
ELSEVIER
SIMULATION MODELLING PRACTICE AND THEORY
Volume 19, issue 6, June 2011
ISSN 1569-190X
International Journal of the Federation of
European Simulation Societies
SIMULATION MODELLING
PRACTICE AND THEORY
Special Issue
Performance of Networks and Ubiquitous Computing Systems:
Techniques and Trends
Guest Editors: Xiaolong Jin, Ahmed Y. Al-Dubai, Geyong Min
and Mohamed Ould-Khaoua
Available online at www.sciencedirect.com
ScienceDirect

Simulation Modelling Practice and Theory 19 (2011) 1427–1444

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Downlink data transmission scheduling algorithms in wireless networks

Keqin Li *

Department of Computer Science, State University of New York, New Paltz, NY 12561, USA

A R T I C L E I N F O

Article history:
Available online 27 November 2010

Keywords:
Average-case performance
Downlink data transmission
Power assignment
Simulation
Transmission scheduling
Wireless network
Worst-case performance

A B S T R A C T

The problem of downlink data transmission scheduling in wireless networks is studied. It is pointed out that every downlink data transmission scheduling algorithm must have two components to solve the two subproblems of power assignment and transmission scheduling. Two types of downlink data transmission scheduling algorithms are proposed. In the first type, power assignment is performed before transmission scheduling. In the second type, power assignment is performed after transmission scheduling. The performance of two algorithms of the first type which use the equal power allocation method are analyzed. It is shown that both algorithms exhibit excellent worst-case performance and asymptotically optimal average-case performance under the condition that the total transmission power is equally allocated to the channels. In general, both algorithms exhibit excellent average-case performance. It is demonstrated that two algorithms of the second type perform better than the two algorithms of the first type due to the equal time power allocation method. Furthermore, the performance of our algorithms are very close to the optimal and the room for further performance improvement is very limited. It is shown that all the above algorithms can be extended to schedule downlink data transmissions with parallel channels. It is also shown that the simple sequential scheduling algorithm is optimal if the total transmission power is equally allocated to the channels. As an extra contribution, an M/G/1 queueing model for the FCFS queueing discipline is established, and it is observed that increasing the number of channels has more impact on the reduction of the average response time than increasing the total transmission power.

1. Introduction

1.1. Motivation

The new generation (3G and beyond) wireless networks provide multiple channels (such as codes, frequency tones, and time slots) through code division multiple access (CDMA), wideband orthogonal frequency division multiplexing (OFDM), and multislot time division multiple access (TDMA). These channels can be allocated to users with different transmission powers and rates. These advancements provide more flexibility in network traffic control and raise new interesting power and channel allocation and data transmission scheduling problems [2,3,11,14,16].

In a wireless network, there is a base station in each cell. The base station handles all requests of data transmission to (downlink) and from (uplink) mobile users in the same cell. It is expected that in emerging wireless networks, data traffic has asymmetrically large downlink demand. Given certain amount of transmission power and certain number of communication channels, and a set of downlink data transmission requests, the downlink data transmission scheduling problem is to

* Tel.: +1 845 257 3534; fax: +1 845 257 3996.
E-mail address: lik@newpaltz.edu

1569-190X/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.simpat.2010.11.010

find a power assignment to the transmission requests and a nonpreemptive schedule of the transmission requests, such that the total time to complete the set of data transmissions is minimized (see Section 2 for a more accurate definition of the problem).

Since both transmission power and communication channels are critical resources in wireless networks, it is an important research problem to find and develop power assignment and transmission scheduling algorithms to effectively and efficiently utilize transmission power and communication channels and to reduce the data transmission times. Unfortunately, there has been little such study in the literature. The motivation of this paper is to investigate heuristic algorithms for downlink data transmission scheduling in wireless networks. To the best of the author's knowledge, this is the first attempt to study nonpreemptive downlink data transmission scheduling algorithms.

1.2. Summary of contributions

It is clear that the downlink data transmission scheduling problem contains two subproblems, namely, *power assignment* and *transmission scheduling*. Hence, every downlink data transmission scheduling algorithm must have two components to solve the two subproblems. In this paper, we consider two types of downlink data transmission scheduling algorithms, depending on which subproblem is solved first. The name of an algorithm is represented as *X–Y*, where *X* is the strategy for power assignment and *Y* is the strategy for transmission scheduling.

In the first type, power assignment is performed before transmission scheduling. In our algorithms, the total transmission power *P* is equally allocated to the channels, i.e., data transmissions are scheduled with *equal powers* (*EP*). We examine the performance of algorithms *EP-LS* and *EP-LPT* which use the well known online *list scheduling* (*LS*) and the offline *longest processing time* (*LPT*) algorithms to schedule downlink data transmissions in wireless networks. It is shown in Sections 3–5 that both algorithms exhibit excellent worst-case performance and asymptotically optimal average-case performance under the condition that the total transmission power is equally allocated to the channels. In general, both algorithms exhibit excellent average-case performance.

In the second type of downlink data transmission scheduling algorithms, power assignment is performed after transmission scheduling. First, data transmissions are scheduled using *virtual times* which give certain measure of actual transmission times. Then, the total transmission power *P* is allocated to the channels such that all the channels complete their data transmissions at the same time. We demonstrate in Section 6 that algorithms *ET-LS* and *ET-LPT*, which schedule data transmissions using algorithms *LS* and *LPT* and assign powers to achieve *equal times* (*ET*), perform better than algorithms *EP-LS* and *EP-LPT*.

Since the preliminary version of the paper [13] was presented, we have extended our work by considering downlink data transmission scheduling with parallel channels. We show that by using the equal channel allocation method, all the above algorithms can be extended to schedule downlink data transmissions with parallel channels and improved performance can be obtained in Section 7. We also show that when parallel channels can be allocated to data transmission requests, the simple sequential scheduling algorithm is optimal if the total transmission power is equally allocated to the channels. As an extra contribution, we establish an M/G/1 queueing model for the first-come-first-served (FCFS) queueing discipline and observe that increasing the number of channels has more impact on the reduction of the average response time than increasing the total transmission power.

2. Preliminaries

2.1. The data transmission model

Assume that the base station has total transmission power *P*. There are *C* wireless channels. The power *P* can be divided and allocated to the channels.

We adopt the channel specifications similar to the original 3G system proposals [3,4,15] for our data transmission model. Each communication channel is characterized by the *signal to interference plus noise ratio* (*SINR*) [8] given by

$$SINR = \frac{gp}{\sigma^2}.$$

In the above equation, p is the power assigned to the channel and σ^2 is the total noise power including interference. Power attenuation of a channel is specified by the parameter

$$g = \frac{S}{d^\alpha},$$

called the physical gain, which is determined by the shadow loss component S, the distance d between a mobile user and the base station, and the distance loss exponent α.

The transmission rate (in bits per second) of a channel is given by the following equation:

$$r = W\log_2\left(1 + \frac{SINR}{\Gamma}\right),$$

where W is the spectral bandwidth used and Γ is dependent on the coding gain from the physical layer error-correcting code [5]. The equation for r can be rewritten as

$$r = W\log_2\left(1 + \frac{S}{\sigma^2\Gamma}\cdot\frac{p}{d^\alpha}\right).$$

Since S, σ^2, and Γ are all constants, they can be combined into one single constant

$$\beta = \frac{S}{\sigma^2\Gamma}.$$

Hence, the equation for r becomes

$$r = W\log_2\left(1 + \beta\cdot\frac{p}{d^\alpha}\right).$$

2.2. The scheduling problem

A downlink data *transmission request* is specified by two parameters (s_i, d_i), namely, the number s_i of bits to be transmitted (called the *size* of the request) and the distance d_i from the base station. If power p_i is allocated to the request, the transmission rate is

$$r_i = W\log_2\left(1 + \beta\cdot\frac{p_i}{d_i^\alpha}\right),$$

and the transmission time is

$$t_i = \frac{s_i}{r_i} = \frac{s_i}{W\log_2\left(1 + \beta\cdot\frac{p_i}{d_i^\alpha}\right)}.$$

A *schedule* for a transmission request is a pair (b_i, j_i), which means that a transmission begins at time $b_i \geqslant 0$ and ends at time $b_i + t_i$ on channel j_i, where $1 \leqslant j_i \leqslant C$. A schedule is nonpreemptive in the sense that a transmission cannot be interrupted and resumed later.

Our *downlink data transmission scheduling* problem can be defined as follows: given total transmission power P and C channels, constants α and β and W, and n downlink data transmission requests specified by (s_i, d_i), $1 \leqslant i \leqslant n$, find a power assignment p_i, $1 \leqslant i \leqslant n$, and a nonpreemptive schedule (b_i, j_i), $b_i \geqslant 0$, $1 \leqslant j_i \leqslant C$, $1 \leqslant i \leqslant n$, for all the n transmission requests, such that the total time to complete the n data transmissions is minimized.

Notice that at any moment, there are at most C transmissions simultaneously and the sum of powers on the C channels cannot exceed P. As pointed out in [3], this is a scheduling problem with both resource (i.e., power) constraint and malleability (i.e., more power results in shorter transmission time). The preemptive version of the problem has been studied in [3], and in this paper, we deal with the nonpreemptive version. It is clear that nonpreemptive scheduling has much less management overhead and easier to implement than preemptive scheduling in real networks, and is definitely worth of investigation.

3. Scheduling with equal powers (EP)

3.1. Power assignment

Our first type of downlink data transmission scheduling algorithms are developed based on the method of *equal power* (*EP*) allocation. If the total transmission power P is equally allocated to the C channels and the n transmission requests, i.e.,

$$p_i = \gamma = \frac{P}{C},$$

for all $1 \leqslant i \leqslant n$, where γ is the average power per channel, the transmission rate of a request is

$$r_i = W\log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right),$$

and the transmission time is

$$t_i = \frac{s_i}{W\log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)},$$

which is solely determined by a data transmission request itself. It is clear that our scheduling problem is now equivalent to the classic multiprocessor scheduling problem [9,10], where a transmission request becomes a task with processing time t_i, and the C channels are C identical processors. The multiprocessor scheduling problem is NP-hard (see p. 238 of [7]).

One effective way to solve NP-hard problems is to use approximation algorithms that produce near-optimal schedules. Let A also denote the length of a schedule produced by algorithm A, and OPT denote the optimal schedule length. We say that algorithm A is *B-approximate* if

$$R^A = \frac{A}{OPT} \leqslant B,$$

for all inputs, and B is called a *worst-case performance bound*.

3.2. Transmission scheduling

3.2.1. Online scheduling

Our downlink data transmission scheduling algorithm *EP-LS* assigns *equal power* (*EP*) γ to all the n transmission requests and then employs the classic *list scheduling* (*LS*) algorithm to schedule the n data transmissions. The *LS* algorithm (and therefore, the *EP-LS* algorithm) is an *online* scheduling algorithm [9], i.e., data transmission requests are scheduled in the order of their arrival times. Given n downlink data transmission requests in any order, the *EP-LS* algorithm schedules them as follows. Whenever a channel becomes available, the next request in the list is scheduled to start transmission on that channel. It is clear that the *EP-LS* algorithm has the same worst-case performance bound as algorithm *LS* for channels with equal powers.

Theorem 1. *If the total transmission power is equally allocated to the channels, the EP-LS algorithm is (2 − 1/C)-approximate.*

Proof. See Ref. [9]. □

Now we further examine the performance of the *EP-LS* algorithm in scheduling downlink data transmission requests. Let OPT_{EP} be the optimal schedule length if all the channels are allocated equal power γ. Lemma 1 is a simple observation of OPT_{EP}.

Lemma 1. *If the total transmission power is equally allocated to the channels, the minimum schedule length is at least*

$$OPT_{EP} \geqslant \frac{1}{CW}\sum_{i=1}^{n}\frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. It is clear that an obvious lower bound for OPT_{EP} is

$$OPT_{EP} \geqslant \frac{1}{C}\sum_{i=1}^{n} t_i.$$

If the total transmission power is equally allocated to the channels, we have

$$t_i = \frac{s_i}{W\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)},$$

for all $1 \leqslant i \leqslant n$. We get the lemma. □

Lemma 2 is a well known observation of the *LS* algorithm (and therefore, the *EP-LS* algorithm).

Lemma 2. *The length of an EP-LS schedule is at most*

$$EP\text{-}LS \leqslant \frac{1}{CW}\sum_{i=1}^{n}\frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)} + \left(1-\frac{1}{C}\right)\max_{1\leqslant i\leqslant n}\frac{s_i}{W\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. Let t_k be the transmission time of the transmission request that is completed last. Assume that all data transmissions are performed in the time interval $[0, EP\text{-}LS]$. It is easy to see that during the time interval $[0, EP\text{-}LS - t_k]$, all channels are busy, i.e., there is no idle channel. Hence, we get

$$\sum_{i=1}^{n} t_i - t_k \geqslant C(EP\text{-}LS - t_k),$$

which gives

$$EP\text{-}LS \leqslant \frac{1}{C}\sum_{i=1}^{n} t_i + \left(1 - \frac{1}{C}\right)t_k.$$

Since

$$t_k \leqslant \max_{1 \leqslant i \leqslant n}(t_i),$$

we have

$$EP\text{-}LS \leqslant \frac{1}{C}\sum_{i=1}^{n} t_i + \left(1 - \frac{1}{C}\right)\max_{1 \leqslant i \leqslant n}(t_i).$$

The rest of the proof follows the value of t_i. □

Combining Lemmas 1 and 2, we reach the following claim.

Theorem 2. *The length of an EP-LS schedule satisfies*

$$R_{EP}^{EP\text{-}LS} = \frac{EP\text{-}LS}{OPT_{EP}} \leqslant B_{EP}^{EP\text{-}LS},$$

where

$$B_{EP}^{EP\text{-}LS} = 1 + \frac{(C-1)\max\limits_{1 \leqslant i \leqslant n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)}}{\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)}},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. The theorem is a straightforward consequence of Lemmas 1 and 2. □

Notice that the worst-case performance bound $(2 - 1/C)$ in Theorem 1 is achieved under the condition that the total transmission power is equally allocated to the channels. To analyze the worst-case performance of algorithm *EP-LS* without such a condition, we need a lower bound for *OPT* in general. We introduce the notation t_i^* which represents the minimal transmission time of the request (s_i, d_i). Clearly, we have $t_i \geqslant t_i^*$ and

$$t_i^* = \frac{s_i}{W\log_2\left(1 + \frac{\beta P}{d_i^{\alpha}}\right)},$$

that is, t_i^* is the transmission time when all of the power P is allocated to (s_i, d_i).

Lemma 3. *The minimum schedule length is at least*

$$OPT \geqslant \frac{1}{CW}\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1 + \frac{\beta P}{d_i^{\alpha}}\right)},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. It is clear that an obvious lower bound for *OPT* is

$$OPT_{EP} \geqslant \frac{1}{C}\sum_{i=1}^{n} t_i^*.$$

The rest of the proof follows the value of t_i^*. □

Combining Lemmas 2 and 3, we have the following result.

Theorem 3. *The length of an EP-LS schedule satisfies*

$$R^{EP\text{-}LS} = \frac{EP\text{-}LS}{OPT} \leqslant B^{EP\text{-}LS},$$

where

$$B^{EP\text{-}LS} = \frac{\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)} + (C-1)\max_{1\leqslant i\leqslant n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)}}{\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta P}{d_i^{\alpha}}\right)}},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. The theorem is a straightforward consequence of Lemmas 2 and 3. □

3.2.2. Offline scheduling

Our downlink data transmission scheduling algorithm *EP-LPT* also assigns equal power γ to all the n transmission requests, but employs the classic *longest processing time* (*LPT*) algorithm to schedule the n data transmissions. The *LPT* algorithm (and therefore, the *EP-LPT* algorithm) is an *offline* scheduling algorithm [10], i.e., all the n transmission requests should be available before a transmission schedule can be produced. The *LPT* algorithm is similar to *LS* except that the n downlink data transmission requests are arranged in nonincreasing order of the t_i's. It is clear that the *EP-LPT* algorithm has the same worst-case performance bound as algorithm *LPT* for channels with equal powers.

Theorem 4. *If the total transmission power is equally allocated to the channels and transmission requests, the EP-LPT algorithm is (4/3 − 1/(3C))-approximate.*

Proof. See Ref. [10]. □

Now we further examine the performance of the *EP-LPT* algorithm in scheduling downlink data transmission requests. Lemma 4 is a simple observation of the *EP-LPT* algorithm.

Lemma 4. *The length of an EP-LPT schedule is at most*

$$EP\text{-}LPT \leqslant \frac{1}{CW}\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)} + \left(1-\frac{1}{C}\right)\frac{s_k}{W\log_2\left(1+\frac{\beta\gamma}{d_k^{\alpha}}\right)},$$

where t_k *is the transmission time of the transmission request that is completed last, for any set of n downlink data transmission requests* (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. The proof follows the same argument as that of Lemma 1, i.e.,

$$EP\text{-}LPT \leqslant \frac{1}{C}\sum_{i=1}^{n} t_i + \left(1-\frac{1}{C}\right)t_k.$$

The rest of the proof follows the values of t_i and t_k. □

Combining Lemmas 1 and 4, we have the following result.

Theorem 5. *The length of an EP-LPT schedule satisfies*

$$R_{EP}^{EP\text{-}LPT} = \frac{EP\text{-}LPT}{OPT_{EP}} \leqslant B_{EP}^{EP\text{-}LPT},$$

where

$$B_{EP}^{EP\text{-}LPT} = 1 + \frac{(C-1)\frac{s_k}{\log_2\left(1+\frac{\beta\gamma}{d_k^{\alpha}}\right)}}{\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)}},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. The theorem is a straightforward consequence of Lemmas 1 and 4. □

Combining Lemmas 3 and 4, we have the following result.

Theorem 6. *The length of an EP-LPT schedule satisfies*

$$R^{EP\text{-}LPT} = \frac{LPT}{OPT} \leqslant B^{EP\text{-}LPT},$$

where

$$B^{EP\text{-}LPT} = \frac{\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta\gamma}{d_i^{\alpha}}\right)} + (C-1)\frac{s_k}{\log_2\left(1+\frac{\beta\gamma}{d_k^{\alpha}}\right)}}{\sum_{i=1}^{n} \frac{s_i}{\log_2\left(1+\frac{\beta P}{d_i^{\alpha}}\right)}},$$

for any set of n downlink data transmission requests (s_i, d_i), $1 \leqslant i \leqslant n$.

Proof. The theorem is a straightforward consequence of Lemmas 3 and 4. □

4. Average-case analysis

When data transmission requests are random variables, both A and OPT are random variables. One average-case performance measure of an approximation algorithm A is $E[R^A]$, namely, $E[A/OPT]$, where $E[\cdot]$ stands for the expectation of a random variable. If

$$\frac{A}{OPT} \leqslant B,$$

then

$$E\left[\frac{A}{OPT}\right] \leqslant E[B],$$

and $E[B]$ is called an *average-case performance bound of type I*. From Theorems 2, 3, 5, and 6, we know that $E[B_{EP}^{EP\text{-}LS}]$, $E[B^{EP\text{-}LS}]$, $E[B_{EP}^{EP\text{-}LPT}]$, and $E[B^{EP\text{-}LPT}]$ give average-case performance bounds of type I for algorithms *EP-LS* and *EP-LPT*, with or without the condition that the total transmission power is equally allocated to the channels. While these expectations are hard to calculate analytically, they can be obtained from statistics of random sampling (see the next section).

The average-case performance of an approximation algorithm can also be measured by $E[A]/E[OPT]$. If

$$\overline{R}^A = \frac{E[A]}{E[OPT]} \leqslant \overline{B},$$

then $\overline{B}$ is called an *average-case performance bound of type II*.

We make the following assumptions for the purpose of deriving average-case performance bounds of type II for algorithms *EP-LS* and *EP-LPT*:

- The s_i's are i.i.d. random variables with a common probability distribution.
- The d_i's are i.i.d. random variables with a common probability distribution.
- The distributions of the s_i's and the d_i's are independent of each other.

The above assumptions imply that the t_i's are also i.i.d. random variables with a common probability distribution. Let $\bar{t}$, σ_t^2, c_t be the mean, variance, and coefficient of variation of the t_i's.

The following result is well known from order statistics.

Lemma 5. *If the* t_i*'s are i.i.d. random variables, we have*

$$E\left[\max_{1\leqslant i\leqslant n}(t_i)\right] \leqslant \bar{t} + \frac{n-1}{\sqrt{2n-1}} \cdot \sigma_t,$$

for all $n \geqslant 1$.

Proof. See p. 62 of [6]. □

The following theorem gives an average-case performance bounds of type II for algorithms *EP-LS* with the condition that the total transmission power is equally allocated to the channels.

Theorem 7. *The length of an EP-LS schedule satisfies*

$$\overline{R}_{EP}^{EP\text{-}LS} = \frac{E[EP\text{-}LS]}{E[OPT_{EP}]} \leqslant \overline{B}_{EP}^{EP\text{-}LS},$$

where

$$\overline{B}_{EP}^{EP\text{-}LS} = 1 + \frac{C-1}{n}\left(1 + \frac{n-1}{\sqrt{2n-1}} \cdot c_t\right),$$

for all $n \geqslant 1$. As $n \to \infty$, we have $\overline{B}_{EP}^{EP\text{-}LS} \to 1$, i.e., algorithm EP-LS is asymptotically optimal if the total transmission power is equally allocated to the channels.

Proof. From Lemma 1, we get

$$E[OPT_{EP}] \geqslant \frac{n\bar{t}}{C},$$

and from Lemmas 2 and 5, we get

$$E[EP\text{-}LS] \leqslant \frac{n\bar{t}}{C} + \left(1 - \frac{1}{C}\right)\left(\bar{t} + \frac{n-1}{\sqrt{2n-1}} \cdot \sigma_t\right).$$

Combining the above two inequalities, we obtain the theorem. □

Similarly, the t_i^*'s are also i.i.d. random variables with a common probability distribution. Let $\overline{t^*}$ be the mean of the t_i^*'s. The following theorem gives an average-case performance bounds of type II for algorithms *EP-LS* without the condition that the total transmission power is equally allocated to the channels.

Theorem 8. *The length of an LS schedule satisfies*

$$\overline{R}^{EP\text{-}LS} = \frac{E[EP\text{-}LS]}{E[OPT]} \leqslant \overline{B}^{EP\text{-}LS},$$

where

$$\overline{B}^{EP\text{-}LS} = \left(1 + \frac{C-1}{n}\right)\frac{\bar{t}}{\overline{t^*}} + \frac{C-1}{n} \cdot \frac{n-1}{\sqrt{2n-1}} \cdot \frac{\sigma_t}{\overline{t^*}},$$

for all $n \geqslant 1$. As $n \to \infty$, we have $\overline{B}_{EP}^{EP\text{-}LS} \to \bar{t}/\overline{t^*}$.

Proof. From Lemma 3, we get

$$E[OPT] \geqslant \frac{n\overline{t^*}}{C}.$$

The rest of the proof is similar to that of Theorem 7. □

The following theorem gives an average-case performance bounds of type II for algorithms *EP-LPT* with the condition that the total transmission power is equally allocated to the channels.

Theorem 9. *The length of an EP-LPT schedule satisfies*

$$\overline{R}_{EP}^{EP\text{-}LPT} = \frac{E[EP\text{-}LPT]}{E[OPT_{EP}]} \leqslant \overline{B}_{EP}^{EP\text{-}LPT},$$

where

$$\overline{B}_{EP}^{EP\text{-}LPT} = 1 + \frac{C-1}{n},$$

for all $n \geqslant 1$. As $n \to \infty$, we have $\overline{B}_{EP}^{EP\text{-}LPT} \to 1$, i.e., algorithm EP-LPT is asymptotically optimal if the total transmission power is equally allocated to the channels.

Proof. For t_k in algorithm *EP-LPT*, we notice that t_k approaches $\min_{1\leqslant i\leqslant n}(t_i)$ as $n\to\infty$. However, we simply use the following bound, $E[t_k]\leqslant\bar{t}$. From Lemma 4, we have

$$E[EP\text{-}LPT]\leqslant\left(\frac{n}{C}+1-\frac{1}{C}\right)\bar{t},$$

which gives rise to the theorem. □

The following theorem gives an average-case performance bounds of type II for algorithms *EP-LPT* without the condition that the total transmission power is equally allocated to the channels.

Theorem 10. *The length of an EP-LPT schedule satisfies*

$$\overline{R}^{EP\text{-}LPT}=\frac{E[EP\text{-}LPT]}{E[OPT]}\leqslant\overline{B}^{EP\text{-}LPT},$$

where

$$\overline{B}^{EP\text{-}LPT}=\left(1+\frac{C-1}{n}\right)\frac{\bar{t}}{\overline{t^*}},$$

for all $n\geqslant 1$. As $n\to\infty$, we have $\overline{B}_{EP}^{EP\text{-}LPT}\to\bar{t}/\overline{t^*}$.

Proof. The proof is similar to that of Theorems 8 and 9. □

It is clear that the average-case performance bounds $\overline{B}_{EP}^{EP\text{-}LS}$, $\overline{B}^{EP\text{-}LS}$, $\overline{B}_{EP}^{EP\text{-}LPT}$, and $\overline{B}^{EP\text{-}LPT}$ of type II given by Theorems 7–10 for algorithms *EP-LS* and *EP-LPT* can be calculated numerically (see the next section).

5. Simulation and numerical data

To show simulation and numerical data for the average-case performance bounds derived in the last section, we need specific probability distributions of the s_i's and the d_i's. We make the following assumptions.

- The s_i's are i.i.d. random variables with a common Pareto distribution, whose probability distribution function and cumulative distribution function are

  $$f_s(x)=\frac{as_0^a}{x^{a+1}},$$

  and

  $$F_s(x)=1-\left(\frac{s_0}{x}\right)^a,$$

  where $x\geqslant s_0$ and $a>2$.
- The d_i's are i.i.d. random variables with a common uniform distribution in the circle with radius d_2 centered at the base station minus the circle with radius d_1 centered at the base station. The probability distribution function and cumulative distribution function are

  $$f_d(x)=\frac{2x}{d_2^2-d_1^2},$$

  and

  $$F_d(x)=\frac{x^2-d_1^2}{d_2^2-d_1^2},$$

  where $d_1\leqslant x\leqslant d_2$.
- The probability distributions for the s_i's and the d_i's are independent of each other.

The mean of the size of a random transmission request is

$$\bar{s}=\frac{a}{a-1}s_0,$$

for $a>1$. The second moment of the size is

$$\overline{s^2}=\frac{a}{a-2}s_0^2,$$

and the variance of the size is

$$\sigma_s^2 = \overline{s^2} - \bar{s}^2 = \frac{a}{(a-2)(a-1)^2} s_0^2,$$

for $a > 2$.

The mean of the transmission time is

$$\bar{t} = \frac{\bar{s}}{W} \int_{d_1}^{d_2} \frac{1}{\log_2(1 + \frac{\beta\gamma}{x^\alpha})} \cdot \frac{2xdx}{d_2^2 - d_1^2},$$

which can be simplified as

$$\bar{t} = \frac{2\bar{s}}{W(d_2^2 - d_1^2)} \int_{d_1}^{d_2} \frac{xdx}{\log_2(1 + \frac{\beta\gamma}{x^\alpha})}.$$

The second moment of the transmission time is

$$\overline{t^2} = \frac{\overline{s^2}}{W^2} \int_{d_1}^{d_2} \frac{1}{(\log_2(1 + \frac{\beta\gamma}{x^\alpha}))^2} \cdot \frac{2xdx}{d_2^2 - d_1^2},$$

which can be simplified as

$$\overline{t^2} = \frac{2\overline{s^2}}{W^2(d_2^2 - d_1^2)} \int_{d_1}^{d_2} \frac{xdx}{(\log_2(1 + \frac{\beta\gamma}{x^\alpha}))^2}.$$

The variance of the transmission time is

$$\sigma_t^2 = \overline{t^2} - \bar{t}^2,$$

and the coefficient of variation is

$$c_t = \frac{\sigma_t}{\bar{t}}.$$

The mean of the minimal transmission time is

$$\overline{t^*} = \frac{2\bar{s}}{W(d_2^2 - d_1^2)} \int_{d_1}^{d_2} \frac{xdx}{\log_2(1 + \frac{\beta P}{x^\alpha})}.$$

To show our simulation data for average-case performance bounds of Type I, we use the following parameters setting: $P = 200$ Watt, $C = 16$, $\alpha = 3$, $\beta = 1$, $\gamma = P/C = 12.5$, $W = 76.8$ kHz, $a = 2.1$, $s_0 = 2$ Kbytes = 16 Kbits, $d_1 = 0.1$, $d_2 = 0.7$. Our parameter setting yields the following values required in our average-case performance bounds:

$$\begin{aligned}
\bar{s} &= 30.55 \text{ (Kbits)},\\
\bar{t} &= 0.058158 \text{ (second)},\\
\sigma_t &= 0.13012,\\
c_t &= 2.23736,\\
\overline{t^*} &= 0.036377 \text{ (second)}.
\end{aligned}$$

In Fig. 1, we display simulation data for average-case performance bounds of type I, i.e., $E[B_{EP}^{EP\text{-}LS}]$, $E[B^{EP\text{-}LS}]$, $E[B_{EP}^{EP\text{-}LPT}]$, and $E[B^{EP\text{-}LPT}]$, where $B_{EP}^{EP\text{-}LS}$, $B^{EP\text{-}LS}$, $B_{EP}^{EP\text{-}LPT}$, and $B^{EP\text{-}LPT}$ are given by Theorems 2, 3, 5, and 6, respectively. For each average-case performance bound and each n = 100, 150, 200, 250, ..., 1000, we report the average of 10,000 random samples of the average-case performance bound, so that all these expectations are obtained with 99% confidence interval of no more than ±1%. The s_i's are generated by the transformation $u = F_s(s_i)$, i.e.,

$$s_i = F_s^{-1}(u) = \frac{s_0}{(1-u)^{1/a}} \in [s_0, \infty),$$

where u is a uniform random variable in $[0, 1)$. The d_i's are generated by the transformation $u = F_d(d_i)$, i.e.,

$$d_i = F_d^{-1}(u) = \sqrt{d_1^2 + u(d_2^2 - d_1^2)} \in [d_1, d_2],$$

where $u \in [0, 1]$. We observe that our average-case performance bounds of type I are decreasing functions of n, and

$$\lim_{n\to\infty} E[B_{EP}^{EP\text{-}LS}] = \lim_{n\to\infty} E[B_{EP}^{EP\text{-}LPT}] = 1,$$

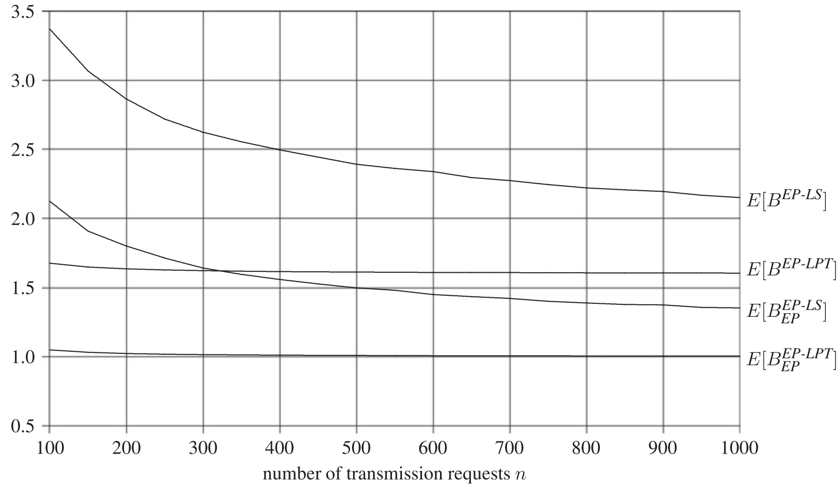

Fig. 1. Simulation data for average-case performance bounds of type I.

and

$$\lim_{n\to\infty} E[B^{EP\text{-}LS}] = \lim_{n\to\infty} E[B^{EP\text{-}LPT}] = \frac{\bar{t}}{t^*} = 1.59878.$$

In Fig. 2, we demonstrate numerical data for $\overline{B}_{EP}^{EP\text{-}LS}$, $\overline{B}^{EP\text{-}LS}$, $\overline{B}_{EP}^{EP\text{-}LPT}$, and $\overline{B}^{EP\text{-}LPT}$ given by Theorems 7–10, respectively, with n = 100, 150, 200, 250, ..., 1000. Again, our average-case performance bounds of type II are decreasing functions of n. By Theorems 7 and 9, we have

$$\lim_{n\to\infty} \overline{B}_{EP}^{EP\text{-}LS} = \lim_{n\to\infty} \overline{B}_{EP}^{EP\text{-}LPT} = 1.$$

By Theorems 8 and 10, we have

$$\lim_{n\to\infty} \overline{B}^{EP\text{-}LS} = \lim_{n\to\infty} \overline{B}^{EP\text{-}LPT} = \frac{\bar{t}}{t^*} = 1.59878.$$

It is clear that our average-case performance bounds of type I are tighter than type II. However, they do not have closed forms yet.

We would also like to mention that the average-case performance of algorithms *EP-LS* and *EP-LPT* should be better than the bounds obtained above since our lower bound for *OPT* given in Lemma 3 is quite loose.

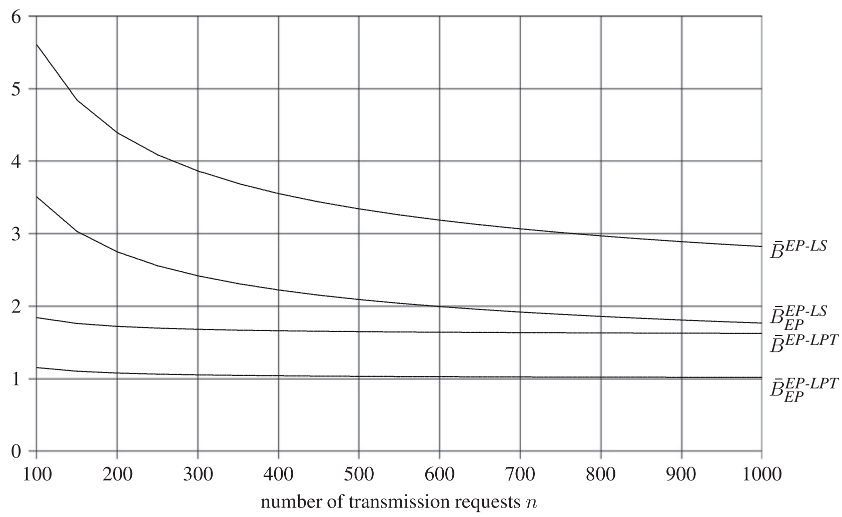

Fig. 2. Numerical data for average-case performance bounds of type II.

6. Power allocation for equal times (ET)

6.1. Power assignment

Our second type of downlink data transmission scheduling algorithms are developed based on the following observation.

Theorem 11. *In an optimal schedule, all the channels 1, 2, ..., C complete their data transmissions simultaneously.*

Proof. Assume that (s_i, d_i) is the last completed transmission request (on channel u) which is allocated power p_i, and (s_j, d_j) is the second last completed transmission request (on channel v) which is allocated power p_j. It is clear that we can always increase p_i and reduce p_j, such that the two transmissions are finished at the same time, and the two channels u and v complete their data transmissions simultaneously, and that the total transmission time is reduced. This process can be repeated until all the channels 1, 2, ..., C complete their data transmissions simultaneously. □

Assume that we have C data transmission requests (s_i, d_i), $1 \leqslant i \leqslant C$, where the transmission for (s_i, d_i) is performed on channel i, for all $1 \leqslant i \leqslant C$. We can always find the p_i's, $1 \leqslant i \leqslant C$, such that all the C transmissions are completed at the same time t, that is,

$$\frac{s_i}{W\log_2\left(1+\beta\cdot\frac{p_i}{d_i^\alpha}\right)} = t,$$

for all $1 \leqslant i \leqslant C$. From the above equation, we get

$$p_i = \frac{d_i^\alpha(2^{s_i/(Wt)}-1)}{\beta}.$$

Therefore, t can be found by solving the following equation,

$$\frac{1}{\beta}\sum_{i=1}^{C} d_i^\alpha(2^{s_i/(Wt)}-1) = P,$$

which comes from the fact that $p_1 + p_2 + \cdots + p_C = P$. By Theorem 11, the time t is the minimum time to complete the C data transmissions. Once t is available, we can find the power assignment, i.e., $p_1, p_2, \ldots, p_C$.

The above power allocation method to achieve *equal times* (*ET*) can be extended to schedule any set S of n data transmission requests. We divide S into C disjoint subsets $S_1, S_2, \ldots, S_C$, such that all the data transmission requests in S_j are scheduled on channel j, where $1 \leqslant j \leqslant C$. Each channel j is allocated power p_j, where $p_1 + p_2 + \cdots + p_C = P$. All the data transmissions in S_j are processed with power p_j. The time required by channel j to complete the transmissions in S_j is

$$T_j(p_j) = \sum_{(s_i,d_i)\in S_j} \frac{s_i}{W\log_2\left(1+\beta\cdot\frac{p_j}{d_i^\alpha}\right)}.$$

Therefore, we need to find $p_1, p_2, \ldots, p_C$ and T such that

$$T_1(p_1) = T_2(p_2) = \cdots = T_C(p_C) = T,$$

by using the methods described below.

The value T can be found numerically in the range $[lb, ub]$ by using the classic bisection method [1], where

$$lb = \max_{1\leqslant j\leqslant C}(T_j(P)) = \max_{1\leqslant j\leqslant C}\sum_{(s_i,d_i)\in S_j} \frac{s_i}{W\log_2\left(1+\beta\cdot\frac{P}{d_i^\alpha}\right)},$$

and

$$ub = \max_{1\leqslant j\leqslant C}\left(T_j\left(\frac{P}{C}\right)\right) = \max_{1\leqslant j\leqslant C}\sum_{(s_i,d_i)\in S_j} \frac{s_i}{W\log_2\left(1+\beta\cdot\frac{P}{Cd_i^\alpha}\right)}.$$

The lower bound lb is easy to justify, since it gives the maximum transmission time of the C channels when each channel is assigned power P. The upper bound ub gives the maximum transmission time of the C channels when the total power P is equally assigned to the C channels. Since it is unlikely that such a power allocation is optimal, T should be less than (or, no more than) ub.

Notice that the search interval $[lb, ub]$ is updated and reduced by half during each repetition of the bisection method. Assume that $[lb, ub]$ is the current search interval and $T = (lb + ub)/2$. We need to find p_j such that $T_j = T$ for all $1 \leqslant j \leqslant C$. Based on the p_j's, the search interval $[lb, ub]$ is adjusted as follows. If $p_1 + p_2 + \cdots + p_C > P$, we set $lb = T$, which means that T should be increased, since power allocation should be reduced. If $p_1 + p_2 + \cdots + p_C < P$, we set $ub = T$, which means that T should be reduced, since there is more power to be allocated.

Given T, the power assignment p_j can also be found numerically in a range $[lb, ub]$ by using the bisection method. The initial value of ub can be P. It is clear that $T_j(ub) = T_j(P) < T$. The initial value of lb can be $P/2^m$, where m is the smallest integer such that $T_j(P/2^m) > T$. Assume that $[lb, ub]$ is the current search interval and $p_j = (lb + ub)/2$. The search interval $[lb, ub]$ is adjusted as follows. If $T_j(p_j) > T$, we set $lb = p_j$, which means that p_j should be increased to reduce $T_j(p_j)$ to T. If $T_j(p_j) < T$, we set $ub = p_j$, which means that p_j should be reduced to increase $T_j(p_j)$ to T.

6.2. Transmission scheduling

We have described our power allocation method to achieve equal transmission times for a given partition of S into C disjoint subsets $S_1, S_2, \ldots, S_C$. It remains to describe how to find the S_j's. It is clear that the S_j's ultimately determine the total transmission time. We consider three transmission scheduling algorithms.

- *ET-EN*: Transmission requests are allocated to the C channels such that all the channels receive *equal number* (*EN*) of transmission requests, namely, $\lfloor n/C \rfloor$ or $\lceil n/C \rceil$.
- *ET-LS*: Transmission requests are allocated to the C channels by using the *LS* algorithm, where each request has a virtual transmission time v_i.
- *ET-LPT*: Transmission requests are allocated to the C channels by using the *LPT* algorithm, where each request has a virtual transmission time v_i.

The virtual transmission time v_i should incorporate both s_i and d_i into consideration. Since the actual transmission time t_i is an increasing function of both s_i and d_i, we set $v_i = s_i d_i$ in this paper.

We would like to mention that since power assignment is performed after transmission scheduling when all the information of the transmission requests are available, all downlink data transmission scheduling algorithms of the second type are offline algorithms.

6.3. Performance comparison

Extensive simulations have been conducted to compare the five downlink data transmission scheduling algorithms proposed in this paper, two of the first type (*EP-LS* and *EP-LPT*) and three of the second type (*ET-EN*, *ET-LS*, and *ET-LPT*).

Assume that we have total power $P = 200$ Watt and $C = 16$ channels. The constants α, β, and W are identical to those in the last section. For each algorithm

$$A \in \{EP\text{-}LS, EP\text{-}LPT, ET\text{-}EN, ET\text{-}LS, ET\text{-}LPT\},$$

we perform the following experiment. For each $n = 100, 200, \ldots, 1000$, we generate n random transmission requests whose probability distribution functions for the s_i's and the d_i's are the same as those in the last section. Then we apply algorithm A to get the schedule length. The above experiment is repeated for 2000 times and the average schedule length is reported. Our simulation data are summarized in Table 1, where the 99% confidence interval is ±5%.

We observe that the performance of the five algorithms can be ranked as *EP-LS*, *ET-EN*, *EP-LPT*, *ET-LS*, *ET-LPT*, with *EP-LS* being the worst and *ET-LPT* being the best. Algorithm *ET-EN* has a very naive transmission scheduling method; however, its performance is better than *EP-LS* due to the equal time power allocation method. Algorithm *ET-LS* uses the *LS* transmission scheduling method based on virtual transmission times; however, due to the equal time power allocation method, it performs better than *EP-LPT* which uses the *LPT* transmission scheduling method based on actual transmission times. As expected, algorithm *ET-LPT* yields even better performance than *ET-LS* due to the improved transmission scheduling method.

We also display the values of a reasonable lower bound (*LB*), namely,

$$LB = \left\lceil \frac{n}{C} \right\rceil \bar{t},$$

Table 1
Average schedule length (in seconds) of various algorithms.

n	*EP-LS*	*ET-EN*	*EP-LPT*	*ET-LS*	*ET-LPT*	*LB*
100	0.64475	0.51896	0.50989	0.45515	0.42548	0.40711
200	1.08013	0.92021	0.87319	0.85462	0.78771	0.75606
300	1.49444	1.31826	1.24132	1.18245	1.13908	1.10501
400	1.89684	1.67803	1.56992	1.55877	1.53609	1.45396
500	2.24731	2.07042	1.95649	1.91223	1.88772	1.86107
600	2.61454	2.41925	2.33329	2.30290	2.24353	2.21002
700	3.06747	2.80678	2.68561	2.67124	2.60823	2.55897
800	3.49306	3.21290	3.01748	3.02132	2.98019	2.90792
900	3.78466	3.59032	3.40694	3.36951	3.32146	3.31503
1000	4.18729	3.92418	3.75658	3.70858	3.69474	3.66399

for the average schedule length of any downlink data transmission scheduling algorithm. We observe that the average schedule length of *ET-LPT* is very close to the lower bound and the room for further performance improvement is very small, especially for large n.

7. Scheduling with parallel channels

It is shown in [3] that if a transmission request (s_i, d_i) is allocated transmission power p_i and c_i parallel channels, the transmission rate is

$$r_i(c_i) = Wc_i\log_2\left(1 + \beta \cdot \frac{p_i}{c_i d_i^\alpha}\right),$$

and the transmission time is

$$t_i = \frac{s_i}{r_i(c_i)} = \frac{s_i}{Wc_i\log_2\left(1 + \beta \cdot \frac{p_i}{c_i d_i^\alpha}\right)},$$

where the power p_i is equally allocated to the c_i channels.

We would like to mention that with the same power allocation, the transmission rate obtained by using parallel channels is greater than the transmission rate by using a single channel, i.e., $r_i(c_i) > r_i(1)$ for all $c_i \geqslant 2$. The reason is that

$$\begin{aligned} r_i(c_i) &= Wc_i\log_2\left(1 + \beta \cdot \frac{p_i}{c_i d_i^\alpha}\right) = W\log_2\left(1 + \beta \cdot \frac{p_i}{c_i d_i^\alpha}\right)^{c_i} = W\log_2\left(\sum_{k=0}^{c_i}\binom{c_i}{k}\left(\beta \cdot \frac{p_i}{c_i d_i^\alpha}\right)^k\right) \\ &= W\log_2\left(1 + \beta \cdot \frac{p_i}{d_i^\alpha} + \sum_{k=2}^{c_i}\binom{c_i}{k}\left(\beta \cdot \frac{p_i}{c_i d_i^\alpha}\right)^k\right) > W\log_2\left(1 + \beta \cdot \frac{p_i}{d_i^\alpha}\right) = r_i(1), \end{aligned}$$

for all $c_i \geqslant 2$. Furthermore, we notice that

$$r_i(c_i) = W \cdot \frac{\beta p_i}{d_i^\alpha} \cdot \log_2\left(1 + \frac{\beta p_i}{c_i d_i^\alpha}\right)^{c_i d_i^\alpha/(\beta p_i)}.$$

Since

$$\left(1 + \frac{\beta p_i}{c_i d_i^\alpha}\right)^{c_i d_i^\alpha/(\beta p_i)}$$

is an increasing function of c_i, we have $r_i(1) < r_i(2) < r_i(3) < \cdots$, that is, given the same power supply, more channels yield greater transmission rate. As a theoretical result, we get

$$\lim_{c_i \to \infty} r_i(c_i) = W \cdot \frac{\beta p_i}{d_i^\alpha} \cdot \log_2 e = 1.442695W \cdot \frac{\beta p_i}{d_i^\alpha},$$

although such transmission rate is hard to achieve practically.

Let the amount of work performed for a transmission request be defined as

$$w_i = c_i t_i = \frac{s_i}{W\log_2\left(1 + \beta \cdot \frac{p_i}{c_i d_i^\alpha}\right)},$$

that is, the channel-time product (certain number of channels devoted for certain amount of time). We notice that for a fixed power, we get increased transmission rate and reduced transmission time by using more channels. However, the amount of work performed is increased.

A schedule for a transmission request (s_i, d_i) is specified as $(b_i, j_1, j_2, \ldots, j_{c_i})$, which means that a transmission for the request starts at time b_i and finishes at time $b_i + t_i$ on channels $j_1, j_2, \ldots, j_{c_i}$, where $1 \leqslant j_1, j_2, \ldots, j_{c_i} \leqslant C$. Our scheduling problem can be generalized as follows: given total transmission power P and C channels, constants α and β and W, and n downlink data transmission requests specified by (s_i, d_i), $1 \leqslant i \leqslant n$, find a power assignment p_i, $1 \leqslant i \leqslant n$, and a nonpreemptive schedule $(b_i, j_1, j_2, \ldots, j_{c_i})$, where $b_i \geqslant 0$, $1 \leqslant j_1, j_2, \ldots, j_{c_i} \leqslant C$, $1 \leqslant i \leqslant n$, for all the n transmission requests, such that the total time to complete the n data transmissions is minimized. As pointed out in [3], this is a multi-dimensional malleable scheduling problem, i.e., more power and/or channel results in shorter transmission time. In addition to the two subproblems of power assignment and transmission scheduling, there is an added subproblem of channel allocation in the problem of downlink data transmission scheduling with parallel channels.

7.1. Equal channel allocation

Let c be an integer that divides C. For each algorithm

$$A \in \{EP\text{-}LS, EP\text{-}LPT, ET\text{-}EN, ET\text{-}LS, ET\text{-}LPT\},$$

we can design an algorithm A_c to solve the problem of downlink data transmission scheduling with parallel channels. Algorithm A_c works as follows. Let the C channels be divided into C/c groups, where each group contains c channels. A group of c channels is treated as a *super-channel*. Super-channels are number as 1, 2, ..., C/c. The subproblem of channel allocation is solved by using the *equal channel allocation* method as follows. Each transmission request is allocated a super-channel, i.e., a group of c parallel channels. The subproblems of power assignment to super-channels and transmission scheduling on super-channels are solved by using algorithm A. When power p_i is assigned to a super-channel j_i for a transmission request (s_i, d_i), where $1 \leqslant j_i \leqslant C/c$, the power p_i is equally allocated to the c channels in the super-channel j_i.

In Table 2, we display the average schedule length of algorithms $EP\text{-}LPT_c$, where c = 1, 2, 4, 8, 16. Assume that we have total power P = 200 Watt and C = 16 channels. The constants α, β, and W are identical to those in Table 1. For each algorithm $EP\text{-}LPT_c$, we perform the following experiment. For each n = 100, 200, ..., 1000, we generate n random transmission requests whose probability distribution functions for the s_i's and the d_i's are the same as those in Table 1. Then we apply algorithm $EP\text{-}LPT_c$ to get the schedule length. The above experiment is repeated for 2000 times and the average schedule length is reported. Our simulation data are in Table 2 have 99% confidence interval of ±5%.

In Table 3, we display the average schedule length of algorithms $ET\text{-}LPT_c$, where c = 1, 2, 4, 8, 16, by using the same method in Table 2.

We observe from Tables 2 and 3 that by using the equal channel allocation method, the performance of algorithm A_c improves as c increases. This means that increased parallelism does reduce the expected data transmission time. However, for large c, such reduction is very limited. Notice that the lower bound in Table 1 does not apply to parallel channels. The performance of algorithm A_c is better than LB when $c \geqslant 2$.

7.2. Equal power allocation

If the total transmission power P is equally allocated to the C channels, (i.e., if a transmission request is allocated c_i channels, it must be supplied with $p_i = c_i\gamma$ power; in other words, channel allocation and power assignment are tied together), the above transmission rate is

$$r_i(c_i) = Wc_i\log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right),$$

Table 2
Average schedule length (in seconds) of algorithm $EP\text{-}LPT_c$.

n	$EP\text{-}LPT_1$	$EP\text{-}LPT_2$	$EP\text{-}LPT_4$	$EP\text{-}LPT_8$	$EP\text{-}LPT_{16}$
100	0.52984	0.40770	0.36668	0.36254	0.36058
200	0.88477	0.75738	0.72848	0.72949	0.72728
300	1.23688	1.17069	1.10042	1.09088	1.09062
400	1.57993	1.49544	1.45514	1.45550	1.45121
500	1.96393	1.83804	1.82342	1.81728	1.82185
600	2.32221	2.20260	2.17774	2.18414	2.18095
700	2.71232	2.58323	2.54790	2.54541	2.54450
800	3.01884	2.94459	2.90824	2.90693	2.90067
900	3.43347	3.31123	3.26575	3.26417	3.27211
1000	3.77715	3.65237	3.63856	3.63653	3.63714

Table 3
Average schedule length (in seconds) of algorithm $ET\text{-}LPT_c$.

n	$ET\text{-}LPT_1$	$ET\text{-}LPT_2$	$ET\text{-}LPT_4$	$ET\text{-}LPT_8$	$ET\text{-}LPT_{16}$
100	0.42347	0.37817	0.36976	0.36446	0.36349
200	0.79790	0.74318	0.73130	0.72775	0.72330
300	1.15932	1.10159	1.09616	1.09084	1.08997
400	1.53784	1.47766	1.45511	1.45415	1.46043
500	1.87284	1.83217	1.82354	1.81610	1.81179
600	2.25005	2.18035	2.22516	2.18373	2.18041
700	2.61290	2.56914	2.54467	2.54254	2.54723
800	2.99704	2.91271	2.90262	2.91287	2.90986
900	3.31546	3.27806	3.26934	3.27999	3.27027
1000	3.68804	3.64376	3.64121	3.63427	3.63154

and the transmission time is

$$t_i = \frac{s_i}{Wc_i \log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)}.$$

The above equation implies that the transmission rate is a linear function of the number of channels, i.e., a data transmission is fully malleable.

The following theorem shows that under the condition that the total transmission power is equally allocated to the channels, the simple sequential scheduling algorithm produces an optimal schedule.

Theorem 12. *If the total transmission power is equally allocated to the channels, the sequential scheduling algorithm is optimal for scheduling with parallel channels.*

Proof. Notice that the amount of work performed for a transmission request is

$$w_i = c_i t_i = \frac{s_i}{W \log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)},$$

which is solely determined by a data transmission request itself. The optimal schedule length is at least

$$\frac{1}{C}\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} \frac{s_i}{WC \log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)} = \sum_{i=1}^{n} \frac{s_i}{r_i(C)}.$$

One easy way to achieve this optimal schedule length is to schedule the n transmission requests sequentially, that is, to schedule the transmission requests one by one and to assign all the C channels to a transmission request. □

7.3. An M/G/1 queueing model

If all the downlink data transmission requests are processed sequentially by using all the channels and all the transmission power, we can study the performance of a wireless network by a queueing model. Such modeling will help us in understanding the impact of the number of channels and the total transmission power on the average response time.

We now give an M/G/1 queueing model for the first-come-first-served (FCFS) queueing discipline.

Assume that the downlink data transmission requests come as a Poisson stream with arrival rate λ. They are processed by the FCFS scheduling algorithm by using all the C channels and total transmission power P for each transmission request, that is,

$$t_i = \frac{s_i}{WC \log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)}.$$

Assume that the s_i's and the d_i's have a joint probability density function $f_{s,d}(x_1,x_2)$, where $x_1 > 0$ and $x_2 > 0$.

Then the mean of the transmission time is

$$\bar{t} = \frac{1}{WC}\int_0^\infty \int_0^\infty \frac{s_i}{\log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)} \cdot f_{s,d}(x_1,x_2)dx_1dx_2.$$

The second moment of the transmission time is

$$\overline{t^2} = \frac{1}{W^2C^2}\int_0^\infty \int_0^\infty \frac{s_i^2}{\left(\log_2\left(1 + \frac{\beta\gamma}{d_i^\alpha}\right)\right)^2} \cdot f_{s,d}(x_1,x_2)dx_1dx_2.$$

The variance of the transmission time is $\sigma_t^2 = \overline{t^2} - \bar{t}^2$, and the coefficient of variation is $c_t = \sigma_t/\bar{t}$.

Our main interest is the average response time (i.e., waiting time plus data transmission time) in this M/G/1 queueing system. By using the well known Pollaczek–Khinchin mean-value formula (see p. 190 of [12]), we get the average response time

$$T = \bar{t}\left(1 + \frac{\rho(1 + c_t^2)}{2(1-\rho)}\right),$$

where $\rho = \lambda\bar{t}$.

We adopt the same assumptions of Section 4, that is, $f_{s,d}(x_1,x_2) = f_s(x_1)f_d(x_2)$, which yields

$$\bar{t} = \frac{2\bar{s}}{WC(d_2^2 - d_1^2)}\int_{d_1}^{d_2} \frac{xdx}{\log_2(1 + \frac{\beta\gamma}{x^\alpha})},$$

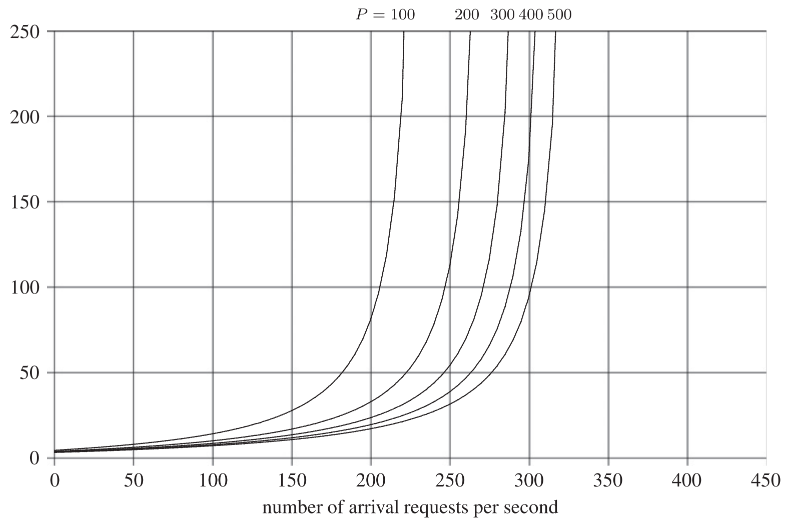

Fig. 3. Average response time (in millisecond) vs. *P*.

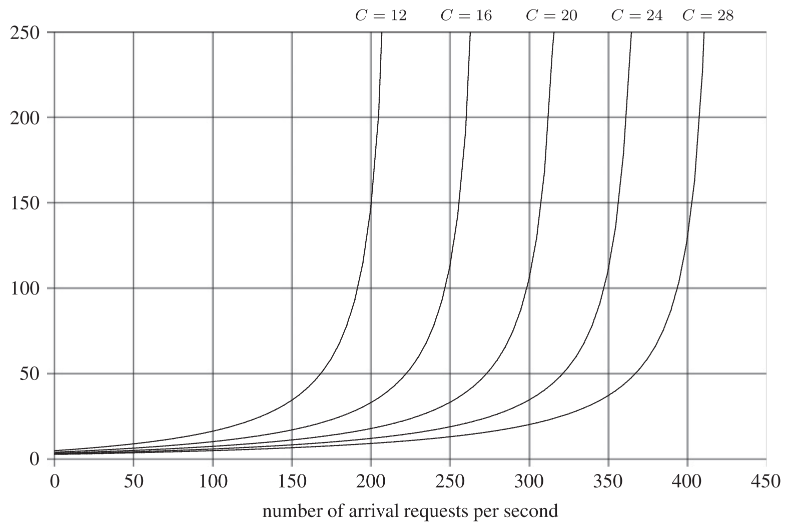

Fig. 4. Average response time (in millisecond) vs. *C*.

and

$$\overline{t^2} = \frac{2\overline{s^2}}{W^2C^2(d_2^2 - d_1^2)} \int_{d_1}^{d_2} \frac{x dx}{\left(\log_2\left(1 + \frac{\beta\gamma}{x^\alpha}\right)\right)^2}.$$

We also use the same parameter setting of Section 4. In Figs. 3 and 4, we display the average response time as a function of the arrival rate. In Fig. 3, where $C = 16$, we show the impact of P, with $P = 100, 200, 300, 400, 500$. In Fig. 4, where $P = 200$, we show the impact of C, with $C = 12, 16, 20, 24, 28$. We observe that increasing the number of channels has more impact on the reduction of the average response time than increasing the total transmission power. For instance, increasing C from 16 to 20 (25% increase) while keeping the same $P = 200$ has about the same effect as increasing P from 200 to 500 (150% increase) while keeping the same $C = 16$.

8. Conclusions

We have studied the downlink data transmission scheduling problem in wireless networks. We pointed out that every downlink data transmission scheduling algorithm must have two components to solve the two subproblems of power assignment and transmission scheduling. We proposed two types of downlink data transmission scheduling algorithms. In the first type (e.g., *EP-LS* and *EP-LPT*), power assignment is performed before transmission scheduling based on equal power allocation. In the second type (e.g., *ET-EN*, *ET-LS*, and *ET-LPT*), power assignment is performed after transmission scheduling based on equal time allocation.

We have analyzed the performance of algorithms *EP-LS* and *EP-LPT*. It is shown that both algorithms exhibit excellent worst-case performance (Theorems 1 and 4) and asymptotically optimal average-case performance (Theorems 2, 5, 7, 9) under the condition that the total transmission power is equally allocated to the channels. In general, both algorithms exhibit excellent average-case performance (Theorems 3, 6, 8, 10). Another advantage of *EP-LS* is that it is an online scheduling algorithm and can be applied to scheduling existing transmission requests without knowing future transmission requests. We demonstrated that algorithms *ET-LS* and *ET-LPT* perform better than algorithms *EP-LS* and *EP-LPT*, due to the equal time power allocation method (Theorem 11). Furthermore, the performance of *EP-LPT* has reached the limit quite closely and the room for further performance improvement is very limited.

We have mentioned that by using the equal channel allocation method, all the above algorithms can be extended to schedule downlink data transmissions with parallel channels and improved performance can be obtained by using parallel channels. When parallel channels can be allocated to data transmission requests, the simple sequential scheduling algorithm is optimal if the total transmission power is equally allocated to the channels (Theorem 12). As an extra contribution, we have established an M/G/1 queueing model for the FCFS queueing discipline, derived its average response time, and observed that increasing the number of channels has more impact on the reduction of the average response time than increasing the total transmission power.

Acknowledgments

Thanks are due to two anonymous reviewers for their comments on improving the paper. A preliminary version of the paper entitled "Power assignment and transmission scheduling in wireless networks" was presented on the 9th International Workshop on Performance Modeling, Evaluation, and Optimization of Ubiquitous Computing and Networked Systems, Atlanta, Georgia, April 23, 2010.

References

[1] http://en.wikipedia.org/wiki/Bisection_method.
[2] J. Bae, R. Berry, M.L. Honig, Power allocation, rate, and coverage for relay-assisted downlink data transmission, Proceedings of IEEE International Conference on Communications 10 (2006) 4451–4456.
[3] L. Becchetti, S. Diggavi, S. Leonardi, A. Marchetti-Spaccamela, S. Muthukrishnan, T. Nandagopal, A. Vitaletti, Parallel scheduling problems in next generation wireless networks, in: Proceedings of the 14th ACM Symposium on Parallel Algorithms and Architectures, 2002, pp. 238–247.
[4] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, A. Viterbi, CDMA/HDR: a bandwidth-efficient high-speed wireless data service for nomadic users, IEEE Communications Magazine 38 (7) (2000) 70–77.
[5] J.M. Cioffi, G.P. Dudevoir, M.V. Eyuboglu, G.D. Forney, MMSE decision-feedback equalizers and coding – Parts I & II, IEEE Transactions on Communications 43 (10) (1995) 2582–2604.
[6] H.A. David, H.N. Nagaraja, Order Statistics, third ed., John Wiley & Sons, Hoboken, New Jersey, 2003.
[7] M.R. Garey, D.S. Johnson, Computers and Intractability – A Guide to the Theory of NP-Completeness, W.H. Freeman, New York, 1979.
[8] O. Goussevskaia, Y.A. Oswald, R. Wattenhofer, Complexity in geometric SINR, in: Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2007, pp. 100–109.
[9] R.L. Graham, Bounds for certain multiprocessing anomalies, Bell System Technical Journal 45 (1966) 1563–1581.
[10] R.L. Graham, Bounds on multiprocessing timing anomalies, SIAM Journal on Applied Mathematics 2 (1969) 416–429.
[11] N. Joshi, S.R. Kadaba, S. Patel, G.S. Sundaram, Downlink scheduling in CDMA data networks, in: Proceedings of the 6th ACM International Conference on Mobile Networking and Computing, 2000, pp. 179–190.
[12] L. Kleinrock, Queueing Systems, Theory, vol. 1, John Wiley & Sons, New York, 1975.
[13] K. Li, Power assignment and transmission scheduling in wireless networks, in: 9th International Workshop on Performance Modeling, Evaluation, and Optimization of Ubiquitous Computing and Networked Systems, 2010.
[14] O. Messas and R. Malhamé, Optimal downlink data transmission scheduling in next generation wireless systems, in: Proceedings of the 13th Mediterrean Conference on Control and Automation, 2005, pp. 707–712.
[15] S. Nanda, K. Balachandran, S. Kumar, Adaptation techniques in wireless packet data services, IEEE Communications Magazine 38 (1) (2000) 54–65.
[16] Y. Zhao, A. Anpalagan, A channel based fair scheduling scheme for downlink data transmission in TD-CDMA networks, Wireless Personal Communications: An International Journal 46 (4) (2008) 469–491.