Effect of Priority Class Ratios on the Novel Delay Weighted Priority Scheduling Algorithm

Vasco QUINTYNE
Department of Computer Science, Physics and Mathematics, University of the West Indies
Cave Hill, St. Michael, Barbados

and

Adrian ALS
Department of Computer Science, Physics and Mathematics, University of the West Indies
Cave Hill, St. Michael, Barbados

ABSTRACT

In this paper, the novel Delay Weighted Priority Scheduling (DWPS) algorithm proposed for the management of cellular network's channel access will be presented. Whereas the simple round-robin algorithm, like that used in the popular Point Coordinated Function (PCF), fails to consider the delay implications on stations with different priority levels, in the DWPS algorithm, mobile stations are queued and assigned priority levels in accordance with the quality of service (QoS) requirements of their applications. Results will show that the DWPS outperforms the simple round-robin algorithm for cell sizes of up to 25 stations.

Keywords: Delay Tolerance Threshold, Delay Weighted Priority Scheduling, Incidence of QoS Degradation, QoS Tolerance Threshold and Ready Ratio

1. INTRODUCTION

Attempts at devising fair channel access schemes for cellular environments have traditionally focused on ensuring guaranteed access to all mobile stations within each cell or ensuring that the available bandwidth is adequately allocated amongst the active stations. Non-contentious scheduling algorithms, particularly the round-robin scheme used within the Point Coordinated Function (PCF) [1-7], have proven to give guaranteed channel access. However, the demand for the highest quality of service (QoS) levels by end-users not only requires guaranteed channel access and adequate bandwidth to be taken into account but also the delay tolerance of the applications being executed on the mobile stations. The differences in the various delay tolerance levels [1] present in a multi-application environment is significant enough to validate the use of priority classes. In the proposed Delay Weighted Priority Scheduling (DWPS) algorithm, mobile stations are queued and assigned priority levels in accordance with the QoS requirements of their applications.

Note that the QoS experienced by a station is partially determined by the queue delay, which is the time a station spends waiting in the queue before it is granted access to the channel. Note that if the scheduling algorithm does not employ prioritisation then the combination of long queue lengths and a high QoS requirement would prove detrimental to stations using delay sensitive applications. Conversely, when prioritization is introduced, the QoS may also be adversely affected due to channel access starvation caused by the preferential treatment of higher priority stations. The treatment of the DWPS algorithm in this paper is aimed at finding the right balance between these extremes.

In section 2, the literature review focuses on the PCF scheduling algorithm. Additionally, the motivation for a new approach to scheduling channel access in cellular networks is presented as a precursor to the novel DWPS algorithm. The DWPS algorithm is then presented in section 3. Afterwards, the methodology along with the assumptions and parameters that were used in the simulations are presented in section 4. The simulation results are then provided in section 5. Finally, the key findings are discussed in section 6.

2. LITERATURE REVIEW

To provide a non-contentious facility for channel sharing the PCF, which employs a round-robin technique to administer channel access, was incorporated into the IEEE MAC protocol [2;4]. More specifically, the IEEE 802.11 MAC protocol, is comprised of the Distributed Coordinated Function (DCF) [2;8] and the PCF algorithms. Together they incorporate the five (5) timing intervals shown in Table 1[5].

Types of Intervals	Duration (μs)
Short Interframe Space (SIFS)	10
Slot time	20
Priority Interframe Space (PIFS)	30
Distributed Interframe Space (DIFS)	50
Extended Interframe Space (EIFS)	Variable but > DIFS

Table 1: IEEE MAC Protocol Timing Intervals

The reader is asked to refer to [2;8] for more information on the DCF algorithm. The PCF protocol is addressed in the next subsection.

PCF Protocol
Figure 1 illustrates the operation of the PCF protocol outlined below:

1) The base station (or access point, in other wireless networks) polls the first mobile station on its list. Note that the polling message may also piggyback additional data for the polled mobile station

2) The polling station then responds by sending its data, otherwise only an acknowledgment (ACK) frame is sent before the SIFS expires.
3) The base station then polls the next station in accordance with one of the following scenarios:
   i) after it receives an ACK from the polled station or
   ii) after it receives any data the polled station has to send or
   iii) after the SIFS expires and before the PIFS has passed.
4) Repeat from step 1 until the polling list is empty or the non-contentious period of the channel access scheduling has ended.

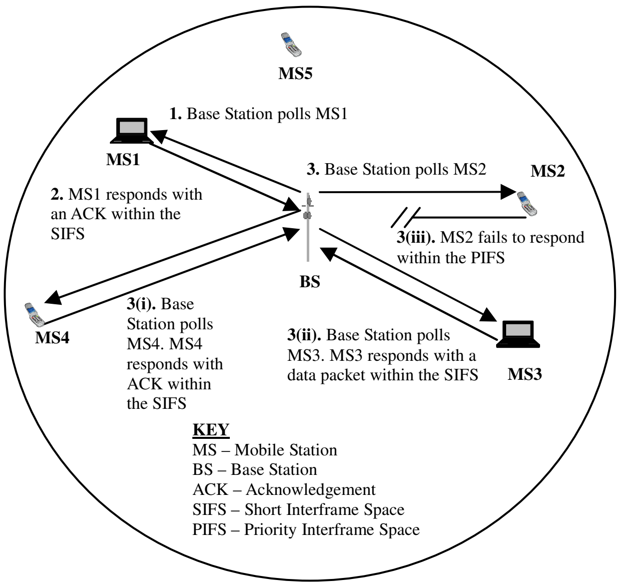

Figure 1 Operation of PCF

The PCF protocol is inefficient as it polls stations that do not require the immediate use of the channel. Moreover, as it does not incorporate any enhancing priority-based schemes, high-priority traffic may be adversely affected. As a result, mobile stations using delay sensitive applications may suffer a significant decrease in the Quality of Service (QoS).

3. THE NOVEL DWPS ALGORITHM

The DWPS algorithm is being proposed as an alternative to the PCF for providing fair channel access to mobile stations within a network. The mechanism by which this is achieved is based on the use of:

1. the application priority levels listed in Table 2,
2. first-in-first-out (FIFO) queues and
3. a ready ratio (RR), defined as the ratio between a station's queue delay and the application's delay tolerance threshold (DTT).

Note that the Delay Tolerance Threshold (DTT) is the application's absolute delay tolerance threshold, which when surpassed results in application function failure. However, the QoS of the applications noticeably degrade before delay reaches the DTT value. For example, the QoS of voice applications degrades after 150ms but can still operate up to 400 ms [1]. For this reason the QoS Delay Threshold (QTT) was devised. This is the preferred upper delay limit of an application. The QTT is used to calculate the RR, which represents the degree of channel access urgency required by a mobile station.

Priority Level	Media Type	Example	DTT	Priority Status
1	Voice	Mobile VoIP	400ms	Very High
2	Video	Video	400ms	High
3	Transaction Services (Low Priority)	Small Messaging Service (SMS)	30s	Medium
4	Bulk Data Transfer	File Transfer	30s	Low

Table 2: Application Priority Levels

The RR is based on an operating system scheduling algorithm called Highest Response Ratio Next (HRRN) [9-11]. The modified equation used for the DWPS is defined as:

$$RR = \frac{\left(QTT_{Px} - W_q\right) + QTT_{Px}}{QTT_{P1}} \qquad (1)$$

where:

W_q = time spent waiting to use the channel
QTT_{Px} = QTT of a station with priority level x
QTT_{P1} = QTT of a station with priority level 1

In the DWPS algorithm, each station wishing to access the shared transmission media is assigned a priority level ranging from 1 (highest) to 4 (lowest) in accordance with the type of application being used on the mobile station.

The priority levels are used to determine which queue the wireless stations will be inserted into and thus by extension, the associated QoS experienced by the application. Mobile stations monitor a control channel to determine when they have been assigned access to the communication channel. In the initial instance the stations are assigned to the channel on a first-come first-served basis. The remaining mobile stations, if any, are then placed into a channel access priority queue as illustrated by the flowchart in Figure 2.

The DWPS algorithm accommodates the use of both priority and multiple queues. The algorithm was used to test the effect of priority ratios on the delay experienced by stations within the cell. The methodology employed to investigate this is described in the next section.

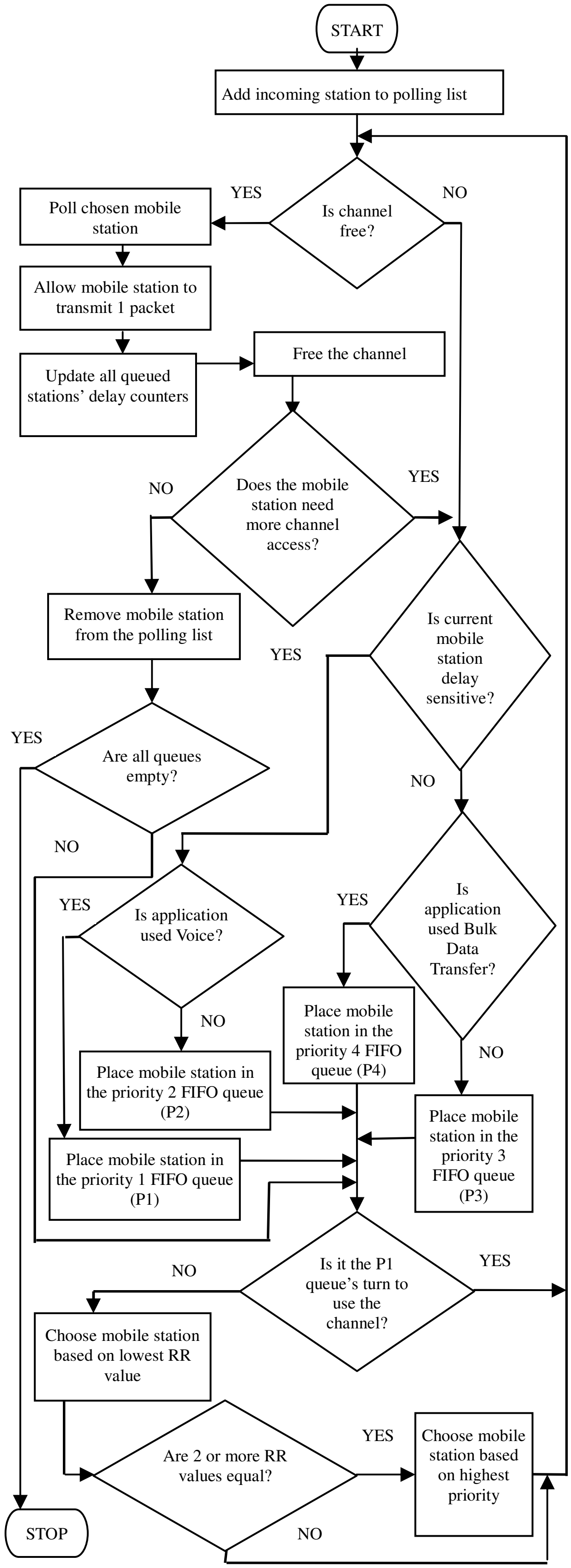

Figure 2: DWPS Algorithm Flowchart

4. METHODOLOGY

The DWPS algorithm was implemented in Java and the simulating environment was based on the following simplifying assertions aimed at reducing processing time:

1. Each mobile station is associated with a single priority level.
2. Communication only occurs with stations in an adjacent cell.
3. The number of mobile stations within the cell is limited to 15, 20 or 25.
4. All priority level 3 mobile stations send 2 packets inclusive of the control packet. All other stations send 61 packets in total.
5. All stations initially have data to send when polled by the base station (BS).
6. When a polled station has finished its communication a control packet is transmitted to the base station and the mobile station is removed from the polling list.

Note that the simulation is not real-time and measurements such as transmission time and propagation delay are calculated. The major simulation parameters used in these calculations are listed in Table 3 below.

Attribute	Measurement
Cell radius	150 m
Link Capacity	2Mbps
Transmission Speed	299792458 ms^{-1}
MS-to-BS distance	100 m

Table 3: Simulation Parameters

Calculation of the Average IDQ in the DWPS Algorithm

An incidence of QoS degradation (IQD) occurs when a particular mobile station's delay surpasses its QTT value. The average IQD for each priority class is calculated by Eq. (2).

$$\overline{IQD_p} = \frac{\sum_{x=1}^{n_p} IQD_p}{n_p}, \quad \text{for } 1 \le p \le 4 \tag{2}$$

where:

p represents the priority level

IQD_p is the number of IQDs for the priority p mobile stations

$\overline{IQD_p}$ is the average IQD for the priority p mobile stations

n_p is the number of priority p stations

5. RESULTS

Figures 3, 4 and 5 show the results of simulating the DWPS algorithm when the total number of stations in the cell is 15, 20 and 25, respectively. In each graph, a plot of the ratio between high priority (P1 and P2) and low priority (P3 and P4) stations against the average IQD for each priority level is presented. Note that the queue length for each priority level is directly proportional to the total number of mobile stations (n_T) within the cell.

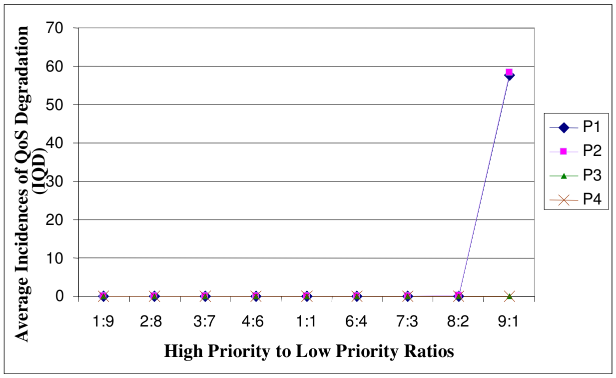

Figure 3: Effect of Priority Ratios on the DWPS Algorithm in a cell containing 15 mobile stations

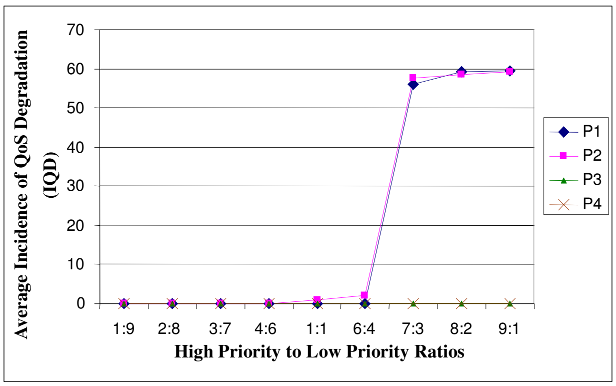

Figure 4: Effect of Priority Ratios on the DWPS Algorithm in a cell containing 20 mobile stations

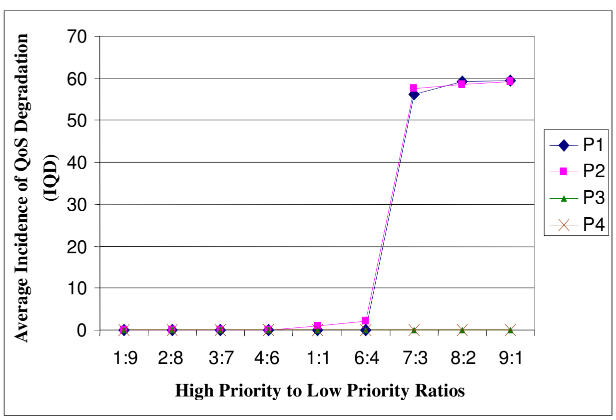

Figure 5: Effect of Priority Ratios on the DWPS Algorithm in a cell containing 25 mobile stations

Observe that in Figure 3, there is a negligible increase of 0.2 in the average IQD value for the priority 2 mobile stations when the ratio is 8:2. This average value increases rapidly to ~58 which is slightly above the average IQD for the priority 1 mobile stations. This trend, which is also present in figures 4 and 5, is due to the favoured treatment of the priority 1 mobile stations over the priority 2 mobile stations by the DWPS algorithm.

Note that the steep increases in the average IQD are an indicator that the algorithm is failing and that the end used may notice degradation in the QoS. In figures 4 and 5 these increases occur near the 6:4 ratios. This lower ratio, compared to the one in figure 3, indicates that as the queue length increases the ability of the DWPS algorithm to successfully manage large numbers of high priority stations decreases. The almost identical graphs produced in figures 4 and 5 suggests that when the cell has 20 stations or more the DWPS algorithm reaches its optimum operating capacity at a high priority to low priority ratio of 4:6. However, this still remains manageable with low IQD values occurring at a ratio of 6:4.

Figure 6 and 7 show the results of simulating the Round-Robin algorithm using the same conditions as those set out in the DWPS simulation. In particular, the behaviour for priority 1 and 2 mobile stations when the total number of stations in the cell is equal to 15 is presented in figure 6. Indeed the results highlight that the algorithm begins to fail after the ratio exceeds 1:1; thus end users may experience degraded QoS levels. In figure 7 all the delay sensitive queues maintain an almost constant average IQD of 60. This indicates that when the number of stations reaches 20, the round-robin algorithm is unable to operate efficiently with regard to the delay sensitive mobile stations. Consequently, the results for a cell containing 25 mobile stations are not presented.

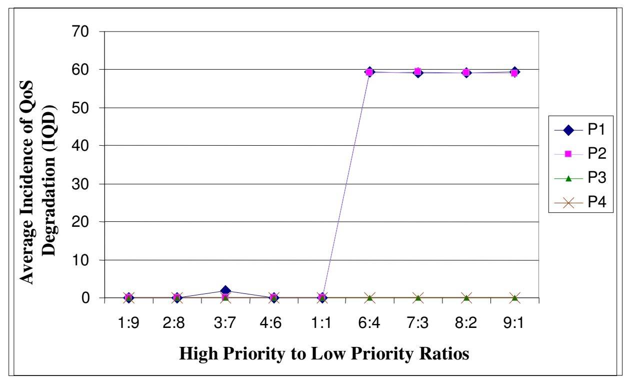

Figure 6: Use of Priority Ratios on the Round-Robin Algorithm in a cell containing 15 mobile stations

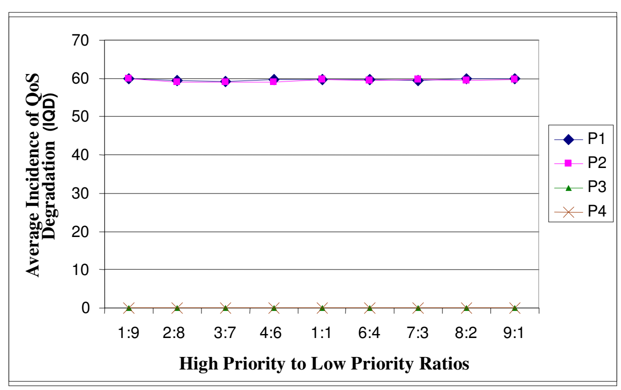

Figure 7: Use of Priority Ratios on the Round-Robin Algorithm in a cell containing 20 mobile stations

Note that all the graphs produced for both the DWPS and round-robin algorithms show an IQD value of zero for the delay tolerant priority 3 and 4 applications. Thus the mobile stations using these applications can operate within their QoS limits for both algorithms as the priority ratios has no affect on these mobile stations when $n_T \leq 25$.

6. CONCLUSION

Results indicate that the ratio of high priority to low priority stations influences the operation of the DWPS algorithm. Indeed the success in the DWPS algorithm can be seen when comparing figures 3 and 6 with respect to the delay in the onset of higher IQD values. Additionally, in comparing figures 4 and 7 the DWPS categorically outperforms the round robin for all ratios.

Optimum ratios for the DWPS algorithm are below 4:6, when the cell contains 20 stations or more. For cells containing 15 stations or lower the optimum ratio can increase to at least 7:3. In essence, to work at optimum levels DWPS requires the ratio of high priority to low priority stations to decrease as the queue sizes increase. These priority ratios could be used to regulate cell admission schemes where QoS is a major concern.

Results show that the DWPS algorithm can be used in preference to the round-robin algorithm as it performs better in each of the cases examined. However, an investigation of the DWPS algorithm when the assumptions are relaxed will be essential to further ascertain its viability in the field.

Note that in the future, more research will be done to include other established non-contentious algorithms such as RETHER [2;6].

7. REFERENCES

[1] A. Tee and J. Cleveland, "A new option proposed for 802.20 requirements on latency and packet error rates", IEEE, 2004. Available at http://www.ieee802.org/ 20/Contribs/C802.20-04-55.ppt

[2] S. Sharma, "Analysis of 802.11b MAC: A QoS, Fairness, and Performance Perspective", 2003 Available at http://www.ecsl.cs.sunysb.edu/tr/wlanrpe.pdf

[3] B. Bensaou, Y. Wang, and C. C. Ko, "Fair medium access in 802_11 based wireless ad-hoc networks", **IEEE/ACM MobiHoc Workshop**, 2000, pp. 99-106.

[4] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11 Wireless Local Area Networks", **IEEE Communications Magazine**, Vol. 35, No. 9, 1997, pp. 116-126.

[5] K. K. Leung, B. McNair, L. J. Cimini, and J. H. Winters, "Outdoor IEEE 802.11 Cellular Networks: MAC Protocol Design and Performance", **IEEE International Conference on Communications**, 2002, pp. 595-599.

[6] T. Chiueh, "RETHER: A Software-Only Real-Time Ethernet for PLC Networks", **Proceedings of the Workshop on Embedded Systems on Workshop on Embedded Systems**, 1999, pp. 6.

[7] L. Chandran-Wadia, "Throughput Performance of the Distributed and Point Coordination Functions of an IEEE 802.11 Wireless LAN", S. Mahajan and S. Iyer, Eds., **Proceedings of the Workshop on Embedded Systems on Workshop on Embedded Systems,** 2002, pp. 36-49.

[8] L. Zheng, A. Dadej, and S. Gordon, "Fairness of IEEE 802.11 Distributed Coordination Function for Multimedia Applications", 2003. Available at http://www.itr.unisa.edu.au/~sgordon/doc/zheng2003-fairness.pdf

[9] C. Ritcjie, **Operating Systems incorporating UNIX & Windows**, 3 ed. New York: Continuum, 1997.

[10] W. Stallings, **Operating Systems**, 2 ed. Upper Saddle River, New Jersey: Prentice Hall, 1995.

[11] W. Stallings, **Operating Systems: Internals and Design Principles**, 6 ed. Upper Saddle River, New Jersey: Pearson Prentice Hall, 2008, pp. 422-423.