Round Opportunistic Fair Downlink Scheduling in Wireless Communications Networks

Mohammad M. Banat and Razan F. Shatnawi

banat@just.edu.jo, reyozan_90@yahoo.com

Department of Electrical Engineering, Jordan University of Science and Technology

ABSTRACT: - Round opportunistic fair (ROF) scheduling is proposed as a heuristic algorithm for improving fairness in wireless downlink transmission scheduling. ROF represents a trade-off between the achievable aggregate throughput on one side, and fairness on the other side. Even though the proposed algorithm is usable with other types of multiple access, a simple single-channel time-slotted system is assumed.

The proposed algorithm takes rounds over the users to be scheduled for downlink transmissions. However, it does not schedule users cyclically as in round robin scheduling; because it prioritizes users with better channel conditions. At the same time, the proposed algorithm does not always schedule the user with the absolute best channel conditions; to allow fairness in user access to system resources.

To keep the presentation of ROF as simple as possible, mathematical analysis that only demonstrate the operational mechanism of the algorithm is included. Fairness and other performance capabilities of ROF are evaluated using computer simulations. Simulation results show that ROF achieves good levels of fairness compared to other well-known algorithms such as proportional fairness scheduling.

KEYWORDS:- Downlink Transmission; Opportunistic Scheduling; Fairness

I. INTRODUCTION

Wireless spectrum efficiency is becoming more and more significant with the increasing demand on wideband wireless services [1]. In a cellular network, channel conditions between the base station (BS) and the different users have generally different and statistically independent random time variation patterns [2]. A signal that is transmitted over a wireless channel can suffer three types of (almost independent) superimposed fading effects: path loss, shadowing and multipath [2]. Path loss is commonly modelled by an inverse power-law function of the received signal power on the distance between the transmitter and receiver. Shadowing takes the form of slow and random variations of the received signal power. Shadowing is mainly due to the presence of large obstacles in the signal transmission path, and is generally independent of the transmitter-receiver distance. Multipath fading takes the form of fast and random variations of the received signal strength. This effect is due to constructive and destructive interference among the received multipath components.

A scheduling policy is a rule, or set of rules, used to specify which user is scheduled to transmit/receive during a time slot. Opportunistic scheduling gives higher transmission priority to users with better channel conditions. Achievable throughput is an increasing logarithmic function of the signal to noise ratio (SNR) [3]. Sending to the user with the best channel conditions (or equivalently to the user with maximum achievable transmission rate) maximizes the downlink throughput [4]. This approach is known as MaxRate scheduling. Several throughput-optimal scheduling algorithms have been presented in [5], [6], [7], and elsewhere. Opportunistic scheduling can cause severe unfairness to users that are far from the BS. Because of their large path losses, such users suffer low probabilities of having good channel conditions, and therefore, can be given only few and far-between transmission opportunities.

Fairness and spectral efficiency are very important issues in resource allocation in multiuser wireless networks. Spectral efficiency is measured by the normalized aggregate throughput in bits/s/Hz. Fairness and spectral efficiency can involve contradicting network behaviours. A tradeoff (that usually depends on the type of services provided by the network) between the two quantities should be sought.

Fair wireless scheduling, especially in time division multiple access (TDMA) systems, has been extensively studied (e.g., in [8], [9], [10]). A major weakness in many previous works is that the channel is classified as either “good” or “bad”. Such a coarse classification is too simple to characterize real wireless channels. Such a classification allows only few degrees of freedom for the purpose of designing a scheduling algorithm.

The opportunistic framework in [11] takes into account three scheduling requirements: temporal fairness, utilitarian fairness, and minimum performance requirement for each user. A scheduling scheme for the

Qualcomm high data rate (HDR) system has been proposed in [12]. This scheduling scheme exploits the time-varying channel conditions and is based on the proportional fairness (PF) concept defined in [13]. More recently, [14] has provided a performance comparison between scheduling policies in the time and frequency domains for the LTE downlink.

The remainder of this paper is organized as follows. In section II, we introduce the channel model and outline our main assumptions. The proposed opportunistic scheduling algorithm is introduced in section III. Simulation results are presented in section IV, along with performance comparisons to other scheduling algorithms. The paper is concluded is section IV.

II. CHANNEL AND SYSTEM MODELS

We assume a downlink Rayleigh fading channel in a single cell system, with path loss, log normal shadowing and multipath fading. The BS is assumed to be at the center of the cell, the radius of which is denoted as D. The BS is assumed to serve, in a time-slotted manner, U fixed users, each equipped with one antenna. The BS is assumed to always have packets to send to all users.

Assuming the transmitted signal power is P_t, the power received by user u is given by

$$P_u = |h_u|^2 P_t \tag{1}$$

where h_u is the channel gain, expressed in the form [15]

$$h_u = \sqrt{c d_u^{-\alpha} s_u} m_u \tag{2}$$

where c is the mean path gain at a reference distance of 1 km, d_u is the distance in km between user u and the BS, α is the path loss exponent (PLE), s_u is the power scaling due to shadow fading and m_u is the phasor sum of the multipath components. α is typically between 2 and 4. The shadow fading power scaling factor s_u is assumed to follow the log-normal distribution. Let

$$s_{u,\mathrm{dB}} = 10 \log_{10} s_u \tag{3}$$

where $s_{u,\mathrm{dB}}$ follows a zero-mean Gaussian distribution with a variance σ_s^2. Typical values of σ_s are around 8 dB. As commonly accepted in the literature [16], we assume shadow fading is exponentially correlated. The multipath fading factor m_u is a zero-mean unit-variance complex Gaussian random variable.

The received SNR of user u is given by

$$Z_u = \frac{P_u}{P_n} \tag{4}$$

where P_n is the received zero-mean additive white Gaussian noise (AWGN) power. Substituting (1) and (2) into (4) yields the SNR of user u as follows:

$$Z_u = c d_u^{-\alpha} \frac{P_t}{P_n} s_u |m_u|^2 \tag{5}$$

Conditioned on s_u, the SNR is an exponentially distributed random variable, with a mean value

$$\begin{aligned} \bar{Z}_u &= \left[Z_u | s_u \right] \\ &= c d_u^{-\alpha} \frac{P_t}{P_n} s_u \end{aligned} \tag{6}$$

The received SNR by a cell edge user u^*, is given by

$$\eta = c D^{-\alpha} \frac{P_t}{P_n} s_{u^*} \left| m_{u^*} \right|^2 \tag{7}$$

Averaging η over shadow fading and multipath fading results in the mean cell edge SNR $\bar{\eta}$ as

$$\bar{\eta} = c D^{-\alpha} \frac{P_t}{P_n} \tag{8}$$

Like in [17], we use $\bar{\eta}$ to represent the acceptable noise level. Substituting (8) into (6) yields

$$\bar{Z}_u = \bar{\eta} \left(\frac{D}{d_u} \right)^{\alpha} s_u \tag{9}$$

Note that a new mean value of the SNR has to be calculated every time s_u changes. Same is true when the distance between the user and the BS changes.

The modulation scheme that is used in the system under consideration is M - QAM. It is well-known that, at a given SNR, both the information bit rate and the bit error rate of M - QAM increase for larger values of M. Therefore, when a BS transmission (BST) is assigned to the scheduled user, M is chosen such that it is higher when the user channel conditions are better. It was shown in [17] that the feasible transmission rate is a logarithmic function of the SNR. Following [15], we adopt the following expression for the achievable normalized throughput (in bits/s/Hz) as a function of the SNR:

$$\xi_u = \log_2 \left(1 + \frac{Z_u}{K} \right) \tag{10}$$

where K is a constant system-efficiency factor that depends on the system design and the target bit error rate.

III. OPPORTUNISTIC SCHEDULING ALGORITHM

In this section, we present the proposed round opportunistic fair (ROF) scheduling algorithm. A brief mathematical background is initially presented for the sake of clarifying the working principles of the algorithm. However, full mathematical analysis will not be attempted here.

ROF is a heuristic opportunistic scheduling algorithm that maximizes the network throughput under various fairness restrictions. The main idea of ROF is to limit assigning BST's to only a dynamic group of users that will be known in this paper as the candidate users. This restriction aims at improving chances of users with generally bad channels conditions to get BST assignments. Hence, good levels of fairness in BST assignment to users are expected to be achieved. Assigning a BST to the candidate user with the best channel conditions achieves maximum throughput under the user candidacy restriction. The number of candidate users controls the trade-off between fairness and network throughput. When there is only one candidate user, ROF performs round-robin scheduling, while when all users are candidates, it functions in a purely opportunistic manner.

The group of "candidate" users consists of users that have the smallest values of an ascendingly sorted "waiting figure". At each scheduling round, ROF grants channel access to the candidate user that has the best channel conditions. The waiting figure of the scheduled user is increased, and the group of candidate users is updated before the following scheduling round.

The role of the scheduling algorithm is to decide which user is to receive data from the BS in a given time slot, based on channel conditions of the users. Therefore, it is assumed that the BS knows the channel conditions of all users. How this information is made available to the BS is beyond the scope of this paper.

Even though the BS can be assumed to use multiple channels to transmit to a number of users in the same time slot, this work is limited to single-channel transmissions. Scheduling users on multiple channels in the same time slot can be the basis of using ROF in multicarrier wireless networks.

It is assumed that a BST consists of a number of symbols that is proportional to the achievable throughput. Quantization to integer numbers of bits is not considered in this paper; and hence whenever throughput is mentioned, what is meant is actually the achievable throughput. In the simulation results below, we measure the achievable throughput normalized to a unit bandwidth.

The duration of a BST is fixed, regardless of the number of transmitted symbols. It is assumed that channel conditions do not change during a BST. However, the channel is assumed to vary independently from one BST to another. As mentioned earlier, all users are assumed to always have data to receive from the BS.

At the beginning of each time slot, the scheduler determines the best candidate user that should receive a BST. The objective is to optimize the network throughput. In the below, we illustrate how the proposed algorithm works.

- Downlink transmissions happen at the ith multiples of the channel coherence time (to make sure that channel conditions do not change during a BST), where $i = 1, 2, \cdots$.
- Each user has a wait figure $w_u(i)$, which controls the number of time slots the user has to wait before competing for a BST. Wait figures of all users are initialized with very small random values. In fact, initial wait figures should all be set to zero. However, this would not provide any means to select the initial set of users that are candidates to get the BST.
- The integer $r_u(i) \in [1, U]$ is used to indicate the rank of user u for the purpose of receiving a BST. Users are ranked in ascending order of their wait figures. Precisely, the rank of user u at time i is equal to

  $$r_u(i) = \sum_{k=1}^{U} 1\left(w_u(i) - w_k(i)\right) \tag{11}$$

  where $1(x)$ is the discrete unit step function, given by

  $$1(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{12}$$

- The number of users that are candidates to receive a BST at a given time instant is denoted as U_x, where $1 \leq U_x \leq U$. Users with ranks $r_u(i) \leq U_x$ are candidates to receive a BST at time i.
- If user v receives a BST at time i, a constant quantity ρ is added to its current wait figure $w_v(i-1)$ to form the new wait figure $w_v(i)$. Mathematically, if user v receives a BST at time i, then

  $$w_v(i) = w_v(i-1) + \rho \tag{13}$$

- The channel gain from the BS to user u at time i is denoted as $h_u(i)$. The set of channel gains $\{h_u(i)\}_{u=1}^{U}$ will be assumed to be statistically independent for different values of u and for different values of i. The channel gain is assumed to take the form in (2).

- The signal to noise ratio $Z_u(i)$ of user u during time slot i is an exponential random variable, the mean of which takes the form in (9). In this paper we use $Z_u(i)$ to calculate the achievable throughput.
- The performance measure of user u is $\xi_u(i)$, as given by (10). Other forms of $\xi_u(i)$ dependence on $Z_u(i)$ can be used to represent a wide range of QoS requirements.
- The user that is scheduled to receive a BST at time i is the one with the largest performance measure from among all candidate users (users with $1 \le r_u(i) \le U_x$). As it will turn out to be the case, U_x is an important parameter in the trade-off between temporal fairness and performance optimization. Note that $U_x = U$ means purely opportunistic transmission (i.e., MaxRate), while $U_x = 1$ means absolutely fair transmission (i.e., RR).
- It is, therefore, useful to define the indicator function

  $$I_u(i) = 1\left(U_x - r_u(i)\right) \tag{14}$$

  Note that user u can be a candidate for receiving a BST at time i only if $I_u(i) = 1$.
- We define the weighted performance measure $\beta_u(i)$ as

  $$\beta_u(i) = \xi_u(i) I_u(i) \tag{15}$$

  Note that $\beta_u(i)$ is equal to zero for all users with higher ranks than U_x. This means that such users are excluded from competition for the BST.
- If v is the identification number of the user that has received a BST at time $i \ge 1$, then the following conditions must be met

  $$1 \le r_v(i) \le U_x \tag{16}$$

  $$I_v(i) = 1 \tag{17}$$

  $$\beta_v(i) = \max_u \left\{\beta_u(i)\right\}_{u=1}^{U_x} \tag{18}$$

IV. SIMULATION RESULTS

To produce the simulation results that are presented below, we have assumed a single-cell cellular system with one BS at the cell center. The system has no co-channel interference, meaning that the system is noise-limited, and that throughput computations are based on the SNR. The cell radius is 1 km, and the PLE is 4. The number of users is 40. User u is assumed to be separated from the BST by a distance equal to

$$d_u = 0.1 + 0.8\left(1 - \frac{u}{U}\right) \tag{19}$$

In other words, user U is closest to the BS and user 1 is farthest, with equal distance increments for the users in between. In throughput calculations using (10), the system efficiency factor K is equal to 8. The number of users that compete for a BST is 10. Each simulation experiment includes 50 runs of 400,000 BSTs (or time slots) each. In each run, measurements like throughput and air time share are taken every 1000 BSTs. This means that 400 readings are taken in every run. Readings are averaged over the 50 runs to produce the experiment results.

The shadow fading component of the channel gain is assumed to change every 100 time slots. Knowing that a random sequence with exponential autocorrelation can be generated by a first order autoregressive model [18], shadow fading updates are performed according to the recursion

$$s_{u,\mathrm{dB}}^{\mathrm{New}} = \eta s_{u,\mathrm{dB}}^{\mathrm{Old}} + (1-\kappa)\varepsilon \tag{20}$$

where $s_{u,\mathrm{dB}}^{\mathrm{New}}$ is the updated shadow fading power scaling factor in dB, $s_{u,\mathrm{dB}}^{\mathrm{Old}}$ is the old shadow fading power scaling factor in dB, κ is the autocorrelation coefficient and ε is a zero-mean white Gaussian sequence with variance σ_ε^2 that is statistically independent of $s_{u,\mathrm{dB}}^{\mathrm{Old}}$. Note that from (20) we should have

$$\sigma_s^2 = \kappa^2 \sigma_s^2 + (1-\kappa)^2 \sigma_\varepsilon^2 \tag{21}$$

Solving (21) for σ_ε^2, one obtains

$$\begin{aligned} \sigma_\varepsilon^2 &= \frac{1-\kappa^2}{(1-\kappa)^2}\sigma_s^2 \\ &= \frac{1+\kappa}{1-\kappa}\sigma_s^2 \end{aligned} \tag{22}$$

The shadowing autocorrelation coefficient is assumed to be 0.8, while its standard deviation is assumed to be 8 dB.

Although many fairness indicators have been proposed and used in the literature, we opt to use the variance of a set of measurements to quantify fairness in the measured quantity. This is motivated mainly by the well-known fact that the variance measures the extent of variation of the measured quantity. When the variance of a measured quantity is lower, the measurements are closer to their average value, and hence the fairness is higher. Obviously, the converse is true as well.

Below we present comparisons between the proposed scheduling algorithm and RR scheduling, MaxRate scheduling and PF scheduling. We also study the effects of path loss and shadowing on the results.

In Figure 1 we have plotted the variance of the numbers of BSTs assigned to all users using ROF and a number of other scheduling schemes. ROF performs very well in this aspect, compared to MaxRate and PF. Note that in the long run, ROF achieves an air time variance that is several orders of magnitude lower than those achieved by PF and MaxRate. Obviously, no scheduling scheme can achieve better air time fairness than the RR scheme, and this is indeed the case according to Figure 1.

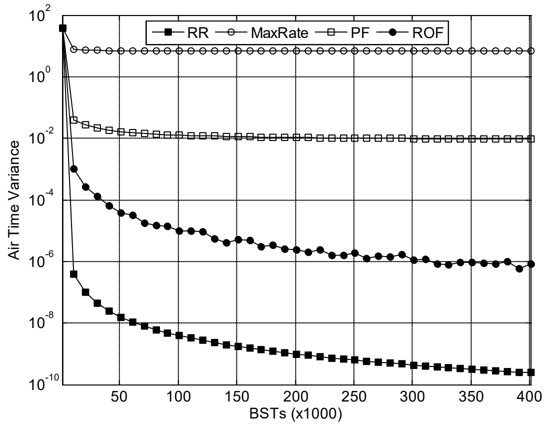

Figure 1: Air time fairness comparisons

In Figure 2 we have plotted the variance of the average normalized throughput achievable by users using ROF and a number of other scheduling schemes. ROF performs better than all three other scheduling schemes in this aspect. It outperforms PF and MaxRate by orders of magnitude in terms of throughput fairness among the users, while also slightly outperforming RR.

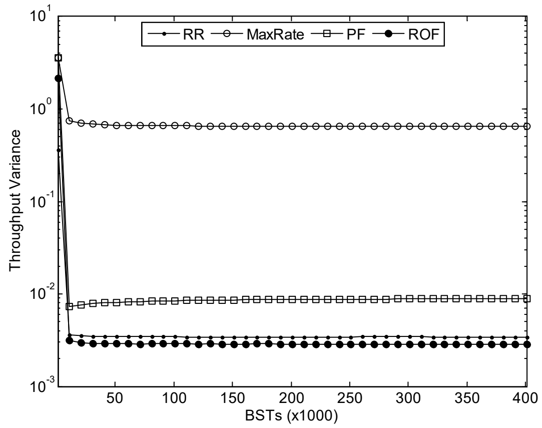

Figure 2: Throughput fairness comparisons

In Figure 3 we have plotted the ROF average normalized throughput versus the number of candidate users for PLE values 2 and 4, and shadowing standard deviation (SSD) value 6 and 10 dB. In all cases, and as expected, the normalized throughput increases with the number of candidate users. As pointed out earlier, when the number of candidates is one (intersection of throughput curve with the vertical axis on the left), ROF performs exactly like an RR algorithm. RR achieves lowest throughput and highest fairness. On the other hard, when the number of candidates equals the number of users (intersection of throughput curve with the vertical axis on the right), ROF performs exactly like a MaxRate algorithm, which is purely opportunistic. MaxRate achieves highest throughput and lowest fairness.

As can be seen from the figure, a higher PLE leads to a higher achievable throughput. This is because a higher PLE causes larger variations in the received SNR, and hence in the achievable throughput. Given the opportunistic nature of the scheduling scheme, larger variations in the achievable throughput can be utilized to increase the overall network throughput.

Similarly, a higher SSD deviations lead to higher achievable throughput. This is because a higher SSD causes larger variations in the received SNR, and hence in the achievable throughput. Given the opportunistic nature of the scheduling scheme, larger variations in the achievable throughput can be utilized to increase the overall network throughput.

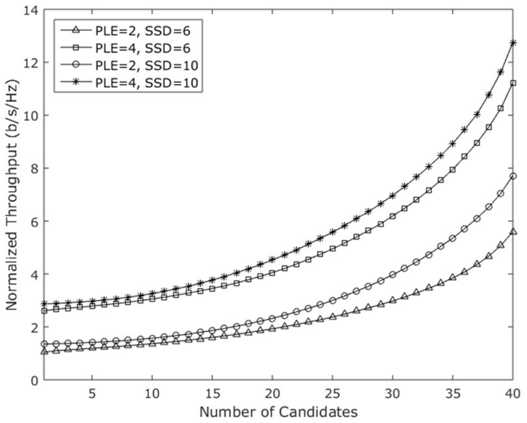

Figure 3: Effect of PLE and shadowing standard deviation on ROF average normalized throughput

V. CONCLUSION

We have presented ROF, a new heuristic algorithm for fair scheduling of opportunistic wireless downlink transmissions. Computer simulations have been used to study throughput and fairness performance of the new algorithm. Results have been compared to those of other well-known scheduling algorithms. Our results indicate that, compared to other algorithms, ROF achieves substantial fairness improvements at the cost of some throughput reduction. This paper has opened up more issues for research. Many ideas can be extended from this work. We can summarize them in the following points: 1. extend our work from single channel to multi-channel. 2. assign different time shares to users. 3. study the delay statistics. 4. working with multiple services classes.

REFERENCES

[1] L. Zhang, M. Xiao, G. Wu, M. Alam, Y. C. Liang and S. Li, "A Survey of Advanced Techniques for Spectrum Sharing in 5G Networks," *IEEE Wireless Communications,* vol. 24, no. 5, pp. 44-51, October 2017.

[2] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, New York: Cambridge University Press, 2005.

[3] A. Goldsmith, Wireless communications, Cambridge University Press, 2005.

[4] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory,* vol. 43, no. 6, pp. 1986-1992, Nov. 1997.

[5] M. Andrews and L. Zhang, "Scheduling algorithms for multi-carrier wireless data systems," in *MobiCom'07*, Montreal, Canada, September 9-14, 2007.

[6] S. Liu, L. Ying and R. Srikant, "Scheduling in multichannel wireless networks with flow-level dynamics," in *SIGMETRICS'10*, New York, NY, USA, 2010.

[7] Y. Chen, X. Wang and L. Cai, "On achieving fair and throughput-optimal scheduling for TCP flows in wireless networks," *IEEE Transactions on wireless communications,* vol. 15, no. 12, pp. 7996-8008, September 2016.

[8] G. Song and Y. G. Li, "Cross-layer optimization for OFDM wireless networks—Part I: theoretical framework," *IEEE Transactions on Wireless Communications,* vol. 4, no. 2, pp. 614-624, Mar. 2005.

[9] T. S. E. Ng, I. Stoica and H. Zhang, "Packet fair queueing algorithms for wireless networks with location-dependent errors," in *INFOCOM 1998*, San Francisco, CA, USA, March 29-April 2, 1998.

[10] V. Bharghavan, S. Lu and T. Nandagopal, "Fair queuing in wireless networks: issues and approaches," *IEEE Personal Communications,* vol. 6, no. 1, pp. 44-53, Feb. 1999.

[11] X. Liu, E. K. P. Chong and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Computer Networks,* vol. 41, no. 4, p. 451–474, 2003.

[12] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana and A. Viterbi, "CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users," *IEEE Communications Magazine,* vol. 38, no. 7, pp. 70-77, Jul. 2000.

[13] A. Jalali, R. Padovani and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *VTC 2000-Spring*, Tokyo, Japan, May 15-18, 2000.

[14] O. Grøndalen, A. Zanella, K. Mahmood, M. Carpin, J. Rasool and O. N. Østerbø, "Scheduling Policies in Time and Frequency Domains for LTE Downlink Channel: A Performance Comparison," *IEEE Transactions on Vehicular Technology,* vol. 66, no. 4, pp. 3345-3360, April 2017.

[15] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment," *IEEE Transactions on Vehicular Technology,* vol. 56, no. 2, pp. 766-778, Mar. 2007.

[16] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electronics Letters,* vol. 27, no. 23, pp. 2145-2146, 7 Nov. 1991.

[17] S. Catreux, P. F. Driessen and L. J. Greenstein, "Data throughputs using multiple-input multiple-output (MIMO) techniques in a noise-limited cellular environments," *IEEE Transactions on Wireless Communications,* vol. 1, no. 2, p. 226–234, Apr. 2002.

[18] W. Wei, Time Series Analysis: Univariate and Multivariate Methods, Addison Wesley, 1994.