FAIRNESS MEASURES FOR BEST EFFORT TRAFFIC IN WIRELESS NETWORKS

Krister Norlund[1], **Tony Ottosson**[2], **Anna Brunstrom**[3]

[1] Ericsson AB, Lindholmspiren 11, 417 56 Göteborg, Sweden, krister.norlund@ericsson.com
[2] Dept. of Signals and Systems, Chalmers Univ. of Technology, 412 96 Göteborg, Sweden, tony.ottosson@s2.chalmers.se
[3] Dept. of Computer Science, Karlstad Univ., 651 88 Karlstad, Sweden, anna.brunstrom@kau.se

Abstract - This paper proposes that fairness in wireless networks should be measured using one of the following new measures: the deterministic unfairness bound called the wireless absolute fairness bound (WAFB) or the statistical unfairness bound called the 99-percentile wireless absolute fairness bound (WAFB_{99}). Compared with previous fairness definitions, the new fairness measures are better suited for measuring fairness of scheduling disciplines that exploit multiuser diversity.

A new scheduling discipline called opportunistic proportional fair scheduling is defined. Numerical results show that the new scheduling discipline has slightly higher throughput and slightly better fairness than proportional fair scheduling.

Keywords - Fairness, scheduling, multiuser diversity.

I. Introduction

Packet scheduling is an important issue for future high-speed wireless networks. In the recent years, a number of scheduling disciplines for scheduling best effort and delay sensitive traffic has been proposed in the literature [1]. In order to evaluate and find the scheduling disciplines with the best performance, appropriate performance measures are required. For best effort users the major concern for the scheduling discipline should be to maximize the total throughput for all users at the same time as each user is served a fair amount of throughput [2]. Therefore one performance measure should be the total throughput achieved and another measure should be the fairness achieved by the scheduling discipline. Absolute delay requirements and queue stability [3] are not considered in this paper.

Fairness is a desirable property of the wireless network as it offers protection between users. This means than the traffic flow of an ill-behaving user cannot affect the traffic flow of another user. Fairness can be defined in many ways. For wireline networks, the classic definition of fairness is the max-min weighted fair share allocation policy. Generalized Processor Sharing (GPS) is an ideal (and unimplementable) scheduling discipline that exactly achieves max-min fairness [4, chapter 9]. One fairness measure used when comparing wireline scheduling disciplines is called the absolute fairness bound [4, chapter 9]. The absolute fairness bound (AFB) measures the maximum difference between what is transmitted if GPS is used and what is transmitted if the scheduling discipline under evaluation is used. In this paper we extend this measure for wireless networks with multiuser diversity.

II. Wireless network characteristics related to fairness

In order to understand what fairness should be in a wireless network, we make a number of observations.

Observation 1: Without modification the wireline max-min fairness definition is not suitable for wireless networks. With max-min fairness all users should receive the same throughput. If user D is situated on the edge of the cell and user C is situated close to the basestation, user D would usually need to use a vast majority of the resources (time slots) in order to achieve a throughput equal to user C. Many people would consider this scheduling to be unjust as the user close to the basestation only would be allowed to use a small amount of the resources. In [2] and [5], the authors propose the use of link quality weights to accommodate for this problem. The link quality weights are used to give each user a throughput that is proportional to the quality of the user's link. Bad quality leads to lower throughput and good quality leads to higher throughput.

For user i backlogged in time interval $[t, t+\tau]$ and any other user j the following relation is true when GPS with link quality weights is used:

$$\frac{G(i,t,t+\tau)}{G(j,t,t+\tau)} \geq \frac{w(i)\cdot C(i,t,t+\tau)}{w(j)\cdot C(j,t,t+\tau)} \qquad (1)$$

$G(i,t,t+\tau)$ is the amount of data transmitted to user i in the time interval $[t, t+\tau]$ if GPS is used as the scheduling discipline. The weight, $w(i)$ is the max-min weight associated to user i. It is a positive number and it can be used to configure the proportional difference in throughput for the users. In this paper, w is always considered to be equal to one for all users. $C(i,t,t+\tau)$ is the link quality weight associated with user i in the time interval $[t, t+\tau]$. The link quality weight can change over time and it can for example be the instantaneous signal to noise ratio or the transmission bit rate given the current signal to noise ratio.

Observation 2: In wireless networks the channel capacity for a single user changes over time both on the short and long time scales. The channel capacity is in this paper defined as the highest possible transmission rate (bits/s), achieving the desired QoS. The channel capacities for different users are independent and varying, these independent variations can be seen as a diversity that can be used to increase the total throughput of the network. This diversity is called multiuser diversity [6], and it is illustrated below.

Example 1: Consider a time slotted wireless network with two users, user C and user D. Fig. 1a depicts the channel capacity during 4 time slots for user C and user D. Fig. 1b depicts the throughput for users C and D if the scheduling discipline called X is used. Scheduling discipline X schedules the users in a round robin fashion without taking advantage of multiuser diversity. Fig. 1c depicts the throughput for users C and D if the scheduling discipline called Y is used. Scheduling discipline Y schedules the users taking advantage of multiuser diversity. Users C and D still use an equal amount of time slots but the total throughput is increased.

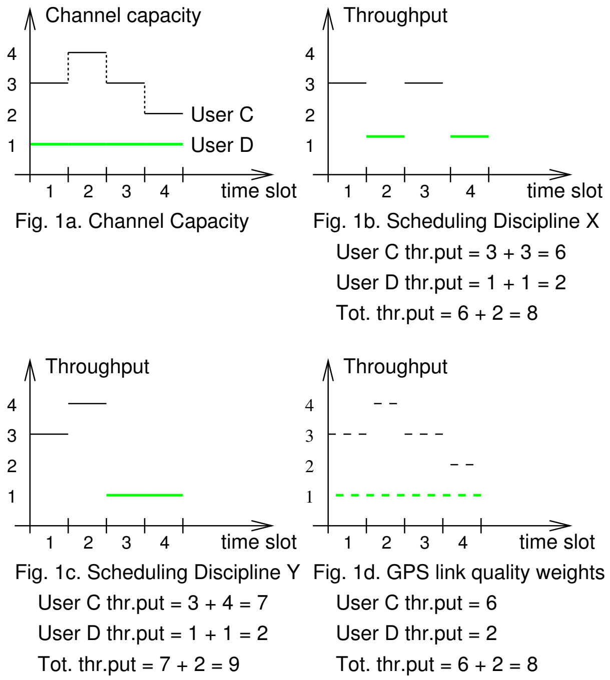

Fig. 1. Scheduling examples

Observation 3: The resources to be shared between the users in a wireless network are typically time and spectrum.

Observation 4: Regardless of how the resources are used the interesting quantity for the users is the bit rate.

Observation 5: A scheduling discipline that takes advantage of multiuser diversity and not necessarily gives all users exactly the same amount of resources can potentially give equal or higher average throughput for all users compared with when GPS with link quality weights is used. Note that if the link quality weights are the channel capacities, GPS with link quality weights will give exactly the same amount of resources (time slots) to all users.

Example 2: Fig. 1d depicts one example on how GPS with link quality weights theoretically could schedule the users in Example 1. The link quality weight for a user in a time slot is equal to the channel capacity for that user in that time slot. The throughput for user C and D in Fig. 1d fulfils the criteria in equation (1). The total throughput for user C in Fig. 1c is higher than in Fig. 1d. The total throughput for user D is the same in Fig. 1c and 1d.

III. Problems with previous fairness definitions

In order to measure the fairness of a scheduling discipline in a wireless network, a number of fairness definitions have been proposed in the literature [7] [2].

In [7], fairness is achieved if all users consume exactly the same amount of resources. Considering observations 3, 4 and 5, it is not necessary to share the resources exactly equally. Hence, this fairness definition is too strict for wireless networks with multiuser diversity.

The authors of [2], argue that it is not necessary to achieve short-term fairness. Therefore they define something they call long-term link-quality-weighted (LT-LQW) outcome-fairness. Considering two backlogged flows over a sufficiently long time interval $[t, t+\tau]$ the LT-LQW outcome-fairness is achieved if:

$$\left| \frac{S(i,t,t+\tau)}{w(i)\cdot \bar{C}(i)} - \frac{S(j,t,t+\tau)}{w(j)\cdot \bar{C}(j)} \right| < \epsilon \ , \qquad (2)$$

where ϵ is a small constant. $S(i,t,t+\tau)$ is the amount of data transmitted to user i in the time interval $[t, t+\tau]$ if the scheduling discipline being evaluated is used. $\bar{C}(i)$ is the average channel capacity for user i.

Using the fairness definition in equation (2), fairness is only achieved if the difference of throughput for the two users is proportional to the difference of average channel capacity for the two users. If the variance of the channel capacity for one user is not the same as the variance for the other users, the scheduling discipline cannot fully utilize the diversity possibilities without decreasing the fairness.

Example 3: In Fig. 1a the variance of the channel capacity for user C is larger than for user D. Using the fairness definition in equation (2) to measure fairness over the time interval [time slot 1, time slot 4], scheduling discipline X in Fig. 1b would be considered fair:

$\left| \frac{3+3}{(3+4+3+2)/4} - \frac{1+1}{(1+1+1+1)/4} \right| = 0$

Scheduling discipline Y in Fig. 1c, would be considered unfair:

$\left| \frac{3+4}{(3+4+3+2)/4} - \frac{1+1}{(1+1+1+1)/4} \right| \neq 0$

Bearing in mind the scenario in Example 1, it seems unjust that scheduling discipline Y should be considered less fair than scheduling discipline X when measuring fairness over time interval [time slot 1, time slot 4]. Even if this example is a short-term fairness example, the problem remains in the long-term case.

Note that the variance of the channel capacity typically is different for different users. A moving user in an open field that is in line of sight of the basestation usually has a smaller variance compared with a user that is moving in an area with obstacles.

IV. New fairness definition and measures

A. A fairness definition for wireless networks

The proposed wireless fairness definition is a modification of the max-min fairness definition to accommodate for observations 1, 3, 4 and 5. The proposed definition of fairness in wireless networks is that, a user has received a fair share of the resources if the user receives a throughput that is equal to or greater than what the user would receive if GPS with link quality weights was used as the scheduling discipline. If a user receives a throughput that is lower than the throughput received using GPS with link quality weights, the user has received an unfairly low share of the resources.

B. Requirements for a new fairness measure

One requirement for a new fairness measure is that it must be possible to express the fairness as a quantitative value. Another very important requirement is that fairness must be measured over a time interval size that is relevant to the best effort applications and users. If fairness is only measured over a time interval that is in the order of ten minutes, a scheduling discipline could schedule data only to user A the first minute and only to user B the second minute and still be considered fair according to the measure. Most best effort users will probably not consider this to be a fair scheduling discipline.

C. Wireless absolute fairness bound

We define the deterministic fairness measure called wireless absolute fairness bound (WAFB) as:

$$\mathrm{WF}(i,t,t+\tau) = \frac{G(i,t,t+\tau)}{w(i)\cdot C(i,t,t+\tau)} - \frac{S(i,t,t+\tau)}{w(i)\cdot C(i,t,t+\tau)}$$

$$\mathrm{WAF}(i,t,t+\tau) = \begin{cases} \mathrm{WF}(i,t,t+\tau)\,, & \mathrm{WF}(i,t,t+\tau) \geq 0 \\ 0 & , \ \mathrm{WF}(i,t,t+\tau) < 0 \end{cases}$$

$$\mathrm{WAFB} = \max_{\forall i,t,\tau} \mathrm{WAF}(i,t,t+\tau)$$

$\mathrm{WF}(i,t,t+\tau)$ is the wireless fairness for user i in time interval $[t,t+\tau]$. $\mathrm{WAF}(i,t,t+\tau)$ is the wireless absolute fairness for user i in time interval $[t,t+\tau]$. WAFB is the wireless absolute fairness bound.

WAFB is measured in bits or bytes and it is always greater than or equal to zero. Just like AFB, WAFB is a deterministic worst case bound on the unfairness of the scheduling discipline. The higher the WAFB or AFB value is, the less fair the scheduling discipline is. To indicate that the unfairness increases with higher values on WAFB, it would be more appropriate to use the name wireless absolute *unfairness* bound rather than wireless absolute fairness bound. However, we choose to be consistent with the absolute fairness bound notation.

The time interval $[t,t+\tau]$ can be any time interval and of any size. With the measure, the time interval and user that gives the worst fairness is the value that gives the WAFB value. Achieving fairness does not mean that each user must consume exactly the same amount of resources (time slots). As long as the users at least receive a throughput equal to the throughput that would be received using GPS with link quality weights as the scheduling discipline, it is considered that the users have been given a fair share of the resources.

WAFB is similar to AFB combined with link quality weights. There is, however, one major difference. If AFB with link quality weights is used as a measure of fairness and a user receives a throughput that is higher than the throughput received when using GPS with link quality weights as scheduling discipline, the scheduling discipline under evaluation is always considered to be unfair. With WAFB, this kind of scheduling disciplines will in many cases be considered to be fair.

One difference between WAFB and LT-LQW, is that LT-LQW measures the throughput difference between different users, while WAFB does not. As WAFB does not compere the throughput for different users, WAFB does not suffer from the problem in Example 3.

D. The 99-percentile wireless absolute fairness bound

The channel capacity for a user in a wireless network is stochastic. WF is a function of the channel capacity and therefore $\mathrm{WF}(i,t,t+\tau)$ is a stochastic process. If τ is equal to infinity, WF and WAFB may be unbounded. Even if WF is less than zero 99.9% of the time, WAFB may be unbounded. Hence, we define the statistical worst case bound called the 99-percentile wireless absolute fairness bound, here denoted WAFB_{99}, as follows:

$$\Pr\{\mathrm{WF}(i,t,t+\tau) \leq x\} = 0.99$$
$$\mathrm{WF}_{99}(i,t,t+\tau) = x$$

$$\mathrm{WAF}_{99}(i,t,t+\tau) = \begin{cases} \mathrm{WF}_{99}(i,t,t+\tau)\,, & \mathrm{WF}_{99}(i,t,t+\tau) \geq 0 \\ 0 & , \ \mathrm{WF}_{99}(i,t,t+\tau) < 0 \end{cases}$$

$$\mathrm{WAF}_{99}(i,\tau) = \max_{\forall t} \mathrm{WAF}_{99}(i,t,t+\tau)$$
$$\mathrm{WAFB}_{99} = \max_{\forall i,\tau} \mathrm{WAF}_{99}(i,\tau)$$

V. Numerical evaluation

In order to verify the usefulness of the new fairness measures we evaluate and compare two different scheduling disciplines using simulation. Two properties are measured: the total throughput (bits/s) and the WAFB_{99} (bits).

A. Scheduling disciplines

The scheduling disciplines are described below, where μ_i is the bit rate supported for user i at time t and $\tilde{\mu_i}$ is the average throughput for user i measured over a relatively long "sliding window". In the simulations the sliding window size was 1 second.

1) Proportional fair scheduling: With the proportional fair (PF) scheduling discipline, we schedule, at time t, the user:

$$k = \arg\max_i \frac{\mu_i(t)}{\tilde{\mu_i}}$$

In a time slotted system the rate is considered constant over one time slot. The PF scheduling discipline has been proposed for HDR [1]. As the PF scheduling discipline exploits multiuser diversity and schedules each user approximately an equal amount of time [6], the throughput for each user will in most cases be slightly higher than the throughput using GPS with link quality weights.

2) Opportunistic proportional fair scheduling: For the purpose of comparing PF scheduling with another scheduling discipline, we define a new scheduling discipline called opportunistic proportional fair (OPF) scheduling. With the OPF scheduling discipline, we schedule, at time t, the user:

$$k = \arg\max_i \left(\frac{c \cdot \mu_i(t)}{\tilde{\mu_i}} + \frac{d \cdot \mu_i(t)}{\max\limits_{\forall j} \mu_j(t)} \right) \quad , \tag{3}$$

where c and d are arbitrary constants larger than or equal to zero. If c is equal to zero and d is larger than zero, then the scheduling discipline works as the max rate scheduling discipline [1]. With the max rate scheduling discipline, we schedule the user with the highest channel capacity at time t. If d is equal to zero and c is larger than zero, then the scheduling discipline is exactly the same as the PF scheduling discipline. In the simulations we used $c = 1.5$ and $d = 1$.

The idea with OPF scheduling is to give each user a throughput that is higher than or equal to the throughput using GPS with link quality weights at the same time as the total throughput is maximized. The PF scheduling discipline usually gives each user a throughput higher than the throughput given using GPS with link quality weights but it does not necessarily maximize the total throughput at the same time. The purpose of the max rate term in OPF scheduling is to maximize throughput.

B. Simulation setup

For the purpose of comparing the scheduling disciplines and evaluating the WAFB_{99} measure, we consider a single carrier time-slotted system. The carrier frequency is 1900 MHz and the bandwidth for the carrier is 200 kHz. Each time slot is 0.667 ms and contains 108 payload symbols. For simplicity we only consider traffic from the basestation to the mobile user.

Only one user receives data in one time slot. The instantaneous signal to interference and noise ratio (SINR) for the upcoming time slot is perfectly predicted and given the instantaneous SINR, the system selects a modulation format for the upcoming time slot. Modulation formats from BPSK up to 256-QAM are used. See [8] for the exact modulation formats and SINR switching levels. The link quality weights are equal to the channel capacities for the given time slot.

One cell containing five best effort users is simulated. All users always have buffered data to transmit. The channel for a user is modeled as a correlated Rayleigh fading channel. Fading is considered to be flat within one time slot. The fading for two different users is uncorrelated. The average SINR per symbol for a user is constant. Starting with user 1 the average SINR per symbol is equal to 8, 12, 16, 20 and 24 dB.

For each scheduling discipline we ran 11 simulations. The simulation time for each simulation is 10 000 seconds. All mobiles have the same Doppler frequency in a simulation. The corresponding Doppler speed in the different simulations is: 3, 5, 10, 20, 40, 60, 80, 100, 120, 140, 160 km/h.

C. Numerical results

Given a scheduling discipline, WAFB_{99} might in some scenarios be bounded and in other scenarios be unbounded. For each simulation we evaluated if WAFB_{99} might be bounded or not. For each user, we verified that the average throughput was higher than the throughput for GPS with link quality weights. If the average throughput would have been lower than the throughput for GPS with link quality weights, then the scheduling discipline would most likely have an unbounded WAFB_{99} for that scenario. Furthermore, the statistics of $\mathrm{WF}_{99}(i, t, t + \tau)$ should indicate that WAFB_{99} is bounded. Fig. 2 depicts the statistics of $\mathrm{WF}_{99}(i, t, t + \tau)$ for user 5 as a histogram. The time interval size, τ, is 0.13 seconds. For this time interval size and user, it at least seems like $\mathrm{WAF}_{99}(i, t, t + \tau)$ is bounded. Finally, $\mathrm{WAF}_{99}(i, \tau)$ must be finite when the time interval size goes to infinity. In Fig. 3a, $\mathrm{WAF}_{99}(i, \tau)$ seems to be zero for increasingly high values of τ.

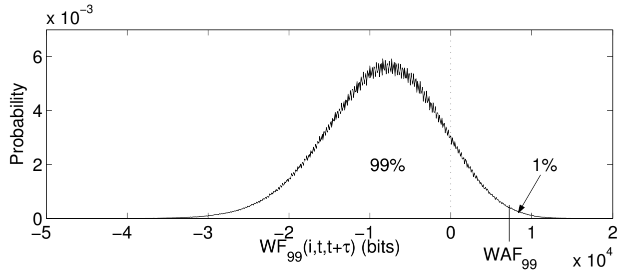

Fig. 2. Histogram WF_{99}. $i = 5$. $\tau = 0.13$ sec. Speed = 100 km/h

Fig. 4a depicts the total throughput for the 22 simulations. The total throughput is depicted as a function of the mobile speed. Comparing OPF scheduling with PF scheduling, it can be seen that the total throughput is slightly higher for OPF scheduling.

Fig. 4b depicts the WAFB_{99} values as a function of the mobile speed for both scheduling disciplines. In Fig. 4b it can be seen that PF scheduling has a slightly higher WAFB_{99} than OPF scheduling. This means that the fairness for PF scheduling is slightly worse than for OPF scheduling.

This was not expected but the reason can be seen in Fig. 3. In Fig. 3b the WAF_{99} curves for users 4 and 5 are lower than in Fig. 3a. Furthermore, the WAF_{99} curves for users 1 and 2 are higher in Fig. 3b than in Fig. 3a. As the maximum WAF_{99} value for the user with the highest maximum WAF_{99} value decreased, WAFB_{99} decreased (the fairness increased). By tuning the constants c and d for the OPF scheduling discipline, there is potential to decrease the WAFB_{99} and increase the total throughput even more.

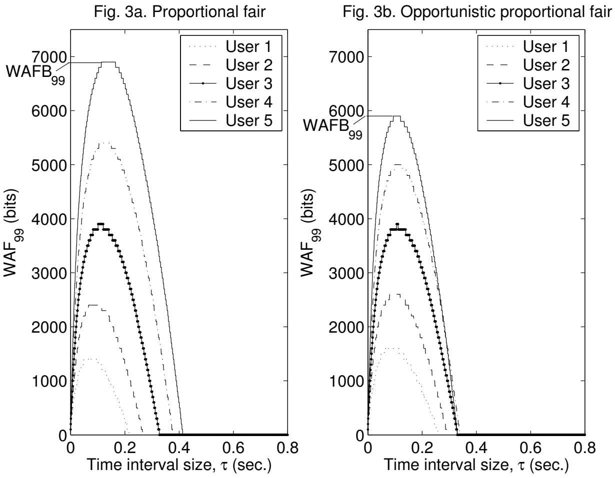

Fig. 3. $\mathrm{WAF}_{99}(i, \tau)$ as a function of τ. Speed = 100 km/h

In Fig. 4b it is easy to understand the magnitude of fairness, as the unit of the fairness measure is bits. As the scheduling disciplines are compared with GPS with link quality weights it is also relatively easy to understand what the fairness is. Fig. 4 gives a good indication of the performance and also the difference in performance of the scheduling disciplines. This means that total throughput and WAFB_{99} are suitable measures when comparing scheduling disciplines for best effort traffic in wireless networks.

VI. Conclusion

In this paper we have defined a deterministic fairness measure and a statistic fairness measure. Both measures are suited for measuring fairness of scheduling disciplines in wireless networks scheduling best effort traffic.

With the proposed fairness measures the size of the time interval leading to the WAFB or WAFB_{99} value is the size that leads to the worst fairness. This means that both long-term and short-term fairness is measured. If, for example, only time intervals larger than 1 second are interesting, the proposed measures can easily be modified to accommodate for this.

Numerical results showed that, given the simulated scenario, there exists at least one scheduling discipline that has both higher throughput and lower WAFB_{99} (unfairness) than the PF scheduling discipline.

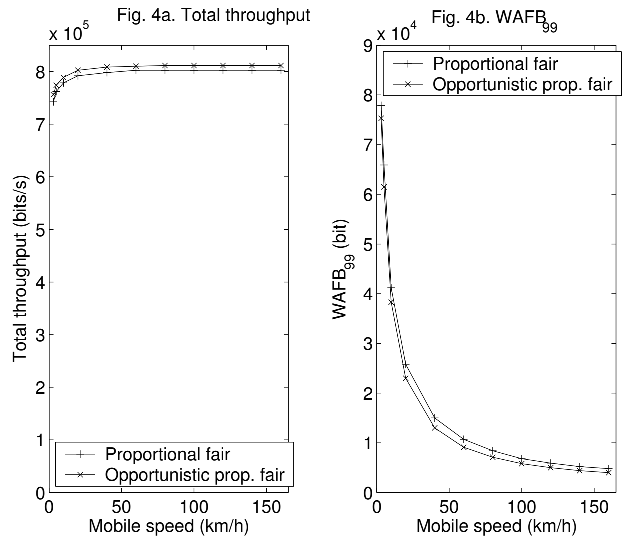

Fig. 4. Total throughput and WAFB_{99}

References

[1] S. Shakkottai and A. Stolyar, “Scheduling algorithms for a mixture of real-time and non-real-time data in HDR,” in *Proc. of the 17th International Teletraffic Congress*, Salvador da Bahia, Brazil, 2001, pp. 793 – 804.

[2] Y. Cao, V. O. K. Li, and Z. Cao, “Scheduling delay-sensitive and best-effort traffics in wireless networks,” in *Proc. IEEE International Conference on Communications*, 2003, pp. 2208 – 2212.

[3] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, “Providing quality of service over a shared wireless link,” *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–154, Feb. 2001.

[4] S. Keshav, *An Engineering Approach to Computer Networking*. Addison Wesley, 1997.

[5] Z. Jiang and N. K. Shankaranarayana, “Channel quality dependent scheduling for flexible wireless resource management,” in *Proc. IEEE Global Telecommunications Conference*, San Antonio, TX USA, 2001, pp. 3633–3638.

[6] P. Viswanath, D. Tse, and R. Laroia, “Opportunistic beamforming using dumb antennas,” *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[7] R. Elliot, “A measure of fairness of service for scheduling algoritms in multiuser systems,” in *Proc. Canadian Conference on Electrical and Computer Engineering*, Winnipeg, Canada, 2002, pp. 1583–1588.

[8] W. Wang, T. Ottosson, M. Sternad, A. Ahlén, and A. Svensson, “Impact of multiuser diversity and channel variability on adaptive OFDM,” in *Proc. Vehicular Technology Conference*, Fall 2003, pp. 547 – 551.