaustive search. The sub-problem (28) is converted into a convex optimization problem after log-transformation. Depending on the strong dual theory, the distributed power update algorithm can converge to the joint and global optimum of sub-problem (28). So, the solution of dual problem (20) is guaranteed. Furthermore, according to the Lagrange dual theory, the solution of dual problem (20) is the lower bound of primal problem (14) due to non-convexity of primal problem (14). Therefore, the proposed algorithm is convergent and sub-optimal. We will further demonstrate the convergence and sub-optimality through the simulation results in the next section.

6 Simulation results and discussions

To evaluate the performance, we compare the proposed distributed sub-optimal resource allocation algorithm (RAA) with the following two algorithms by MATLAB simulation:

1. The centralized optimal RAA: The centralized algorithm is proposed in Section 4. This algorithm can find the global optimal solutions of NUM problem (14) for different ρ.

2. Conventional RAA (Huang *et al.*, 2011): Nodes jointly optimize the rate, power, and channel allocation without consideration of outdated CSI.

Furthermore, we do the performance comparison under three different states: perfect CSI, $\rho = 0.75$, and $\rho = 0.35$. The performance is evaluated in terms of network utility, fairness index, and energy efficiency. Network utility is defined by $U(f_s) = \log f_s$ in Mo and Walrand (2000). The fairness index (Jain *et al.*, 1984) is measured by $\frac{(\sum_{s\in S} f_s)^2}{(S\sum_{s\in S} f_s)^2}$. Power efficiency is defined as $\frac{\sum_{s\in S} f_s}{\sum_{l\in L} P_l}$.

In the simulated model, there are 15 wireless mesh routers which are randomly located in a 700 m×700 m square field and the one nearest to the center acts as the gateway. Each router is equipped with three radios and three orthogonal frequency channels are available in the network. Each non-gateway node transmits a flow through the shortest path to the gateway. In fact, we simulate a time varying dynamic environment where speed dependent Doppler spectrum is included (using Jakes' isotropic scattering model). The channel gain has a small-scale Rayleigh fading component and a large-scale path loss component with a path loss exponent of 4. The non-gateway nodes move randomly at a speed of 50 km/h and results are averaged over 4000 time slots. The parameters used in the simulations are listed in Table 1. We select the step-size parameters by using the method described in Section 1.2.1 of Bertsekas (1995).

Table 1 List of simulation parameters

Parameter	Value
Received noise power, n_l	1.0×10^{-11} W
Signal wavelength, λ	0.0517 m
Channel bandwidth, B	2 Mb/s
Processing gain, K	128
Step size of congestion price update, ξ	0.01
Step size of power price update, ε	0.01
Step size of power update, η	0.006
The maximum power constraint, $P_n^{\max}$	0.5 W

In Fig. 1, we plot the network utility curves of the optimal RAA, proposed distributed RAA, and conventional RAA with perfect CSI, $\rho = 0.75$, and 0.35, respectively. Several characteristics are noteworthy. First, as we can see, the network utility curves converge to within a neighborhood of the optimal values for different ρ. Thus, our proposed algorithm can lead to a near-optimal solution for NUM problem (14). Second, as expected, network utility decreases, as correlation coefficient ρ decreases, as shown by lines labeled as perfect CSI, $\rho = 0.75$, and $\rho = 0.35$. Third, for the perfect CSI situation, the proposed RAA and conventional RAA all use perfect CSI to optimize resource allocation; thus, they show the same performance. However, for $\rho = 0.75$ and $\rho = 0.35$, the proposed RAA shows obviously better performance than the conventional one. This is because nodes may assign an even higher (or lower) rate when the real links have bad (or good) channel conditions in the outdated-CSI-dependent conventional RAA. Unfortunately, both conditions lead to lower available f_s. In contrast, the proposed distributed RAA takes outdatedness of CSI into consideration to decrease the probability of misallocation, which improves network utility. In addition, with lower ρ, the performance improvement between the proposed distributed RAA and the conventional RAA is more obvious. Since higher ρ means that the outdated CSI is closer to the actual CSI, the allocated results by the conventional algorithm are more reasonable, so the possible compensated part by the proposed algorithm is limited.

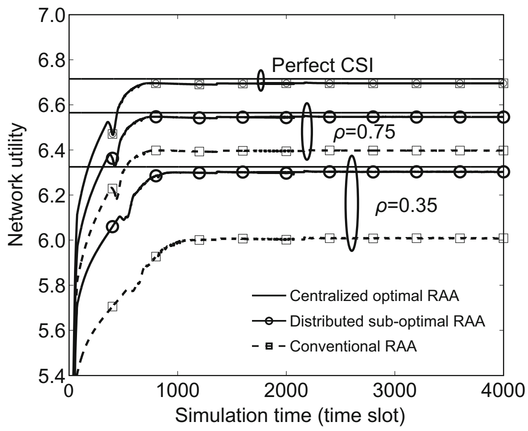

Fig. 1 Evolution of network utility

Fig. 2 shows the evolution of the fairness index. The fairness index may also converge to sub-optimal solutions. However, the conventional RAA takes outdated CSI as instantaneous CSI to allocate rates to users. The allocated rates may deviate from what are actually required, so the real fairness cannot be guaranteed. In contrast, the proposed RAA uses the conditional average capacity to replace the capacity computed with outdated CSI. By this means, the allocation error can be reduced and the results are made more in line with the actual situation. Thus, the proposed RAA shows a better fairness index than the conventional one in the context of outdated CSI.

The energy efficiency curves are shown in Fig. 3. The performance of the proposed RAA is obviously superior to that of the conventional one. The performance gaps are induced by two kinds of misjudgments. First, the conventional algorithm misjudges a good channel condition as a bad one, so the transmitted power is increased according to the power control algorithm. As a result, the available link capacity is

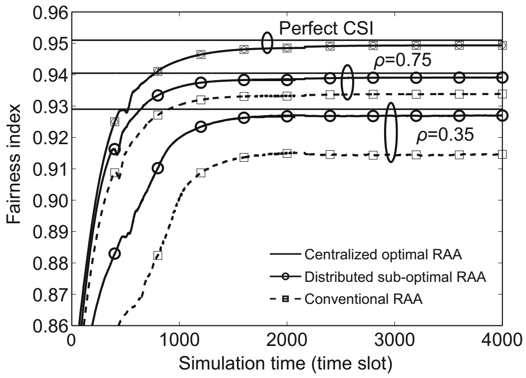

Fig. 2 Evolution of the fairness index

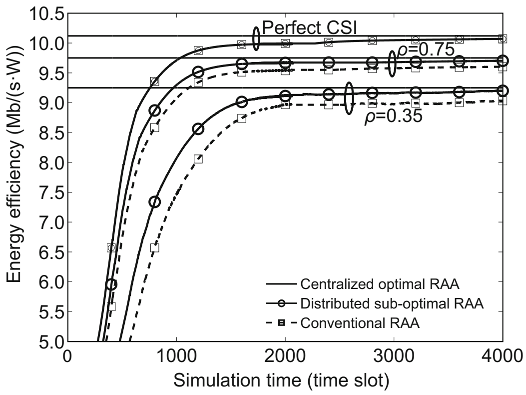

Fig. 3 Evolution of energy efficiency

increased. However, the actual flow rate through the link has been set to a relatively small value due to the supposed bad channel condition. As a consequence, the allocated power is useless and wasted. Second, the conventional algorithm misjudges a bad channel condition as a good one, so the transmitted power is decreased. Unfortunately, the link capacity cannot even satisfy the requirements of the allocated flow rate. Both misjudgments degrade the energy efficiency. In contrast, the proposed RAA decreases the misallocation and thus increases the energy efficiency.

We further investigate the performance dependence on node speed (Fig. 4). In the simulations, nodes move at speeds from 50 to 200 km/h, and results are averaged over 2000 time slots. As expected, the gaps between the perfect CSI cases and the compensated cases increase since the error after compensation increases for higher speeds. However, the proposed RAA outperforms the conventional RAA in all cases.

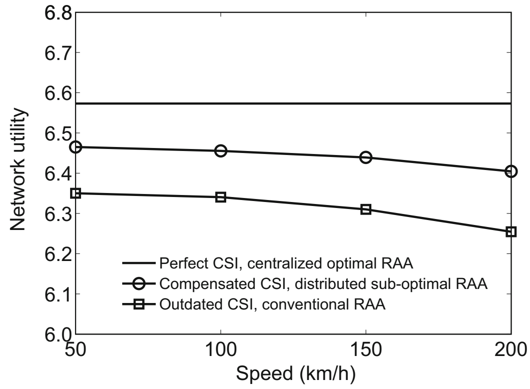

Fig. 4 Network utility with increasing speed

7 Conclusions

In this paper, we have first studied the problem of resource allocation in wireless multi-hop networks with outdated channel sensing under channel conditions where there is correlation between the outdated CSI and current CSI. The closed-form expression of the expectation of conditional capacity is obtained. According to this, both a centralized optimal and a distributed sub-optimal joint power, channel, and rate allocation algorithm are proposed in wireless multi-hop networks with outdated CSI. Based on the simulation results, we have verified that the proposed RAA can allocate the resource reasonably and achieve significantly higher network utility than the conventional one.

References

Aggarwal, R., Assaad, M., Koksal, C.E., *et al.*, 2011. Joint scheduling and resource allocation in the OFDMA downlink: utility maximization under imperfect channel-state information. *IEEE Trans. Signal Process.*, **59**(11):5589-5604. [doi:10.1109/TSP.2011.2162953]

Ahmad, A., Assaad, M., 2013. Optimal power and subcarriers allocation in downlink OFDMA system with imperfect channel knowledge. *Optim. & Eng.*, **14**(3):477-499. [doi:10.1007/s11081-011-9181-z]

Akyildiz, I.F., Wang, X.D., 2008. Cross-layer design in wireless mesh networks. *IEEE Trans. Veh. Technol.*, **57**(2):1061-1076. [doi:10.1109/TVT.2007.911615]

Augusto, C.H.P., Carvalho, C.B., da Silva, M.W.R., *et al.*, 2011. REUSE: a combined routing and link scheduling mechanism for wireless mesh networks. *Comput. Commun.*, **34**(18):2207-2216. [doi:10.1016/j.comcom.2011.01.008]

Bertsekas, D.P., 1995. Nonlinear Programming. Athena Scientific, Belmont, Massachusetts.

Bin, Z., Zhang, Z., Wei, H., *et al.*, 2013. Performance analysis for amplify-and-forward relay selection with outdated channel state information. *China Commun.*, **10**(6):100-110. [doi:10.1109/CC.2013.6549263]

Biyanwilage, S., Gunawardana, U., Liyanapathirana, R., 2011. Power allocation for nonregenerative OFDM relay links with outdated channel knowledge. Proc. of the 11th Int. Symp. on Communications and Information Technologies, p.428-432. [doi:10.1109/ISCIT.2011.6089964]

Boyd, S.P., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press, Cambridge, UK.

Chen, J.J., Wu, S.L., Lin, W.Y., 2013. A cross-layer design for energy efficient sleep scheduling in uplink transmissions of IEEE 802.16 broadband wireless networks. *Adv. Intell. Syst. Appl.*, **20**(1):635-644. [doi:10.1007/978-3-642-35452-6_64]

Chiang, M., 2005. Balancing transport and physical layers in wireless multihop networks: jointly optimal congestion control and power control. *IEEE J. Sel. Areas Commun.*, **23**(1):104-116. [doi:10.1109/JSAC.2004.837347(410)23]

Chiang, M., Low, S.H., Calderbank, A.R., *et al.*, 2007. Layering as optimization decomposition: a mathematical theory of network architectures. *Proc. IEEE*, **95**(1):255-312. [doi:10.1109/JPROC.2006.887322]

Cui, H., Song, L., Jiao, B., 2013. Weighted amplify-and-forward relay selection with outdated channel state information. *IEEE Wirel. Commun. Lett.*, **2**(6):651-654. [doi:10.1109/WCL.2013.091013.130529]

Gradshteyn, I.S., Ryzhik, I.M., Jeffrey, A., 2000. Table of Integral, Series, and Products. Academic Press, San Diego, USA.

Huang, X., Feng, S.L., Zhuang, H.C., 2011. Jointly optimal congestion control, channel allocation and power control in multi-channel wireless multihop networks. *Comput. Commun.*, **34**(15):1848-1857. [doi:10.1016/j.comcom.2011.05.004]

Jain, R., Chiu, D.M., Hawe, W.R., 1984. A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer Systems. Technical Report, No. DEC-TR-301, Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA.

Kim, H., Wang, H., Lim, S., *et al.*, 2012. On the impact of outdated channel information on the capacity of secondary user in spectrum sharing environments. *IEEE Trans. Wirel. Commun.*, **11**(1):284-295. [doi:10.1109/TWC.2011.112311.110307]

Li, X.C., Cao, F.M., Dong, X.H., *et al.*, 2011. Power control for delay constrained multi-channel communications using outdated CSI. *Wirel. Commun. Mob. Comput.*, **11**(9):1302-1311. [doi:10.1002/wcm.928]

Li, Y., Yin, Q., Xu, W., *et al.*, 2011. On the design of relay selection strategies in regenerative cooperative networks with outdated CSI. *IEEE Trans. Wirel. Commun.*, **10**(9):3086-3097. [doi:10.1109/TWC.2011.072511.110077]

Lim, S., Wang, H., Kim, H., *et al.*, 2012. Mean value-based power allocation without instantaneous CSI feedback in spectrum sharing systems. *IEEE Trans. Wirel. Commun.*, **11**(3):874-879. [doi:10.1109/TWC.2012.020712.102197]

Mo, J., Walrand, J., 2000. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. Network.*, **8**(5):556-567. [doi:10.1109/90.879343]

Nguyen, M.V., Duong, T.Q., Hong, C.S., *et al.*, 2012. Optimal and sub-optimal resource allocation in multi-hop cognitive radio networks with primary user outage constraint. *IET Networks*, **1**(2):47-57. [doi:10.1049/iet-net.2011.0028]

Papandriopoulos, J., Evans, J.S., 2006. Low-complexity distributed algorithms for spectrum balancing in multiuser DSL networks. Proc. IEEE Int. Conf. on Communications, p.3270-3275. [doi:10.1109/ICC. 2006.255311]

Qu, Q., Milstein, L.B., Vaman, D.R., 2010. Cross-layer distributed joint power control and scheduling for delay-constrained applications over CDMA-based wireless ad-hoc networks. *IEEE Trans. Commun.*, **58**(2):669-680. [doi:10.1109/TCOMM.2010.02.060069]

Rad, A.H.M., Wong, V.W.S., 2008. Cross-layer fair bandwidth sharing for multi-channel wireless mesh networks. *IEEE Trans. Wirel. Commun.*, **7**(9):3436-3445. [doi:10.1109/TWC.2008.070238]

Raniwala, A., Chiueh, T., 2005. Architecture and algorithms for an IEEE 802.11-based multi-channel wireless mesh network. Proc. 24th Annual Joint Conf. of the IEEE Computer and Communications Societies, p.2223-2234. [doi:10.1109/INFCOM.2005.1498497]

Shi, Y., Hou, Y.T., Kompella, S., *et al.*, 2011. Maximizing capacity in multihop cognitive radio networks under the SINR model. *IEEE Trans. Mob. Comput.*, **10**(7):954-967. [doi:10.1109/TMC.2010.204]

Xu, W., Dong, X.D., 2012. Optimized one-way relaying strategy with outdated CSI quantization for spatial multiplexing. *IEEE Trans. Signal Process.*, **60**(8):4458-4464. [doi:10.1109/TSP.2012.2197209]

Appendix A: Derivation for $\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X}, \boldsymbol{P})$

We first use the following equation (Ahmad and Assaad, 2013, Eq. (33)):

$$E_{\gamma|\hat{\gamma}}\left[\log(1+\kappa\gamma)\right] = \alpha + \beta\left[\log(\kappa\theta) + \psi(a)\right], \tag{A1}$$

where $\psi(\cdot)$ is the Euler psi function (Gradshteyn *et al.*, 2000, Section 8.360), $\alpha = \log(1+z) - \frac{z}{1+z}\log z$, $\beta = \dfrac{z}{1+z}$, $z = \kappa\hat{\gamma}$, $\theta = (\rho^2\hat{\gamma} + \dfrac{\sigma_w^2}{1-\rho^2})^2/a$, $a = (\dfrac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}+1)^2/(2\dfrac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}+1)$.

Because different channels (for different links (m, n)) are independent and identically distributed (i.i.d.) as defined in Section 2.2, we have

$$\begin{aligned}
&E\left[B\log_2\left(1+\frac{KP_l}{\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i\gamma_{il}+1}\gamma_{ll}\right)|(\hat{\gamma}_{1l},\hat{\gamma}_{2l},\cdots,\hat{\gamma}_{|L|l})\right]\\
&=\frac{B}{\log 2}\left\{\alpha_l+\beta_l\psi(a_l)+\beta_l E\left[\log\left(\frac{KP_l\theta_l}{\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i\gamma_{il}+1}\right)|(\hat{\gamma}_{1l},\hat{\gamma}_{2l},\cdots,\hat{\gamma}_{(l-1)l},\hat{\gamma}_{(l+1)l},\cdots,\hat{\gamma}_{|L|l})\right]\right\}\\
&=\frac{B}{\log 2}\alpha_l+\frac{B}{\log 2}\beta_l\psi(a_l)+\frac{B}{\log 2}\beta_l\log(KP_l\theta_l)\\
&\quad-\frac{B}{\log 2}\beta_l E\left[\log\left(\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i\gamma_{il}+1\right)|(\hat{\gamma}_{1l},\hat{\gamma}_{2l},\cdots,\hat{\gamma}_{(l-1)l},\hat{\gamma}_{(l+1)l},\cdots,\hat{\gamma}_{|L|l})\right],
\end{aligned}\tag{A2}$$

where $\alpha_l = \log(1+z_l) - \dfrac{z_l}{1+z_l}\log z_l$, $\beta_l = \dfrac{z_l}{1+z_l}$, $z_i = \dfrac{KP_l\hat{\gamma}_{ll}}{\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i\hat{\gamma}_{il}+1}$, $\theta_l = \left(\rho^2\hat{\gamma}_{ll} + \dfrac{\sigma_w^2}{1-\rho^2}\right)^2/a_l$, $a_l = \left(\dfrac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{ll}+1\right)^2/\left(2\dfrac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{ll}+1\right)$.

Let $M' = \{1', 2', \cdots, m'\} \subset L$ denote the interference link set of link l. We obtain $\sum_{i\neq l}\boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}P_i\gamma_{il}+1 = \sum_{i'\neq l, i'\in M'}P_{i'}\gamma_{i'l}+1$. So, based on Eq. (A2), Eq. (11) can be written as

$$\begin{aligned}
&E\left[B\log_2\left(1+\frac{KP_l}{\sum_{i\neq l}\boldsymbol{x}_i^*(\boldsymbol{x}_l^*)^{\mathrm{T}}P_i\gamma_{il}+1}\gamma_{ll}\right)|(\hat{\gamma}_{1l},\hat{\gamma}_{2l},\cdots,\hat{\gamma}_{|L|l})\right]\\
&=\frac{B}{\log 2}\alpha_l+\frac{B}{\log 2}\beta_l\psi(a_l)+\frac{B}{\log 2}\beta_l\log(KP_l\theta_l)\\
&\quad-\frac{B}{\log 2}\beta_l E\left[\log\left(\sum_{i'\neq l,i'\in M'}P_{i'}\gamma_{i'l}+1\right)|(\hat{\gamma}_{1'l},\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right].
\end{aligned}\tag{A3}$$

And the equation

$$\begin{aligned}
&E\left[\log\left(\sum_{i'\neq l,i'\in M'}P_{i'}\gamma_{i'l}+1\right)|(\hat{\gamma}_{1'l},\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]\\
&=E\left\{\log\left[\left(1+\sum_{i'\neq l,i'\neq 1',i'\in M'}P_{i'}\gamma_{i'l}\right)\left(1+\frac{P_1\gamma_{1'l}}{1+\sum_{i'\neq l,i'\neq 1',i'\in M'}P_{i'}\gamma_{i'l}}\right)\right]|(\hat{\gamma}_{1'l},\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right\}\\
&=E\left[\log\left(1+\sum_{i'\neq l,i'\neq 1',i'\in M'}P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]\\
&\quad+E\left[\log\left(1+\frac{P_{1'}\gamma_{1'l}}{1+\sum_{i'\neq l,i'\neq 1',i'\in M'}P_{i'}\gamma_{i'l}}\right)|(\hat{\gamma}_{1'l},\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]
\end{aligned}\tag{A4}$$

$$= E\left[\log\left(1+\sum_{i'\neq l, i'\neq 1', i'\in M'} P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]$$

$$+\alpha_{1'}+\beta_{1'}\psi(a_{1'})+\beta_{1'}\log(P_{1'}\theta_{1'})-\beta_{1'}E\left[\log\left(1+\sum_{i'\neq l, i'\neq 1', i'\in M'} P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]$$

$$=\alpha_{1'}+\beta_{1'}\psi(a_{1'})+\beta_{1'}\log(P_{1'}\theta_{1'})-(\beta_{1'}-1)E\left[\log\left(1+\sum_{i'\neq l, i'\neq 1', i'\in M'} P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]$$

holds. Let $\tau_i=\alpha_i+\beta_i\psi(a_i)+\beta_i\log(P_i\theta_i)$. Thus, Eq. (A4) can be simplified as follows:

$$\begin{aligned}
&E\left[\log\left(\sum_{i'\neq l, i'\in M'} P_{i'}\gamma_{i'l}+1\right)|(\hat{\gamma}_{1'l},\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]\\
&=\tau_{1'}-(\beta_{1'}-1)E\left[\log\left(1+\sum_{i'\neq l, i'\neq 1', i'\in M'} P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{2'l},\cdots,\hat{\gamma}_{m'l})\right]\\
&=\tau_{1'}-(\beta_{1'}-1)\left\{\tau_{2'}-(\beta_{2'}-1)E\left[\log\left(1+\sum_{i'\neq l, i'\neq 1', i'\neq 2', i'\in M'} P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{3'l},\cdots,\hat{\gamma}_{m'l})\right]\right\}\\
&=\tau_{1'}-\tau_{2'}(\beta_{1'}-1)+(\beta_{1'}-1)(\beta_{2'}-1)E\left[\log\left(1+\sum_{i'\neq l, i'\neq 1', i'\neq 2', i'\in M'} P_{i'}\gamma_{i'l}\right)|(\hat{\gamma}_{3'l},\cdots,\hat{\gamma}_{m'l})\right]\\
&=\tau_{1'}-\tau_{2'}(\beta_{1'}-1)+\tau_{3'}(\beta_{1'}-1)(\beta_{2'}-1)-\cdots\\
&\quad+(-1)^{m-1}(\beta_{1'}-1)(\beta_{2'}-1)...(\beta_{(m-1)'}-1)E\left[\log(1+P_{m'}\gamma_{m'l})|\hat{\gamma}_{m'l}\right]\\
&=\tau_{1'}-\tau_{2'}(\beta_{1'}-1)+\tau_{3'}(\beta_{1'}-1)(\beta_{2'}-1)-\cdots\\
&\quad+(-1)^{m-1}\tau_{m'}(\beta_{1'}-1)(\beta_{2'}-1)...(\beta_{(m-1)'}-1)+(-1)^{m}(\beta_{1'}-1)(\beta_{2'}-1)...(\beta_{m'}-1),
\end{aligned} \tag{A5}$$

where $\alpha_{i'}=\log(1+z_{i'})-\dfrac{z_{i'}}{1+z_{i'}}\log z_{i'}$, $\beta_{i'}=\dfrac{z_{i'}}{1+z_{i'}}$, $z_{i'}=\dfrac{KP_{i'}\hat{\gamma}_{i'l}}{\sum_{j'\in\{i'+1,i'+2,\dots,m'\}}P_{j'}\hat{\gamma}_{j'l}+1}$, $\theta_{i'}=\dfrac{1}{a_{i'}}\left(\rho^2\hat{\gamma}_{i'l}+\dfrac{\sigma_w^2}{1-\rho^2}\right)^2$, $a_{i'}=(\dfrac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{i'l}+1)^2/(2\dfrac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{i'l}+1)$, $i'\in M'$.

Finally, plugging Eq. (A5) into Eq. (A2), the expectation of conditional capacity under the SINR model can be derived as

$$\begin{aligned}
\bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{X},\boldsymbol{P})=\frac{B}{\log 2}\big[&\tau_l-\tau_{1'}\beta_l+\tau_{2'}\beta_l(\beta_{1'}-1)-\tau_{3'}\beta_l(\beta_{1'}-1)(\beta_{2'}-1)+\cdots+(-1)^m\tau_{m'}\beta_l\\
&\cdot(\beta_{1'}-1)(\beta_{2'}-1)...(\beta_{(m-1)'}-1)+(-1)^{m+1}\beta_l(\beta_{1'}-1)(\beta_{2'}-1)...(\beta_{m'}-1)+\beta_l\log K\big],
\end{aligned} \tag{A6}$$

where $\tau_i=\alpha_i+\beta_i\psi(a_i)+\beta_i\log_2(P_i\theta_i)$, $\alpha_i=\log_2(1+z_i)-\frac{z_i}{1+z_i}\log_2 z_i$, $\beta_i=\frac{z_i}{1+z_i}$,
$z_i=\begin{cases}\frac{KP_i\hat{\gamma}_{il}}{\sum_{j\in\{i+1,i+2,\dots,m'\}}P_j\hat{\gamma}_{jl}+1}, & \boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}=1, i\neq l, i\in M',\\ \frac{KP_i\hat{\gamma}_{ii}}{\sum_{j\in\{1,2,\dots,m'\}}P_j\hat{\gamma}_{ji}+1}, & i=l,\end{cases}$
$a_i=\begin{cases}\frac{\left(\frac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{il}+1\right)^2}{2\frac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{il}+1}, & \boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}=1, i\neq l, i\in M',\\ \frac{\left(\frac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{ii}+1\right)^2}{2\frac{\sigma_w^2}{1-\rho^2}\rho^2\hat{\gamma}_{ii}+1}, & i=l,\end{cases}$
$\theta_i=\begin{cases}\frac{1}{a_i}\left(\rho^2\hat{\gamma}_{il}+\frac{\sigma_w^2}{1-\rho^2}\right)^2, & \boldsymbol{x}_i\boldsymbol{x}_l^{\mathrm{T}}=1, i\neq l, i\in M',\\ \frac{1}{a_i}\left(\rho^2\hat{\gamma}_{ii}+\frac{\sigma_w^2}{1-\rho^2}\right)^2, & i=l,\end{cases}$ $\forall i\in M'\cup l$.

Appendix B: Derivation for $\nabla_l L(\tilde{\boldsymbol{P}}, \boldsymbol{\zeta})$

Rewrite Eq. (29) here:

$$L(\tilde{\boldsymbol{P}}, \boldsymbol{\zeta}) = \sum_{n\in N} \sum_{l\in L_n^{\rm in}} \xi_l \bar{C}_{l,\gamma|\hat{\gamma}}(\boldsymbol{P}) - \sum_{n\in N} \zeta_n \sum_{l\in L_n^{\rm in}} \left(\mathrm{e}^{\tilde{\boldsymbol{P}}_l} - P_n^{\max}\right). \tag{B1}$$

Let

$$f(\widetilde{\boldsymbol{P}}) = \sum_{n\in N} \sum_{l\in L_n^{\rm in}} \xi_l \bar{C}_{l,\gamma|\hat{\gamma}}(\widetilde{\boldsymbol{P}}) = \overbrace{\xi_l \bar{C}_{l,\gamma|\hat{\gamma}}(\widetilde{\boldsymbol{P}})}^{\text{for link } l} + \overbrace{\sum_{n\in N} \sum_{i\neq l, l\in L_n^{\rm in}} \xi_i \bar{C}_{i,\gamma|\hat{\gamma}}(\widetilde{\boldsymbol{P}})}^{\text{for link } i\neq l} = f_l(\tilde{\boldsymbol{P}}) + f_i(\tilde{\boldsymbol{P}}). \tag{B2}$$

According to Eq. (A2), we can obtain $\nabla_l f_l(\tilde{\boldsymbol{P}}) = \xi_l B\beta_l/\log 2$ with respect to $\tilde{P}_l$. Next, we need to obtain the $\nabla_l f_i(\tilde{\boldsymbol{P}})$ with respect to $\tilde{P}_l$. It is obvious that $\nabla_l f_i(\tilde{\boldsymbol{P}}) \neq 0$ only if link l belongs to the interference link set of link i. So, we define the interference link set of link i as $N' = \{l, 2', 3', \cdots, n'\} \subset L$. By plugging Eq. (A5) into Eq. (A2), we can derive that

$$\begin{aligned}
&\bar{C}_{i,\gamma|\hat{\gamma}}(\tilde{\boldsymbol{P}}) = E\left[B\log_2\left(1 + \frac{K\mathrm{e}^{\tilde{P}_i}\gamma_{ii}}{\sum_{k\neq i} \boldsymbol{x}_k^*(\boldsymbol{x}_i^*)^{\rm T}\mathrm{e}^{\tilde{P}_k}\gamma_{ki} + 1}\right) |(\hat{\gamma}_{1i}, \hat{\gamma}_{2i}, \cdots, \hat{\gamma}_{|L|i})\right] \\
&= \frac{B}{\log 2}\alpha_i + \frac{B}{\log 2}\beta_i\psi(a_i) + \frac{B}{\log 2}\beta_i \log(K\mathrm{e}^{\tilde{P}_i}\theta_i) \\
&\quad - \frac{B}{\log 2}\beta_i \left\{\alpha_l + \beta_l\psi(a_l) + \beta_l \log(\mathrm{e}^{\tilde{P}_l}\theta_l) - (\beta_l - 1)E\left[\log\left(1 + \sum_{j'\neq i, j'\neq l, j'\in N'} \mathrm{e}^{\tilde{P}_j}\gamma_{j'i}\right) |(\hat{\gamma}_{2'i}, \cdots, \hat{\gamma}_{n'i})\right]\right\}.
\end{aligned} \tag{B3}$$

Here, α,β,a,θ are constants and can be calculated using Eq. (A6). So, we plug Eq. (B3) into $f_i(\tilde{\boldsymbol{P}})$ and take the derivative of $f_i(\tilde{\boldsymbol{P}})$ with respect to $\tilde{P}_l$. As a consequence, we obtain

$$\nabla_l f_i(\tilde{\boldsymbol{P}}) = \sum_{n\in N} \sum_{i\neq l, l\in L_n^{\rm in}} \boldsymbol{x}_l^*(\boldsymbol{x}_i^*)^{\rm T} \frac{\xi_i B\beta_i\beta_{li}}{\log 2}, \tag{B4}$$

where $\beta_{li} = \beta_l$ is used to distinguish β_l in Eq. (B3) from that in $\nabla_l f_l(\tilde{\boldsymbol{P}})$. So far, it is straightforward that

$$\nabla_l L(\tilde{\boldsymbol{P}}, \boldsymbol{\zeta}) = \frac{\xi_l B\beta_l}{\log 2} - \sum_{n\in N} \left(\sum_{i\neq l, l\in L_n^{\rm in}} \boldsymbol{x}_l^*(\boldsymbol{x}_i^*)^{\rm T} \frac{\xi_i B\beta_i\beta_{li}}{\log 2} - \zeta_n \mathrm{e}^{\tilde{P}_l}\right), \tag{B5}$$

where β_l and β_i can be derived from Eq. (A2), and β_{il} can be obtained from Eq. (A5).