Optimal Radio Channel Allocation for Fair Queuing in Wireless Data Networks

Teerawat Issariyakul and Ekram Hossain
University of Manitoba and TR*Labs*
Winnipeg, MB, Canada.
Email: {ekram, teerawat}@ee.umanitoba.ca.

***Abstract*— In this paper, the problem of fair scheduling in a wireless network is formulated as an assignment problem and an *Optimal Radio Channel Allocation (ORCA)* strategy is proposed for fair bandwidth allocation in a centralized manner. Simulation results show that the performance improvement due to ORCA can be significant compared to other wireless fair-queuing mechanisms proposed in the literature such as WPS (Wireless Packet Scheduling).**

Keywords- wireless fair queuing, assignment problem, Hungarian method.

I. Introduction

Providing fair share of channel bandwidth among the different data flows is one of the key issues in provisioning QoS (Quality of Service) in wireless data networks. The fairness is achieved by using a scheduling protocol to allocate bandwidth among the mobiles in proportion to their weights.

Due to the location-dependent and bursty channel errors, the wireless fair scheduling algorithms need to take channel condition into account. The mobile perceiving a *dirty* channel should defer the transmission and let another mobile with a *clean* channel to transmit data. The scheduler should compensate for that when the channel of the deferred mobile becomes *clean* again. A mobile deferring the transmission is considered to be *lagging*, while a mobile receiving extra allocation is considered to be *leading*. General wireless fair queuing algorithms consist of error-free service, lead and lag model, compensation model, slot queues and packet queue, and channel monitoring and prediction [1].

II. Fair Queuing Algorithms

A. Weighted Round Robin (WRR)

WRR allocates channel bandwidth to each mobile in proportion to the corresponding weight [1]. For example, with three flows, if the weight vector is {1,3,2}, the allocation of time slots among the flows in a frame would be $\{x_1, x_2, x_2, x_2, x_3, x_3\}$.

B. Wireless Fair Queuing (WFQ)

Channel allocation due to WFQ is similar to that of WRR in that WFQ gives allocation in proportion to the mobiles' weights. In addition, WFQ spreads out the allocation in order to reduce the severity of burst errors, especially for mobiles with larger weights. WFQ algorithm is very simple and has time-complexity of $O(n)$, where n is the number of mobiles.

In WFQ, the weight of each mobile is assumed to be an integer. The frame size in terms of the number of time slots per frame is calculated by the summing up the weights of all mobiles. The i^{th} mobile is assigned with a score and initially it is $1/w_i$. Then in the next time slot the channel will be assigned to the mobile with a minimum score. As the i^{th} mobile is assigned the time slot, the corresponding score is increased by $1/w_i$. The process repeats by assigning the next time slot to the mobile with the minimum score. The procedure continues until the frame ends. For the above example, the allocation would be $\{x_2, x_3, x_2, x_1, x_2, x_3\}$.

C. Wireless Packet Scheduling Protocol (WPS)

This algorithm is a derivative of WFQ. WPS assumes perfect knowledge about channel condition [2]. In other words, given a channel state matrix (whose row and column represents mobile and time slot), WPS tries to find the best channel allocation among all mobiles.

More specifically, it first performs WFQ regardless of channel condition. Then, given the simulated channel conditions, it checks the time slot allocations in the WFQ solution in the chronological order of the time slots. For each time slot, if the corresponding owner perceives a *dirty* channel, the scheduler will try to swap the allocation. The swapping algorithm will look in the future and search for a clean slot. Then it checks if the owner of the future time slot perceives a clean channel in the current time slot. If so, the allocation of both the mobiles will be swapped.

If it is not possible to find another eligible mobile, the current time slot will be given to another mobile for which the channel during the current slot is perceived to be *clean*, and the current mobile will be put off. The procrastination of the current mobile will be compensated in the next frame by means of lead-lag counter. In particular, the mobile relinquishing a time slot will be considered *lagging* while the mobile which is given an extra time slot will be considered *leading*. In the implementation, the lead and lag might be represented by a positive and negative number. At the beginning of each time frame, before scheduling, a set of *effective weights* is calculated by subtracting the lead counter from the original weights of each mobile. All the scheduling computations afterward will be performed using the effective weights rather than the original weights [2].

The channel state information along with the use of the swapping algorithm improve the performance significantly. The complexity of the swapping algorithm is $O(n)$ [1]. Therefore, the complexity of WPS is increased to $2 * O(n)$. Note that, the compensation by means of lead-lag counter is able to improve the performance without having any effect on the time-complexity of the algorithm.

III. An Assignment Problem and Its Solution, The Hungarian Method

Given n individuals, n jobs, as well as a set of costs $c_{ij} \ \epsilon \ I^{0+}$ corresponding to the assignment of j^{th} job to the i^{th} individual, the *assignment problem* is to achieve minimum cost for the assignment of all the jobs among the individuals such that each individual does exactly one job and each job is done by exactly one person [3]. Mathematically,

$$minimize \quad C_T = \sum_{i=1}^{n}\sum_{i=1}^{n} c_{ij} \cdot x_{ij} \tag{1}$$

subject to

$$x_{ij} = \begin{cases} 1, & \text{individual } i \text{ is assigned to job } j, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

$$\sum_{i} x_{ij} = 1 \quad \text{and} \tag{3}$$

$$\sum_{j} x_{ij} = 1. \tag{4}$$

A solution to this problem can be obtained by using the *Hungarian Method* which is related only to the manipulation of the matrix **C** $\epsilon \ \Re^{n\times n}$ whose elements are c_{ij} [3]. Since the objective is to minimize the cost function, the solution is located where the cost is minimum, i.e., x_{ij} = 1 if c_{ij} = 0. Note that, if c_{ij} = 0, x_{ij} will not necessarily be equal to 1.

We illustrate the Hungarian method [3],[4] by using the following cost matrix **C**:

$$\mathbf{C} = \begin{pmatrix} 5 & 2 & 3 & 4 \\ 7 & 8 & 4 & 5 \\ 6 & 3 & 5 & 6 \\ 2 & 2 & 3 & 5 \end{pmatrix} \tag{5}$$

- **Step1:** Subtract each element in a row by the minimum value in the row (i.e., $c_{ij} = c_{ij} - min_i, \forall_j$).
- **Step2:** Subtract each element in a column by the minimum value in the column (i.e., $c_{ij} = c_{ij} - min_j, \forall_i$).

After step 2, the matrix **C** will be as followed:

$$\mathbf{C} = \begin{pmatrix} 3 & 0 & 1 & 1 \\ 3 & 4 & 0 & 0 \\ 3 & 0 & 2 & 2 \\ 0 & 0 & 1 & 2 \end{pmatrix} \tag{6}$$

- **Step 3:** Use minimum lines drawn[1] through all zeros in the matrix **C**. If the number of lines is equal to the matrix dimension (i.e., n), go to step 5. Otherwise, go to step 4.

$$\mathbf{C} = \begin{pmatrix} 3| & 0| & 1 & 1 \\ \underline{3}| & \underline{4}| & \underline{0} & \underline{0} \\ 3| & 0| & 2 & 2 \\ 0| & 0| & 1 & 2 \end{pmatrix} \tag{7}$$

- **Step 4:** Find the minimum min_{ij} among the elements that are not drawn through by any line (in this example, min_{ij} is 1). Subtract min_{ij} from each element which does not has any line drawn through and add min_{ij} to each element which has two lines drawn through. Go back to step 3.

$$\mathbf{C} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 4 & 5 & 0 & 0 \\ 3 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{8}$$

Step 3:

$$\mathbf{C} = \begin{pmatrix} 3 & 0| & 0| & 0| \\ 4 & 5| & 0| & 0| \\ 3 & 0| & 1| & 1| \\ \underline{0} & \underline{0}| & \underline{0}| & \underline{1}| \end{pmatrix} \tag{9}$$

Now, the number of lines needed to cover all zeros is four, which is equal to the dimension of the matrix, and therefore, we proceed to step 5.

- **Step 5:** Now the solution can be identified on the lines as a set of zeros, such that every row and column has exactly one selected zero.

$$\mathbf{C} = \begin{pmatrix} 3 & 0 & 0 & \boxed{0} \\ 4 & 5 & \boxed{0} & 0 \\ 3 & \boxed{0} & 1 & 1 \\ \boxed{0} & 0 & 0 & 1 \end{pmatrix} \tag{10}$$

$$x_{ij} = \begin{cases} 1, & (i,j) = \{(4,1),(3,2),(2,3),(1,4)\}, \\ 0, & otherwise. \end{cases} \tag{11}$$

Therefore, the minimum cost will be achieved if we assign the 1^{st} job to 4^{th} individual, 2^{nd} job to 3^{rd} individual, 3^{rd} job to 2^{nd} individual, and 4^{th} job to 1^{st} individual. With this assignment the total cost is 2+3+4+4 = 13.

IV. Optimal Radio Channel Allocation (ORCA)

The wireless scheduling problem can be formulated as an assignment problem where each individual and job represents a mobile and a time slot, respectively. Each element c_{ij} in the

[1] A line drawn through a row is represented by lines under all elements in the row. A line drawn through a column is represented by | on the right of each element in the column. In (7), lines are drawn through row 2, column 1, and column 2.

cost matrix **C** represents channel state of mobile i during time slot j and can be defined as follows:

$$c_{ij} = \begin{cases} 0, & \text{slot } j \text{ is clean for mobile } i, \\ 1, & \text{slot } j \text{ is dirty for mobile } i. \end{cases} \quad (12)$$

The constraint (3) exhibits the fact that there can be only one mobile transmitting at a time to avoid collision, while (4) restricts the number of time slots in a frame for each mobile to be one. This would be the case where all mobiles have equal weight. In the case that the mobiles have different weights, say 1, 2, and 3, we can put in the cost matrix 1, 2, and 3 identical rows for 1^{st}, 2^{nd}, and 3^{rd} mobile, respectively. Note that the weights must be integer. If real-number weights are given, they must first be converted to integer.

In the Hungarian method, a square cost matrix is required to satisfy (3) and (4). In fact, if the number of columns is less than the number of rows, some mobiles might not get any allocation. On the other hand, if the number of columns is more than the number of rows, some time slots might be left empty. In order to make the number of columns equal to that of the rows, number of time slots per scheduling frame should be determined as $\sum_k w_k$. After the cost matrix is formed, the optimal allocation can be found in the same way as the assignment solution.

In practice, system frame size might be different from the scheduling frame size. If the scheduling frame is smaller than the system frame, several scheduling will be needed to fill a system frame. If the scheduling frame size is larger than the system frame size, on the other hand, the ORCA solution can fill up the system frame. The part of the solution which is not used to fill the system frame will fill the beginning of the next system frame.

The solution of the assignment problem (i.e. (1)) might allocate some dirty slots to a mobile. After the solution is obtained, the compensation using lead-lag counter as in WPS [2] is utilized to avoid useless transmission. Rather than using original weights, the scheduler utilizes effective weights calculated by subtracting the original weight by the current lead counter.

V. Performance Analysis

A. Performance Measures

To analyze the performance of the different fair scheduling schemes, the following parameters are defined:

- **Average Data Efficiency (γ)**:

$$\gamma = \frac{1}{n} \sum_{i=1}^{n} \frac{N_c^{(i)}}{N_t^{(i)}}. \quad (13)$$

- **Fairness** (F):

$$F = \left(\frac{1}{n} \cdot \sum_{i=1}^{n} \left(\frac{N_c^{(i)}}{\sum_{j=1}^{n} N_c^{(j)}} - w_i \right)^2 \right)^{\frac{1}{2}}. \quad (14)$$

- **Average Delay** (D):

$$D = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{\sum_{j=1}^{N_c^{(i)}} d_{ij}}{N_c^{(i)}}. \quad (15)$$

where $N_c^{(i)}$ is the number of packets transmitted by mobile i during *clean* time slots, $N_t^{(i)}$ is the total number of packets transmitted by mobile i, n is the total number of mobiles, w_i is the normalized weight of mobile i, and d_{ij} is the delay corresponding to the j^{th} successfully transmitted packet from mobile i.

Each mobile is assumed to have a one-packet buffer and no new packet is generated until the packet in the buffer has been transmitted. Delay is measured by the number of time slots required to transmit the packet from the buffer. Note that, the fairness measure (F) is based on the average Euclidean distance from actual allocation to the mobiles' weights. The ideal scheduling algorithm would yield minimum delay, data efficiency $\gamma = 1$, and fairness $F = 0$.

B. Wireless Channel Model

Wireless channel is modelled by a two-state Markov chain [5], where the steady-state packet error probability (P_E) is given by

$$P_E = \frac{1-p}{(1-p)+(1-q)} \quad (16)$$

in which p and q are the transition probabilities from *clean* to *clean* and from *dirty* to *dirty* state, respectively. We refer to $(1-p)+(1-q)$ as the *transition degree*. The perfect knowledge of channel is assumed to eliminate the effect of imprecise prediction and to clearly show the behavior of each scheduling protocol in presence of ideal prediction. Note that, given an average packet error rate P_E, the channel state transition probabilities can be obtained using normalized Doppler frequency $f_d T$ [6], where f_d is the maximum Doppler shift given by $f_d = \frac{v}{c} \cdot f_c$ (v is the mobile speed, c is the light speed, f_c is the carrier frequency) and T is the packet length. When $f_d T$ is small, the fading process is more correlated, while for higher values of $f_d T$, channel fading is more independent. In the simulation results presented in this paper, rather than using $f_d T$, we use different values of *transition degree* to investigate the impact of channel error correlation on the scheduler performance.

C. Simulation Parameters and Methodology

In the paper, the proposed ORCA is simulated in comparison to WFQ [1] and WPS [2]. The values used for P_E and *transition degree* are from the set {0.1, 0.15, 0.3} and {0.1, 0.3, 0.6, 1}, respectively. The number of mobiles is 10, all with the weight of 2, and the length of the simulation is 100000 time slots.

A mobile experiencing a *dirty* channel for a long time potentially has a large *lag* value. This mobile receives compensation at the beginning of every scheduling.

During the simulation, the compensation is set to one. In other words, the lead-lag counter is reset after the effective weights have been calculated. We also assume that, as in [2], the lead-lag counter is bounded between -4 to +4.

To observe both the short-term and long-term fairness and delay performances, different measurement windows are used. For the long-term and the short-term performance measures, window sizes of 4000 and 200, respectively, are used. The separation between two successive measurement windows is assumed to be 2000 and 100 time slots for the long-term and short-term case, respectively.

D. Results and Discussions

1) Data efficiency under different channel conditions: WFQ does not exploit the channel status information. As a result, the performance of WFQ degrades as the channel becomes more error-prone (Fig. 1). For WPS and ORCA, the data efficiency is always 1, because these schemes are aware of the channel state and capable of avoiding transmissions in the *dirty* channels.

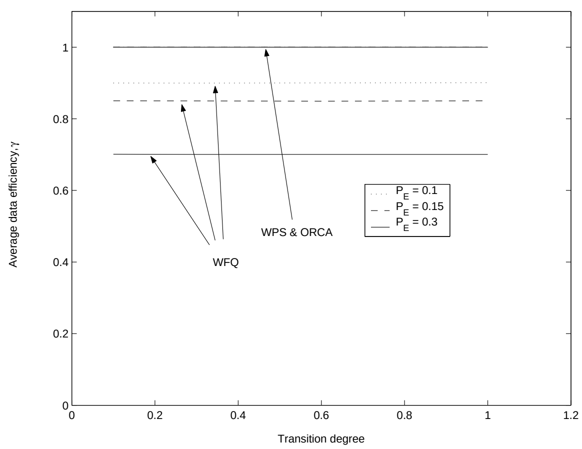

Fig. 1. Long-term average data efficiency under different channel conditions.

Note that, for all of the schemes, the *transition degree* does not have any impact on data efficiency.

2) Fairness under different channel conditions: As expected, for all the schemes, better fairness is achieved as the average packet error rate decreases (Figs. 2 and 3). Fairness of WFQ is not affected by *transition degree*, because the channel state does not have any effect on the allocation calculated by WFQ.

For WPS and ORCA, the fairness deteriorates as the channel states become more correlated. For a certain average packet error rate, if channel states are more correlated, it is more probable that a mobile will experience *dirty* channel for a long time. In the worst case, the channel might be *dirty* during all the time slots in a frame. In this case, the solution to satisfy fairness within a frame cannot be found. Since the compensation memory is restricted to one frame, a mobile experiencing *dirty* channel for two consecutive frame time will not be able to get the compensation.

WPS does not allow backward swapping. Therefore, it cannot give a better solution, like ORCA. At the same level of error rate, ORCA is always perform better than WPS.

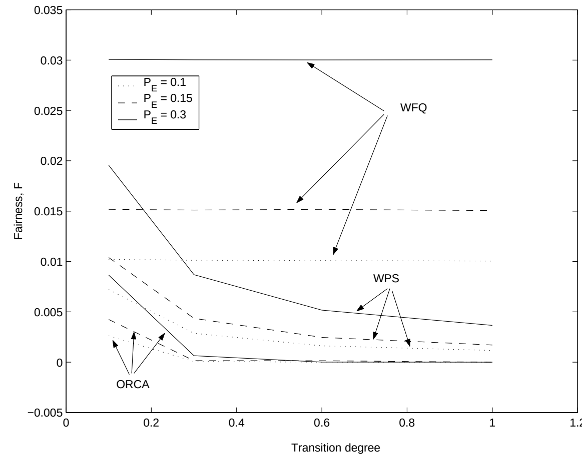

Fig. 2. Long-term fairness under different channel conditions.

Especially in the low transition degree environment, the improvement of ORCA from WPS is tantamount to that of WPS from WFQ. Note that ORCA brings about large improvement when a channel changes slowly. This is a favorable feature because the prediction of channel becomes more inaccurate in a fast-changing channel. However, in the worst case, time-complexity of the Hungarian method can be as high as O(n^3) [7] , where n, the number of rows in the cost matric, is a function of the number of simultaneous users and the variation of the weight of each user. In practice, the number of the simultaneous is limited to a small value. If variation of the client's weight is sufficiently small, the complexity can be limited to a reasonable degree.

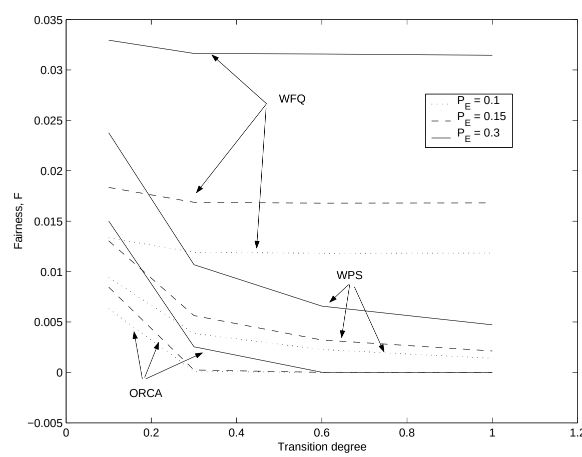

Fig. 3. Short-term fairness under different channel conditions.

In general, both the algorithms have two main procedures: intra-frame scheduling and inter-frame compensation. WPS and ORCA employs the swapping algorithm and the Hungarian method as the former procedure. The lead-lag compensation is exercised as the latter procedure for both the approaches.

Although leading to long-term fairness, the intra-frame compensation leads to short-term unfairness. If the number of compensation frames and lead-lag bound are set to infinity, a mobile will definitely be able to obtain the compensated allocation but such a compensation might be far in the future. When a mobile experiences a *dirty* channel for a long time. After some time, its effective weight will grow without limit. The other mobiles perceiving *clean* channels will possess minimum effective weights. After the channel of the deferred mobile becomes better, the mobile may 'log' the channel, thereby deteriorating the fairness (especially short-term fairness).

The other extreme is the case when we limit the number of compensation frames or the lead-lag bound to zero. If a *clean* channel cannot be found for a mobile within a scheduling frame, the scheduler will not compensate for the allocation that has been put off at all.

ORCA has the potential to find a better solution for intra-frame allocation than WPS. Therefore, it is expected that better short-term fairness can be achieved with ORCA regardless of the number of compensation frames. With large number of compensation frames, the long-term fairness for both schemes might become comparable. Note that, the long-term fairness (Fig. 2) will be better than the short-term fairness (Fig. 3).

3) Average packet delay under different channel conditions: The trend of average packet delay is similar to that of fairness. In other words, it decreases as transition degree increases.

The performance comparison between WPS and ORCA in terms of the long-term average packet delay is shown in Fig. 4. Both algorithms perform better as *transition degree* increases, for the same reason as described in the previous section. ORCA always outperforms WPS under the same channel condition. Other than causing more delay, the impact of channel condition on short-term delay is similar to that on long-term delay, thus being omitted in this paper.

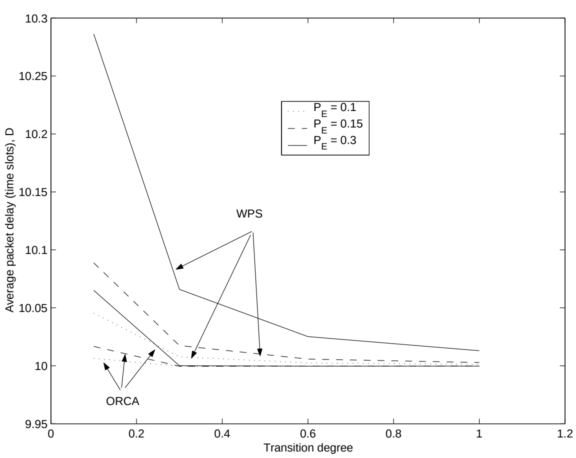

Fig. 4. Comparison between WFQ and ORCA in terms of long-term average packet delay.

VI. SUMMARY

Fair scheduling schemes such as WFQ or WPS do not perform optimization, and therefore, the best solution might not be found in some cases. We have formulated the scheduling problem for fair bandwidth allocation among the mobile nodes as an assignment problem. The optimization using the Hungarian method and lead-lag compensation constitute the *Optimal Radio Channel Allocation* (ORCA) protocol. Simulation results have shown that, due to such optimization the ORCA scheme outperforms both the WFQ and WPS scheme.

Impacts of system parameters (e.g., number of mobiles, delay bound for packets from each mobile, compensation memory) on the performance of the proposed ORCA scheme are currently being investigated. The performance of ORCA in the presence of imperfect channel estimation will be also investigated. Also, in the proposed formulation, the binary condition of the cost function might be relaxed. The non-binary value of cost element c_{ij} might be based on how bad the channel is.

Instead of the assignment problem, a transportation or linear programming problem, where the solution is not restricted to binary, might be employed. Under these circumstances, the solution might exhibit the level at which each mobile should transmit. For example, such a solution can be applied to adjust the power level as well as the transmission rate of each mobile, in which case not only fairness but also energy efficiency can be improved. Also, a constraint on delay bound could be added to the objective function in order to support the QoS requirements of different applications.

REFERENCES

[1] T. Nandagopal, S. Lu, and V. Bhargharvan, "A unified architecture for the design and evaluation of wireless fair queuing algrithms," *Proc. of ACM Mobicom'99*, pp. 132-142, Aug. 1999.

[2] S. Lu, V. Bharghavan, and R. Skikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Trans. on Networking*, vol. 7, No. 4, pp. 473-489, Aug. 1999.

[3] H. W. Khun, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, Q.2, pp. 83-97, 1955.

[4] L. Cooper and D. Steinberg, *Methods and Applications of Linear Programming*, Philadelphia: Saunders, 1974.

[5] H. S. Wang and N. Moayeri, "Finite-state markov channel - a useful model for radio communication channels," *IEEE Transaction on Vehicular Technology*, vol. 44, no. 1, pp. 163-171, Feb. 1995.

[6] A. Chockalingam, M. Zorzi, L. B. Milstein, and P. Venkataram, "Performance of a wireless access protocol on correlated Rayleigh fading channels with capture," *IEEE Trans. Commun.*, vol. 46, pp. 644-655, May 1998.

[7] C. H. Papadimitriou and K.Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*, New Jersey: Prentice-Hall, 1982.