Front. Electr. Electron. Eng. China 2007, 2(4): 381–387
DOI 10.1007/s11460-007-0071-2

RESEARCH ARTICLE

CHEN Wei, CAO Zhigang, FAN Pingyi, Khaled B. Letaief

Cross layer optimization of downlink power allocation in multi-user wireless communication systems

© Higher Education Press and Springer-Verlag 2007

Abstract In multi-user wireless communication systems, dynamic power allocation is an important means to deal with the time-varying nature of the physical and network layers. However, the current layer optimization approach to power allocation cannot achieve the global optimum of the overall system performance. To solve this problem, a cross-layer optimization framework is presented for downlink power allocation, which takes both the channel and buffer states into account. A cross-layer optimization problem is formulated to optimize the total throughput with queue length and power constraints. An analytical solution and a low complexity dynamic programming algorithm, which are referred as water-filling in cellar (WFIC) policy, are presented to optimize the downlink power allocation. Finally, simulation results are presented to demonstrate the potential of the proposed method.

Keywords wireless network, power allocation, fading channel, cross-layer design

1 Introduction

In wireless communication systems, the channel time-variation and the power constraint are two fundamental characteristics. The recent research on wireless network mainly focuses on the problems caused by these two basic characteristics. The channel time-variation happens because of the multi-path nature of wireless environments. The power constraint is due to the following two facts. First, as frequency is reused in cellular networks, the signals transmitted by one base station may cause interference to the adjacent cells. To limit this adjacent cell interference, the transmission power of a base station should be bounded. Second, in some transmitters such as the satellite, the total available power is essentially limited. Now, dynamic power allocation is a fundamental means to increase the resource efficiency in time-varying wireless networks.

Resource allocation in wireless communications systems has received much attention in recent years. In the point-to-point scenario, Goldsmith and Varaiya [1] first obtained the optimal power allocation scheme and the ergodic capacity of fading channels. To minimize the average packet-delay, Collins and Cruz further studied joint channel-aware and queue-aware power control policies in two-state fading channels [2]. A more complicated and general cross-layer model is investigated by Berry and Gallager [3]. They proposed the asymptotic optimal tradeoff between average power and average delay. In Ref. [4], a low-complexity scheduler was proposed to achieve the suboptimal performance in terms of delay-power tradeoff. Recent work by Bettesh and Shamai adopted dynamic programming to calculate the optimal power allocation as the buffer size is relatively large [5]. Being interested in the single-link scenario, the above work has not considered the user domain power allocation as well as the total power constraint.

In multi-user downlink transmission, however, the total power constraint should be taken into account. Otherwise, each user will determine its required transmission power individually, despite the inherent power constraint in the base station. As a result, the total power required by the users may exceed the upper bound of the available power in the base station. To solve this problem, dynamic power allocation over user domain is extensively studied. Knopp and Humblet [6] presented a simple power control scheme, which is later referred to multi-user diversity. By utilizing multi-user diversity, the total power and the entire bandwidth are allocated to the best-channel user. This can maximize the sum-capacity of symmetric multi-user system with frequency non-selective channels. To maximize the stable region of queues, Neely et al. [7] proposed a discrete power allocation scheme based on Lyapunov function. A survey on cross layer power allocation, which maximizes the stable region, can be found in Ref. [8]. An iterative algorithm was given in Ref. [9] to minimize the energy in multi-user systems where randomly arrived packets should be transmitted in a given time

Received January 31, 2007; accepted April 3, 2007

CHEN Wei, CAO Zhigang (✉), FAN Pingyi, Khaled B. Letaief
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
E-mail: czg-dee@tsinghua.edu.cn

duration. In Refs. [10] and [11], dynamic resource allocation in orthogonal frequency division multiple access (OFDMA) system was considered.

In this paper, the downlink power allocation of multi-user wireless communication systems is the focus in this paper. It will be shown that the queue length determines the requirement of maximum capacity. If the power allocated to a user is more than needed to transmit all packets in the buffer, the redundant power will be wasted. Therefore, the constraints of the physical layer and the network layer should be considered simultaneously in power allocation. To that end, a cross-layer optimization approach is adopted to downlink power allocation. In this paper, the optimal power allocation problem, which considers the physical and network layer states simultaneously, will be presented. An analytical solution and a dynamic programming algorithm are proposed to solve this cross-layer optimization problem.

The conventional water filling policy, which only considers the physical layer constraint, will not be the optimal one and can be regarded as a special case of the proposed scheme. Throughout this paper, a base of 2 will be used in the logarithm function.

2 System model

Consider a multi-user wireless communication system, where centralized power allocation is performed in the downlink transmission, as shown in Fig. 1. Transmission bandwidth is shared by users in an orthogonal-channel manner. Each user's wireless link is time-varying with independent identical distributed fading process. Power control and adaptive modulation and coding (AMC) are adopted to adjust the transmission rate under the given fading state and bit error rate (BER) requirement. Each user has an individual buffer with independent identical distributed arrival process. The base station is assumed to have the perfect knowledge on the channel and buffer state information [10,11].

2.1 Physical layer

Block fading channel has been used to model the slowly varying flat-fading channel. Here N is considered as active users in system communicating over N independent discrete-time block-fading channels with independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN). It is assumed that N users' fading coefficient remained fixed during one time slot and M base-band symbols are transmitted by each user in one time slot. Here, M is determined by channel bandwidth W and the duration of each time slot T via Nyquist's theorem, i.e. $M = 2WT$. Let $h_i(k)$ denote the base-band complex channel gain of user i in the kth time slot, namely, $[kT,(k+1)T]$, which captures the effects of path-loss and multi-path fading. Let $P_i(k)$ denote the transmitted signal power allocated to user i in kth time slot. Assume each user's noise power in receiver is σ^2. The signal to noise ratio in receiver is then given by $\mathrm{SNR}_i(k) = |h_i(k)|^2\ P_i(k)/\sigma^2$. For simplicity, the equivalent noise power of user i in the kth time slot is defined as $\sigma_i^2(k) = \sigma^2 \big/ |h_i(k)|^2$, which represents the link quality rather than channel gain. According to Shannon's formula, the per-symbol capacity of user i in the kth time slot is given by

$$C_i(k) = \frac{1}{2}\log\left(1 + \frac{P_i(k)}{\sigma_i^2(k)}\right) \tag{1}$$

2.2 Network layer

The queue model in discrete-time continuous-length form is adopted in this part. The model can be obtained by sampling the fluid model with sample rate $f_s = 1/T$. The user i's queue length variation in the kth time slot is considered. It is assumed that the transmitter removes $s_i(k)$ bits from user i's buffer in time kT and encodes these into rate $s_i(k)/M$ codeword in order to transmit all in the kth time slot. Let $u_i(k)$ and $a_i(k)$ denote

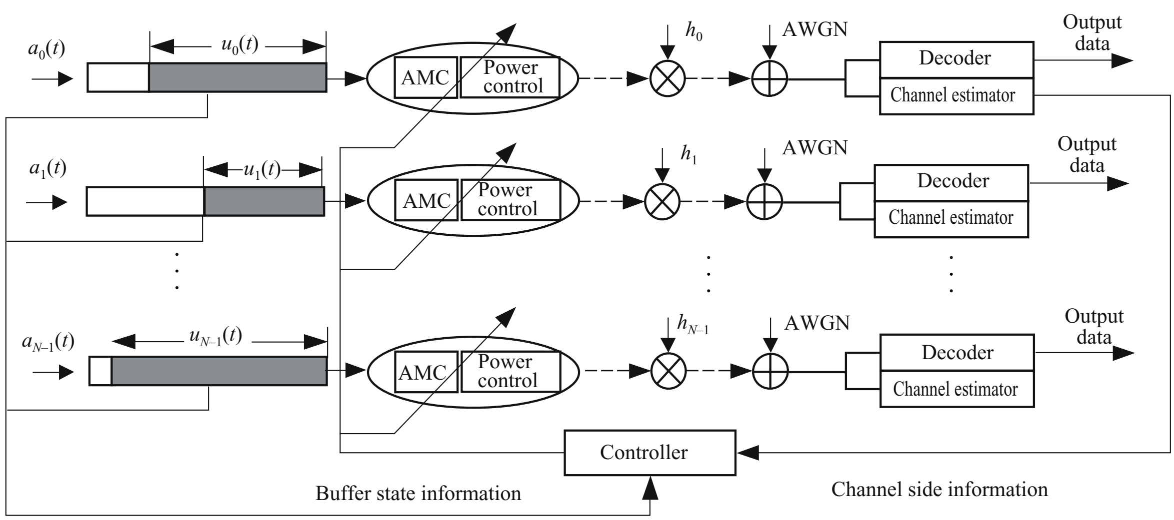

Fig. 1 Multi-user wireless communication system

the unfinished work of user i at the start of kth time slot and the traffic arriving at user i's buffer in kth time slot. In this case, the unfinished work of user i at the start of $(k+1)$th time slot is given by

$$u_i(k+1) = u_i(k) + a_i(k) - s_i(k) \tag{2}$$

2.3 Cross layer optimization problem

By integrating the block-fading channel model and discrete-time continuous-length queue model, a cross-layer model of system is presented in this part to describe the relationship between the physical layer and the network layer. Still the queue length variation of user i in the kth time slot is considered. The amount of bits removed from buffer is equal to that of bits transmitted. The information-theoretic optimal coding in the physical layer is considered. Therefore, the throughput and the capacity are related by

$$s_i(k) = 2WTC_i(k). \tag{3}$$

Because modulation and encoding scheme should be determined and the amount of bits in accordance with the scheme should be moved from buffer to encoder at the start of one time slot, only the traffic which arrived before time kT can be served in the kth time slot. Therefore, the amount of user i's traffic transmitted in the kth time slot should not be more than the amount of traffic in the buffer at the start of this time slot, i.e. $s_i(k) \leqslant u_i(k)$. Thus the capacity required by user i in this time slot is upper-bounded by

$$C_i(k) \leqslant \frac{u_i(k)}{2WT}. \tag{4}$$

The upper-bound of the capacity required by user i in the kth time slot is denoted by

$$R_i(k) = \frac{u_i(k)}{2WT}. \tag{5}$$

Equations (3) and (5) describe the relationship between the physical layer and the network layer. Because the physical layer capacity required by a user is upper-bounded by its buffer state, thus allocating redundant power to the user without enough data in its buffer cannot increase the total throughput, but on the contrary it will result in a waste of power. A reasonable optimal power allocation should take both the channel fading states and the buffer states of N users into account.

In most power allocation problems, the total available power is upper-bounded. Throughout this paper, the upper bound of available power in one time slot is denoted by $P_{\max}$. Thus the first kind of constraint, which is caused by the upper bound of available power, is given by

$$\sum_{i=0}^{N-1} P_i(k) \leqslant P_{\max}. \tag{6}$$

According to the above discussion on cross-layer model, the power allocated to one user should not exceed one's requirement determined by buffer state. Otherwise, the redundant power will be wasted. Consequently, the second kind of constraint, which is caused by the user's buffer state, is given by

$$\frac{1}{2}\log\left(1+\frac{P_i(k)}{\sigma_i^2(k)}\right) \leqslant R_i(k),\ \ \forall i. \tag{7}$$

The two kinds of constraints determine the feasible region of the power allocation problem in the kth time slot. Because users in the system belong to the same class, each user's average capacity over time domain is maximized if and only if the total capacity is maximized in each time slot. Our aim is to maximize the total capacity of N users by power allocation. In order to simplify the expression, the index k is dropped in the following discussion about power allocation in one time slot.

Based on the above, cross-layer optimization problem of power allocation can be formulated as follows

$$\max \sum_{i=0}^{N-1} \log\left(1+\frac{P_i}{\sigma_i^2}\right)$$
$$\text{subject to: } \begin{cases} \sum_{i=0}^{N-1} P_i \leqslant P_{\max} \\ \log\left(1+\frac{P_i}{\sigma_i^2}\right) \leqslant 2R_i,\ \ \forall i \end{cases}. \tag{8}$$

Note that Eq. (8) is a convex optimization problem. It has a unique optimal solution.

3 Optimal resource allocation

Let $\vec{P}^* = \left[P_0^*, P_1^*, \ldots, P_{N-1}^*\right]$ denote the optimal solution to the cross layer power allocation problem. Moreover, the objective function is denoted by $f(P_0, P_1, \ldots, P_{N-1})$. If the constraints of the maximal required capacity are removed from Eq. (8), the optimal solution corresponds to a water-filling power allocation [12], which is given by

$$P_i^* = \left(\frac{1}{2\theta} - \sigma_i^2\right)^+ \tag{9}$$

where θ denotes the value of Lagrange multiplier and for any real number x, $(x)^+ := \max(0, x)$. However, as each user's capacity is upper-bounded, the water-filling power allocation may not be optimal. The following theorem gives the necessary condition to optimal power allocation.

Theorem 1 The optimal solution to Eq. (8), $\vec{P}^*$, satisfies the following equality.

$$\left(\sum_{i=0}^{N-1} P_i^* - P_{\max}\right)\left\{\sum_{i=0}^{N-1}\left[\log\left(1+\frac{P_i^*}{\sigma_i^2}\right) - 2R_i\right]^2\right\} = 0 \tag{10}$$

Proof If Eq. (10) is not satisfied, then $\sum_{i=0}^{N-1} P_i^* < P_{\max}$, and there exists a user i, satisfying $\log\left(1+P_i^*/\sigma_i^2\right) < 2R_i$. This implies that there is a $\vec{P}^{**}$ in the feasible region, which satisfies $P_i^{**} > P_i^*$ and $P_j^{**} = P_j^*$, $\forall j \neq i$. Because $f(P_0, P_1, \ldots, P_{N-1})$ is a monotonously increasing function of P_i, $\forall i$, the value of objective function determined by $\vec{P}^{**}$ is greater than that determined by $\vec{P}^*$. It contradicts to the assumption that $\vec{P}^*$ is the optimal solution. Thus, Eq. (10) should be satisfied.

In order to solve Eq. (8), two lemmas are first presented as follows.

Lemma 1 For a given i, if $P_{\max} \geqslant \sum_{n=0}^{N-1} (2^{2R_i}\sigma_i^2 - \sigma_n^2)^+$, the optimal power allocation to user i can be given by $P_i^* = (2^{2R_i} - 1)\sigma_i^2$.

Proof Let $P_i^{\max}$ denote the maximal required power of user i. Because $P_i^{\max} = (2^{2R_i} - 1)\sigma_i^2$, it has $P_i^{\max} + \sigma_i^2 = 2^{2R_i}\sigma_i^2$. As $P_{\max} \geqslant \sum_{n=0}^{N-1} (2^{2R_i}\sigma_i^2 - \sigma_n^2)^+$, $P_n' = \left(P_i^{\max} + \sigma_i^2 - \sigma_n^2\right)^+$ can be allocated to user n. Therefore, if the allocated power of user i is less than $P_i^{\max}$, e.g. $P_i' = P_i^{\max} - \Delta_i$, $\Delta_i > 0$, there must exist a nonempty set J, for $\forall j \in J$, $P_j' \geqslant \left(P_i^{\max} + \sigma_i^2 - \sigma_j^2\right)^+ + \Delta_j$, $\Delta_j > 0$ and $\sum_{j \in J} \Delta_j = \Delta_i$. Also note that $(x)^+ \geqslant x$. As a result, $\forall j \in J$, $P_j' + \sigma_j^2 - \Delta_j \geqslant P_i^{\max} + \sigma_i^2$.

As

$$\left.\frac{\partial f}{\partial P_k}\right|_{\vec{P}=(P_0,P_1,\ldots,P_{N-1})} = \frac{\ln 2}{P_k + \sigma_k^2}$$

by differential mean value theorem, it has

$$f(\ldots, P_k + \Delta, \ldots) - f(\ldots, P_k, \ldots) = \frac{\ln 2}{\bar{P}_k + \sigma_k^2}\Delta$$

where $\bar{P}_k \in [P_k, P_k + \Delta]$.

Let $\Delta f\Big|_{(\ldots,P_k,\ldots)}^{(\ldots,P_k+\Delta,\ldots)} = f(\ldots, P_k + \Delta, \ldots) - f(\ldots, P_k, \ldots)$, it has

$$\Delta f\Big|_{(\ldots,P_i,\ldots)}^{(\ldots,P_i+\Delta_j,\ldots)} - \Delta f\Big|_{(\ldots,P_j,\ldots)}^{(\ldots,P_j-\Delta_j,\ldots)} = \left[\frac{\ln 2}{\bar{P}_i + \sigma_i^2} - \frac{\ln 2}{\bar{P}_j + \sigma_j^2}\right]\Delta_j.$$

Becaues for $\forall j \in J$,

$$\bar{P}_j + \sigma_j^2 \geqslant P_i^{\max} + \sigma_i^2 > \bar{P}_i + \sigma_i^2,$$

it follows that

$$\Delta f\Big|_{(\ldots,P_i,\ldots)}^{(\ldots,P_i+\Delta_j,\ldots)} - \Delta f\Big|_{(\ldots,P_j,\ldots)}^{(\ldots,P_j-\Delta_j,\ldots)} > 0.$$

Thus it means the capacity increment of user i is larger than the capacity decrement of user j when moving Δ_j from user j's power to user i's. Because there is $\sum_{j \in J} \Delta_j = \Delta_i$, it guarantees to have enough power to be reallocated to user i until its maximal required power $P_i^{\max}$ is achieved, and the total capacity is enlarged.

Lemma 2 If $\tilde{P}_n = (2^{2R_n} - 1)\sigma_n^2$ and $\tilde{P}_n$ is the nth entry of optimal solution vector to Eq. (8), let $\left(P_0^*, \ldots, P_{n-1}^*, P_{n+1}^*, \ldots, P_{N-1}^*\right)$ denote the optimal solution vector to the reduced-dimensional (RD) problem given by

$$\text{(RD)} \quad \begin{array}{l} \max \sum_{i=0, i\neq n}^{N-1} \log\left(1 + \frac{P_i}{\sigma_i^2}\right) \\ \text{subject to: } \begin{cases} \sum_{i=0, i\neq n}^{N-1} P_i \leqslant P_{\max} - \tilde{P}_n \\ \log\left(1 + \frac{P_i}{\sigma_i^2}\right) \leqslant 2R_i, \forall i \neq n \end{cases} \end{array}. \quad (11)$$

Then, the optimal solution vector to Eq. (8) can be obtained by $\left(P_0^*, \ldots, P_{n-1}^*, \tilde{P}_n, P_{n+1}^*, \ldots, P_{N-1}^*\right)$.

Proof From the constraints of Eq. (11), it has $\sum_{i=0,i\neq n}^{N-1} P_i^* \leqslant P_{\max} - \tilde{P}_n$ and $\log\left(1 + P_i^*/\sigma_i^2\right) \leqslant 2R_i$, $\forall\ i \neq n$, it has $\sum_{i=0,i\neq n}^{N-1} P_i^* + \tilde{P}_n \leqslant P_{\max}$. It is also noted that $\tilde{P}_n = (2^{2R_n} - 1)\sigma_n^2$, i.e. $\log\left(1 + \tilde{P}_n/\sigma_n^2\right) = 2R_n$. Thus, it has $\log\left(1 + \tilde{P}_n/\sigma_n^2\right) \leqslant 2R_n$ and $\log\left(1 + P_i^*/\sigma_i^2\right) \leqslant 2R_i$, $\forall i \neq n$. As a result, $(P_0^*, \ldots, P_{n-1}^*, \tilde{P}_n, P_{n+1}^*, \ldots, P_{N-1}^*)$ lies in the feasible region of Eq. (9).

The objective function of Eq. (9) can be rewritten as

$$f(P_0, P_1, \ldots, P_{N-1}) = \sum_{i=0, i\neq n}^{N-1} \log\left(1 + \frac{P_i}{\sigma_i^2}\right) + \log\left(1 + \frac{P_n}{\sigma_n^2}\right).$$

It is easy to see that the first term is maximized because $\left(P_0^*, \ldots, P_{n-1}^*, P_{n+1}^*, \ldots, P_{N-1}^*\right)$ is the optimal solution to Eq. (11). Also note that user n's required capacity is bounded by $R_n = \log\left(1 + \tilde{P}_n/\sigma_n^2\right)/2$, it follows that the second part is also maximized when allocating power $\tilde{P}_n$ to user n. Consequently, $\left(P_0^*, \ldots, P_{n-1}^*, \tilde{P}_n, P_{n+1}^*, \ldots, P_{N-1}^*\right)$ is the optimal solution to Eq. (8).

By Lemmas 1 and 2, a dynamic programming algorithm is presented to solve the power allocation problem (see Algorithm 1). The basic idea is to remove one user with minimal $2^{2R_i}\sigma_i^2$ from the current user set and to formulate a new RD problem.

Algorithm 1

initialize $P_{\max}$, σ_i^2, R_i, user_set $= \{0, 1, \ldots, N-1\}$;
$i = \arg\min_{k \in \text{user\_set}} 2^{2R_k}\sigma_k^2$;
while $\left(P_{\max} - \sum_{n \in \text{user\_set}} \left(2^{2R_i}\sigma_i^2 - \sigma_n^2\right)^+ \geqslant 0\right)$ do
  $P_i^* = (2^{2R_i} - 1)\sigma_i^2$;
  $P_{\max} = P_{\max} - P_i^*$;
  user_set = user_set$/\{i\}$;
  $i = \arg\min_{k \in \text{user\_set}} 2^{2R_k}\sigma_k^2$;
end while
if ($P_{\max} > 0$ and user_set $\neq \phi$) then

Traditional water-filling power allocation;

end if

Calculate each user's capacity and the sum capacity C^*;

End.

This algorithm is called "water-filling in cellar", see Fig. 2. Because the upper bound of each user's capacity is given, water is not filled into a pond with irregular bottom, but filled into a cellar with both irregular bottom and irregular ceil. This can be considered as an extension of traditional water-filling. The computation complexity of this algorithm is $O(N^2)$.

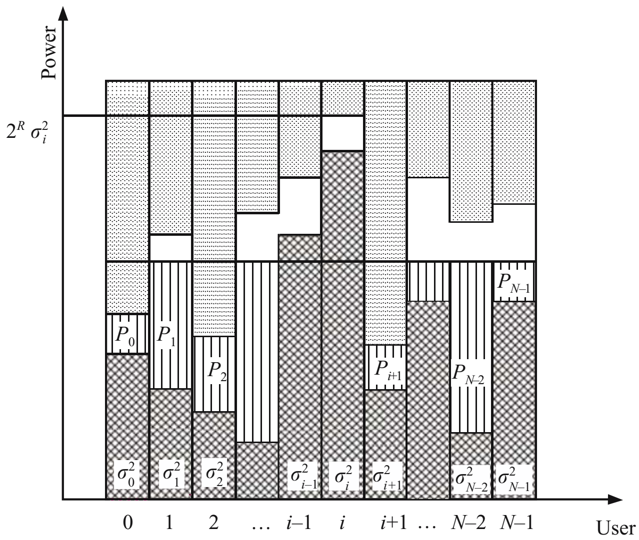

Fig. 2 Water-filling in cellar

According to the dynamic programming algorithm, the analytic solution to Eq. (8) can be obtained by

$$P_i^* = \min\left\{(2^{2R_i}-1)\sigma_i^2,\quad \left(\frac{1}{2\theta'}-\sigma_i^2\right)^+\right\}$$

where θ' denotes the value of Lagrange multiplier corresponding to the traditional water-filling power allocation in Algorithm 1. Here, it should be pointed out that Lemma 1 presents the necessary but not sufficient condition for the optimal power allocation. A stronger theorem is presented without detailed proof.

Theorem 2 The ith element of optimal solution to Eq. (8) is $P_i^* = (2^{2R_i}-1)\sigma_i^2$, if and only if

$$P_{\max} \geqslant \sum_{n=0}^{N-1} \min\left\{\left(2^{2R_i}\sigma_i^2 - \sigma_n^2\right)^+,\ \left(2^{2R_n}-1\right)\sigma_n^2\right\}. \qquad (12)$$

Proof By Algorithm 1, this conclusion is obvious.

4 Simulation results

In this section, the performance of optimal cross-layer power allocation with perfect channel and buffer state information is evaluated via the numerical experiments. Some simulations on comparison of our proposed water-filling in cellar scheme (WFIC) to two other familiar power allocation policies including traditional water-filling power allocation (TWF), (i.e. Eq. (9)) and constant power allocation (CPA), (i.e. $P_i = P_{\max}/N$) are presented.

4.1 Parameter settings

Consider a typical multi-user wireless communication system with parallel block fading channels, for example, a single FDMA cell. Each user employs a block fading channel with bandwidth $W = 2$ MHz and the central frequency of the system $f_c = 1\ 900$ MHz. Assume each user moving in the low speed $v \approx 1.5$ m/s. Then the channel coherence time can be obtained by

$$T_{\text{coh}} = c/vf_c \approx 105 \text{ ms}.$$

Throughout the simulation, a Rayleigh fading channel is assumed. Accordingly, the channel gain g_i of each user has an exponential distribution given by

$$p(g_i) = \bar{g}^{-1}\exp(-g_i/\bar{g}), \quad g_i > 0$$

where $\bar{g}$ denotes the average channel gain $E[|h_i|^2]$. Taking path-loss into account, we assume $\bar{g} = 0.1$. In addition, the average power of AWGN in receiver is assumed to be $\sigma^2 = 0.1$ for each user. In this case, the average SNR in receiver $\bar{\gamma}$ is the only parameter for physical layer when the channel model is given. In this section, two sample traffic models are considered, namely, Poisson arrival process and Constant arrival process. The arrival rate is denoted by λ.

4.2 Simulation results

The three schemes are first compared by presenting their average queue length against the arrival rate. Let the number of the active users be $N = 50$, and the average SNR in the receiver be $\bar{\gamma} = 10$ dB. In this case, the average transmitting power of each user is $\bar{P}_i = \bar{\gamma}\sigma^2/\bar{g} = 10$ for any i. Clearly, the total available power is $P_{\max} = 500$. The arrival rate λ is considered varying from 0.5 to 6 Mbps.

Figure 3 presents the comparison of average queue length of three different policies. The simulation results illustrated that the average queue length of WFIC is the shortest for arbitrary arrival rate in stable region. It can be found that the average queue length corresponding to TWF is the longest in slight loaded scenario. It is also found that the average queue length of the CPA will be the longest in heavy loaded scenario. The proposed WFIC, which considers both physical and network layer status, can not only avoid the power waste, but also increase the power efficiency as much as possible. Accordingly, the overall performance of our proposed scheme is optimal in both heavy and slight loaded scenario. When the arrival rate λ approaches the boundary of the stable region, the average queue lengths of TWF and WFIC are approximately equal. This is because the constraints of the maximal required power are always released as the queues are long enough.

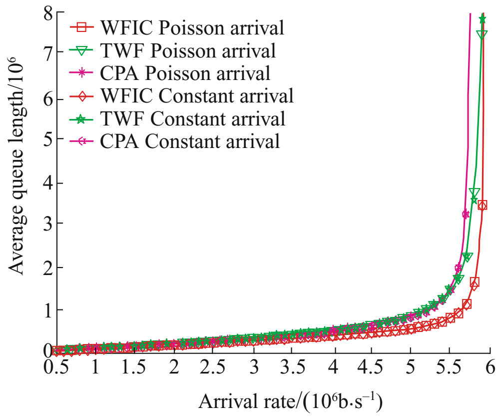

Fig. 3 Average queue length against arrival rate

Next, the average queue length against the number of the users will be presented. Consider a slight loaded system with arrival rate $\lambda = 1$ Mbps. Let the number of the user N vary from 1 to 30. The simulation parameters of the physical layer are the same as that used above. Hence, $P_{\max} = 10$ N.

Figure 4 presents the simulation results. It can be seen that the three schemes have completely different behaviors. As the number of active users increases, the average queue length corresponding to WFIC decreases and approaches an asymptotic value. This implies that larger number of users will result in higher power efficiency for WFIC. Moreover, if the traffic load is not too slight and $N > 20$, the average queue length can approximately approach the asymptotic value corresponding to $N \to \infty$. This implies that our proposed power allocation scheme can approximately achieve the performance limitation even in the practical system without very large quantity of users.

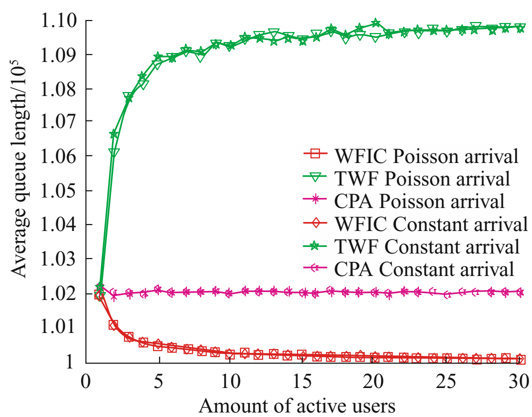

Fig. 4 Average queue length against the number of users

In contrast to WFIC, TWF presents the opposite performance that the average queue length grows larger as N increases. This is simply because that the increment of N will increase the probability of the short queue meeting good channel state. As a result, the service rate will decrease despite that each user's $\overline{\gamma}$ remains constant.

It is also perceivable that the average queue length of CPA is a constant as the number of active user increases. This is due to the fact that with this scheme, each user's allocated power and the average service rate do not depend on the number of the users.

In addition, the average queue length of all these three schemes is the same when the system serves only one active user. This can validate the correctness of the simulation results. From the simulation results above, it can be seen that WFIC can achieve higher performance gain over TWF or CPA when the system serves more users.

Finally, the average queue length against the average SNR is compared. Consider a system with $N = 50$ active users and arrival rate $\lambda = 2$ Mbps. Let $\overline{\gamma}$ vary from 3 to 30 dB. Figure 5 presents the simulation results. To achieve the same average queue length and/or delay, the WFIC can provide at least 5 dB SNR gain over TWF, and at least 3 dB SNR gain over CPA.

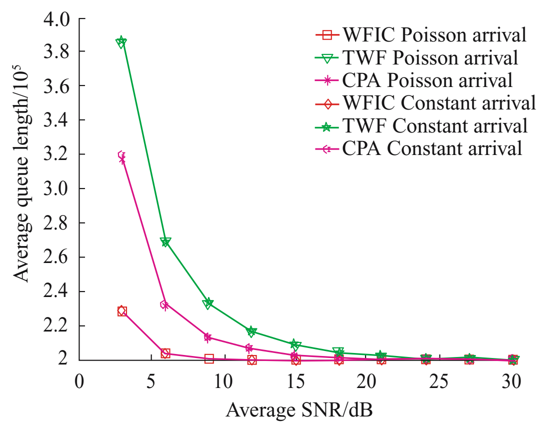

Fig. 5 Average queue length against the average SNR

5 Conclusion

In this paper, a cross-layer optimization approach for downlink power allocation in multi-user wireless communication systems is presented. In the proposed scheme, not only the channel fading state, but also the buffer state is considered simultaneously. A cross layer optimization problem on downlink power allocation is formulated on the basis of a cross-layer model consisting of both the physical layer and the network layer. An analytical optimal solution to the cross-layer model is proposed. In addition, a low complexity dynamic programming algorithm called "water-filling in cellar", which can calculate the optimal power allocation results with computation complexity $O(N^2)$, is presented. It is also shown that the conventional water-filling is a special case of the proposed scheme as the queue length is infinite. Finally, the sample simulation results are presented to demonstrate the potential of the proposed scheme. It can be shown that WFIC can provide at least 5 dB SNR gain over TWF and at least 3 dB SNR gain over CPA.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 60472027).

References

1. Goldsmith A J, Varaiya P P. Capacity of fading channels with channel side information. IEEE Transactions on Information Theory, 1997, 43: 1,986–1,992
2. Collins B, Cruz R. Transmission policies for time varying channels with average delay constraints. In: Proceedings of Allerton Conference on Communication, Control, and Computing. Monticello: IL, 1999
3. Berry R A, Gallager R G. Communication over fading channels with delay constraints. IEEE Transactions on Information Theory, 2002, 48: 1,135–1,149
4. Rajan D, Sabharwal A, Aazhang B. Delay-bounded packet scheduling of bursty traffic over wireless channels. IEEE Transactions on Information Theory, 2004, 50: 125–144
5. Bettesh I, Shamai S. Optimal power and rate control for minimal average delay: The single-user case. IEEE Transactions on Information Theory, 2006, 52: 4,115–4,141
6. Knopp R, Humblet P A. Information capacity and power control in single-cell multiuser communications. In: Proceedings of IEEE ICC, 1995, 1: 331–335
7. Neely M J, Modiano E, Rohrs C E. Power allocation and routing in multibeam satellites with time-varying channels. IEEE/ACM Transactions on Networking, 2003, 11: 138–152
8. Berry R A, Yeh E M. Cross-layer wireless resource allocation. IEEE Signal Processing Mag, 2004, 21: 59−68
9. Biyikoglu E, Gamal A El. On adaptive transmission for energy efficiency in wireless data networks. IEEE Transactions on Information Theory, 2004, 50: 3,081–3,094
10. Song G, Li Y. Utility-based resource allocation and scheduling in OFDM-based wireless broadband networks. IEEE Commun Mag, 2005, 43: 127–134
11. Letaief K B, Zhang Y J. Dynamic multiuser resource allocation and adaptation for wireless systems. IEEE Wireless Communication, 2006, 13: 38–47
12. Cover T M, Thomas J A. Elements of Information Theory. New York: Wiley, 1991