Simulation Analysis of Packet Scheduling Algorithm for WWW and Video Streaming Service in UMTS Downlink FDD Mode

MARKO PORJAZOSKI, BORISLAV POPOVSKI
Department of Telecommunications, Faculty of Electrical Engineering
St.s Cyril and Methodius University
Karpos b.b., 1000, Skopje
FYR Macedonia

Abstract: - In this paper, we focus on data transmission in WCDMA systems using packet scheduling for DCH and DSCH, in case when only video streaming service users and WWW service users are engaged. Network exploiting the DCH and DSCH as a data transport channels can provide higher throughput than a network without DSCH, if a good combination of resource sharing between the DCH and DSCH is selected.

Key-Words: - UMTS, WCDMA, Packet Scheduling, DCH, DSCH, Streaming Video, Interactive WWW Service

1 Introduction
WCDMA is a 3G standard, aiming to provide high speed wireless data services with different Quality of Services (QoS) guarantees [1][2]. One of the features of WCDMA is the reach set of both dedicated channels (e.g. Downlink Physical Data Channel (DPDCH)) and shared channels (e.g. Downlink Shared Channel (DSCH)). These channels can support circuit and packet-switched services to enable multimedia applications ranging from voice to best-effort data.

2 Transport Channels
Two transport channels caring a downlink transmission are [3]:

a) Dedicated Channel (DCH). DCH is available exclusively to a specific user. Dedicated Channel is devoted to services with stringent transfer delay requirements. The transmission rate of a DCH can be changed every 10 ms.

b) Downlink Shared Channel (DSCH). The Downlink Shared Channel is defined to support flexible multiplexing of bursty data traffic in WCDMA. DSCH is usually used to support best-effort data services, as it cannot provide guarantees for the service quality. Its Transmission capacity is divided up among several users. The number of users multiplexed on DSCH varies with time. Since the base station may transmit to many users at one time, co-channel interference exists.

Depending on the type of service to be provided, transport channels should be managed and allocated appropriately. In this paper we are focusing on interactive and streaming services, whose quality requirements deal with the achieved bit rate, the percentage of lost packets, delay and the jitter of the delay (for streaming service).

In order to differentiate quality levels for streaming services, we assume two layered video application that is characterized by two different flows: a basic layer, with the minimum requirements for a proper visualization, and an enhancement layer, that contains additional information to improve the quality of the received images. Basic layer will be transmitted through DCH channel while the enhancement layer will be transmitted only if there is capacity in the DSCH channels. The DCH channel operates at a fixed bit rate equal to the source bit rate, which means that a fixed number of transport blocks should be transmitted in each Transmission Time Interval (TTI). It is assumed that the possible retransmissions of the basic layer can be carried out in the DSCH channel together with the enhancement layer, and having a higher precedence than the latter.

In case of interactive (WWW) service users, data is transmitted in DCH. DCH will operate at variable bit rate for this users.

3 Interference and Code Management
Radio Resource Management (RRM) is responsible for utilization of the air interface resources. RRM is needed to guaranty Quality of Services (QoS), to maintain the planned coverage area and to offer high capacity [1]. RRM strategies comprise of several algorithms and they all have in common the monitoring of the cell load factor for adopting the algorithms decisions. In downlink direction RRM functions include: admission control, congestion control, packet scheduling and code management.
Code management is devoted to manage Orthogonal Variable Spreading Factor (OVSF) code tree used to allocate physical channel orthogonality among different users. For services with tolerant delay requirements different user’s flows can be scheduled to use shared channels. Decision about who should transmit and its transmission parameters (transport format and power level) are the responsibility of the packet scheduler.

For the \( n \) users transmitting simultaneously at a given cell, the following inequality for \( i \)-th user must be satisfied [4]:

\[
\frac{P_{h,i}}{L_{p}(d_{i})} \times SF_{i} \geq \frac{E_{b}}{N_{a}} \times \frac{P_{r}}{L_{p}(d_{i})} + \rho \times \frac{P_{r} - P_{h,i}}{L_{p}(d_{i})} - I_{i-oth} \geq \sum_{i=1}^{n} P_{r,i} - \rho \times \frac{P_{r} - P_{h,i}}{L_{p}(d_{i})}
\]

(1)

\( P_{r} \) – base station transmitted power; \( P_{h,i} \) – power devoted to \( i \)-th user; \( I_{i-oth} \) – intercell interference observed by \( i \)-th user; \( L_{p}(d_{i}) \) – \( i \)-th user path loss; \( r \) – channel coding rate; \( P_{h} \) – background noise; \( SF \) – spreading factor, relating the bit duration to the chip duration; \( \rho \) - orthogonality between codes used in downlink direction.

It is obtained that:

\[
P_{h,i} \geq \frac{L_{p}(d_{i})}{SF_{i}} \left( \frac{P_{r}}{P_{h,i}} \right) + \frac{P_{r} - P_{h,i}}{L_{p}(d_{i})} \frac{E_{b}}{N_{a}} \times r + \rho \times \frac{P_{r} - P_{h,i}}{L_{p}(d_{i})}
\]

(2)

Adding all \( n \) (users) inequalities, the total power transmitted by the base station inequalities can be expressed as:

\[
P_{r \text{ max}} \geq P_{r} = \frac{P_{h}}{(1 - \eta_{DL})} \sum_{i=1}^{n} \frac{L_{p}(d_{i})}{SF_{i}} \left( \frac{E_{b}}{N_{a}} \times r + \rho \times \frac{P_{r} - P_{h,i}}{L_{p}(d_{i})} \right)
\]

(3)

where \( \eta_{DL} \) is load factor defined as [1]:

\[
\eta_{DL} = \sum_{i=1}^{n} \left( \frac{L_{p}(d_{i})}{SF_{i}} \left( \frac{E_{b}}{N_{a}} \times r + \rho \times \frac{P_{r} - P_{h,i}}{L_{p}(d_{i})} \right) \right) < 1
\]

(5)

Beside interference control, packet scheduler should manage dynamical allocation of OVSF codes. The system has \( SF_{\text{max}} \) orthogonal codes with maximum spreading factor \( SF_{\text{max}} = 512 \). According to the properties of these codes, there availability is guaranteed when the Kraft’s inequality is satisfied [5]:

\[
\sum_{i=1}^{n} \frac{R_{b,i}}{R_{b}} \leq SF_{\text{max}}
\]

(6)

where \( R_{b,i} \) is number of bits in transport block (TB) for \( i \)-th user and \( R_{b} \) is minimal number of bits in TB (corresponding to spreading factor \( SF_{\text{max}} = 512 \)).

4 Algorithm description

The scheduling strategy used in our simulations is presented in [4][6]. Algorithm is divided in two parts:

- Prioritization. All users intended to transmit information must be classified according to a certain prioritization criteria. First scheduler will order requests depending on service class they belong to, from highest to lowest priority level. Second prioritization rule is based on number of basic layer TB in users buffer waiting for retransmission. When two or more users have the same number of TB to be retransmitted, a third prioritization level based on the Service Credit (SCr) is considered.

The SCr is associated with an active link or user and it computes the difference between the bit rate requested by the user and the bit rate that the system has provided to him. So it accounts for the amount of service the system owes to the user. The SCr value of each active connection must be updated every TTI, following the expression:

\[
SCr(k) = SCr(k-1) + \frac{RG}{TB} - NumTx(k-1)
\]

(7)

where SCr(k) is the Service Credit for TTI=k, SCr(k-1) is the Service Credit in the previous TTI, RG is the guaranteed bit rate measured in bits/TTI, TB is the number of bits of a Transport Block for the considered RAB and NumTx(k-1) is the number of successfully transmitted Transport Blocks in the previous TTI.

- Resource Allocation and Availability Check

Once requests are ordered, the next step consists in deciding whether or not they are accepted for transmission and which is the accepted TF. The limitations dealing with interference and code availability are taken into account in this phase. To this end, it is required to estimate the expected load factor and transmitted power level once all the
requests are accepted. Then, the expected load factor whenever there are \( n \) transmissions in the system in frame \( t \) is:

\[
\bar{\eta}(n,t) = \sum_{i=1}^{n} \frac{P_y I_{0,n}(t-1) \times L_y(d_i)}{P_y}.
\]  

(8)

The expected power is given by:

\[
\bar{P}_y(n,t) = \frac{P_y}{(1-\bar{\eta}(n,t))} \sum_{i=1}^{n} \frac{L_y(d_i)}{SF_i} + \rho
\]  

(9)

Figure 1. Resource allocation process

Some differences between real and estimated load factor value can occur, as a consequence of inaccuracies in the measurement of the other-to-own-cell interference and path loss.

Algorithm execution follows the flow shown on Figure 1. Assuming a \( n \) already granted transmissions and initially selected TF for \( n+1 \) request, the Kraft’s inequality is evaluated, the expected load factor is compared with a threshold \( \phi \) and expected transmission power should be below the maximum transmitted power. If all tree conditions are satisfied, transmission is granted for this request during one TTI, otherwise, the transport format is reduced by one (i.e. transmission bit rate is reduced). If this is not possible, the request should wait for the next frame.

5 Simulation scenario

The system model includes 7 omnidirectional base stations. Distance between base stations is 1 km. Mobile users are uniformly distributed in the scenario moving with speed of 50 km/h. Log-normally distributed shadowing with standard deviation of 10 dB is considered. Maximum transmitted power by the base station is 43 dBM.

Traffic generation model proposed in [7] is used, including following parameters:
- Interactive Service (WWW traffic): Session arrival process – Poisson process; Number of packet call requests per session – geometrically distributed with a mean value 5 calls within a session; Reading time between packet calls – geometrically distributed with a mean 33 [s]; Number of packets within a packet call – geometrically distributed with a mean value 25; Inter arrival time between packets – geometrically distributed with a mean 0.0625 [s] (for 64 kbps); Packet size – Pareto Distribution (with cut-off), max packet size 66666 bytes with parameters \( \alpha=1.1, k=81.5 \).

We considered eight possible transport formats for interactive class listed below in Table 1.
- Streaming Basic and Enhancement Service: CBR model; bit rate 32 kbps (each 40 ms packet with 1280 bits is generated); TTI 40 ms; Transport Block Size 320 bits; Basic Layer has two TFs: TF0 – no transmission and TF1 transmission of 4 TB in TTI.

Table 1. DCH TFs for 64 kbps WWW traffic and DSCH TFs for Streaming video enhance-ment layer

<table>
<thead>
<tr>
<th>TB sizes, bits</th>
<th>320 bits (payload) + 16 bits (MAC/RLC header)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFS</td>
<td>TH0, bits</td>
</tr>
<tr>
<td></td>
<td>TF1, bits</td>
</tr>
<tr>
<td></td>
<td>TF2, bits</td>
</tr>
<tr>
<td></td>
<td>TF3, bits</td>
</tr>
<tr>
<td></td>
<td>TF4, bits</td>
</tr>
<tr>
<td></td>
<td>TF5, bits</td>
</tr>
</tbody>
</table>

| TTI, ms       | 40 |

In our simulations we have considered two different cases:
- number of video users is 140 and number of WWW users is changing from 120 to 200
- number of video users is 160 and number of WWW users is changing from 120 to 200.
6 Results
System parameters are dependable from load factor and number of users in the system. Load Factor and power estimation, as it's shown in equation (8) and (9), is based on interference measured in previous TTI. For higher numbers of users difference between estimated and real value of load factor is becoming greater, so base station transmitted power is reaching faster its maximum value and as a result less users can transmit their data (especially in DSCH). Algorithm best performances (user bit rate, packet lost and delay jitter) are achieved for load factor = 0.95.

7 Conclusions
In this paper, the performance of the WCDMA network area with seven cells, exploiting the DCH and DSCH was studied by simulations. It was shown that as the traffic load increased, the network resources utilization increased until it reached the maximum. If load is further increased, the resource utilization saturates and starting decreasing again. The results suggest that there is an optimal set of parameters, by which the network resources utilization reaches the maximum.

References:
[2] 3GPP TS 23.107, Quality of Service (QoS) concept and architecture
[3] 3GPP TS 25.211, Physical channels and mapping of transport channels onto physical channels (FDD)
[7] TR 101 112-UMTS 30.03, Selection procedures for the choice of radio transmission technologies of the UMTS