
Hongfei Du

Submitted for the Degree of
Doctor of Philosophy
from the
University of Surrey

UNIVERSITY OF SURREY

Centre for Communication System Research
School of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, UK

© Hongfei Du, June 2007
This thesis is dedicated to my parents and my wife, with love and thanks for all they have done for me throughout my life as well as during the life of my PhD studies.
Acknowledgments

I would like to take the opportunity and express my sincere gratitude to my principal supervisor Prof. Barry Evans, director of the centre for communication systems research (CCSR), University of Surrey, for his support and guidance throughout the research work during my PhD. My special thanks are to my supervisor Dr. Linghang Fan for his guidance of my research to this topic, and devoted a lot of his valuable time and attention to review my papers. My gratitude goes also to University of Surrey and European Union for supporting financially my PhD studies at CCSR.

I would like to express my sincere thanks to: Victor Kueh, for guiding and helping me significantly during initial stage of my PhD, and always supporting and advising me; Upendra Mudugamuwa for his valuable help on the initial step of the ns2 simulator and many technical discussions.

I would like to be grateful to the members of the CCSR for the precious discussion moments with whom I shared with: Atta Qudus, Merkourios Karaliopoulos, Guray Acar, Wei Koong Chai, Malumbo Chipeta, Aduwati Sali, Yaning Wang, Dan He.

Many thanks also to Liz James, Anne Ruben, Stephanie Evans, Chris Clerk, Mary Oduko and Terry Roberts, for their secretarial and computing support in CCSR.

I am extremely grateful to my family, my father and mother, and my sister, for their continuous patience and support during my studies aboard in U.K.

I leave my deepest gratitude to my beloved wife Lin Zhang, for her love and patience, which make life easier and joyful during the difficult but memorable times. My sincere thanks to her parents for offering me such a wonderful and beautiful wife, whose love and encouragement mean invaluable to me; and whose advice and artistry inspire me to create miracle and make my dream true.

Hongfei Du

Guildford, June 2007
Abstract

In recent years, there has been tremendous growth in digital multimedia technologies, from voice to data to video, and the recent, but growing, demand of supporting diverse quality of service (QoS) guarantees. It places new demands for future wireless networks in utilising the available radio resource in a more efficient and effective way. The key to this demand is the involvement of efficient radio resource management (RRM), to provide various QoS support for multimedia service delivery. Due to the unique broadcast nature and ubiquitous coverage of satellite communication system, the synergy between satellite networks and terrestrial networks provides new opportunities for delivering point-to-multipoint (or one-to-many) multimedia content to a large audience spread over extensive geographical area. It is expected that the satellite component will play a complementary, but essential, role in delivering multimedia data to those areas where the terrestrial high-bandwidth communication infrastructures are, either economically or technically, unreachable. The emphasis throughout this research is on the potential optimization techniques pertinent to the RRM to facilitate multimedia broadcast/multicast service (MBMS) delivery over the satellite digital multimedia broadcasting (SDMB) system, which has emerged as one of the most promising approaches for this mission.

We concentrate on the algorithm development and performance evaluation of RRM strategies implemented at the radio access layer in the SDMB system, aimed at the efficient delivery of multimedia applications to mobile users at satisfactory QoS.

Firstly, we have developed a novel two-level channel multiplexing scheme for the radio resource allocation (RRA), which is capable of optimally performing both logical and transport channel multiplexing via two new optimization algorithms, namely optimum estimation algorithm (OEA) and power-oriented adaptation (POA) algorithm. Simulation results prove that the proposed scheme can effectively improve the performance in terms of channel utilisation, power consumption as well as transmission capacity.

Secondly, we have investigated the optimization of packet scheduling algorithms via the various adaptations of a proportional differentiation model. By taking into account multiple performance measures, e.g., buffer occupancy, queuing delay and data rate, several novel algorithms, i.e. buffer-length related queue (BLRQ), delay differentiation queue (DDQ) and combined delay and rate differentiation (CDRD), are introduced for performing the packet scheduling task in SDMB. Their performance has been evaluated via simulation means and compared with existing schemes.
It is demonstrated that the proposed proportional differentiation packet scheduling schemes can achieve significant performance improvements in queuing delay, jitter and channel utilisation.

Finally, we further optimize the packet scheduling schemes by using cross-layer design and adaptive optimization mechanisms. Cross-layer joint priority queue (CJPQ) and adaptive multi-dimensional QoS-based (AMQ) packet scheduling algorithms are introduced within this context and investigated via simulations. It is shown that these proposals can significantly improve the QoS performance amongst heterogeneous competing flows in terms of both scheduling efficiency and fairness, offering better flexibility and scalability features.

**Key Words:** SDMB, MBMS, RRM, RRA, channel mapping/multiplexing, packet scheduling, cross-layer design.
# Table of contents

ACKNOWLEDGMENTS .................................................................................................................. II

ABSTRACT ................................................................................................................................ III

TABLE OF CONTENTS ................................................................................................................ V

LIST OF FIGURES .......................................................................................................................... IX

LIST OF TABLES ............................................................................................................................ XII

GLOSSARY OF TERMS .................................................................................................................... XV

1 INTRODUCTION ............................................................................................................................. 1
  1.1 Background .......................................................................................................................... 1
  1.2 Motivation and scope ......................................................................................................... 3
  1.3 Major contributions .......................................................................................................... 5
  1.4 Thesis organisation ......................................................................................................... 7

2 OVERVIEW OF SATELLITE DIGITAL MULTIMEDIA BROADCASTING ............................................. 9
  2.1 Digital multimedia broadcasting ...................................................................................... 9
  2.2 Multimedia Broadcast Multicast Service ......................................................................... 10
  2.3 Satellite Digital Multimedia Broadcasting ...................................................................... 11
    2.3.1 Convergence of satellite and terrestrial networks .................................................... 11
    2.3.2 The SDMB system architecture .............................................................................. 13
    2.3.3 SDMB access scheme and air interface ................................................................... 16
  2.4 QoS provisioning ............................................................................................................. 17
  2.5 Cross-layer optimization ................................................................................................. 19

3 RADIO RESOURCE MANAGEMENT FOR SDMB ........................................................................ 21
  3.1 Background ...................................................................................................................... 21
    3.1.1 Radio resource allocation ....................................................................................... 22
    3.1.2 Packet scheduling ................................................................................................. 23
    3.1.3 Admission control ................................................................................................... 24
    3.1.4 Interaction between RRM entities .......................................................................... 25
  3.2 Literature review on radio resource allocation .................................................................. 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.1</td>
<td>Channel mapping and multiplexing in SDMB</td>
<td>26</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Related work and problem analysis</td>
<td>27</td>
</tr>
<tr>
<td>3.3</td>
<td>Literature review on packet scheduling techniques</td>
<td>30</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Packet scheduling for general wired systems</td>
<td>30</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Packet scheduling schemes in wireless systems</td>
<td>34</td>
</tr>
<tr>
<td>3.4</td>
<td>Packet scheduling in SDMB</td>
<td>36</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Overview</td>
<td>36</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Assessment of existing packet scheduling schemes in SDMB</td>
<td>37</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>NOVEL RADIO RESOURCE ALLOCATION SCHEME</td>
<td>40</td>
</tr>
<tr>
<td>4.1</td>
<td>Overview</td>
<td>40</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction</td>
<td>40</td>
</tr>
<tr>
<td>4.3</td>
<td>Feasibility analysis</td>
<td>42</td>
</tr>
<tr>
<td>4.4</td>
<td>Proposed radio resource allocation strategy</td>
<td>45</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Overview</td>
<td>45</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Two-stage bin-packing with optimum estimation algorithm</td>
<td>46</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Power-oriented adaptation algorithm</td>
<td>53</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Derivation of the TFCS</td>
<td>56</td>
</tr>
<tr>
<td>4.5</td>
<td>Performance evaluation</td>
<td>59</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Evaluation metrics/simulation scenarios</td>
<td>59</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Evaluation results</td>
<td>61</td>
</tr>
<tr>
<td>4.6</td>
<td>Summary</td>
<td>73</td>
</tr>
<tr>
<td>5</td>
<td>PROPORTIONAL DIFFERENTIATION BASED PACKET SCHEDULING</td>
<td>75</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>75</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Overview</td>
<td>75</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Design of packet scheduling algorithm</td>
<td>76</td>
</tr>
<tr>
<td>5.2</td>
<td>Buffer length related queuing</td>
<td>78</td>
</tr>
<tr>
<td>5.3</td>
<td>Delay differentiation queuing</td>
<td>80</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Overview</td>
<td>80</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Description of DDQ scheduling algorithm</td>
<td>80</td>
</tr>
<tr>
<td>5.4</td>
<td>Performance evaluation of scheduling algorithms</td>
<td>82</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Performance evaluation of BLRQ</td>
<td>82</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Performance evaluation of DDQ</td>
<td>84</td>
</tr>
<tr>
<td>5.5</td>
<td>Summary</td>
<td>88</td>
</tr>
</tbody>
</table>
# Table of contents

6 Packet Scheduling with Cross-layer Design ................................................................. 89

6.1 Introduction ................................................................................................................. 89
6.2 Combined delay and rate differentiation queuing ......................................................... 90
  6.2.1 Motivation ............................................................................................................... 90
  6.2.2 Description of CDRD scheduling algorithm ......................................................... 90
6.3 Cross-layer Joint Priority Queuing packet scheduling ............................................... 95
  6.3.1 Overview of CJPQ Scheme ................................................................................... 95
  6.3.2 Description of CJPQ scheduling algorithm ........................................................... 97
6.4 Dynamic resource allocation ...................................................................................... 101
  6.4.1 Background .......................................................................................................... 101
  6.4.2 Dynamic rate matching ....................................................................................... 102
  6.4.3 DRM-based resource allocation ........................................................................... 103
6.5 Performance evaluation ............................................................................................. 106
  6.5.1 Simulation scenarios ............................................................................................ 106
  6.5.2 Performance evaluation for DDQ with DRA ...................................................... 107
  6.5.3 Performance evaluation for CDRD with DRA .................................................. 111
  6.5.4 Performance evaluation for CJPQ with DRA ...................................................... 120
6.6 Summary .................................................................................................................... 123

7 Adaptive Multi-dimensional QoS-based Packet Scheduling .................................... 124

7.1 Introduction ................................................................................................................ 124
7.2 Motivation for the AMQ scheduling scheme ............................................................. 125
7.3 Adaptive multidimensional QoS-based packet scheduling ........................................ 127
  7.3.1 Overview .............................................................................................................. 127
  7.3.2 Adaptive service prioritization ............................................................................ 129
  7.3.3 Adaptive resource allocation .............................................................................. 132
  7.3.4 Flexibility and scalability analysis ....................................................................... 136
7.4 Performance Analysis .............................................................................................. 137
  7.4.1 Simulation scenario ............................................................................................. 137
  7.4.2 Queuing delay evaluation .................................................................................... 139
  7.4.3 Delay threshold flexibility ................................................................................... 146
  7.4.4 Physical channel utilisation .................................................................................. 148
  7.4.5 Fairness analysis .................................................................................................. 149
7.5 Summary .................................................................................................................... 152

8 Conclusions and Future Work .................................................................................... 153
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Thesis conclusions</td>
<td>153</td>
</tr>
<tr>
<td>8.2 Future work</td>
<td>155</td>
</tr>
<tr>
<td>8.2.1 Cross-layer RRM investigating TCP-MAC interactions</td>
<td>155</td>
</tr>
<tr>
<td>8.2.2 SDMB with return links supporting interactive applications</td>
<td>156</td>
</tr>
<tr>
<td>PUBLICATIONS</td>
<td>157</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>159</td>
</tr>
<tr>
<td>APPENDIX A: STREAMING SERVICE CHARACTERISATION</td>
<td>166</td>
</tr>
<tr>
<td>APPENDIX B: RADIO BEARER CONFIGURATION</td>
<td>170</td>
</tr>
<tr>
<td>APPENDIX C: TRANSPORT CHANNEL MULTIPLEXING AND TRANSPORT FORMAT COMBINATION</td>
<td>182</td>
</tr>
<tr>
<td>APPENDIX D: DESCRIPTION OF THE SDMB SYSTEM-LEVEL SIMULATOR</td>
<td>185</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1-1. Structure of the Thesis ........................................... 7
Figure 2-1: SDMB system architecture ........................................ 14
Figure 2-2. SDMB reference architecture and relevant radio interface .................................................. 16
Figure 3-1: Interaction between RRM entities in SDMB system ........................................................................ 25
Figure 3-2. WCDMA channel mapping and relevance to SDMB ........................................................................ 26
Figure 3-3. Illustration of transport channel multiplexing in SDMB [18] .................................................. 28
Figure 3-4. Residual capacities in single-level channel multiplexing .................................................. 29
Figure 3-5. Packet scheduling strategy procedures ........................................................................... 36
Figure 4-1. Channel mapping options considering (a) Transport channel multiplexing (b) Logical and transport channel multiplexing .............................................................................. 41
Figure 4-2. MAC-m architecture in SDMB: (a) SDMB RAN side, (b) UE side [97] ........................................ 43
Figure 4-3. Illustration of logical channel multiplexing in SDMB ........................................................................ 44
Figure 4-4: Radio resource allocation framework ................................................................................. 46
Figure 4-5. Inter-relationship between the two multiplexing stages .................................................. 48
Figure 4-6: Optimum estimation algorithm for estimating required FACHs ........................................ 49
Figure 4-7. Power-oriented adaptation algorithm .............................................................................. 55
Figure 4-8. Channel mapping structure utilising 2-stage bin-packing algorithm combination for scenario 1A ................................................................................................................................. 62
Figure 4-9. Required number of FACHs for single-level and 2-level channel multiplexing under different traffic mixes ................................................................................................................................. 64
Figure 4-10. Total transmission capacity for single-level and 2-level channel multiplexing under different traffic mixes ................................................................................................................................. 65
Figure 4-11. Transmission capacity utilisation for single-level and 2-level channel multiplexing under different traffic mixes ................................................................................................................................. 66
Figure 4-12. Required number of FACHs for single-level and 2-level channel multiplexing under different S-CCPCH capacities ................................................................................................................................. 67
Figure 4-13. Total transmission capacity for single-level and 2-level channel multiplexing under different S-CCPCH capacities ................................................................................................................................. 67
Figure 4-14. Transmission capacity utilisation for single-level and 2-level channel multiplexing under different S-CCPCH capacities ................................................................................................................................. 68
Figure 4-15. Comparison of TFCS size for different S-CCPCHs between single-level and 2-level channel multiplexing under various scenarios ................................................................................................................................. 70
List of Figures

Figure 4-16. Total physical channel transmit power under different S-CCPCHs.............................71
Figure 4-17. Total physical channel transmit power under different traffic mixes...........................72
Figure 5-1: Packet scheduling procedure ...............................................................................................77
Figure 5-2: Buffer length for MLPQ and BLRQ scheduling ..............................................................83
Figure 5-3: Packet mean delay at RLC buffers for different PS algorithms .....................................85
Figure 5-4: Packet mean jitter at RLC buffers for different PS algorithms .......................................86
Figure 5-5: Code channel utilisation for MLPQ and DDQ ..................................................................87
Figure 6-1: Proposed packet scheduling framework ...........................................................................92
Figure 6-2. Architecture of CJPQ packet scheduling scheme .............................................................95
Figure 6-3. Illustration of the layer interactions of the proposed cross-layer packet scheduling scheme in the SDMB GW..........................................................96
Figure 6-4. The relationship between Rate Matching and Transport Channel multiplexing [14] 101
Figure 6-5. Procedures of DRM-based resource allocation algorithm .............................................104
Figure 6-6. Flowchart for the DRA algorithm .....................................................................................105
Figure 6-7. Queuing delay of different scheduling schemes ............................................................108
Figure 6-8. Queuing jitter of different scheduling schemes ..............................................................110
Figure 6-9. Throughput of different packet scheduling schemes .....................................................110
Figure 6-10. Queuing delay CDF distribution for streaming performance evaluation at RLC layer. .......................................................................................................................111
Figure 6-11. Queuing delay variation performance evaluation at RLC layer. ....................................113
Figure 6-12: PDF delay for MLPQ scheduling ...................................................................................114
Figure 6-13: PDF delay for CDRD scheduling ..................................................................................114
Figure 6-14. Queuing delay performance evaluation under different delay thresholds. ................115
Figure 6-15: Mean code channel utilisation for MLPQ and CDRD scheduling ..............................116
Figure 6-16: Instantaneous code channel utilisation for MLPQ (a) and CDRD (b) scheduling. 117
Figure 6-17: Mean FACH channel utilisation for MLPQ and CDRD scheduling ............................118
Figure 6-18. Performance evaluation on channel throughput/utilisation achieved at the physical layer: (a) Instantaneous data rate of FACH with MLPQ scheduling, (b) Instantaneous data rate of FACH with CDRD scheduling, (c) Instantaneous physical channel utilisation for MLPQ and CDRD scheduling using SRM and DRM techniques.........................................................119
Figure 6-19. Comparisons on queuing delay performance with different scheduling schemes... 120
Figure 6-20. Comparison on physical channel utilisation with different scheduling schemes..... 121
Figure 6-21. Comparison on packet loss rate with different scheduling schemes......................... 122
Figure 7-1. Proposed AMQ packet scheduling framework ...............................................................128
Figure 7-2. Proposed adaptive resource allocation procedures. ......................................................134
List of Figures

Figure 7-3. An example of TFC selection mechanism under different resource allocation schemes. ................................................................................................................................. 135

Figure 7-4. Queuing delay experienced at the RLC buffer in the satellite hub under different scheduling schemes for Scenario I: (a) mean queuing delay, (b) CDF queuing delay for MLPQ, (c) CDF queuing delay for DDQ, (d) CDF queuing delay for AMQ. .............................................. 141

Figure 7-5. Queuing delay performance under different scheduling schemes for Scenario II: (a) mean queuing delay, (b) PDF delay for MLPQ, (c) PDF delay for DDQ, (d) PDF delay for AMQ. ...................................................................................... 144

Figure 7-6. Queuing delay performance at the RLC buffer in the satellite hub under different scheduling schemes for Scenario III: (a) mean queuing delay, (b) mean queuing jitter for different scheduling schemes ....................................................................................................... 145

Figure 7-7. Evaluation of queuing delay threshold variations for AMQ scheme: (a) Scenario II, (b) Scenario III ............................................................................................................................... 146

Figure 7-8. PDF distribution of throughput ratios for Scenario III: (a) MLPQ-based scheduling, (b) DDQ-based scheduling, (c) AMQ-based scheduling .................................................................................. 150

Figure 7-9. Variance of throughput ratios for Scenario III .................................................................................................................................................. 151

Figure 8-1. Comparison of queuing delay for different scheduling schemes for Scenario III ............................................. 154

Figure B-1: Channel mapping structure utilising FF+BF bin-packing algorithm combination for 80% streaming–20% download traffic mix. ................................................................. 174

Figure C-1. Data exchange through protocol layers .............................................................................................................................. 182

Figure D-1. Functional models of system-level SDMB simulator ................................................................................................. 186
List of Tables

Table 4-1. Simulation scenario reference table .....................................................................................61
Table 4-2. Derived channel mapping configuration: Scenario 1A (kbps) ...........................................62
Table 4-3. Derived channel mapping configuration: Scenario 1B (kbps) ...........................................63
Table 4-4. Derived channel mapping configuration: Scenario 2B (kbps) ...........................................63
Table 4-5. TFCS size for respective S-CCPCHs in single-level channel multiplexing. ...................69
Table 4-6. TFCS size for respective S-CCPCHs in Two-level channel multiplexing. .......................69
Table 4-7. Mapping derived for scenario 1A: (a) POA-based scenario; (b) Worst-case scenario. 71
Table 4-8. Mapping derived for scenario 2B: (a) POA-based scenario; (b) Worst-case scenario. 73
Table 5-1. Radio bearer mapping for the evaluation of BLRQ. ...........................................................82
Table 5-2. Mean queuing delay for MLPQ and BLRQ (seconds) ......................................................82
Table 5-3: Simulation multiplexing scenario ........................................................................................84
Table 6-1. Factors considered in the joint priority function ...............................................................97
Table 6-2: Channel multiplexing configuration ..................................................................................107
Table 7-1. Radio bearer mapping configuration (kb/s) for Scenario I .................................. 138
Table 7-2. Radio bearer mapping configuration (kb/s) for Scenario II ...........................................138
Table 7-3. Radio bearer mapping configuration (kb/s) for Scenario III ..........................................138
Table 7-4. Mean physical channel utilisation for different PS schemes (%) ...................................148
Table A-1: Streaming service characterisation for 80% streaming -20% download given 3x384kbps S-CCPCH ..........................................................166
Table A-2: Streaming service characterisation for 50% streaming -50% download given 3x384kbps S-CCPCH ..........................................................167
Table A-3: Streaming service characterisation for 80% streaming -20% download given 3x384kbps S-CCPCH ..........................................................167
Table A-4: Streaming service characterisation for 80% streaming -20% download given 3x128 kbps S-CCPCH ..........................................................167
Table A-5: Streaming service characterisation for 50% streaming -50% download given 3x128 kbps S-CCPCH ..........................................................167
Table A-6: Streaming service characterisation for 20% streaming -80% download given 3x128 kbps S-CCPCH ..........................................................168
Table A-7: Streaming service characterisation for 80% streaming -20% download given 1x384 kbps S-CCPCH ..........................................................168
| Table A-8: Streaming service characterisation for 50% streaming -50% download given 1x384 kbps S-CCPCH | 168 |
| Table A-9: Streaming service characterisation for 20% streaming -80% download given 1x384 kbps S-CCPCH | 168 |
| Table A-10: Streaming service characterisation for 80% streaming -20% download given 1x384kbps + 3x128 kbps S-CCPCH | 169 |
| Table A-11: Streaming service characterisation for 50% streaming -50% download given 1x384kbps + 3x128 kbps S-CCPCH | 169 |
| Table A-12: Streaming service characterisation for 20% streaming -80% download given 1x384kbps + 3x128 kbps S-CCPCH | 169 |
| Table B-1: Required streaming MTCHs and FACHs for 80% streaming-20% download given 3x384 kbps S-CCPCH | 170 |
| Table B-2: Required streaming MTCHs and FACHs for 50% streaming-50% download given 3x384 kbps S-CCPCH | 170 |
| Table B-3: Required streaming MTCHs and FACHs for 20% streaming-80% download given 3x384 kbps S-CCPCH | 171 |
| Table B-4: Required streaming MTCHs and FACHs for 80% streaming - 20% download given 3x128 kbps S-CCPCH | 171 |
| Table B-5: Required streaming MTCHs and FACHs for 50% streaming - 50% download given 3x128 kbps S-CCPCH | 171 |
| Table B-6: Required streaming MTCHs and FACHs for 20% streaming - 80% download given 3x128 kbps S-CCPCH | 171 |
| Table B-7: Required streaming MTCHs and FACHs for 80% streaming - 20% download given 1x384 kbps S-CCPCH | 172 |
| Table B-8: Required streaming MTCHs and FACHs for 50% streaming - 50% download given 1x384 kbps S-CCPCH | 172 |
| Table B-9: Required streaming MTCHs and FACHs for 20% streaming - 80% download given 1x384 kbps S-CCPCH | 172 |
| Table B-10: Required streaming MTCHs and FACHs for 20% streaming-80% download given 1x384kbps+3x128kbps S-CCPCH | 172 |
| Table B-11: Required streaming MTCHs and FACHs for 50% streaming-50% download given 1x384kbps+3x128kbps S-CCPCH | 173 |
| Table B-12: Required streaming MTCHs and FACHs for 20% streaming-80% download given 1x384kbps+3x128kbps S-CCPCH | 173 |
Table B-13: Mapping derived for 80% streaming – 20% download given 3x384 kbps S-CCPCH .......................................................... 173
Table B-14: Mapping derived for 50% streaming – 50% download given 3x384 kbps S-CCPCH .......................................................... 174
Table B-15: Mapping derived for 20% streaming - 80% download given 3x384 kbps S-CCPCH .......................................................... 175
Table B-16: Mapping derived for 80% streaming – 20% download given 3x128 kbps S-CCPCH .......................................................... 176
Table B-17: Mapping derived for 50% streaming – 50% download given 3x128 kbps S-CCPCH .......................................................... 176
Table B-18: Mapping derived for 20% streaming – 80% download given 3x128 kbps S-CCPCH .......................................................... 177
Table B-19: Mapping derived for 80% streaming – 20% download given 1x384 kbps S-CCPCH .......................................................... 178
Table B-20: Mapping derived for 50% streaming – 50% download given 1x384 kbps S-CCPCH .......................................................... 178
Table B-21: Mapping derived for 20% streaming – 80% download given 1x384 kbps S-CCPCH .......................................................... 179
Table B-22: Mapping derived for 80% streaming – 20% download given 1x384+3x128 kbps S-CCPCH .................................................. 180
Table B-23: Mapping derived for 50% streaming – 50% download given 1x384+3x128 kbps S-CCPCH .................................................. 180
Table B-24: Mapping derived for 20% streaming – 80% download given 1x384+3x128 kbps S-CCPCH .................................................. 181
Table D-1. System simulation parameters ................................................................................................................................. 188
## Glossary of Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G / 3G</td>
<td>2nd / 3rd Generation</td>
</tr>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project</td>
</tr>
<tr>
<td>AC</td>
<td>Admission Control</td>
</tr>
<tr>
<td>ACK</td>
<td>ACKnowledgment</td>
</tr>
<tr>
<td>AMC</td>
<td>Adaptive Modulation and Coding</td>
</tr>
<tr>
<td>AMQ</td>
<td>Adaptive Multi-dimensional QoS-based</td>
</tr>
<tr>
<td>APF</td>
<td>Adaptive Proportional Fairness (chapter 3)</td>
</tr>
<tr>
<td>ARA</td>
<td>Adaptive Resource Allocation</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat reQuest</td>
</tr>
<tr>
<td>ASP</td>
<td>Adaptive Service Prioritization</td>
</tr>
<tr>
<td>BC/MC</td>
<td>Broadcast/Multicast</td>
</tr>
<tr>
<td>BCCH</td>
<td>Broadcast Common Control CHannel</td>
</tr>
<tr>
<td>BCH</td>
<td>Broadcast CHannel</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>BF</td>
<td>Best Fit</td>
</tr>
<tr>
<td>BLER</td>
<td>Block Error Rate</td>
</tr>
<tr>
<td>BLRQ</td>
<td>Buffer Length Related Queuing</td>
</tr>
<tr>
<td>BM-SC</td>
<td>Broadcast Multicast Service Center</td>
</tr>
<tr>
<td>C/I</td>
<td>Carrier-to-Interference ratio</td>
</tr>
<tr>
<td>CBP</td>
<td>Call Blocking Probabilities</td>
</tr>
<tr>
<td>CBQ</td>
<td>Class Based Queuing</td>
</tr>
<tr>
<td>CBR</td>
<td>Constant Bit Rate</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code Division Multiple Access</td>
</tr>
<tr>
<td>CDP</td>
<td>Call Dropping Probabilities</td>
</tr>
<tr>
<td>CDRD</td>
<td>Combined Delay and Rate Differentiation</td>
</tr>
<tr>
<td>CIF</td>
<td>Channel-condition Independent Fair</td>
</tr>
<tr>
<td>CIF-Q</td>
<td>Channel-condition Independent Fair Queuing</td>
</tr>
</tbody>
</table>


xv
<table>
<thead>
<tr>
<th>Glossary of Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CJPQ</strong></td>
</tr>
<tr>
<td><strong>CPICH</strong></td>
</tr>
<tr>
<td><strong>CRC</strong></td>
</tr>
<tr>
<td><strong>CS</strong></td>
</tr>
<tr>
<td><strong>CSI</strong></td>
</tr>
<tr>
<td><strong>DAB</strong></td>
</tr>
<tr>
<td><strong>DDQ</strong></td>
</tr>
<tr>
<td><strong>DMB</strong></td>
</tr>
<tr>
<td><strong>DPF</strong></td>
</tr>
<tr>
<td><strong>DRA</strong></td>
</tr>
<tr>
<td><strong>DRM</strong></td>
</tr>
<tr>
<td><strong>DTX</strong></td>
</tr>
<tr>
<td><strong>DVB</strong></td>
</tr>
<tr>
<td><strong>DVB-H</strong></td>
</tr>
<tr>
<td><strong>DVB-RCS</strong></td>
</tr>
<tr>
<td><strong>DVB-S</strong></td>
</tr>
<tr>
<td><strong>EDF</strong></td>
</tr>
<tr>
<td><strong>EDGE</strong></td>
</tr>
<tr>
<td><strong>ESA</strong></td>
</tr>
<tr>
<td><strong>ETSI</strong></td>
</tr>
<tr>
<td><strong>EU</strong></td>
</tr>
<tr>
<td><strong>FACH</strong></td>
</tr>
<tr>
<td><strong>FDD</strong></td>
</tr>
<tr>
<td><strong>FDMA</strong></td>
</tr>
<tr>
<td><strong>FEC</strong></td>
</tr>
<tr>
<td><strong>FF</strong></td>
</tr>
<tr>
<td><strong>FIFO</strong></td>
</tr>
<tr>
<td><strong>FLO</strong></td>
</tr>
<tr>
<td><strong>FQ</strong></td>
</tr>
<tr>
<td><strong>FSMC</strong></td>
</tr>
<tr>
<td><strong>FTP</strong></td>
</tr>
<tr>
<td><strong>GEO</strong></td>
</tr>
<tr>
<td>Term</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>GoS</td>
</tr>
<tr>
<td>GPRS</td>
</tr>
<tr>
<td>GPS</td>
</tr>
<tr>
<td>GSM</td>
</tr>
<tr>
<td>GW</td>
</tr>
<tr>
<td>HEO</td>
</tr>
<tr>
<td>HLS</td>
</tr>
<tr>
<td>HSDPA</td>
</tr>
<tr>
<td>HTTP</td>
</tr>
<tr>
<td>HPD</td>
</tr>
<tr>
<td>IEE</td>
</tr>
<tr>
<td>IEEE</td>
</tr>
<tr>
<td>IETF</td>
</tr>
<tr>
<td>IMR</td>
</tr>
<tr>
<td>IP</td>
</tr>
<tr>
<td>IPv4</td>
</tr>
<tr>
<td>IPv6</td>
</tr>
<tr>
<td>ISDN</td>
</tr>
<tr>
<td>ISP</td>
</tr>
<tr>
<td>IST</td>
</tr>
<tr>
<td>ITU</td>
</tr>
<tr>
<td>IWFQ</td>
</tr>
<tr>
<td>JJF</td>
</tr>
<tr>
<td>JPF</td>
</tr>
<tr>
<td>kbps</td>
</tr>
<tr>
<td>LAN</td>
</tr>
<tr>
<td>LEO</td>
</tr>
<tr>
<td>LOS</td>
</tr>
<tr>
<td>MAC</td>
</tr>
<tr>
<td>MAESTRO</td>
</tr>
</tbody>
</table>
| Glossary of Terms

<table>
<thead>
<tr>
<th>MBMC</th>
<th>MBM Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBMS</td>
<td>Multimedia Broadcast/Multicast Service</td>
</tr>
<tr>
<td>MCCCH</td>
<td>MBMS point-to-multipoint Control CHannel</td>
</tr>
<tr>
<td>MICH</td>
<td>MBMS Notification Indicator Channel</td>
</tr>
<tr>
<td>MLPQ</td>
<td>Multi-Level Priority Queuing</td>
</tr>
<tr>
<td>MoDiS</td>
<td>Mobile Digital broadcast Satellite</td>
</tr>
<tr>
<td>ms</td>
<td>millisecond</td>
</tr>
<tr>
<td>MS</td>
<td>Mobile Station</td>
</tr>
<tr>
<td>MSCH</td>
<td>MBMS point-to-multipoint Scheduling CHannel</td>
</tr>
<tr>
<td>MTCH</td>
<td>MBMS point-to-multipoint Traffic CHannel</td>
</tr>
<tr>
<td>NF</td>
<td>Next Fit</td>
</tr>
<tr>
<td>NLOS</td>
<td>Non-line-of-sight</td>
</tr>
<tr>
<td>NRT</td>
<td>Non Real Time</td>
</tr>
<tr>
<td>OEA</td>
<td>Optimum Estimation Algorithm</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency Division Multiplexing</td>
</tr>
<tr>
<td>P-CCPCH</td>
<td>Primary Common Control Physical CHannel</td>
</tr>
<tr>
<td>PDA</td>
<td>Personal Digital Assistant</td>
</tr>
<tr>
<td>PDCP</td>
<td>Packet Data Convergence Protocol</td>
</tr>
<tr>
<td>PDD</td>
<td>Proportional Delay Differentiation</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PDU</td>
<td>Protocol Data Unit</td>
</tr>
<tr>
<td>PF</td>
<td>Proportional Fair</td>
</tr>
<tr>
<td>PHY</td>
<td>Physical Layer</td>
</tr>
<tr>
<td>PLR</td>
<td>Packet Loss Rate</td>
</tr>
<tr>
<td>POA</td>
<td>Power-Oriented Adaptation</td>
</tr>
<tr>
<td>PRIQ</td>
<td>PRIority Queuing</td>
</tr>
<tr>
<td>PS</td>
<td>Packet Scheduling</td>
</tr>
<tr>
<td>p-t-mp</td>
<td>Point-to-multipoint</td>
</tr>
<tr>
<td>p-t-p</td>
<td>Point-to-point</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RAB</td>
<td>Radio Access Bearer</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>RAN</td>
<td>Radio Access Network</td>
</tr>
<tr>
<td>RLC</td>
<td>Radio Link Control</td>
</tr>
<tr>
<td>RM</td>
<td>Rate Matching</td>
</tr>
<tr>
<td>RMTP</td>
<td>Reliable Multicast Transport Protocol</td>
</tr>
<tr>
<td>RNC</td>
<td>Radio Network Controller</td>
</tr>
<tr>
<td>RR</td>
<td>Round Robin</td>
</tr>
<tr>
<td>RRA</td>
<td>Radio Resource Allocation</td>
</tr>
<tr>
<td>RRC</td>
<td>Radio Resource Control</td>
</tr>
<tr>
<td>RRM</td>
<td>Radio Resource Management</td>
</tr>
<tr>
<td>RTT</td>
<td>Round-Trip Time</td>
</tr>
<tr>
<td>s</td>
<td>seconds</td>
</tr>
<tr>
<td>Sat-GW</td>
<td>Satellite Gateway</td>
</tr>
<tr>
<td>SATIN</td>
<td>SATellite UMTS IP based Network</td>
</tr>
<tr>
<td>SatNEx</td>
<td>Satellite Communications Network of Excellence</td>
</tr>
<tr>
<td>S-CCPCH</td>
<td>Secondary Common Physical CHannel</td>
</tr>
<tr>
<td>SCED</td>
<td>Service Curve-based Earliest Deadline</td>
</tr>
<tr>
<td>SCH</td>
<td>Synchronisation CHannel</td>
</tr>
<tr>
<td>SDMB</td>
<td>Satellite Digital Multimedia Broadcasting</td>
</tr>
<tr>
<td>SDU</td>
<td>Service Data Unit</td>
</tr>
<tr>
<td>SF</td>
<td>Spreading Factor</td>
</tr>
<tr>
<td>SMS</td>
<td>Short Message Service</td>
</tr>
<tr>
<td>SMTP</td>
<td>Simple Mail Transfer Protocol</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>SPF</td>
<td>Static Priority Function</td>
</tr>
<tr>
<td>S-RAN</td>
<td>Satellite Radio Access Network</td>
</tr>
<tr>
<td>SRM</td>
<td>Static Rate Matching</td>
</tr>
<tr>
<td>STFQ</td>
<td>Start-Time Fair Queuing</td>
</tr>
<tr>
<td>S-UMTS</td>
<td>Satellite UMTS</td>
</tr>
<tr>
<td>TB</td>
<td>Transport Block</td>
</tr>
<tr>
<td>TBFQ</td>
<td>Token Bank Fair Queuing</td>
</tr>
<tr>
<td>TBS</td>
<td>Transport Block Size</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TCTF</td>
<td>Target Channel Type Field</td>
</tr>
<tr>
<td>T-DAB</td>
<td>Terrestrial- Digital Audio Broadcasting</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>TDD</td>
<td>Time Division Duplex</td>
</tr>
<tr>
<td>TDM</td>
<td>Time Division Multiplex</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time Division Multiple Access</td>
</tr>
<tr>
<td>T-DMB</td>
<td>Terrestrial Digital Multimedia Broadcasting</td>
</tr>
<tr>
<td>TELNET</td>
<td>TELetype NETwork</td>
</tr>
<tr>
<td>TF</td>
<td>Transport Format</td>
</tr>
<tr>
<td>TFC</td>
<td>Transport Format Combination</td>
</tr>
<tr>
<td>TFCI</td>
<td>Transport Format Combination Indicator</td>
</tr>
<tr>
<td>TFCS</td>
<td>Transport Format Combination Set</td>
</tr>
<tr>
<td>TIA</td>
<td>Telecommunications Industry Association</td>
</tr>
<tr>
<td>TrCH</td>
<td>Transport CHannel</td>
</tr>
<tr>
<td>TTI</td>
<td>Transmission Time Interval</td>
</tr>
<tr>
<td>T-UMTS</td>
<td>Terrestrial UMTS</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>UM</td>
<td>Un-acknowledged mode</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunication System</td>
</tr>
<tr>
<td>USRAN</td>
<td>UMTS Satellite Radio Access Network</td>
</tr>
<tr>
<td>UTRA</td>
<td>UMTS Terrestrial radio Access</td>
</tr>
<tr>
<td>UTRAN</td>
<td>UMTS Terrestrial Radio Access Network</td>
</tr>
<tr>
<td>VBR</td>
<td>Variable Bit Rate</td>
</tr>
<tr>
<td>VC</td>
<td>Virtual Clock</td>
</tr>
<tr>
<td>VoIP</td>
<td>Voice over IP</td>
</tr>
<tr>
<td>VSAT</td>
<td>Very Small Aperture Terminal</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband CDMA</td>
</tr>
<tr>
<td>WFQ</td>
<td>Weighted Fair Queuing</td>
</tr>
<tr>
<td>WG</td>
<td>Working Group</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
<tr>
<td>WRR</td>
<td>Weighted Round Robin</td>
</tr>
<tr>
<td>WTP</td>
<td>Waiting Time Priority</td>
</tr>
<tr>
<td>WWW</td>
<td>World Wide Web</td>
</tr>
</tbody>
</table>
Chapter 1

1 Introduction

1.1 Background

Satellite communications was first predicted by Arthur C. Clarke in the *Wireless World* magazine in 1945, the first satellite was launched in October 1957 called *Sputnik I* by the USSR [1], which had no communications capacity but demonstrated that a satellite could be placed in earth orbit. The first true communication satellites, *Telstar I and II*, were built by *Bell Telephone Laboratories* and launched in July 1962 and May 1963. They used C-band transponders adapted from terrestrial microwave link equipment and carried solar cells and batteries that allowed the continuous use of the single transponder. The successful demonstrations via *Telstar* satellites of live TV links and multiplexed telephone circuits across the Atlantic Ocean proved the feasibility of satellite communications [2]. Nowadays the use of satellites in communication systems plays an increasingly important role in everyday life, forming an essential part of global telecommunication systems for the delivery of voice, data, and multimedia content worldwide.

Since mobile users are more likely to be widely separated geographically over very large areas of the earth, the satellite can form the star point of a communication net simultaneously linking together these terminals. One of the features of satellite communication is that it provides communication links to remote communities in sparsely populated areas which are practically inaccessible by other means of communications. From an economic point of view, a feature of any satellite system is that the cost is *distance insensitive*, meaning that it costs almost the same to provide a satellite communication link over a short distance as it does over a large distance, thus the use of a satellite in communications is especially economical when the cost can be reasonably spread over a large number of users in wide geographical coverage [3]. An overview of the role of satellite in communications is given in [4]-[5].

As the broadband wireless revolution has unfolded, digital multimedia broadcasting (DMB) is rapidly becoming the *sine qua non* of future high-speed high-quality data provisioning. With the phenomenal growth of digital multimedia technology and the growing demand of supporting diverse quality of service (QoS) guarantees, new demands are placed on future wireless networks to utilise the available resources in a more efficient way. The answer to this demand is how to
efficiently delivery multimedia services to an extensive audience range, providing both spectrum efficiency as well as QoS guarantees.

A variety of initiatives, such as the Multimedia Broadcast/Multicast Services (MBMS), Digital Video Broadcasting-Handheld (DVB-H), and terrestrial/satellite-DMB (T/S-DMB), have been envisaged to provide one-to-many content distribution to mobile users. The 3rd Generation Partnership Project (3GPP) MBMS framework [6] defines a unidirectional point-to-multipoint (p-t-mp) mode for the provisioning of one-to-many multimedia services and thereby optimizing the available capacity in cellular networks. DVB-H [7] as initiated by the DVB forum implements additional features based on the DVB-Terrestrial standard to address the specific constraints associated with mobile handheld terminals in terms of power consumption, reception, mobility and etc. Another alternative is T-DMB, as a modified version of the Eureka 147 (T-DAB) standard, uses Orthogonal Frequency Division Multiplexing (OFDM) technology. At the same time, Media Forward Link Only (FLO) [8] air interface specification developed by Qualcomm was recently approved by the Telecommunications Industry Association (TIA) as a new air interface standard for multicast delivery, aimed at offering cost-efficient and high-quality multimedia services to markets in the U.S..

As a complementary alternative to 3G terrestrial mobile networks, the satellite digital multimedia broadcasting (SDMB) system is attracting a great deal of attention within the satellite community [9], [10] as a cost-effective approach for delivering MBMS services over satellite networks. Being defined by the European Telecommunications Standards Institute (ETSI), the system provides datacast capacity for various mobile operators. Based on its broadcast nature, the SDMB system offers extensive coverage, low transmission cost for large numbers of terminals as well as high QoS guarantees for real-time multimedia applications. By employing the wideband code-division multiple access (WCDMA) with frequency division duplexing (FDD), the system can be closely integrated with existing mobile cellular networks, and minimize potential cost impacts on both 3G cellular terminals and network operators.

Due to the unidirectional nature of the SDMB system and the p-t-mp services it provides, given the absence of the power control mechanism and lack of channel-state information, the design of efficient radio resource management (RRM) strategies implemented at the SDMB access layer, is challenging. Within the scope of WCDMA, the functionalities of RRM can be classified into three main sub-streams: radio resource allocation (RRA), packet scheduling (PS), and admission control (AC), which inter-operate with each other during the resource allocation procedures. In this thesis, we focus on the optimization techniques of the former two functions, namely RRA and PS, which, given the unique system characteristics, have substantial impact on the overall achieved performance and QoS for the SDMB system.
The design of an RRA scheme is one of the most important issues in RRM strategy in that RRA is responsible for estimating the required number of logical/transport/physical channels, and mapping them together with the actual transport format combination set (TFCS) for each physical channel [11].

Another key element of RRM strategy is the design of an efficient packet scheduling algorithm, which is the major subject considered in this research. An efficient packet scheduling algorithm in SDMB demands particular consideration in that it is not only required to satisfy the QoS requirements of competing service flows, but also has to optimize the transmission power setting of each physical channel on the basis of the required reception QoS level and under the constraint that the total available transmission power for all the physical channels within a satellite beam is fixed.

1.2 Motivation and scope

As already mentioned, in SDMB, the RRM functionalities comprise three main parts: PS, RRA, and AC. These functionalities cooperate interactively during the resource allocation procedures.

From the viewpoint of RRM, given the lack of real-time interaction between user and satellite radio access network (S-RAN), the downlink power control becomes irrelevant in SDMB. Therefore, the S-RAN has to perform the resource management without any knowledge of the user-side measurements regarding the downlink channel quality, which is normally expressed in terms of signal-to-noise ratio (SNR) in terrestrial mobile networks. In the unidirectional unicast terrestrial Universal Mobile Telecommunications System (T-UMTS) network, this channel status can be effectively exploited by its short-term resource management function (i.e., packet scheduling) to optimize the resource utilisation efficiency. In conclusion, the unidirectional system nature and the p-t-mp service it provides render the channel state information (CSI) irrelevant to the S-RAN.

In SDMB, the mobile users are typically distributed over a wide geographical area, receiving both line-of-sight (LOS)\(^1\) and non-line-of-sight (NLOS)\(^2\) signals from either -or both- satellite or terrestrial repeater, i.e. Intermediate Module Repeater (IMR). It is not easy to identify the "worst case channel conditions" for particular user, since the performance metrics associated with each

\(^1\) Line-of-sight propagation link refers to electro-magnetic radiation travelling in a straight line, without any obstructions blocked in its propagation path.

\(^2\) None-line-of-sight propagation link describes a radio transmission across a path that is obstructed by physical objects, e.g., buildings, hills, tree, and etc.
user are largely different, for example, the SNR, the delay spread, the Doppler spread, the direct-to-diffuse power ratio [12] and channel propagation statistics.

Previous research on the channel mapping has used a conventional single-stage bin-packing algorithm [13], which assumes that the MBMS point-to-multipoint Traffic CHannel (MTCH) logical channels are mapped one-to-one onto the Forward Access CHannel (FACH) transport channels, which are subsequently multiplexed onto the Secondary Common Control Physical CHannel (S-CCPCH). The problem with this simple one-to-one mapping at the transport channel is that there exists residual capacity on the transport channels which is not utilised when the bit rate of the logical channel does not exactly match the corresponding bit rate of the transport channel, i.e. the MTCH rate is less than the FACH rate.

The unique features of the SDMB system require comprehensive consideration of the design of a spectrum-effective and resource-friendly packet scheduling scheme. Firstly, as a satellite communication system, the SDMB system suffers from the long transmission delay, which make the instantaneous feedback information via a return link and the fast power control mechanism unfeasible. Besides, its unidirectional nature accompanied with the built-in broadcast/multicast (BC/MC) mode further aggravates the design difficulty on a feasible packet scheduling scheme. Finally, modern real time multimedia demands more stringent QoS guarantees. To this end, the design of an effective and efficient packet scheduling scheme in the SDMB system requires both consideration of the system constraints and the satisfaction of session QoS demands.

With respect to satellite networks, previous studies [13] have systematically addressed the RRM problems in SDMB via traditional packet scheduling schemes, namely multi-level priority queuing (MLPQ) and weighted fair queuing (WFQ). However, both of these feature major weaknesses in the provisioning of QoS differentiated multimedia services with respect to both efficiency and fairness. Firstly, MLPQ always processes packets from those non-empty queues with the highest priority; as a result, packets waiting in the lower priority queues may suffer from considerably long queuing delay. Furthermore, it is generally agreed that background applications do not have stringent delay constraints, and the only requirement for applications in this category is that information should be delivered to the user essentially error free. In fact, background applications still need a delay constraint (at least an upper bound), since data can be effectively useless if it is received too late for any practical purpose. Finally, MLPQ deals with queues having the same priority in a round-robin fashion. As a consequence, no priority differentiation is made between traffic flows within the same QoS rank. However, this is not an efficient tactic, since there are other essential QoS metrics (e.g. delay and guaranteed data rate) which must also be taken account of in the scheduling decisions.
Based on the above problems encountered in the SDMB RRM framework, the aim of this research is to propose efficient RRM scheme in order to optimize the system performance and maximize the transmission capacity.

By taking stock of the previous research in RRM schemes, to further optimize the resource utilisation, the objectives of this research are listed as follows:

- Design a new RRA scheme to resolve the inefficiency in resource utilisation incurred in traditional single-stage channel multiplexing.

- Optimization of an RRA scheme to improve the system performance in terms of the channel utilisation, transmission capacity and power consumption.

- Design of new packet scheduling algorithms that can optimally utilising resource and efficiently scheduling traffic whilst guaranteeing the prescribed QoS demands for individual users.

- Exploit cross-layer concepts with respect to packet scheduling, by means of utilising useful information from other layers (e.g., application layer, transport layer, and physical layer); to further improve the packet scheduling performance and guarantee a high level of QoS requirements.

- Adopt the most appropriate packet scheduling policy in a dynamic and adaptive way so as to track the instantaneous queuing performance induced by heterogeneous multimedia traffic and respective QoS preferences, and thereby achieve a more intelligent scheduling task.

1.3 Major contributions

The contributions made in this thesis are summarised as follows:

1. Advancement on RRA scheme via novel 2-level channel multiplexing scheme:

- Proposal of a two-level channel multiplexing scheme as a promising efficient option for the radio bearer configuration over the S-RAN via a novel two-stage bin-packing solution derived from the classical bin-packing theory.

- Analytical modeling of the radio bearer configuration procedures, comparing the proposed 2-level channel multiplexing scheme with the existing single-level channel multiplexing scheme and performance evaluation through the assessment for extensive scenarios of satellite radio bearer configuration.
Chapter 1. Introduction

• Development of a 5-step radio bearer configuration procedure associated with the 2-level channel multiplexing scheme, based on the existing 3-step approach in single-level channel multiplexing.
• Development and performance evaluation of an Optimum Estimation Algorithm (OEA) that is capable of deriving the most resource-friendly channel multiplexing option in terms of required number of transport channels given the constraints of service rate requirements and system capacity.
• Development and performance evaluation of a Power-Oriented Adaptation (POA) algorithm that considers the power requirements from both the application and system side so as to increase the overall power utilisation efficiency and minimize the radio resource consumption.
• Development of a new approach for the TFCS derivation in accordance with 2-level channel multiplexing scheme and thereby reduce the TFCS size.

2. Optimization of packet scheduling algorithms via the adaptations of proportional differentiation model:

• Buffer Length Related Queue (BLRQ) is proposed as an enhanced version of MLPQ, aimed at balancing all the traffic flows with regards to their respective queue length, which effectively prevent excess packet loss and unfavorable buffer behaviours.
• Delay Differentiation Queue (DDQ) is developed based on the Hybrid Proportional Delay (HPD) scheduling scheme, it performs service prioritization dynamically depending on the QoS rank and the waiting time/queuing delay experienced by packets in each FACH queue.
• Combined Delay and Rate Differentiation (CDRD) takes into account multiple key performance parameters (i.e. required data rate, and maximum acceptable delay) for better satisfaction of service QoS requirements. It balances all service flows in order to achieve a high-grade QoS satisfaction and improves overall system performance.

3. Development of comprehensive packet scheduling schemes with cross-layer optimization and adaptive mechanism:

• Cross-layer Joint Priority Queue (CJPQ) employs a joint priority function to utilise the cross-layer correspondence through the layered protocol stacks, the proposed scheme exploits the QoS requirements at both application layer and transport layer and seeks to provide better system performance by dynamically adapting to the queuing behaviours induced by heterogeneous multimedia traffics at the data link layer.
Chapter 1. Introduction

- To further optimize the packet scheduling performance by utilising physical layer rate criterion, based on the Dynamic Rate Matching (DRM) technique proposed in [14], we have considered the instantaneous data rate information at the physical layer derived via DRM technique as an important criterion of the cross-layer packet scheduling framework.

- **Adaptive Multi-dimensional QoS-based (AMQ)** packet scheduling scheme consists of an Adaptive Service Prioritization (ASP) algorithm and an Adaptive Resource Allocation (ARA) algorithm. The former takes into account multiple essential QoS attributes in order to prioritize different sessions and adaptively shares the available resources among heterogeneous sessions based on their respective QoS-satisfaction degree. The latter allocates the resources adaptively according to the current QoS satisfaction degree of each session, and prevents excess degraded QoS on a single session, and thereby maintaining a reasonable level of QoS for all competing sessions.

The work conducted in this research has been input to EU IST FP6 European Satellite Communications Network of Excellence (SatNEx) and Mobile Applications & sErvices Satellite & Terrestrial inteRwOrking (MAESTRO) projects. More details relevant to this work can be found at Publications in the end of the thesis.

1.4 Thesis organisation

This thesis is structured as illustrated in Figure 1-1.

![Figure 1-1. Structure of the thesis.](image-url)
Chapter 1 provides the background information, motivations and major contributions of this thesis. Chapter 2 presents an overview of SDMB, emphasising the system description, QoS demand and cross-layer aspects. In Chapter 3, a brief introduction of the basic idea of RRM functionalities implemented in the SDMB system is given, by introducing existing literature on RRM, we outline the main challenges and new opportunities for the design of an efficient RRM scheme. We then proceed to present our development and evaluation carried out on 2-level channel mapping schemes in Chapter 4. In Chapter 5, several novel packet scheduling algorithms considering different aspects of successful QoS provisioning (e.g., buffer status, queuing behaviours) are proposed based on a proportional differentiation model. Their performance is evaluated and discussed via comprehensive simulation results, compared with existing schemes. Then we examine the cross-layer design of a more efficient packet scheduling scheme considering cross-layer correspondence across multiple protocol layers in Chapter 6. By applying adaptive sharing mechanisms to both service prioritization and resource allocation procedures, an AMQ packet scheduling scheme is investigated in Chapter 7, offering great improvements on the system efficiency and flexibility. We conclude our work in Chapter 8 and explore future work pertinent to this research.
Chapter 2

2 Overview of Satellite Digital Multimedia Broadcasting

2.1 Digital multimedia broadcasting

Recent advances in the digital multimedia broadcasting (DMB) industry have offered the network operator a platform to deliver multimedia services to a mass-market in a spectrum-efficient and cost-effective way. It is widely expected that voice service revenues will decrease over the next decade and mobile operators therefore must increase data revenue in a competitive landscape via multimedia applications [15]. One of the most popular ways to deliver multimedia applications to mobile terminals is via broadcasting or multicasting mode. Broadcast and multicast services for other digital wireless communication networks, such as digital TV and DVB-T, are also considered in [16].

The provisioning of p-t-mp multimedia services over the 3rd generation (3G) terrestrial mobile cellular networks in broadcast and multicast mode has been regarded as a key to the efficient use of the precious wireless resources, and is currently under investigation and standardization within the 3GPP Multimedia Broadcast Multicast Services (MBMS) framework [17]. MBMS data is mapped onto radio network bearers and is transmitted over air in parallel to traditional unicast data. However, serious concerns are expressed as to whether T-UMTS can cope with the additional requirements of MBMS delivery on top of the other point-to-point (p-t-p) T-UMTS services, due to the spectrum limitations and very limited means to improve the spectrum efficiency. At the same time, satellites are advertised as a promising platform for MBMS delivery due to their unique wide area coverage capabilities. Detailed description for MBMS data transfer over satellite UMTS (S-UMTS) air interface can be found in [18]. In Europe, much research effort [9] has been devoted to the SDMB system, where the satellite is used in the delivery of p-t-mp multimedia services to 3G handsets.

With the increasing employment of high bandwidth applications in 3G mobile systems, especially with a large number of users receiving the same high data rate services, efficient and fair
information distribution is essential. Thus, broadcast and multicast are techniques to decrease the amount of transmitted data within the network and to use resources more efficiently.

Broadcast/multicast is a method for transmitting datagrams from a single source to several destinations. Due to the broadcasting nature of the radio channel, this method is efficient for sessions sharing the same (or even common) content. If the nature of the offered service lends itself to spatial and temporal bundling of the demands into one transmission, the benefit of multicast and broadcast is therefore that the data is sent just once by the network and transmitted to all the designated users, located in the same cell, over a single common channel without clogging up the air interface with multiple transmissions of the same data as caused by multiple usage of unicast sessions.

Broadband satellite networks offer high-speed information access and multimedia services, such as data multicasting and interactive video. Due to its broadcast nature and ubiquitous coverage, satellite systems may become a very efficient complement to terrestrial mobile networks, providing them with far more p-t-p equivalent capacity for far less investment cost on both the network operators and mobile devices. Examples of future broadband satellite communication systems are Astrolink, Cyberstar, Spaceway, SkyBridge, Teledesic, iSky [19]-[24]. Each of these systems is designed to exploit the emerging market for provisioning high-quality multimedia information services to mobile users with relatively small terminals over extensive coverage. A survey of future broadband multimedia satellite system is provided in [25].

Even though the broadcast and multicast delivery mode is able to give many benefits for certain application areas such as inherently 'non-interactive' applications, e.g. video/audio streaming and file downloading applications in an area with high user density (stadiums, trade shows, etc.), there are still many challenging issues to be solved in order to provide an efficient use of the spectrum resources and a required QoS level for the broadcast and multicast services. The main challenging tasks lie in the resource management functionality, which is desired to consider QoS constraints for all members in a particular BC/MC group.

### 2.2 Multimedia Broadcast Multicast Service

In the wake of increasing use and rapid penetration of high-speed and high-quality multimedia applications in the past few years, to efficiently distribute the multimedia services over a large number of users, BC/MC transmission is advertised as an attractive approach for achieving less unnecessary transmission and better resource utilisation. Point-to-multipoint services allow data from a single source entity to be transmitted to multiple recipients in a unidirectional BC/MC transmission mode. MBMS is defined as a BC/MC service where data is transmitted over a
unidirectional p-t-mp bearer to multiple recipients within the intended BC/MC user group. It was initially produced by 3GPP and aimed at facilitating higher data rate multimedia services via a p-t-mp transmission mode. The MBMS service is split into two categories: MBMS bearer service and MBMS user service. The former delivers traffic flows via IP multicast addresses in either multicast- or broadcast- mode, offering the possibility of sharing the transmission resources between the core and radio networks. Whilst the latter operates at the MBMS service layer and defines “streaming” and “download” delivery methods, offering the continuous streaming transmission and “download and play, or push-and-store” transmission approaches. MBMS is currently under development by various work groups (WGs) in 3GPP, and is expected to be available from 2007. The MBMS technology provides a brand new opportunity for 3G mobile networks as a promising approach to support mobile TV and video broadcasting services. Besides, it requires relatively small changes to the underlying 3G standards. Rather than maintaining a dedicated p-t-p connection to every handset, only a single broadcast channel is needed in each cell. Another major benefit of MBMS is that it is capable of achieving the same level of coverage without additional spectrum, licensing or base-station, and can provide mobile operators with a stronger control of the mobile TV and video mass market than they would have if they uses other technologies, such as DVB-H, DMB.

2.3 Satellite Digital Multimedia Broadcasting

2.3.1 Convergence of satellite and terrestrial networks

The integration and convergence of the satellite component to the terrestrial network have become a promising approach for the delivery of MBMS. The level of integration between the terrestrial and satellite system is still under discussion, ranging from the high-level integration where the satellite component is considered as an independent network, to a lower level one where the satellite interface is embedded in the terrestrial network making the maximum possible reuse of available resources and infrastructures.

Regarding the integration of the satellite component to the terrestrial network, the S-UMTS poses an essential complement to the T-UMTS network. Satellite communication infrastructure provides high bandwidth, wide area coverage, reconfigurability, broadcast/multicast capability. When these unique features are combined with the terrestrial network infrastructure, they will provide immense new opportunities for both network operators and mass markets. The S-UMTS system was investigated as an integral part of the T-UMTS networks under the EU IST FP5 projects VIRTUOUS [26] and FUTURE [27], focusing on the non-access stratum functions and S-RAN specific issues. In EU IST MOBILITY project [28], the provisioning of live DVB-S service was
further envisaged via Ku-band (10.7-12.75 GHz) DVB-S broadcasting satellite to maritime scenarios, where the satellite will be the adequate solution.

The geostationary (GEO) satellites orbit is achieved around 35786 km directly above the Earth Equator, which yields unavailable long propagation delays. Typical round trip time (RRT) between earth station and GEO satellite is around 500 ms, which is unacceptable for many communication applications without specific techniques, e.g. echo cancellation\(^3\). Also, satellite channels experiences a BER at the level of 10e-4, which effectively affects the SNR received at the earth stations. In comparison to extreme low BER infrastructure, e.g., 10e-9 for fiber-optic cables, the high BER poses another constraint for the satellite communications. However, the use of forward error correction (FEC) can be employed to compensate the BER performance in satellite link and achieve comparable BER performance to optical fiber.

Being positioned from 700 to 2000 km above the earth surface, LEO-satellite constellations can offer some advantages for global communications. The RRT experienced in a LEO at a 1000 km attitude is around 7 ms. However, the main limitation of LEO is that the satellites are fast moving relative to the earth surface and can involve frequent hand-over control and traffic reservations.

Highly elliptical orbit (HEO) can be used in which the inclination angle to the Equator can be chosen to achieve synchronism with the rotation of the earth. This means that the satellite will appear at the same point each day and there are multiple satellites (3-6) can be used to achieve continuous coverage. The advantage of HEOs is that they provide higher elliptical angles across the coverage area and hence improved coverage and channel conditions for urban area. They offer the advantage of less terrestrial infrastructures, e.g., the number of IMRs. However, in the SDMB system, we have assumed the one of the move conventional GEO satellite.

By taking stock of the lessons learnt from existing mobile satellite communications, the following experiences and trends can be summarised as [30]:

- Low earth orbit (LEO) and HEO constellations have proved too expensive to compete with GEOs or cellular systems, so there is now a return to GEOs.
- Satellites can only economically provide niche services to areas inaccessible to cellular; hence, for mass market services there needs to be integration, not to compete but to collaborate with cellular.

\(^3\) In telephone network, echo cancellation is used to remove echo from a voice communication link in order to improve the voice quality, details of echo cancellation techniques can be found at [29]. Sophisticated echo cancellation technologies have permitted GEOs to be successfully for both voice and high-speed data applications.
Based on the above considerations, the SDMB system was proposed within EU as one of the most attractive options for this integration, providing wide coverage, low cost and high QoS guarantees. In Europe, the SDMB concept was previously investigated under the name “S-UMTS”, which was the subject of two EU IST FP5 projects, namely SATIN (SATellite UMTS IP based Network) and MoDiS (MOobile DIGital broadcasting Satellite). The main objective of SATIN was to elaborate the concept, definition of a radio access scheme that features maximum commonalties with WCDMA air interface and to evaluate main aspects involved in this interface via simulation means [31]. Project MoDiS was launched to demonstrate the whole range of system concepts via a test-bed satellite emulator integrated with a 3G terrestrial network [32]. It is noted that the “SDMB” is the generic name given to the system researched in EU projects by Alcatel who have proposed the operational system.

In the meantime, various aspects of the SDMB system was studied separately through a series of ESA ARTES programs, from the mobile terminal to radio propagation [33]. The successful validation and demonstration of the SDMB concept carried out within the EU IST project MoDiS [34] has pushed the system towards an operational stage. The whole range of issues pertinent to the SDMB system, from system definition to standardization, is addressed in the EU IST project MAESTRO 4 [35]. The research presented in this thesis is mainly carried out within the MAESTRO and SatNEx projects between 2005 and 2008. The primary goal of the research of excellence SatNEx is to achieve long-lasting integration of the European research in satellite communications [36].

2.3.2 The SDMB system architecture

As illustrated in Figure 2-1, the SDMB system defines a hybrid satellite-terrestrial communication system, featuring a unidirectional geostationary satellite component that is responsible for the delivery of the p-t-mp MBMS services and provides a European coverage by multiple umbrella cells. The system implements a satellite based broadcast layer over 2.5G and 3G terrestrial mobile cellular networks. It takes advantage of inherent satellite capacity in providing BC/MC services

4 The integrated project MAESTRO (Mobile Application and sErvices based on Satellite and Terrestrial inteRworking) is partly funded by European Commission under the 6th research framework program, http://ist-maestro.dyndns.org/maestro/
Chapter 2. Overview of satellite digital multimedia broadcasting

The SDMB system intends to implement a BC/MC layer over unicast T-UMTS mobile cellular networks for the efficient delivery of interactive BC/MC digital multimedia services to mobile terminals. The scheme features maximum commonalities with the UMTS Terrestrial Radio Access Frequency-Division Duplex (UTRA-FDD) air interface standardized within 3GPP initiative, which is better known as WCDMA [37]. The user equipment (UE) applies the standard 3G terminal enriched with SDMB-enabling functions, which, given the unidirectional nature, are very limited. The SDMB-enabled broadcast/multicast service center (BMSC), is enhanced with SDMB-specific functions from the standard 3GPP MBMS BMSC [6]. The main benefits from this satellite based BC/MC scheme lie in the UE side. The commonality of the transmission waveform between SDMB air interface and T-UMTS air interface enables the highest possible degree of reusing the standard 3G terminal hardware with significant advantages in terms of terminal size, power consumption as well as market penetration.

Figure 2-1 illustrates the overall SDMB system architecture, in which the high-power GEO satellite and limited number of low-power terrestrial gap fillers (i.e. Intermediate Module Repeater (IMR)) can provide outdoor and in-building coverage, with national wide umbrella cells maximising the potential audience. In this way, a handheld mobile terminal (i.e. UE) can receive the MBMS data through the satellite and/or the IMR, who performs one-way repeater functionality. The IMR module plays a mandatory role in supporting in-building/in-urban...
coverage and moderate transmission rate of BC/MC services. Direct satellite path is the preferred communication link; however, the communication would be sustained via IMRs when the direct satellite path is unavailable or blocked. It is noteworthy that no direct satellite return link is envisaged under the baseline SDMB infrastructure, where the return path may be provided via the T-UMTS network if applicable. WCDMA is adopted as the air interface between satellite and UE and between IMR and UE. Unlike the conventional unicast scenario, the BMSC is introduced to support the BC/MC data transmission.

It is assumed that in SDMB, MBMS services are intended for transmission to UEs in either a broadcast or a multicast mode. In the latter case, service is only delivered to the UEs within a specific multicast group. Packets from the BMSC are firstly buffered at the satellite hub (or Node B) in a first-in-first-out (FIFO) manner before being scheduled for transmission over satellite link. Being closely integrated into the baseline architecture of 2.5G/3G mobile cellular networks, the system enjoys maximum reuse of technology and infrastructure and minimum system development cost. The hybrid system takes advantage of the satellite inherent broadcast capability to provide efficient delivery of MBMS content to extensively mass mobile market.

For MBMS delivery in SDMB, the information transfer from the source (i.e. the content provider) through the geostationary satellite to a multitude of mobile receivers (i.e. UEs) is a challenging issue. From the viewpoint of information theory, this type of information transfer can be regarded as a particular case of broadcast channel [38], where the propagation channels are largely diversified within the mobile users. The following types of users are investigated in SDMB [9]:

- Static users in LOS to the satellite;
- Mobile users with various speeds in LOS;
- User in NLOS receiving signals from the terrestrial repeaters, moving at various speeds;
- User in LOS to the satellite and to a group of IMRs at the same time;
- Users in indoor propagation conditions.
2.3.3 SDMB access scheme and air interface

Figure 2-2 shows the architectural position of the SDMB components integrated within the packet-switched domain of UMTS with respect to the 3GPP MBMS reference architecture. The UMTS satellite radio access network (USRAN), interfaces with the UMTS core network through lu interface. The satellite gateway (Sat-GW) is composed of two physically separated functional nodes: RNC and Node B, which communicates with each other through lub interface. The discrepancies between SDMB and UMTS are related to the interfaces Sf5, Sui6 and Sii7 [18]. WCDMA is adopted as the air interface Sui between Satellite and UE and between IMR and UE. The IMR can be operated in either a “low power” or “high power” mode to avoid the electromagnetic coupling effects between its transmitting and receiving parts depending on whether the operational frequency of the interface between UE-IMR matches the corresponding frequency between IMR and satellite. Unlike the conventional unicast scenario, the BMSC is introduced as an additional functional entity in MBMS, supporting BC/MC data transmission operated in Packet-Switched domain.

5 Sf: The SDMB feeder link between satellite and gateway, which can be at Ku band.
6 Sui: The SDMB air interface Sui between IMR-UE and SAT-UE is WCDMA.
7 Sii: The SDMB air interface between SAT-IMR can be either the same (Sui) or different (Sii) with the IMR-UE air interface, depending on the frequency of operation.
2.4 QoS provisioning

Supporting diverse quality of service in modern telecommunication systems is gaining importance in recent years; the focus of the service provisioning is on the final delivered service quality level. QoS has become one of the key differentiators used not only to improve the potential market share of mobile multimedia industry, but also to gain the customers' retention and loyalty. UMTS has, for the first time, given serious emphasis on the so-called "end-to-end QoS", where a common end-to-end framework has been established and formalized within its standard body 3GPP. To improve revenue and profit and offer business advantages over traditional wireless information service providers (ISPs), it is vital for the future wireless operators or service providers to provide consistent and diverse QoS across the network. However, QoS degradation is inevitable due to the growing growth in demand for data and multimedia applications. An ISP has to manage the resource allocation under the performance trade-offs amongst connections with diverse traffic priorities. Sometimes, the lower priority connections may even have to be forcibly disconnected so as to carry the higher priority connections [39].

One of the key elements in providing the overall delivered QoS level is the degree of efficiency and effectiveness of managing and utilising available radio resources. The objective of resource management is to effectively utilise the available bandwidth in networks to carry users' connections, and to avoid QoS degradation. The authors in [40] argue that the QoS degradation is a major mechanism to decrease the Call Blocking Probabilities (CBP) and Call Dropping Probabilities (CDP), whilst the QoS upgrades could be used to increase the bandwidth usage if bandwidth becomes re-available, but it will increase the CBP and CDP. Research aiming at achieving better QoS provisioning from RRM strategies over WCDMA network is discussed in [41].

UMTS supports four basic service classes [42]: conversational, streaming, interactive and background. It is worth noticing that there is no one-to-one correspondence between service class and traffic class, e.g. streaming service can use a bearer of the conversational class if the traffic class demands more stringent QoS [43]. However, for the aforementioned basic service classes, there are fundamental characteristics for QoS requirements as outlined in the followings:

Conversational class includes voice, interactive games, two-way control telemetry, telnet and etc, their QoS requirements are:

- Preserve time relation (variation) between information entities of the stream;
- Stringent and low delay conversational pattern.
Supporting the conversational class in a GEO satellite system has proven to be difficult as the long delay requires use of echo cancellation technique. However, it is more likely to support conversational class in LEO constellations.

Interactive service class is applied when the online requesting data from a remote server occurs. The information is to be transferred with low bit error rate (BER) performance. To support interactive non-real-time application, the following QoS are required [44]:

- Support request response pattern;
- Preserve the payload content.

Streaming service class is mainly unidirectional with short-time variations between information entities of the flow (i.e. packet). Packets in this category are expected to be transferred to the receiver in a steady and continuous stream. Fundamental characteristics for streaming are:

- Unidirectional continuous stream;
- Preserve the time variation between information entities of the stream.

Background service is characterised by a transparent transfer without explicit limit on delay or delay variation. Thereby, the following characteristics for this category of service class can be identified:

- There is no explicit delay requirement for the data transfer;
- The information should be delivered with a very low BER.

In fact, the only requirement for background service is that the information should be delivered to the user essentially error free. Example of background services are software packages, video clips. Due to its one-way delay-tolerant feature, background service is easy to support on satellite networks.

In the SDMB system, the service types considered can be divided into two main classes: “streaming” and “download”, which corresponds to UMTS streaming and background QoS class. The latter can be further sub-categorised into two sub-classes according to its sensitivity to delay, namely “hot download” and “cold download”.

1) Streaming defines a data transfer technology that allows large multimedia files to be played at the receiver in a steady and continuous fashion, display is allowed even before the completion of transmission, data can be stored temporarily at the receiver buffer without influencing the current display. Streaming is becoming increasingly important for real time multimedia applications, such as TV or video broadcasting. Service in this category requires explicit upper bound on queuing delay and delay variation.
2) *Hot download* supports a content delivery form that the data can be stored at the receiver for their offline access. Compared with streaming, the hot download service has more tolerant demand on delay and jitter but more stringent demands on packet loss. Examples of service in this category are emergency broadcasting, stock exchange updates.

3) *Cold download* defines a transmission mode that requires the least demand on delay but the most stringent demand on packet loss, services in this category are often transmitted as a time-independent individual file, such as software package, video clips, images or text messages.

### 2.5 Cross-layer optimization

In recent years, cross-layer design has been addressed as a popular subject in research and development activity in the area of communication networks. Although layered architectures have performed well for wired networks, when they are applied to wireless networks, the performance is far from optimum. The reason for the above statement is that the wireless medium allows richer modalities of communication than wired networks. Performance gains are proven to be achievable by actively exploiting the inter-relationship between protocol layers. There have been numerous different interpretations of cross-layer design in the literature, a more unified definition is given in [45], describing cross-layer design as a violation of layered architecture: “Protocol design by the violation of a reference layered communication architecture is cross-layer design with respect to the particular layered architecture.”

The so-called violation can be considered, for example as creating new entities interfacing different layers, redefining the boundaries of protocol layers, or joint cooperation amongst different layers including protocol designs and parameter tunings.

There have been various cross-layer proposals in the literature, which has been made independently by researchers from different backgrounds aiming at improving specific targets. In general, cross-layer design evolves around five protocol layers: application layer, transport layer, network layer, medium access control (MAC) layer and physical layer. Preliminary literature survey within this context is given in [46]-[50], which differ from each other according to the layers that are jointly optimized.

However, most of these existing proposals can be classified into a simple taxonomy which defines most basic ways that the layered architecture can be violated:

- Creation of new interfaces
- Merging of adjacent layers
- Design coupling without new interfaces
Vertical calibration across layers

The above classification was proposed in [45], where it is also argued that creation of new interfaces can be categorised into three subcategories depending on the direction of information flow along the newly defined interfaces:

- Upward: from lower to a higher layer
- Downward: from higher layer to a lower layer
- Back and forth: iterative flow between two layers

As far as our research is concerned, the aim is to achieve more efficient resource management, which is located in the data link layer. A great deal of cross-layer optimization on resource management can be found in [46]-[54].

Given the particular considerations of the system architecture and responsibilities of respective functional entities, in this research, we employ the most commonly used and feasible approach for our cross-layer design, i.e. design the protocol at the data link layer based on the parameters from other layers. We consider the cross-layer correspondence to data link layer via both “upward” and “downward” approaches. Cross-layer optimization on packet scheduling scheme is presented in Chapter 6.
Chapter 3

3 Radio Resource Management for SDMB

3.1 Background

Radio resource management functions allocate and manage the radio resource in wireless communication systems. Cellular mobile communications are dynamic in nature; the dynamism arises from multiple dimensions, namely propagation conditions, traffic generation conditions, interference conditions, etc. Therefore, the dynamic network evolution calls for a radio resource management in a dynamic way, which is carried out by the RRM mechanism with series of associated parameters that need to be chosen, measured, analysed and optimized. Since the RRM strategies are not subject to standardization, to improve overall system performance and reduce operator infrastructure cost, RRM functions can be implemented via many different algorithms. Research and development in this important field is especially demanding and challenging. The main problems of designing appropriate RRM functions faced in satellite systems are explained in [44].

SDMB allows a user or an application to negotiate the characteristics of the network at the connection set up phase. The network may check whether sufficient resources are available, and returns the results to the application, which can accept or deny the connection request according to an AC scheme. After admission of the connection request, the network should keep the performance of the connection as contracted. The above rules also apply for broadcast and multicast scenarios. By admitting the connection request the access network has to make a choice for the type of the radio access bearer taking into account several conditions such as attributes of the requested service, number of group members in the cell, current load conditions and etc. In contrast to unicast, i.e., p-t-p service delivery, the p-t-mp network has to select the type of the transport channel, namely if a common channel should be used or a dedicated channel is used instead. For instance, if there is only one multicast member in the cell, it is not worth using a common channel since a common channel needs additionally a return link dedicated channel to maintain the quality of the connection, i.e., measurement control/report, power control and the error correction due to its unidirectional nature. In other words, usage of a common channel is not always more effective than that of dedicated channels. Therefore, well defined criteria for...
selecting the transport channel type among others is necessary for optimally utilising system capacity, e.g., the minimum number of members in the multicast group, momentary load condition, current/predictable channel condition, QoS constraints of the session, etc. Moreover, since the number of members in a multicast session can be dynamically changing, there should be another criterion for the appropriate timing when a Radio Access Bearer (RAB) re-assignment is necessary.

In SDMB, the RRM functionalities comprise three main parts: PS, RRA and AC. These functionalities cooperate interactively during the resource allocation procedures.

The PS operates periodically in each transmission time interval (TTI) of the radio bearers. It time-multiplexes flows with diverse QoS into physical channels and adjusts the transmit power for physical channels on the basis of the packets to be served and the required reception quality of the service in terms of the target block error rate (BLER)\(^8\).

The RRA entity is responsible for the radio bearer configuration at the beginning of each session, which includes the estimation of the required number of logical/transport/physical channels along with their mappings for each physical channel through the scheme layers/sub-layers.

The admittance decision of each incoming MBMS session is handled by the admission control function during the phase of service establishment/re-negotiation, aimed at preserving the required QoS while making efficient utilisation of resources.

3.1.1 Radio resource allocation

Due to the unidirectional nature of the SDMB system and the p-t-mp services it provides, channel mapping becomes an important issue in the Radio Resource Allocation (RRA) scheme, which is responsible for the estimation of the required number of logical/transport/physical channels, and mapping them together with the actual Transport Format Combination Set (TFCS) for each physical channel.

As mentioned above, in order to resolve the inefficient resource utilisation in traditional single-stage channel multiplexing, a 2-level channel multiplexing framework has been proposed in [55]. This approach performs channel mapping at both logical and transport channels. At the first level

---

\(^8\) It is noted that the error rate of a communication link at the bit level by means of the BER (Bit Error Rate) or at the TB (transport block) level by means of the BLER (transport block error rate), depending on the service under consideration. For instance, for error-tolerant services operating in Transparent Mode or Unacknowledged Mode RLC mode, it is usual to measure performance by means of the residual BER in the delivered transport blocks, while in the case of services using the Acknowledged Mode RLC mode allowing for retransmissions, the throughput is usually limited by the BLER [47].
of multiplexing, multiple MTCHs are mapped onto a single FACH (logical channel multiplexing), whereas the mapping of several FACHs onto the S-CCPCH (transport channel multiplexing) is regarded as the second level of multiplexing. In order to minimize both the residual capacity on FACHs and the required number of FACHs, a 2-stage bin-packing algorithm with optimum estimation of FACH has been further developed in [56]. The objective of this algorithm is to achieve the highest possible degree of utilisation and efficiency in available system capacity whilst meeting the target QoS requirements.

3.1.2 Packet scheduling

After the assignment of a certain RAB to the multicast session, the network should maintain the contracted performance throughout the session. In practice, it is considered that the network has to maintain, not only this multicast session, but also other multicast sessions as well as other unicast sessions, which have their own requirements in terms of delay, throughput, jitter, priority, etc. Moreover, especially for the satellite network, it is also considered that the group members are distributed at large distances from each other. Hence, the selection of an appropriate Transport Format (TF) has a strong impact on the performance of connections, not only the multicast session itself, but also on the other active sessions due to the generated interference level. According to the WCDMA channel sharing technique, for each TTI, we have to decide how to accommodate a datagram over channels by choosing an optimal TF combination, for the currently active sessions. This TF selection has to be done dynamically according to the dynamically changing load condition, the number of multicast members and the radio propagation condition.

Given that there is no real-time interaction between the user and the S-RAN in the considered baseline architecture, the role of the packet scheduler is therefore different than in the traditional T-UMTS networks [57]. The packet scheduler in the unidirectional satellite system will have to decide on allocations without knowledge of the state of individual channels, i.e., channel state-dependent scheduling is not possible. Given that the packet scheduler in SDMB cannot exploit this cross-layer information on the radio channel conditions perceived by the users from the physical layer, the design of the various scheduling policies in SDMB relies exclusively on the traffic characteristics (maximum transfer delay, guaranteed rate, etc) and considerations of buffer size, etc. In any case, even if such information were available, it would have to be exploited in a complex way due to the p-t-mp nature of the services, i.e., the decisions regarding the scheduling of a single service data flow need to consider the state of multiple links corresponding to all the users that have activated the service in each (multicast) group.

The packet scheduler interacts primarily with RRA module and AC module over the SDMB radio interface. The scheduler treats independently each physical channel on a TTI-scale. The exact
Transport Format Combination (TFC) corresponds to a certain number of bits passed from the scheduler to Layer 1, upper-limited by the maximum allowed data rate of the physical channel [58]. The task of the scheduler is to select every TTI and for each S-CCPCH some appropriate TFC, featuring a certain TBS size for each one of the FACH channels mapped onto it. The actual context of the term “appropriate” is dictated by several factors, such as the service QoS requirements and the physical channel utilisation efficiency, and this differentiates one scheduler from another [59].

3.1.3 Admission control

Admission control is one of the most critical RRM functions in QoS enabled IP-based wireless networks. Future high-speed multimedia mobile networks such as those relying on WCDMA radio technology require advanced and sophisticated functions to decide on the acceptance or rejection of new session requests. In interference limited wireless packet networks the design of efficient AC schemes becomes a rather complex and challenging issue [59]. The admission control module has tight interaction with the packet scheduler. In order to allow the packet scheduler to handle with satisfactory QoS provisioning for ongoing sessions, the admission control module limits the traffic load so as to preserve the required QoS or Grade of Service\(^9\) (GoS) requirements.

In the SDMB system, admission control is responsible for the following tasks;

- Decides the admittance/rejection of new session requests;
- Limits the traffic load carried over the radio interfaces;
- Guarantees the link level performance/quality of all established radio connections;
- Fulfils the appropriate QoS targets;

Unlike packet scheduling, admission control is responsible for the long-term resource management at session-level, determining the admittance of new session requests according to the system power/load constraints and service QoS requisites.

\(^9\) In telephone networks, Grade of Service (or Blocking Probability Percentage) is normally defined in terms of the portion of call attempts rejected because there is insufficient equipment (i.e. congestion), although delay criteria such as delay to receive dial tone may be used [60].
3.1.4 Interaction between RRM entities

![Diagram of RRM entities interaction](image)

Figure 3-1: Interaction between RRM entities in SDMB system

As illustrated in Figure 3-1, the interaction between RRM entities can be described as follows:

When there is a new session request, the AC will check the service QoS constraint and load/power of the system to determine whether to accept or reject the request. Also the AC needs to predict the long-term load of the network and interference level that may arise.

If the request is accepted, the RRA module is triggered to estimate the radio bearer configuration according to the traffic characteristics and specific scenarios (e.g. available rate of FACHs and S-CCPCHs). The radio bearer will be reconfigured whenever there is a new session request admitted by AC, or an existing session completes.

During the RB reconfiguration, TFCS is derived for each S-CCPCH according to the service characteristics. A generic approach for deriving this TFCS can be found in [61]. The mappings of the MTCHs/FACHs to S-CCPCHs, as well as the TFCS available to S-CCPCHs, are passed to PS for the short-term selection of TF(C).

Packet scheduler time-multiplexes service flows with different QoS requirements into physical channels, in such a way as to satisfy these requirements and adjusts the transmit power of the physical channel on the basis of the required reception quality of the service under the constraint that the total available power for all the physical channels within a beam is fixed.

AC predicts the total system load based on the information regarding the current system status (number of admitted flows, requested QoS etc) and on the declared QoS requirements of the incoming traffic. PS calculates the actual system load resulting from the per-TTI scheduling decisions.
3.2 Literature review on radio resource allocation

3.2.1 Channel mapping and multiplexing in SDMB

Due to the unidirectional nature of the SDMB system and the p-t-mp services it provides, channel mapping becomes an important issue within the Radio Resource Allocation (RRA) entity, which is responsible for estimating the required number of logical/transport/physical channels, and mapping them together with the actual TFCS for each physical channel.

![WCDMA channels relevant to SDMB](image)

Figure 3-2. WCDMA channel mapping and relevance to SDMB.

The radio access scheme in SDMB draws heavily on the reference WCDMA air interface in order to achieve maximum commonalities with terrestrial UMTS, and its functions overlap on two layers:\footnote{Note that the layer numbering is in line with the UMTS radio access network architecture and layer numbering. It simply reflects their relative positions in the SDMB radio interface, without strict mapping on the respective layers of the ISO protocol stack [57].}

- The radio link layer (or layer 2) of the radio interface;
- Part of the radio network layer (or layer 3) of the radio interface.

Layer 2 in turn consists of three sub-layers, the MAC, radio link control (RLC) and the Packet Data Convergence Protocol (PDCP) sub-layers. The first two exist in both the user-plane and the control-plane, whereas PDCP is only relevant to the user-plane [11]. The network layer is part of the control plane and is also organised in sub-layers, which include the Radio Resource Control (RRC) sub-layer. Both user-plane and control-plane layers of the SDMB access scheme are
summarised in Figure 3-2. The channel mappings concern both the user and control planes, whereby the RRA has to decide whether and how to multiplex logical channels (MTCHs/MSCHs/MCCHs) onto transport channels (FACHs) and then multiplex transport channels onto physical channels (S-CCPCHs).

The set of channels relevant to SDMB radio interface, as shown in Figure 3-2, is a subset of the WCDMA channels; only the downlink common channels are relevant due to the unidirectional nature of the system and the p-t-mp services it provides. The SDMB logical channel sets include the MTCH, the MBMS point-to-multipoint Control Channel (MCCH), the MBMS point-to-multipoint Scheduling Channel (MSCH) and the Broadcast Common Control Channel (BCCH). The SDMB-relevant WCDMA transport channels are the FACH and the Broadcast Channel (BCH). Due to the limited capacity/bandwidth of satellite channels, the MTCH and FACH have been selected as the logical/transport channel pair for data transmission, with maximum user bit rate of 384kbps. As for the WCDMA physical channels, only the Primary Common Control Physical Channel (P-CCPCH), the S-CCPCH, the Synchronisation Channel (SCH), the Common Pilot Channel (CPICH), and the MBMS Notification Indicator Channel (MICH) are applicable in SDMB, whilst the S-CCPCHs are used for carrying MTCHs and FACHs channels.

3.2.2 Related work and problem analysis

As aforementioned, within the scope of the previous research [13], only single-level channel multiplexing (i.e. transport channel multiplexing) was considered for the support of delivering MBMS services. An example of the transport channel multiplexing adopted in the previous European IST SATIN project is shown in Figure 3-3. As shown in Figure 3-3, multiple FACH transport channels are multiplexed onto a single S-CCPCH physical channel. Data carried by different transport channels are mixed together via interleaving/segmentation and delivered in the same physical channel. Hence, a common transmission power is considered for all transport channels multiplexed on the S-CCPCH when performing the radio resource allocation task.

11 The project SATIN (SAtellite UMTS IP-based network) was funded by the European Commission under the 5th research framework program.
In order to perform the single-level channel multiplexing, a single-stage bin-packing algorithm was proposed in [62], which splits the RRA task into three main steps as follows:

- **Step 1: Estimation of the required number and bit rate of MTCHs/FACHs** - The RRA first derives the expected traffic load for each service based on the service characterisation and target blocking probabilities. In order to guarantee the blocking probabilities targeted for each streaming service, the RRA invokes well-known results [63] of classical multi-server queuing theory to estimate the required number of MTCHs/FACHs.

- **Step 2: Mapping of MTCHs/FACHs onto S-CCPCHs** - This task is to map the derived MTCHs/FACHs onto the available S-CCPCHs. The number of available S-CCPCHs and their maximum capacity are estimated a priori from the link budget and link-layer simulation results.

- **Step 3: Derivation of the TFCS for each S-CCPCH** - During the session initialization process, TFCS is derived for each S-CCPCH in this step. A TFCS is defined as a set of Transport Format Combinations (TFCs) to be used by an UE, which is going to be chosen from a separate TFC for every S-CCPCH when the packet scheduler performs short-term scheduling task.
In SDMB, a Transport Block (TB) is the minimum element that can be accepted by the physical layer to be jointly encoded, a set of TBs that are exchanged between MAC and physical layer at the same TTI using the same transport channel is defined as a TB set (TBS). For each TTI, the MAC will choose an appropriate Transport Format (TF) or TF set (TFS) from each multiplexed transport channel. TFS is a set of TFs allowing a transport channel to support different instantaneous bit rates. In the presence of the transport channel multiplexing, the combination of the selected TFs/TFSs for all the transport channels forms a TFC. The total TFCs that an UE is permitted to transmit in each physical channel during each TTI are included in a list called the TFCS. There is no specific rule on the derivation of the TFCS to specify which TFC is included in the TFCS and which is not.

As stated in the above context, the problem with the single-level channel multiplexing is that there exists residual capacity (i.e. bandwidth) on the transport channel when the bit rate of the logical channel does not exactly match the corresponding bit rate of the transport channel, i.e. the MTCH bit rate is less than the FACH bit rate. Figure 3-4 shows an illustration of channel multiplexing in terms of bit rate matching. It is noted that there are residual capacities in FACHs when the bit rate of MTCH is less than the bit rate of FACH, which lead to inefficient usage of radio resources.

Consequently, 2-level channel multiplexing emerges as a promising solution to the problem faced by its single-level counterpart. However, several issues have to be considered in the 2-level channel multiplexing in order to perform resource-friendly and power-efficient resource allocation.

Firstly, for a given set of MTCHs, there are arbitrary possible sets of required FACHs with respect to their bit rate constraints. In order to minimize the required number of FACHs and reduce their residual capacities, optimization should be performed for the 2-level channel.
multiplexing scheme. The optimized algorithm should be capable of selecting the best resource-friendly/bandwidth-efficient multiplexing solutions given the MTCHs/FACHs bit rate constraints.

Secondly, one of the most important tasks in SDMB is the optimal usage of the transmission power. In SDMB, all services multiplexed onto a given S-CCPCH will be transmitted based on the worst-case power requirement (i.e. the transmission power is set according to the service that requires the highest transmission power). Therefore, an appropriate way has to be formulated to group all the services according to their respective power requirements in a power-effective manner. In the single-level channel multiplexing scheme, “power-aware” was considered to minimize $E_b/N_0$ requirements for each physical channel, whereby services featuring similar power requirements were suggested to be grouped together into a single physical channel, so that the total transmission power was minimized. This approach improves the overall performance of the SDMB system in terms of power saving, however, it does not take account of optimizing total transmission power, i.e., no rule is formulated for the best power saving estimation. To this end, rather than grouping the services with similar $E_b/N_0$ requirement in a qualitative way, an algorithm should be developed to be capable of deriving the most power-efficient multiplexing solution under given bit rate and power constraints. In this way, the multiplexing decision is executed in accordance with power consumption at the beginning of each new session starts, making it an attractive solution for a power-constraints system. The proposed power-efficient multiplexing can be adaptively applied in a real-time scenario; the solution will derive the most power-efficient multiplexing scheme at the admittance of each new session transmission.

Finally, no strict rules or algorithms are defined for deriving the TFCS for each S-CCPCH, and only general rules were specified in the previous research in SDMB RRM [62]. However, if several key factors (e.g. guaranteed bit rate, maximum bit rate) are considered when deriving the TFCS of each physical channel, the packet-level dynamics can be better tracked and resource utilisation can be greatly improved.

### 3.3 Literature review on packet scheduling techniques

#### 3.3.1 Packet scheduling for general wired systems

Generally speaking, the main functionality of packet scheduling operates at the MAC sublayer of the data link layer, and is aimed at coordinating the access among competing flows arriving at the queuing buffer in the RLC sublayer of the data link layer. The decision in the packet scheduler is made in coordination with some specific criteria in terms of fairness and service requirements,
which varies between one scheduling algorithm to another, effectively impacting the overall QoS guarantees and the network performance. Therefore, the packet scheduling should comply with the following objectives:

- to coordinate the serving order of contending flows, aimed towards the highest possible degree of resource utilisation and spectrum efficiency.
- to minimize the transmission power consumption so as to meet the system power constraints.
- to achieve the best possible QoS satisfaction in terms of different performance criteria, e.g., delay, data rate, on the basis of the service prescribed requirements.
- to track the instantaneous traffic dynamics of contending flows and thereby maintain a certain level of performance requirements.

In the existing literature, there are numerous packet scheduling algorithms targeted at achieving diverse effectiveness, creating crucial hurdles for the design of efficient packet scheduling. The following summarises representative basic packet scheduling schemes in wired networks, emphasizing on the respective performance trade-offs and QoS guarantees.

- **First In First Out.** The simplest and most widely deployed packet scheduling discipline in communication system today is First In First Out (FIFO) (or First Come First Served), where packets in the queue are served in the order they arrive. It must be noted that this policy describes the packet serving behaviour in a single queue dimension. This scheduling discipline is unable to provide any guarantee on delay or bandwidth, and no immunity can be assured for contending flows.

- **Round Robin (RR).** Another simple scheduling algorithm is RR, where queues are served recursively in their order in a non-preemptive manner. This scheduling discipline is non-priority based, thereby offering no differentiation between differentiated service classes. Moreover, the round-robin scheduling is insensitive to packet size; queues with large packet size would be favored over other queues. This algorithm is capable of providing both long-term and short-term fairness and is easy to implement. However, it does not consider any differentiation among users; the overall system throughput is fairly low. To ensure a minimum bandwidth allocation and distribution, being regarded as a well-known variation of RR, the Weighted Round Robin (WRR) [64] assigns a weight to each class. In proportion to the prescribed weights, the available bandwidth is allocated to each class in a round robin manner. The weight assigned to each class can be regarded as a tunable parameter that can effectively determine the overall performance of each class, e.g., delay, throughput.
• **Generalised Processor Sharing (GPS).** GPS [65] assumes an idealised scheduling discipline to share the resources in an efficient, flexible and fair manner based on a fluid model (i.e., infinitesimal packet sizes) that is often not practically realizable. End-to-end guarantees (i.e., bounded delay) among classes can be provided given that the traffic characteristics of the classes are known. Moreover, it can ensure fair allocation of bandwidth among all backlogged sessions which is essential for supporting best-effort and link-sharing services. Due to its desired properties, GPS has become a foundation of QoS network architecture and a benchmark against which the performance of other packet-based service disciplines can be effectively measured and predicted.

• **Priority Queuing (PRIQ).** To offer service differentiation between two or more classes, PRIQ is introduced aimed at providing the best performance for the highest priority class. Packets are maintained in separate queues according to their priorities, queues are served from the highest priority class to the lowest priority class whilst packets in the each queue are served in a FIFO manner. Packets in lower priority queues will not be served until all the higher priority queues are empty. Extra complexity is involved for maintaining prioritized queues. The PRIQ scheduling is capable of offering high priority queues with the highest possible throughput/bandwidth and lowest possible queuing delay. However, lower priority queues are served at the mercy of the higher priority queues, care should be taken so as not to starve lower priority classes when higher priority classes saturate the scheduler. Therefore, this type of scheduling discipline cannot provide consistently satisfied performance on real time scenarios.

• **Virtual Clock (VC).** VC [66] timestamps packets with a virtual clock time obtained from the packet size and data rate of the queue. Packets are scheduled based on their timestamps on the header packet of each queue. VC is capable of controlling the average transmission rate of data flows, and enforcing user’s resource usage according to reservations. VC achieves worst-case delay for a leaky-bucket controlled session; however, VC cannot bound short-term unfairness.

• **Fair Queuing (FQ).** To allow contending flows to fairly share the available link capacity, FQ [67] estimates the finishing time of all the packets at the head of all non-empty queues based on the arrival time of the packet and the packet size, and selects the queue with the minimum finishing time. The data rate achieved by FQ will not be affected by the packet size. Regarded as a packet approximation of GPS, FQ is capable of providing both bounded delay and fairness in wired networks.

• **Weighted Fair Queue (WFQ).** To introduce proportional weights to FQ, WFQ was proposed [68], where the priority is given to the competing flows inversely proportional to
the required bandwidth. Unlike FQ, WFQ allows sessions to have a different share of the resource; it can be regarded as a simple approximation of GPS, where the scheduling allows different sessions to occupy proportional resources. By regulating the weight associated to each session, it is possible to guarantee QoS performance, e.g., end-to-end delay bound and guaranteed data rate.

- **Class Based Queuing/Hierarchical Link Sharing (CBQ/HLS).** CBQ/HLS [69] scheduling algorithm aims to provide flexible link sharing and support multiple queues or classes with bandwidth guarantees. Each class of traffic queue is associated with a link-sharing bandwidth which is aimed to be guaranteed during the scheduling decision. CBQ queues are classified in a hierarchical style, child queues are assigned with some portion of the root queues’ bandwidth, i.e., resource in root queues are split amongst their child queues. In this case, priority may be assigned to each queue based on their required bandwidth.

- **Earliest Deadline First (EDF).** EDF [70] associates each packet with a deadline, which is obtained by the sum of the packet arrival time and its associated delay bound for the class to which the packet belongs. It serves queues in ascending order of deadlines and therefore provides the QoS guarantees for competing sessions. The objective of EDF is to minimize the maximum lateness of packets.

- **Service Curve-based Earliest Deadline (SCED).** As one of the optimized variations of EDF, the SCED [71] first policy employs service curves to provide a wide spectrum of service characterisation by specifying the service using a function. It can be regarded as a generalised policy in that by setting appropriate specification of the service curves, other well-known policies such as VC and the EDF can be mapped as special cases. With the ability for allocating and guaranteeing service curves with arbitrary shapes, SCED is shown to have greater capability to support end-to-end delay-bound requirements than other known scheduling policies.

Of late, cross-layer design has been addressed as a popular subject in research and development activity in the area of communication networks. Although layered architectures have performed well for wired networks, when they are applied to wireless networks, the performance is far from optimum. The reason is that the wireless medium allows richer modalities of communication than wired networks [45]. Performance gains are achievable by actively exploiting the dependence between protocol layers. Toward this end, a survey of cross-layer design for packet scheduling in wireless systems is presented in the following section, and an optimized packet scheduling scheme utilising the cross-layer concept is suggested for a satellite broadcast system.
3.3.2 Packet scheduling schemes in wireless systems

The aforementioned packet scheduling algorithms can be regarded as universal methodologies which can be applied to any scheduling decision problem. However, when applied directly to wireless networks, their performance is far from optimum, in that wireless transmission features burst-like channel errors and location-dependent link states conditioned by time-varying interference, fading and shadowing impairments. Therefore, in wireless environments, mobile stations may experience degraded channel conditions and therefore be unable to transmit data effectively. Considering the wireless error-prone characteristics, the design of packet scheduling techniques suited for wireless systems has attracted much research efforts in the literature and some most popular trends are outlined as follows:

- **Channel-state dependent scheduling.** To apply packet scheduling to wireless networks, compensation is used for offering differentiated treatments for different channel conditions [72]. Priority is given to users who experience bad channel conditions during the scheduling decision period. This type of scheduling classifies the wireless channel into two states, namely BAD and GOOD states, representing the error and error-free channel conditions, respectively. One of the most popular models for emulating the channel state transition procedure is Finite-state Markov channel (FSMC) model [73] with specified error probability associated to wireless channel.

- **Fair queuing related scheduling.** Most of the existing literature on wireless fair queuing algorithms suggested using the well-known wire line fair queuing algorithms for their error-free service model [72]. Representatives of the channel-state dependent scheduling are the Idealised Wireless Fair Queuing (IWFQ) [74], the Channel Condition Independent Fair Queuing (CIF-Q) [75], the Service Based Fairness Approach [76] and the Wireless Fair Service scheduler [77]. They all apply a compensation model on top of classical wireline queuing algorithms, for example, IWFQ uses WFQ or its variants to compute its error-free service, while the CIF-Q simulates the error-free service by applying a compensation model on top of the Start-time Fair Queuing (STFQ) [78], which can be regarded as an enhanced variation of WRR.

- **Location-dependent scheduling.** One of the key difficulties experienced in wireless networks is that a multimedia session can experience location-dependent channel errors, which may have significant impact on the amount of data the session can effectively transmit. Representative contributions in this subject are Channel-condition Independent Fair (CIF) algorithms proposed in [75], where delay and throughput are guaranteed for error-free sessions and both long-term and short-term fairness are considered for error.
Chapter 3. Radio Resource Management for SDMB system

sessions. A token bank fair queuing (TBFQ) scheduling is proposed in [79] for broadband point-to-multipoint WLAN, considering both throughput and fairness under location-dependent channel error conditions.

- **Max C/I scheduling.** In this scheme, the wireless channel quality in terms of the carrier-to-interference ratios (C/I values) is estimated by the receiver and reported back to the transmitter via a feedback channel. A most proper modulation and coding scheme is derived for each user based on the reported C/I and system capacity specifications. Max C/I scheduling technique [79] ranks the mobile users in terms of their respective channel quality, users with the best C/I value have the highest rank and resources are allocated to users according to some predefined criteria. This approach is easy to implement and capable of providing an upper bound on system capacity. However, the performance of Max C/I scheme depends on the distance between mobile users and base station, and the “starvation problem” is more severe for those users near the edge of cell. Therefore, it can be regarded as one of the most unfair schemes for wireless cellular networks.

- **Proportional fair scheduling.** Proportional fair (PF) [81], [82] packet scheduling is applied to wireless communication systems by scheduling the radio resource according to the pre-assigned priority associated to each user. This scheme provides better fairness than Max C/I and better throughput than Round Robin. However, the PF does not necessarily provide a good overall system throughput, e.g., it provides a poor delay profile compared to Max C/I [83]. It is also shown that the PF could provide a fair output for the wireless end-users as time elapses [82].

- **Cross-layer scheduling.** The methodology introduced for this type of scheduling uses cross-layer information for the scheduling decision. As proposed in [84], in order to achieve more efficient scheduling for diverse QoS guarantees, the interactive queuing behaviours induced by heterogeneous traffic and the dynamic variation of wireless channels are considered in the scheduler design. One of the most popular schemes in this area is to design an adaptive modulation and coding (AMC) scheme at the physical layer in conjunction with the packet scheduling procedure at the data link layer to guarantee the prescribed QoS and achieve efficient bandwidth utilisation simultaneously [85]. For example, in [86], the authors utilise the CSI estimated at the receiver, and select the most appropriate modulation-coding pair, which is sent back to the transmitter through a feedback channel for updating the AMC mode.

- **Opportunistic scheduling.** The basic idea of opportunistic scheduling is to allocate resources to links according to their experienced channel conditions, which is also referred to as channel aware scheduling. Opportunistic scheduling introduces an important trade-off...
between system performance and fairness among users. Numerous research activities on this topic can be found at [87]. Although the opportunistic scheduling may be desirable for best-effort traffic in the physical layer, it provides no QoS guarantees for end users.

- **Utility-based scheduling.** The utility concept has been introduced to the packet scheduling for over a decade [88]. A utility function was proposed to map the service delivered into the performance of the application, aimed at maximising the performance of the applications. In wireless networks, the utility can be represented by specific functions, e.g., delay [67], to reflect various performance criteria.

### 3.4 Packet scheduling in SDMB

#### 3.4.1 Overview

In SDMB, the nonavailability of a return link via satellite penalizes the system effectiveness and efficiency on short-term resource allocation. No fast power control mechanism is applicable in such a system; therefore, the packet scheduling algorithm becomes the focus of efficient resource allocation.

More specifically, in SDMB, the packet scheduler operates periodically in each TTI of the radio bearers, being responsible for two important tasks:

- Time-multiplexing of service flows with different QoS requirements into physical channels with fixed spreading factor (SF), such as to satisfy these requirements.

- Adjusting the transmit power of the physical channel carrying the data flows on the basis of the required reception quality of the service in terms of the target BLER, and under the constraint that the total available power for all the physical channels within a satellite beam is fixed.

---

**Check the capacity request for each FACH.** Based on the radio bearer configuration, determine the respective requests for each S-CCPCH.

**Service prioritization:** derive the serving order on the basis of specific priority criteria, e.g. performance metrics, QoS attributes.

**Resource allocation:** fulfill the bit rate and transmit power assignments within the specific resource allocation interval (i.e. one TTI).

**Scheduled packets are passed to physical layer for radio frame transmission, e.g. CRC, Turbo coding, rate matching, interleaving, ...**

---

**Figure 3-5. Packet scheduling strategy procedures**
As shown in Figure 3-5, the packet scheduling strategy can be encapsulated in following two steps:

- Service prioritization: The incoming service requests are ordered according to the priority criteria. In selecting the respective criteria, the service attributes are considered to provide dynamic intelligent scheduling task whilst considering multiple essential QoS factors that have crucial impact on system performance.

- Resource allocation: Once all of the multiplexed sessions are prioritized, resource is allocated accordingly to these sessions, which consist of bit rate and transmit power assignments within the specific resource allocation interval (i.e., one TTI).

With the unique nature of the SDMB system and growing demand of QoS requirements, the design of a packet scheduler in SDMB becomes a particular challenging task for the following reasons:

- Given the lack of a return satellite link, the scheduler has to perform the resource allocation without the knowledge of the state of individual channels, i.e., channel state dependent scheduling is not possible. Even if such information were available, it would have to be exploited in an unconventional manner considering the p-t-mp nature of the supported services.

- Wireless multimedia applications feature stringent and diverse QoS requirements, hence the design of a packet scheduler has to take into account both the differentiation and fulfillment of these requirements.

- The total available power for all the physical channels within a single satellite beam is limited, the packet scheduler has to be designed such as to satisfy the power constraint and minimize the unnecessary power consumption.

### 3.4.2 Assessment of existing packet scheduling schemes in SDMB

Given the unidirectional nature of the SDMB system and the p-t-mp services it provides, aimed at maximising spectrum efficiency and satisfying diverse QoS requirements whilst preserving the radio resources is the aim of the RRM. The design of RRM functionalities implemented at the SDMB access layer proves to be an especially challenging task, and the packet scheduling algorithm, which is the single function performing short-term resource allocation, is the focus of such efficient resource allocation.

Although numerous studies on packet scheduling schemes have been proposed in the literature for both wire- and wireless- network [89]-[93], they cannot be easily applied to the SDMB system
because of its unique architecture constraints. One interesting approach in this context is delay differentiated scheduling, where waiting time and queuing delay are considered in the packet scheduling, as waiting time priority (WTP) and proportional delay differentiation (PDD) schemes as proposed in [89] for terrestrial differentiation networks. Another possibility to achieve is to exploit the channel quality of the fast-varying wireless link for more efficient packet scheduling. For example, adaptive proportional fairness (APF) scheduling was proposed in high-speed downlink packet access (HSDPA) system [90], considering QoS demands for multimedia applications, where the CSI information for individual user can be tracked via return channel. However, given the unidirectional nature and long propagation delay, the SDMB system is incapable of tracking real-time CSI from the mobile terminal side, which renders infeasible channel-state dependent scheduling.

Previous studies [91] have systematically addressed the packet scheduling problems in the SDMB system via classical packet scheduling schemes, namely multi-level priority queuing (MLPQ) and weighted fair queuing (WFQ). However, both of these schemes feature major weaknesses in QoS-differentiated multimedia services provisioning with respect to both efficiency and fairness.

WFQ-based scheduling was motivated and developed in the SDMB system based on the well-known WFQ scheme, being capable of guaranteeing a minimum bandwidth per bearer or per set of bearers grouped together for traffic handling purposes [91]. The WFQ-based scheduler is more specifically based on the Virtual Spacing policy that uses the notion of Virtual Time [94]. The weights are primarily set according to the data rates of the multiplexed service flows rather than its priority. The weight distribution amongst flows can be adapted in response to new acceptances of a service or variation of channel mapping. The serving orders of the queues are computed depending on the time-stamp of the head packet of each queue, queues with the lowest time-stamp on its head packet will be served first. The non-priority nature of this scheduling policy leads to unacceptably long queuing delays in higher priority queues. The performance of WFQ is worse than that of MLPQ in terms of both delay and delay variation [13].

Multi-Level Priority Queuing-based scheduling [91] is effectively the adaptation of the multi-level, non-pre-emptive priority discipline [95] to the SDMB context. MLPQ always processes packets starting from those non-empty queues having the highest priority first, with queues having the same priority served in a round robin fashion.

Firstly, MLPQ employs a strict QoS-based prioritization scheme, in which a lower-priority service may suffer from considerably longer queuing delays. It always processes packets from those non-empty queues with the highest priority; as a result, packets waiting in the lower priority queues may suffer from considerably longer queuing delays. This scheme favours the high
priority classes, assuring a delay bound for their packets, whilst it provides no guarantees for lower priority classes.

Furthermore, it is generally agreed that background applications have no stringent delay constraint, and the only requirement for application in this category is that information should be delivered to the user essentially error free. In fact, background applications still need a delay constraint (at least an upper bound), since data can effectively be useless if it is received too late for practical purposes.

Finally, MLPQ deals with queues having the same priority in a round-robin fashion. Consequently, there is no differentiation made between queues with the same QoS rank. However, this is not an efficient mechanism. Rather than prioritizing queues in a strict manner, other essential QoS metrics (e.g., delay tolerance and guaranteed data rate) should also be considered in the scheduling discipline design.

Although MLPQ and WFQ have advantages in computational and implementation complexity, their efficiency and fairness are far from optimum.

### 3.5 Summary

In this chapter, basic ideas of RRM in the SDMB system are presented in a holistic manner. The RRM is the main subject throughout this thesis, where we focus on two of its most critical functions in the unidirectional broadcasting scenario: radio resource allocation and packet scheduling. The rest of the thesis will investigate the possible optimization techniques pertinent to these functions, aimed at efficient resource utilisation and comprehensive QoS assurance. We do so by considering the existing optimization schemes within this context in both wired, wireless and satellite environments, seeking to use the best from each to suit the specific SDMB system.
Chapter 4

4 Novel Radio Resource Allocation Scheme

4.1 Overview

The provisioning of multimedia applications requires wireless networks to optimally allocate radio resource and efficiently adapt to dynamically changing wireless environments. Radio Resource Allocation (RRA) strategy is an important issue in the design of the SDMB system, aimed at providing more efficient and cost-effective delivery of BC/MC services to a wide range of audience. Based on the single-level channel multiplexing framework, the existing RRA strategy only performs transport channel multiplexing and assumes one-to-one ideal mapping between logical and transport channels. However, the bit rates of the logical channels do not always exactly match those of the transport channels, and this leads to inefficient usage of radio resource. To solve this problem, in this chapter, a novel RRA strategy, which is capable of optimally performing both logical and transport channel multiplexing, is proposed for the SDMB system. In order to improve the resource utilisation and reduce the power consumption, two new algorithms, namely Optimum Estimation Algorithm (OEA) and Power-Oriented Adaptation (POA) algorithm, are developed to optimize the proposed 2-level channel multiplexing scheme. In comparison to the existing single-level channel multiplexing scheme, simulation results show that the proposed 2-level channel multiplexing scheme achieves significantly improved performance in terms of transport/physical channel utilisation, transmission power consumption as well as total transmission capacity.

4.2 Introduction

In recent years, satellite networks have gained considerable interest for the delivery of multimedia applications due to their inherent ubiquity of wide coverage, native support of BC/MC and relatively high achievable speeds. One of the emerging candidates is the SDMB system, which implements a satellite based broadcast layer over 2.5G and 3G terrestrial mobile cellular networks aimed at the efficient delivery of the MBMS [6]. Due to the scarcity of satellite capacity and bandwidth resource, with the absence of a direct satellite return link and the p-t-mp nature in the SDMB system, the design of the SDMB radio access scheme is becoming a particularly
challenging task. In this chapter, we focus on one of the RRM entities relevant to the SDMB radio access scheme: the RRA, which is responsible for estimation of the required number of logical/transport/physical channels, and mapping them together with the actual TFCS onto each physical channel.

Previous research on RRA performs single-level channel multiplexing across the protocol layers [62]. It assumes that the MTCH logical channels are mapped one-to-one onto the FACH transport channels, which are subsequently multiplexed onto the S-CCPCH physical channel. This type of channel mapping which only considers single-level channel multiplexing (i.e. transport channel multiplexing) is shown in Figure 4-1 (a). Therein a one-to-one correspondence between the MBMS services and the Radio Access Bearers (RABs) is assumed, whereby the MBMS services are mapped one-to-one onto the logical channels (i.e. MTCHs) at the RLC sub-layer at the beginning of each session starts. The logical channels are then mapped, in a one-to-one manner, onto the transport channels (i.e. FACHs) at the MAC sub-layer, and transport channels are multiplexed onto the physical channels (i.e. S-CCPCH) at physical layer. However, with this simple one-to-one mapping between logical and transport channels, there exist residual capacities (i.e. bandwidth) on the transport channels when the bit rate of the logical channel does not exactly match the corresponding bit rate of the transport channel, i.e. the MTCH bit rate is less than the FACH bit rate.

In order to resolve this inefficiency in resource utilisation, based on our previous research [55], [56], an optimized 2-level channel multiplexing scheme has been proposed. This scheme performs both logical and transport channel multiplexing as shown in Figure 4-1 (b). At the first level of multiplexing, multiple MTCHs are multiplexed onto a single FACH (i.e. logical channel multiplexing), whereas the multiplexing of several FACHs onto a single S-CCPCH (i.e. transport channel multiplexing) is regarded as the second level of multiplexing. In order to achieve the highest possible degree of utilising scarce wireless resource whilst meeting the target QoS
requirements for differentiated service groups, several optimization processes have been introduced and developed for the proposed 2-level channel multiplexing scheme.

Firstly, a new algorithm, namely Optimum Estimation Algorithm (OEA), is proposed and designed to minimize the required number of FACHs and reduce their residual capacities for a given set of MBMS services, whose service characteristics and blocking probabilities are known a priori. In this algorithm, a 2-stage optimized "bin-packing" approach has been developed to perform the channel multiplexing. It is noted that, from the viewpoint of bin-packing, the MTCHs will be referred to as items to be packed; the S-CCPCHs as bins; while the FACHs will be denoted as intermediate-bins [55]. In this study, we consider how to perform optimum packing from items to intermediate-bins with their bit rate constraints in order to improve resource utilisation. Secondly, the power requirements are considered in designing the 2-level channel multiplexing, and a new Power-Oriented Adaptation (POA) algorithm is developed to perform the transmission power optimization. In this scheme, the power consumption of individual MTCH/FACH candidate is considered as a critical dimension. By recursively searching the tentative multiplexing scenarios, the POA algorithm is capable of effectively reducing the total required transmission power at the physical layer.

Simulation has been conducted to demonstrate the effectiveness and efficiency of the proposed scheme. In comparison to the conventional single-level channel multiplexing scheme, the proposed 2-level channel multiplexing scheme achieves significantly better performance in terms of channel utilisation, transmission power consumption as well as total transmission capacity.

The remainder of this chapter is organised as follows. The feasibility of the logical channel multiplexing is investigated in Section 4.3. Section 4.4 presents the proposed 2-level channel multiplexing scheme, including the description of the OEA algorithm and the POA algorithm. In Section 4.5, the proposed 2-level channel multiplexing scheme is evaluated via means of simulation and the performance enhancement is discussed. Finally, the summary is given in Section 4.6.

4.3 Feasibility analysis

Since the 2-level channel multiplexing has introduced the logical channel multiplexing, this section will discuss the feasibility of the logical channel multiplexing in the RRA strategy for the SDMB system.

The logical channel multiplexing in SDMB is controlled by its MAC sub-layer, which is modified and adopted based on the UMTS MAC sub-layer, and the UMTS MAC has specified entities and functionalities to support the efficient delivery of MBMS services. For example, the newly
defined MAC-m entity and the introduction of MBMS-id field in MAC header make the logical channel multiplexing a feasible solution in the design of an efficient RRA multiplexing scheme.

Figure 4-2 illustrates the data delivery between logical channel and transport channel that is standardized within 3GPP [96]. In order to support p-t-m type of logical channels, the MAC header is modified from the standard UTRAN format. UE-id field, which was used for identifying different logical channels multiplexed onto the same transport channel, is herein replaced by the MBMS-id field that fulfils the equivalent task for MBMS services. As illustrated in Figure 4-2 (a), the SDMB-RAN utilises this field to distinguish between MBMS services, while from the UE-side in Figure 4-2 (b), the MBMS-id allows the user to identify specific MBMS service. Target Channel Type Field (TCTF) implies the type of the logical channel with the presence of logical channel multiplexing [59]. To this end, the logical channel multiplexing is feasible and workable in the SDMB system.

Figure 4-2. MAC-m architecture in SDMB: (a) SDMB RAN side, (b) UE side [97].
An example of logical channel multiplexing considered in SDMB is shown in Figure 4-3. Firstly, RLC performs segmentation/reassembly of variable-length upper layer Packet Data Units (PDUs) into/from smaller size-adjustable RLC PDUs. Concatenation/padding may be applicable depending on the sizes of the PDUs between higher layer and RLC layer. The TCTF field and MBMS-id field in the MAC header are used to identify the logical channel type and MBMS service respectively. The logical channel multiplexing is performed at the MAC layer, and the MAC header includes MBMS-id field that identifies a specific logical channel when several logical channels are multiplexed onto the same transport channel. From the viewpoint of resource utilisation, the services carried on the logical channels within the same transport channel should feature similar characteristics and QoS requirements. Similarly, the services carried on the transport channels with similar characteristics and QoS requirements are suggested to be multiplexed onto the same physical channel; otherwise, the radio resource would be used inefficiently. In order to efficiently track the packet-level dynamics of the competing flows, the size of TFCS should be proportional to the number of multiplexed transport channels at the physical layer (i.e. size of TFCs). By employing logical channel multiplexing, the number of transport channels, namely TFCs’ size is dramatically reduced; therefore the essential benefit from logical channel multiplexing is the reduction in the size of TFCS.
4.4 Proposed radio resource allocation strategy

4.4.1 Overview

Based on our previous research on 2-level channel multiplexing [55], [56], an innovative and more efficient approach is developed. It has several key features: firstly, according to the upper layer MTCHs' bit rate and the available bit rate of FACHs, a novel OEA algorithm is designed for the 2-stage bin-packing algorithm so as to minimize both the required number of FACHs as well as their residual capacities, thereby maximising the radio resource utilisation of the system. Secondly, the power-aware concept has been enhanced and implemented in the design of 2-level channel multiplexing and a new POA algorithm is proposed aimed at achieving the most power-efficient channel multiplexing. Finally, a specific rule has been formulated for the derivation of TFCS for each physical channel in the 2-level channel multiplexing scenario. The objectives of our approach are listed as follows:

- Minimize the residual capacity on FACHs;
- Minimize the required number of FACHs;
- Reduce the size of TFCS;
- Improve transmission capacity;
- Improve channel utilisation;
- Optimize total transmission power.

In comparison to the aforementioned three-step single-level channel multiplexing procedure, we define a five-step procedure to perform the proposed 2-level channel multiplexing task, which includes:

- Estimation of the required number and bit rate of MTCHs based on their service characteristics and target blocking probabilities;
- Estimation of the required number and bit rate of FACHs using the OEA algorithm;
- Mapping of derived FACHs onto S-CCPCHs;
- Power optimization using the POA algorithm;
- Derivation of TFCS for each S-CCPCH.
Two optimization algorithms are proposed for the 2-level channel multiplexing: the OEA algorithm and the POA algorithm. The purpose of the OEA algorithm is to minimize the required number of FACHs and reduce their residual capacities. While the POA algorithm is designed to minimize the required total transmission power. These two algorithms will be addressed in detail in the following sections.

4.4.2 Two-stage bin-packing with optimum estimation algorithm

Figure 4-4 depicts the proposed two-level channel multiplexing framework. The purpose of the two-stage bin-packing with OEA algorithm is to minimize the required number of FACHs and MTCHs, and reduce the residual capacity of the transport channels. The first stage of channel mapping uses the proposed OEA algorithms, whilst the second stage of channel mapping uses the Best-Fit (BF) bin-packing algorithm.

Previous research [62] on single-stage bin-packing algorithms splits the RRA task into three main steps, i.e. estimation of required MTCHs, mapping of logical/transport/physical channels, and derivation of TFCS for each physical channel. In the context of the proposed 2-level channel multiplexing, five key steps as described above are identified to perform the configuration of radio bearer allocation and mapping. In this section, the first three steps relating to the channel mapping using the proposed OEA algorithm will be elaborated respectively, the steps regarding the power optimization and TFCS derivation will be presented in the following section.

**Step 1: Estimation the required number and bit rate of MTCHs**

The aim of the RRA in mode A \(^{13}\) [61] is to perform the dimensioning of the system, on the basis of the traffic mix; the assumption is that there is adequate, minimum characterisation of individual

---

\(^{13}\) In SDMB, two modes of operation are identified with the RB configuration. Mode A defines a fixed RB configuration over some interval of time, over which the traffic mix remains the same. The mappings of
services in terms of arrival rate $\lambda_i$, duration $\mu_i$ and requested rate for each type of service $R_i$. Hence the method described herein is only applicable for the streaming services portion in the traffic mix.

This step is slightly different from the conventional single-level channel multiplexing. Let $K$ be the set of different streaming services. A streaming service can be characterised by the 3-tuple - arrival rate $\lambda_i$, duration $\mu_i$ and requested rate for each type of service $R_i$. The cardinality of the service set is $N$, i.e. $|K| = N$. No assumption is made for the flow burstiness; the flow might be of constant bit rate (CBR) or variable bit rate (VBR), but in the latter case the $R_i$ value is set to the mean/guaranteed rate attribute. Each element $s_i$ corresponds to a member of the streaming service sets.

Let $P_{bl}$ be a vector of size $N$ corresponding to blocking probabilities targeted for each streaming service, i.e. there is one-to-one correspondence between $s_i$ and $P_{bl}$.

Then the required MTCHs for each $s_i$ can be derived via well-known results of classical queuing theory [63]:

- From the m-server loss queuing system, for each service type $s_i$ separately, i.e. invocation of the M/M/m/m formula $N$ times.
- From the extension of the Erlangian formula to the multiple services scenario over all types of flows $s_i$ requesting the same rate $R_i$, irrespective of the arrival rates or service durations of the individual services. The respective formula [98] is applicable under the complete-sharing (CS) assumption: FACHs can be fully shared among services requesting the same rates (i.e. as long as the derived TFCS can cope with possible discrepancies at the packet level).

In both cases, the required number of MTCHs is the number of servers that will guarantee the target blocking probability $P_{bl}$.

**Step 2: Estimation of the required number and bit rate of FACHs using the OEA algorithm**

logical/transport/physical channels are drawn once per reconfiguration interval, the pre-configured MTCHs/FACHs can be at active/idle status depending on whether a service flow is carried over them or not. In mode B, the RB mapping is drawn in an ad-hoc manner by the AC without any prior configuration, AC decides the admittance of the service request on the basis of power/load constraints, the task of RRA is to re-map/re-configure MTCH/FACHs upon S-CCPCHs upon a newly admitted service request. In mode B, the S-CCPCH serves only active channels.
This is the additional step for the extension from the single-level channel multiplexing to the 2-
level channel multiplexing algorithm. From a bin-packing point of view, the aim is to estimate the
required intermediate-bins (i.e. FACHs) based on the given items (i.e. MTCHs). There are two
proposed approaches to estimate and determine the required number of FACHs in the 2-level
channel multiplexing.

A general approach for performing this task is to specify the intermediate-bin a priori according to
the available bit rate of FACHs satisfying the following rules [55]:

- The total bit rate of the FACHs should not be less than the total bit rate of the MTCHs so
  that all the MTCHs can be carried by the FACHs.
- A wide range of bit rate of the FACHs should be made available to be chosen from so as to
  provide the maximum possible degree of selectivity and satisfaction for the different rates
  of the MTCHs. This is to allow the different mapping schemes to be employed in the next
  step.

The MTCHs are then mapped onto the predefined FACHs using different bin-packing algorithm,
as will be described in Step 3. The inter-relationship and inter-dependence between two stages can
be illustrated in Figure 4-5:

![Figure 4-5. Inter-relationship between the two multiplexing stages.](image)

As shown in Figure 4-5, after the first stage of multiplexing, the FACHs are filled with multiple
MTCHs, whilst the S-CCPCHs are filled with multiple FACHs. this will lead to better resource
utilisation and less residual capacity on both FACHs/S-CCPCHs. However, given the fixed
MTCH and available bit rate of FACHs (i.e. 256, 128, 64, 32kbps), the multiplexing scenario
from MTCHs to FACHs is uncertain, and this uncertainty also applies to the multiplexing from
FACHs to S-CCPCHs. Therefore, rules/algorithms have to be developed to make this
multiplexing task resource-friendly.

According to the specific service characteristic and mapping scenario, aimed at minimizing both
the number of required FACHs as well as their residual capacities, the required number and bit
rates of FACHs are optimally estimated based on the bit rates required by the MTCHs and the
available bit rates supportable for the FACHs. This proposed optimization algorithm, which is called “optimum estimation algorithm (OEA)”, can be described as follows.

Let $R$ be a vector of size $N$, corresponding to a set of MTCHs’ bit rates, i.e. $\{R_1, R_2, \ldots, R_N\}$. The set of available FACHs’ bit rates is $S = \{S_1, S_2, \ldots, S_m\}$, which in our case is $\{256\text{kbps}, 128\text{kbps}, 64\text{kbps}, 32\text{kbps}\}$, following the 3GPP MBMS standardization [6]. The required number of FACHs will be set according to both $R$ and $S$.

![Diagram of the optimum estimation algorithm for estimating required FACHs](image)

Firstly, the elements in both $R$ and $S$ are reordered in decreasing order of bit rates, and vector $D = \{\}$ is predefined as the output of the FACH bit rate set, as shown in Figure 4-6.

Then, the bit rates of MTCHs are selected from $R$, one by one, and compared to the available FACHs’ bit rate $S_j$ in sequence:

- If $R_i = S_j$, which means the bit rate of the selected MTCH exactly matches the bit rate of FACH, then the bit rate $S_j$ is assigned into the required FACH set $D$;
- If $R_i > S_j$, which means the MTCHs cannot be accommodated by any available FACH, then the required session cannot be supported to transmission.

• If \( R_i < S_j \), in order to achieve the minimum number of required FACHs, the residual capacity of FACHs after the mapping is considered. Then \( S_{j+1} \), which is the first lower bit rate next to \( S_j \) in the FACH set \( S \), is selected and checked if it can accommodate the MTCH.

Note that the residual capacity of FACHs after the first-stage of mapping is an important parameter when selecting the best-suited FACHs. Let \( C_{res} \) be a parameter which is defined as the difference between the bit rate of FACH \( S_j \) and the summation of the bit rates of the remaining MTCHs:

\[
C_{res} = S_j - \sum_{k=i}^{N} R_k \tag{4.1}
\]

• If \( C_{res} = 0 \), it means that the MTCHs are exactly fitted in the FACH, then \( D = \{ D, S_j \} \), and all the remaining MTCHs are assigned to the FACH;

• If \( C_{res} > 0 \), it means that there is residual capacity on the chosen FACH. Then another FACH with smaller rate \( S_{j+1} \) is selected for accommodating the remaining MTCHs in order to achieve minimum residual capacity;

• If \( C_{res} < 0 \), it means that the chosen FACH is fully filled by the MTCHs, and there are remaining MTCHs need to be assigned to new FACHs, then \( D = \{ D, S_j \} \) and shifts to the next MTCHs that is not assigned.

The procedure is repeated recursively for each one of the remaining MTCHs until all the MTCHs are assigned.

It must be noted that the first stage of channel multiplexing (i.e. logical channel multiplexing) is actually performed by the OEA algorithm, which minimizes both the residual capacity as well as the required number of the FACHs, and this in effect leads to the maximum channel utilisation and transmission capacity on both FACHs and S-CCPCHs. From the implementation point of view, the optimum estimation algorithm features nonlinear (with both inner and outer loop) and unpredictable (undeterministic variables) characteristics. In order to evaluate the computational complexity of the novel algorithm, we consider the worst case scenarios where the processing time is the most expensive among all possible scenarios. The "Big O notation" [99] is cited hereby to examine the asymptotic complexity of the algorithm. With the input size of \( n \) (i.e. the total number of MTCHs), the computational time complexity (running time) is derived as \( O(n^2) \), whilst the worst case computational complexity of the proposed algorithm features quadratic statistics.
Chapter 4. Novel Radio Resource Allocation Scheme

Step 3: Mapping of the MTCHs – FACHs – S-CCPCHs

Previous research at this step used the single-stage bin-packing channel mapping scheme with no logical channel multiplexing, whereby MTCHs logical channels were mapped to FACH transport channels in a one-to-one manner, whilst the classical bin-packing algorithm is employed when FACH transport channels are multiplexed onto S-CCPCH physical channels.

There are several alternative bin-packing algorithms [100] that can be chosen for the channel mapping.

- Next-Fit (NF) bin-packing algorithm: each item is assigned to the current bin if it fits; otherwise, it is assigned to the next new bin, which becomes the current bin.
- First-Fit (FF) bin-packing algorithm: the items are considered according to increasing indices and each item is assigned to the lowest indexed initialized bin into which it fits. Only when the current item cannot fit into any initialized bin is a new bin introduced.
- Best-Fit (BF) bin-packing algorithm: the current item is assigned to a feasible bin (if any) having the smallest residual capacity.

Exiting literature [55] indicates that the BF bin-packing algorithm achieves the highest possible degree of utilising the lower layer channel capacity under certain mapping condition, where both items and bins are predefined as fixed vectors before mapping is performed.

After the first stage of channel mapping/multiplexing, the derived FACH transport channel becomes a fixed set of intermediate-bins to be assigned, whilst the number of available S-CCPCHs and their maximum capacities are assumed to be known a priori and is herein predefined as a fixed set of bins for a certain scenario. Thereby BF bin-packing algorithm, where the current item is assigned to a feasible bin (if any) having the smallest residual capacity, is chosen as the mapping algorithm in the second stage of channel mapping/multiplexing (transport channel multiplexing) in that it achieves the best performance under scenarios with fixed sets of MTCH/FACH data rates.

On the other hand, for the dimensioning of background QoS class, which cannot be characterised in terms of 3-tuple \( \{ \lambda, \mu, R \} \), the capacities not used for streaming FACHs will be used to carry download (i.e. push & store) services. After the estimation of required FACHs for streaming services in Step 2, more residual capacities are released in the physical channels and can be used for carrying download FACHs. In order to fully fill the capacity gaps in the physical channels after allocating streaming services, given its best performance on fixed state bin-packing scenario, the BF bin-packing algorithm is employed to estimate the number of the maximum allowable
download services that can be carried on the residual capacities of the physical channels. The aim of this estimation is to carrying download services using all residual capacities remaining in the physical channels.

Mathematical formulation of conventional single-stage bin-packing approach, which does not take into account the power requirement ($E_p/N_0$) of individual services, could be expressed as:

$$\text{minimize: } Z = \sum_{j=1}^{N} y_j$$

subject to: $$\sum_{j=1}^{N} R_i x_{ij} \leq c_j y_j; \quad j \in \{1...N\}$$

$$\sum_{j=1}^{N} x_{ij} = 1; \quad i \in \{1...N\}$$

where $y_j = 1$, if bin $j$ is used, 0 otherwise; and $x_{ij} = 1$, if item $i$ is assigned to bin $j$, 0 otherwise. $N$ is the number of items to be assigned, the total number of bins is $Z$, and their capacity $c_j$. This is the single-stage bin-packing problem solution, i.e. given the items (MTCHs/FACHs), and packs them into the minimum possible number of bins (S-CCPCHs). A feasible solution of the problem could be: $\{ Z \leq N \}$.

In this section, the conventional single-stage bin-packing approach is modified and developed for adapting the 2-level channel multiplexing. The first stage of channel multiplexing (logical channel multiplexing) is performed by the proposed the OEA algorithm as described in Step 2, which achieves both the minimum number of FACHs as well as the minimum residual capacity on those FACHs, and this in effect leads to the maximum utilisation on both FACHs and S-CCPCHs. After the first stage of multiplexing, the derived FACH transport channel becomes a fixed set of intermediate-bins to be assigned, whilst the S-CCPCH physical channel is predefined as a fixed set of bins in a certain scenario. Thereby Best-Fit bin-packing algorithm [100], where the current item is assigned to a feasible bin (if any) having the smallest residual capacity, is chosen as the mapping algorithm in the second stage of channel multiplexing (transport channel multiplexing) in that it achieves the best performance under certain condition bin-packing mapping.

Given that the number of MTCHs is $N$, the number of FACHs is $M$, and their capacity is $c_j$, the number of available S-CCPCHs is $K$ and their maximum capacity $c$, are known a priori. Let $Y$ be the total number of required FACHs and $w$ be the total residual capacities on those FACHs. In comparison to (4.2), the mathematical formulation that gives the 2-stage bin-packing solution could be extended as:
\[ \text{minimize: } Z = \sum_{k=1}^{K} z_k; \quad Y = \sum_{j=1}^{M} y_j; \quad w = \sum_{j=1}^{M} r_j \]

subject to:
\[ \sum_{j=1}^{M} R_j x_j = c_j y_j + r_j; \quad j \in \{1...M\} \]
\[ \sum_{j=1}^{M} c_j y_j \leq c z_k \tag{4.3} \]
\[ \sum_{j=1}^{M} x_j = 1; \quad i \in \{1...N\} \]
\[ \sum_{k=1}^{K} y_{jk} = 1; \quad j \in \{1...M\} \]

where \(z_k = 1\) if bin (S-CCPCH) \(k\) is used, 0 otherwise; \(y_j = 1\) if intermediate-bin (FACH) \(j\) is used, 0 otherwise; and \(x_j = 1\), if item \(i\) is assigned to intermediate-bin \(j\), 0 otherwise; and \(y_{jk} = 1\), if intermediate-bin \(j\) is assigned to bin \(k\), 0 otherwise. \(r_j\) is the residual capacity of each intermediate-bin \(j\). The condition for a feasible solution of the problem should be: \(Z \leq M \leq N\).

Given the conventional linear objective function of the optimization problem, adaptation of the classical bin-packing problem is used to obtain this solution.

4.4.3 Power-oriented adaptation algorithm

In this step, the POA algorithm is performed as an independent optimization step, aimed at deriving the best possible channel multiplexing solution which can ultimately lead to the minimum physical channel transmission power. This enables enhancement of the power-aware concept from the single-level channel multiplexing counterpart and allows choice of the best-case mapping from all possible candidates, rather than only grouping those MTCHs/FACHs with the similar power requirements.

Following Steps 1 and 2, the sets of required MTCHs/FACHs have been determined, whilst there can be arbitrary possible channel multiplexing solutions from MTCHs to FACHs, i.e. which MTCH is mapped onto which FACH. In fact, each solution leads to independent total transmission power in the physical channels, essentially depending on the worst case power requirement rather than the sum of individual service powers. To this end, the power-oriented adaptation algorithm evaluates the most power-efficient channel multiplexing solutions from MTCHs to FACHs/S-CCPCHs under fixed MTCHs/FACHs’ bit rate sets.

In SDMB, the downlink power requirement for a given service \(j\) is determined according to the minimum \(E_b/N_0\) requirement for this service. By assuming the uniform traffic distribution and
constant propagation conditions over a spot-beam coverage, the link quality for S-UMTS FDD downlink can be expressed via the classical WCDMA Downlink Pole Equation as [115]:

\[
\left(\frac{E_b}{N_o}\right)_i = \left(\frac{W}{R_i}\right) \times \left(\frac{P_i}{P_n \times L_i + I_{own} + I_{oth}}\right)
\]

(4.4)

where \(W\) is the chip rate, \(R_i\) is the \(i^{th}\) MTCH instantaneous bit rate, \(P_i\) is the received power by a user for \(i^{th}\) service, \(P_n\) is the thermal noise, \(L_i\) is the path loss from satellite to the \(i^{th}\) UE, \(I_{own}\) is the own beam interference, \(I_{oth}\) is the other beam interference.

Furthermore, the \(I_{own}\) can be expressed by using the orthogonality factor (\(\alpha\)) and with the total transmission power (\(P\)) of a satellite spot beam as:

\[
I_{own} = (1 - \alpha) \times P
\]

(4.5)

Given the following assumptions for SDMB [101]:

- The other cell interference can be considered negligible as the adjacent spot beams use different base-band frequencies;

- Orthogonality between channels can be assumed to be perfect (i.e. \(\alpha=1\));

- Therefore, equation (4.4) can be simplified for a given service as:

\[
P_i = \left(\frac{R_i}{W}\right) \times \left(\frac{E_b}{N_o}\right)_i \times P_n \times L_i
\]

(4.6)

Therefore the required transmit power for the \(i^{th}\) service \(P_i\) is directly proportional to \(\left(\frac{E_b}{N_o}\right)_i\), and the session bit rate \(R_i\). In the following, the \(E_b/N_o\) requirement is used as the reflecting factor of the power requirement for individual services.

As shown in Figure 4-7, the POA algorithm can be described as follows:
Chapter 4. Novel Radio Resource Allocation Scheme

Let $R, S$ be the derived sets of MTCHs' bit rates and required FACHs' bit rates based on the POA algorithm. Assign the $E_b/N_0$ requirement for each service $i$ as a vector of $N$ corresponding to the set of MTCHs, i.e. $P = \{P_1, P_2, ..., P_N\}$.

Find all the mapping solution candidates which can fit into the derived MTCH-FACH set recursively, checking the feasibility of the derived MTCH and FACH in terms of their bit rates $(R_i^{MTCH}, R_j^{FACH})$ satisfying:

$$R_j^{FACH} = \sum_{i=1}^{I} R_i^{MTCH}$$  \hspace{1cm} (4.7)

where $I$ is the total number of MTCHs multiplexed onto the $j^{th}$ FACH. The derived $m$ feasible mapping scenario will be candidates for the following steps.

Figure 4-7. Power-oriented adaptation algorithm.
The total transmission power $P_{\text{total}}$ in the physical layer will be finally determined by the worst case transmission power of all multiplexed traffic flows as (4.8). The required power for each FACH $P_{i}^{\text{FACH}}$ is the maximum required power amongst all MTCHs multiplexed onto it, whilst the power requirement for each S-CCPCH $P_{i}^{\text{SCCPCH}}$ is the maximum power required amongst all FACHs multiplexed onto it.

\[
P_{i}^{\text{FACH}} = \max \{ P_{j}^{\text{MTCH}} \}
\]

\[
P_{i}^{\text{SCCPCH}} = \max \{ P_{j}^{\text{FACH}} \}
\]

\[
P_{\text{total}} = \sum P_{i}^{\text{SCCPCH}}
\]  

Evaluate the total transmission power of scenario $x$ ($P_{\text{total}}^{x}$) for all the $m$ rate-feasible mapping scenarios. The scenario which requires the minimum total transmission power $P_{\text{total}}^{\min}$ as defined in (4.9) is the selected power-optimum scenario with POA total transmission power ($P_{\text{POA}}^{\text{optimal}}$).

\[
P_{\text{total}}^{\text{POA}} = P_{\text{total}}^{\min} = \min (P_{\text{total}}^{x} : x = 1, \ldots, m)
\]  

However, this proposed POA algorithm introduces extra computational complexity due to its recursive adaptation nature. With the input size of $n$ (i.e. the total number of MTCHs), the complexity function is derived as $f(n) = 5n + 7$, whilst the involved computational time complexity (running time) of POA algorithm is at worst $O(n)$ featuring linear statistics. Given relatively small number of MTCHs, computational complexity does not introduce significant constraints.

### 4.4.4 Derivation of the TFCS

The TFCS should be broad enough to capture the packet-level dynamics of the expected services over some future time interval. The chosen transport block size is in line with the packet sizes defined by the applications in order to minimize the overheads (headers and padding).

The method described above is only applicable for streaming services. On the other hand, for the dimensioning of background services (i.e. download services), which cannot be characterised in terms of $\{\lambda, \mu, R_i\}$, the residual capacities on FACHs are mapped all together down to S-CCPCHs, i.e. the FACHs are spread over the available S-CCPCHs and the residual rate of FACHs is difficult to be further utilised. Therefore, the network makes available only the remaining capacities in S-CCPCHs for download services, whilst the residual capacity in FACHs is unable
to be used and thus wasted. Negotiations on the acceptable rate for the download services can be made in accordance with the service priority level for optimum utilisation.

This task performs the derivation of TFCS for each S-CCPCH during the session initialization process; the TFCS derived is passed to the packet scheduler for its short-term resource allocation. For every TTI, the packet scheduler performs a time-multiplexing procedure by choosing a separate TFC for every S-CCPCH from the corresponding TFCS list. A generic approach for deriving the TFCS in traditional single-level channel multiplexing can be found in [61]. Based on this generic approach, we proposed a new derivation algorithm in the 2-level channel multiplexing scenario which is described in the following.

As mentioned above, in order to derive the TFCS for the transport channel multiplexing, the TF/TFS for each transport channel should be derived a priori. According to its definition, a TF consists of the following two parts [102]:

- The dynamic part \( \{ \text{TB}_\text{size}(S^{TB}), \text{TBS}_\text{size}(S^{TBS}) \} \)
- The semi-static part \( \{ \text{TTI}(T), \text{coding rate } (\delta), \text{rate matching ratio } (\sigma), \text{CRC size } (\alpha) \} \)

The TFCs consist of TBSs of different transport channels which are time-multiplexed onto a single S-CCPCH channel and this indicates the total amount of data that can be transferred from the MAC to the physical layer. The size of a TB(S) can be defined as:

\[
S^{TBS} = S^{TB} \times N = (S^{TB}_{\min} \times K) \times N
\]  (4.10)

Where \( S^{TB}_{\min} \) is the minimum TB size (possible values in bits are 320, 160, 80, 40, 122, 244 etc) and \( K \) is the integer number of steps into one TB and \( N \) is the number of TBs in one TBS.

So the task of selecting TFS is focused on the appropriate combination of \( K \) and \( N \), under the constraint of:

\[
S^{TBS}_{\min} < S^{TBS} < S^{TBS}_{\max}
\]  (4.11)

The minimum TBS size \( S^{TBS}_{\min} \) is based on the guaranteed bit rate and the value of TTI, and the maximum TBS size \( S^{TBS}_{\max} \) is determined by the maximum bit rate and value of TTI. The guaranteed bit rate is the guaranteed information bit rate of the incoming service in bps, the maximum bit rate is the maximum information bit rate of the service in bps, and TTI is the value in seconds (possible values 0.01, 0.02, 0.04, 0.08).

Unlike the single-level channel multiplexing scenario, an additional step must be considered in order to derive the TB size of each FACH \( i \). As illustrated in Figure 4-3, the logical channel
multiplexing is actually performed in the form of concatenation of MAC Service Data Units (SDUs), the TB size of each FACH \( i \) can be defined as the addition in size of corresponding RLC PDU (i.e. \( S^{R_{L_{C_{-PDU}}} (j)} \)) and MAC header (i.e. \( S^{MAC_{-header}} \)):

\[
S^{TB} (i) = \sum_{j} ^{J} [S^{R_{L_{C_{-PDU}}} (j)} + S^{MAC_{-header}} ]
\]

\[
= \sum_{j} ^{J} [S^{MAC_{-SDU}} (j) + S^{MAC_{-header}} ]
\]

\[
S^{MAC_{-header}} = S^{MBMS_{-id}} + S^{TCTF}
\]

where \( j \) is the \( j \)th MTCH and \( J \) is the total number of the MTCHs that are multiplexed onto the corresponding FACH.

Since different spreading factors (SFs) are assigned for different slot formats and each correspond to a specific channel bit rate, the SF for an S-CCPCH can be expressed as \( 256/2^k \) (\( k = 0, 1, \ldots , 6 \)) and ranges from 4 to 256. For a given SF, channel bit rate is given by [103]:

\[
R^{S-CCPCH} (kbps) = \frac{(20 \times 2^k) \times 15}{SF^{S-CCPCH} \times 10}
\]

where 20 corresponds to the minimum number of bits in a given time slot within a frame, 15 corresponds to the number of slots within one frame and 10 represents the length of the frame in ms.

The TF size of each FACH \( i \) (\( S^{TF} (i) \)) can be expressed as the combination of TB size (\( S^{TB} \)), CRC size (\( \alpha(i) \)), Rate Matching Ratio (\( \sigma(i) \)), Coding Rate (\( \delta(i) \)), frame interval (\( \Lambda \)), TTI (\( T \)) which is given in the following equation:

\[
S^{TF} (i) = \frac{[S^{TB} (i) + \alpha(i)] \times \delta(i) \times \sigma(i) \times N}{T \times \Lambda}
\]

We must then select the possible combinations that satisfy the following equation:

\[
\sum_{i=1}^{N} \frac{S^{TF} (i)}{T \times \Lambda} \leq \frac{(20 \times 2^k) \times 15}{SF^{S-CCPCH} \times 10}
\]

where \( i \) is the \( i \)th FACH and \( N \) is the total number of the FACHs. The frame interval (i.e. 100 = 1/10 ms) is the number of frames per second. The factor \( (20 \times 256 \times 15 / SF) \) comes from the fact that \( T_{slot} = 2560 \) chips, \( 20 \times 2k \) bits per slot and 15 slots per radio frame.
Following the analysis described above the initial RAB configuration has been produced in the form of a TFCS matrix. At the second stage, this TFCS matrix is subject to admission control and then the final RAB configuration is performed.

### 4.5 Performance evaluation

Comprehensive evaluation and simulation of the performance of the proposed schemes are carried out over a wide range of scenarios, aimed at observing and demonstrating the following:

- Evaluate and compare the performance of the proposed algorithms under different S-CCPCH
- Evaluate and compare the performance of the proposed algorithms under different traffic mixes
- Impact on the overall network performance enhancement

Streaming service characterisation and evaluation of a number of required FACHs are the initial step in RRA, forming the input of the channel multiplexing. Detailed results of this evaluation can be found in Appendix A and B.

#### 4.5.1 Evaluation metrics/simulation scenarios

Simulations have been conducted via an extensive range of scenarios, reflecting the real SDMB system configuration. In the simulation, streaming services are carried by the MTCHs and first mapped onto the FACHs via the OEA algorithm. The FACHs are then mapped to the S-CCPCHs via the classical Best-Fit bin-packing algorithm.

As mentioned, the types of service supported within the SDMB system are streaming and download, which correspond to UMTS QoS class (streaming, background). Download service allocation can be handled in either a non-priority-based or priority-based way; a general approach performing priority-based download service dimensioning within MBMS was investigated in [105]. Priority handling for download service allocation is achieved by selecting a Transport Format Combination for which high priority data is mapped with a "high bit rate" TF, at the same time allowing lower priority data to be mapped with a "low bit rate" (could be zero bit rate) TF. This priority-based download service allocation can be applied in future research, as described in the following section. In this chapter, a non-priority-based download service allocation scheme has been applied in the simulations.
The pre-requisite of the simulation is to estimate the streaming service characterisation. The characterisation for each streaming service can be calculated from the 3-tuple of the streaming service: (bit rate, system arrival rate, service duration). The evaluation results of service characterisation and load estimation for typical scenarios is given in Appendix A, and this will be used as the input parameters for the estimation of required MTCHs and FACHs.

In order to evaluate the performance of a 2-level channel multiplexing, simulation has been carried out for a wide range of different traffic mixes and physical channel capacities. The traffic mixes herein refer to the capacity allocated (reserved for) to each type of services (streaming, download) assuming implicitly a fixed boundary for the capacity.

The following typical scenarios are illustrated in this chapter, for different traffic mixes, in which we take the scenario with 3 S-CCPCHs of 384 kbps each, with a traffic mix of:

- 80% for streaming and 20% for download
- 50% for streaming and 50% for download
- 20% for streaming and 80% for download

For different S-CCPCH, traffic mix of 80% streaming-20% download, 50% streaming-50% download are selected, the following scenarios are considered:

- 3 S-CCPCHs of 384 kbps each
- 3 S-CCPCHs of 128kbps each,
- 1 S-CCPCHs of 384 kbps each
- 1 S-CCPCH of 384 kbps and 3 S-CCPCHs of 128kbps each

For the above scenarios, the MBMS streaming services (MTCHs) are first mapped to the FACHs via the optimum estimation algorithm and power-oriented adaptation algorithm. The derived FACHs are then mapped to the S-CCPCHs via the classical Best-Fit bin-packing algorithm. A non-priority-based download service allocation scheme has been used to assign the download services to the residual capacity of S-CCPCHs according to a specific traffic mix.

The simulation scenarios demonstrated in this chapter are summarised in Table 4-1.
Table 4-1. Simulation scenario reference table

<table>
<thead>
<tr>
<th>Traffic mix</th>
<th>80% streaming - 20% download</th>
<th>50% streaming - 50% download</th>
<th>20% streaming - 80% download</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-CCPCH configuration(kbps)</td>
<td>Scenario 1A</td>
<td>Scenario 1B</td>
<td>Scenario 1C</td>
</tr>
<tr>
<td>3 x 384</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 x 128</td>
<td>Scenario 2A</td>
<td>Scenario 2B</td>
<td>Scenario 2C</td>
</tr>
<tr>
<td>1 x 384</td>
<td>Scenario 3A</td>
<td>Scenario 3B</td>
<td>Scenario 3C</td>
</tr>
<tr>
<td>1 x 384 + 3 x 128</td>
<td>Scenario 4A</td>
<td>Scenario 4B</td>
<td>Scenario 4C</td>
</tr>
</tbody>
</table>

It is noted that, in our simulation, the traffic mix remains fixed during each radio bearer reconfiguration, namely the entire duration of the multimedia session transmission. Radio bearer reconfiguration is initiated by AC based on power/load constraints, and is not in the scope of this chapter. In the event that the present traffic does not perfectly match the FACHs dimension, the packet scheduler, which is intermingled with RRA, is responsible for shaping the traffic dynamics according to their QoS requirements (i.e. guaranteed data rate and delay requirements, e.g. maximum delay/jitter allowable).

4.5.2 Evaluation results

Five aspects of performance evaluation are presented in this section. Section 4.5.2.1 shows the channel multiplexing configuration derived for 2-level channel multiplexing under different physical channel scenarios and different traffic mixes. Section 4.5.2.2 presents the comparison between single-level and 2-level channel multiplexing under the same physical channel capacity and different traffic mixes. Section 4.5.2.3 presents the comparison between single-level and 2-level channel multiplexing under the same traffic mix and different physical channel capacities. Section 4.5.2.4 analyses the impact of 2-level channel multiplexing on the size of TFCS. Finally, Section 4.5.2.5 shows the performance improvement of adapting the POA algorithm in the developed 2-level channel multiplexing for different scenarios.

4.5.2.1 Mapping derivation of 2-level channel multiplexing under different scenarios and traffic mixes.

Three typical scenarios are presented in this section:

---

The channel mapping derivation of more scenarios can be found in Appendix B.
- Scenario 1A: S-CCPCH of 3x384kbps with traffic mix of 80%-20%
- Scenario 1B: S-CCPCH of 3x384kbps with traffic mix of 50%-50%
- Scenario 2B: S-CCPCH of 3x128kbps with traffic mix of 50%-50%

Table 4-2. Derived channel mapping configuration: Scenario 1A (kbps)

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs</td>
<td>256x1;128x1</td>
<td>256x1</td>
<td>256x1</td>
</tr>
<tr>
<td>Download FACHs</td>
<td>-</td>
<td>-</td>
<td>128x1</td>
</tr>
<tr>
<td>Streaming + download FACHs</td>
<td>-</td>
<td>128x1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4(^{15})</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>256</td>
<td>128</td>
<td>256</td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs</td>
<td>256x1</td>
<td>64x1;32x2</td>
<td>256x1</td>
<td>32x1</td>
<td>128x1;64x2</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>64x1;32x1</td>
<td>-</td>
<td>64x2</td>
</tr>
</tbody>
</table>

Figure 4-8. Channel mapping structure utilising 2-stage bin-packing algorithm combination for scenario 1A.

Table 4-2 shows the channel mapping configuration derived for Scenario 1A. The overall channel mapping configuration for this scenario is illustrated in Figure 4-8. After the mapping for the streaming services (MTCHs) has been performed, there is no residual capacity on streaming FACHs, whilst there exist residual capacities on S-CCPCHs, which are assigned to download FACHs for carrying download services. For instance, the 128 kbps residual capacity on S-CCPCH 3 is allocated to one download FACH 6, which in turn is assigned equally in capacity to 2 MTCHs of 64 kbps each, so as to accommodate two download applications.

\(^{15}\) FACHs with shadow effect denote those FACHs carrying both streaming and download MTCHs.
As mentioned above, priority handling can be applied for the download service allocation. From this point of view, the “hot download” service will be given priority treatment due to its stringent delay requirement, and thus a higher-rate MTCH will be allocated to the higher priority download service. For example in Table 4-2, given a residual capacity of 96 kbps in S-CCPCH 2 after allocating streaming, this capacity can be assigned according to the priority of the download MTCHs, i.e. “hot download” service is assigned to a MTCH of 64 kbps, whilst “cold download” service is assigned to a MTCH of 32 kbps. The objective of this priority-based download service allocation is to perform better utilisation of physical channels for QoS-differentiated download services.

Table 4-3. Derived channel mapping configuration: Scenario 1B (kbps)

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs</td>
<td>256x1, 64x1</td>
<td>256x1</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs</td>
<td>64x1</td>
<td>128x1</td>
<td>256x1, 128x1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>256</td>
<td>64</td>
<td>64</td>
<td>256</td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs</td>
<td>256x1</td>
<td>32x2</td>
<td>-</td>
<td>128x1</td>
<td>64x2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs</td>
<td>-</td>
<td>-</td>
<td>32x2</td>
<td>-</td>
<td>64x2</td>
<td>128x2</td>
<td>64x2</td>
</tr>
</tbody>
</table>

Table 4-4. Derived channel mapping configuration: Scenario 2B (kbps)

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs</td>
<td>128x1</td>
<td>64x1</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs</td>
<td>-</td>
<td>64x1</td>
<td>128x1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>128</td>
<td>64</td>
<td>64</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs</td>
<td>128x1</td>
<td>32x2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs</td>
<td>-</td>
<td>-</td>
<td>32x2</td>
<td>64x2</td>
</tr>
</tbody>
</table>
Chapter 4. Novel Radio Resource Allocation Scheme

The channel mapping configuration derived for Scenario 1B is shown in Table 4-3. It can be noticed that all the FACHs are fully occupied by both streaming MTCHs (on FACHs 1, 2, 4) and download MTCHs (on FACHs 3, 5, 6, 7); hence zero residual capacity in FACHs is achieved. The total transmitted streaming capacity (256kbps + 64kbps + 256kbps = 576kbps) is equal to the total transmitted download capacity (64kbps + 128kbps + 256kbps + 128kbps = 576kbps), which corresponds to the specific 50% streaming - 50% download traffic mix. Note that non-priority-based download service allocation scheme is applied for both Scenarios 1B and 2B.

Table 4-4 shows the channel mapping configuration derived for Scenario 2B. This is a simpler mapping scenario compared to Table 4-2 and Table 4-3, whereby the same total rate is allocated to the download and streaming services, given a traffic mix of 50%-50%. The results show that all the FACHs are fully occupied by streaming MTCHs (on FACHs 1, 2) and download MTCHs (on FACHs 3, 4) with zero residual capacity carried on the S-CCPCH physical channel.

4.5.2.2 Performance comparison for single-level and 2-level channel multiplexing under different traffic mixes.

In this section, the following scenarios are evaluated and compared:

- Scenario 1A: S-CCPCH of 3x384kbps with traffic mix of 80%-20%;
- Scenario 1B: S-CCPCH of 3x384kbps with traffic mix of 50%-50%;
- Scenario 1C: S-CCPCH of 3x384kbps with traffic mix of 20%-80%.

<table>
<thead>
<tr>
<th>Traffic Mix</th>
<th>Required Number of FACHs</th>
</tr>
</thead>
<tbody>
<tr>
<td>80% streaming 20% download</td>
<td>12</td>
</tr>
<tr>
<td>50% streaming 50% download</td>
<td>14</td>
</tr>
<tr>
<td>20% streaming 80% download</td>
<td>16</td>
</tr>
</tbody>
</table>

Figure 4-9. Required number of FACHs for single-level and 2-level channel multiplexing under different traffic mixes.
Figure 4-9 shows the comparison of the required number of both streaming and download FACHs between the single-level and 2-level channel multiplexing. As seen from the bar chart, with the same physical channel capacity, the total numbers of required FACHs are greatly reduced by employing 2-level channel multiplexing. For example, the total required number of FACHs is reduced from 14 to 7 in Scenario 1B (i.e. traffic mix of 50%-50%), which means 7 extra FACHs can be saved and further utilised, thereby improving the overall utilisation of FACHs. By having fewer FACHs, this in effect means that there are less FACHs to be multiplexed to a single S-CCPCH, and therefore the size of the TFCS is reduced.

![Figure 4-9. Comparison of required FACHs between single-level and 2-level channel multiplexing.](image)

**Figure 4-9.** Comparison of required FACHs between single-level and 2-level channel multiplexing.

One important task for RRM is to increase the transmission capacity within the limited physical channel capacity. From this point of view, 2-level channel multiplexing enjoys significant benefits over its single-level counterpart. In Figure 4-10, by employing single-level channel multiplexing, the total MTCH transmission capacity (streaming and download) is 448kbps and 480kbps for Scenario 1A and 1B, respectively. However, when the proposed 2-level channel multiplexing is applied, the total MTCH transmission capacity over the same lower layer FACHs/S-CCPCHs is significantly increased to 1152kbps and 1088kbps, respectively, which in effect leads to better radio resource utilisation.

![Figure 4-10. Total transmission capacity for single-level and 2-level channel multiplexing under different traffic mixes.](image)

**Figure 4-10.** Total transmission capacity for single-level and 2-level channel multiplexing under different traffic mixes.
Chapter 4. Novel Radio Resource Allocation Scheme

Physical channel utilisation for S-CCPCH 3x384 kbps under different traffic mixes

![Figure 4-11](image)

Figure 4-11. Transmission capacity utilisation for single-level and 2-level channel multiplexing under different traffic mixes.

The proposed 2-level channel multiplexing achieves the maximum utilisation on FACHs by assigning more MTCHs to it. As seen from Figure 4-11, there is a fairly large amount of residual capacity, which could not be further utilised and appears as a waste of capacity on the FACHs/S-CCPCHs, remaining in FACHs in single-level channel multiplexing for both scenarios. As shown in the bar chart, these residual FACH capacities for Scenario 1A and 1B are: 704kbps and 640kbps, respectively, which correspond to 61.1% and 55.6% of total FACH capacities. However, by applying 2-level channel multiplexing, these residual capacities are better utilised and zero residual FACH capacities have been achieved for all scenarios.

4.5.2.3 Performance comparison for single-level and 2-level channel multiplexing under different S-CCPCH capacities.

The following scenarios are investigated in this section:

- Scenario 1B: S-CCPCH of 3x384kbps with traffic mix of 50%-50%;
- Scenario 2B: S-CCPCH of 3x128kbps with traffic mix of 50%-50%;
- Scenario 3B: S-CCPCH of 1x384kbps with traffic mix of 50%-50%;
- Scenario 4B: S-CCPCH of 1x384+3x128kbps with traffic mix of 50%-50%. 

Figure 4-12. Required number of FACHs for single-level and 2-level channel multiplexing under different S-CCPCH capacities.

Figure 4-13. Total transmission capacity for single-level and 2-level channel multiplexing under different S-CCPCH capacities.

Figure 4-12 shows that, by employing 2-level channel multiplexing, the total numbers of required FACHs under different physical channel capacities are reduced for all the physical channel capacities. The 2-level channel multiplexing achieves the maximum utilisation of FACHs and thereby the required number of FACHs is greatly reduced irrespective of physical channel capacity, e.g. the required number of FACHs is reduced from 14 to 7 in Scenario 1B and from 12 to 5 in Scenario 4B. As shown in Figure 4-13, by employing 2-level channel multiplexing, the total MTCH transmission capacity is increased for all physical channel scenarios. When using
single-level channel multiplexing, the total transmission capacity (streaming and download) is 480kbps and 192kbps respectively for Scenario 1B and 2B. However, when the proposed 2-level channel multiplexing is applied, the total transmission capacity over the same lower layer FACHs/S-CCPCHs is significantly increased to 1152kbps and 384kbps, respectively, which in effect leads to better radio resource utilisation. Therefore, the proposed 2-level channel multiplexing yields a significant performance improvement on the required number of FACHs and total MTCH transmission capacity under the same traffic mix with various physical channel capacities.

![Figure 4-14. Transmission capacity utilisation for single-level and 2-level channel multiplexing under different S-CCPCH capacities.](image)

Figure 4-14 illustrates that, by utilising the proposed 2-level channel multiplexing for various physical layer capacities, the wasted capacities of FACHs are better reutilised. As shown in the bar chart, these residual FACH capacities for Scenario 1B and 3B are: 640kbps and 256kbps, respectively, which correspond to 66.7% and 55.6% of total FACH capacities. However, by applying 2-level channel multiplexing, these residual capacities are further fully utilised and zero residual FACHs capacities have been achieved for all scenarios.

4.5.2.4 **TFCS size analysis**

The performance enhancement for the 2-level channel multiplexing with respect to the reduction of TFCS size is discussed in this section. The size of the TFCS is defined as the number of possible TFC candidates it contains, and it is initialized by the RRA module at the beginning of each session for each S-CCPCH. During each session, the exact TFC will be chosen from the predefined TFCS set by the packet scheduler module, which performs the short-term resource
allocation task on a TTI-scale. The reduction in the size of TFCS leads to better radio resource utilisation on the physical channel.

Table 4-5 and Table 4-6 show the TFCS sizes under different scenarios by using single-level and 2-level channel multiplexing, respectively. As shown in the tables, by using the 2-level channel multiplexing scheme, the TFCS sizes have been reduced in almost all the scenarios. For example, in the scenario of S-CCPCH configuration at 3x384kbps with 80%-20% traffic mix, the size of TFCS is 60 for S-CCPCH 1 by using single-level channel multiplexing, and it is reduced to 20 by using 2-level channel multiplexing scheme. The maximum reduction on the TFCS size is up to 95.4% for all scenarios, whilst the overall average reduction on the TFCS size for all the considered scenarios is 59.4%.

These results are also well illustrated in Figure 4-15, where the TFCS sizes are significantly reduced for all the S-CCPCHs irrespective of their capacities and traffic mixes. This indicates that a higher number of multiplexed logical channels at transport layer will lead to fewer required transport channels and thereby greater reductions on the size of TFCS.

Table 4-5. TFCS size for respective S-CCPCHs in single-level channel multiplexing

<table>
<thead>
<tr>
<th>S-CCPCH configuration (kbps)</th>
<th>3 x 384</th>
<th>3 x 128</th>
<th>1 x 384</th>
<th>1 x 384 + 3 x 128</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic mix</td>
<td>TFCS</td>
<td>TFCS</td>
<td>TFCS</td>
<td>TFCS</td>
</tr>
<tr>
<td>80%-20%</td>
<td>384</td>
<td>384</td>
<td>384</td>
<td>122</td>
</tr>
<tr>
<td>50%-50%</td>
<td>384</td>
<td>384</td>
<td>384</td>
<td>122</td>
</tr>
<tr>
<td>20%-80%</td>
<td>384</td>
<td>384</td>
<td>384</td>
<td>122</td>
</tr>
</tbody>
</table>

Table 4-6. TFCS size for respective S-CCPCHs in Two-level channel multiplexing

<table>
<thead>
<tr>
<th>S-CCPCH configuration (kbps)</th>
<th>3 x 384</th>
<th>3 x 128</th>
<th>1 x 384</th>
<th>1 x 384 + 3 x 128</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic mix</td>
<td>TFCS</td>
<td>TFCS</td>
<td>TFCS</td>
<td>TFCS</td>
</tr>
<tr>
<td>80%-20%</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>50%-50%</td>
<td>45</td>
<td>20</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>20%-80%</td>
<td>64</td>
<td>20</td>
<td>20</td>
<td>4</td>
</tr>
</tbody>
</table>

16 No specific rules can be conducted on the derivation of the TFCS; hereby the TFCS sizes are derived for the TFCSs that contain all the combinations of the possible transport formats of the transport channels.
4.5.2.5 Performance enhancement by adopting POA algorithm

Performance evaluation of the adopting POA algorithm in 2-level channel multiplexing scenario is investigated in this section, herein the different traffic mixes and physical channel scenarios are considered and both streaming and download services are evaluated.

The following scenarios are used to evaluate the performance of POA under the same physical channel capacity but with different traffic mixes:

- Scenario 1A: S-CCPCH of 3x384kbps with traffic mix of 80%-20%;
- Scenario 1B: S-CCPCH of 3x384kbps with traffic mix of 50%-50%;
- Scenario 1C: S-CCPCH of 3x384kbps with traffic mix of 20%-80%;

The following scenarios are used to evaluate the performance of POA under the same traffic mix but with different physical channel capacities:

- Scenario 1B: S-CCPCH of 3x384kbps with traffic mix of 50%-50%;
- Scenario 2B: S-CCPCH of 3x128kbps with traffic mix of 50%-50%;
- Scenario 3B: S-CCPCH of 1x384kbps with traffic mix of 50%-50%;
- Scenario 4B: S-CCPCH of 1x384+3x128kbps with traffic mix of 50%-50%.
Chapter 4. Novel Radio Resource Allocation Scheme

Figure 4-16. Total physical channel transmit power under different S-CCPCHs.

Figure 4-16 shows the simulation results of total transmit power under the same physical channel capacity (i.e. 3x384kbps) scenario with different traffic-mixes in Scenario 1A, 1B and 1C. Compared with the worst-case channel multiplexing scenario, the POA-based scheme achieves the optimum power consumption under same physical channel capacity but with different traffic mix scenarios. As seen from Figure 4-16, compared with the worst-case scenario, the total transmit power is greatly reduced when applying the POA algorithm, numerically, the transmit power is reduced by 14.8%, 8.6% and 9.9% for Scenario 1A, 1B and 1C, respectively.

Table 4-7. Mapping derived for scenario 1A: (a) POA-based scenario; (b) Worst-case scenario

<table>
<thead>
<tr>
<th>S-CCPCH Bit rate (kbps)</th>
<th>384</th>
<th>384</th>
<th>384</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>256x1</td>
<td>256x1</td>
<td>256x1</td>
</tr>
<tr>
<td></td>
<td>128x1</td>
<td>128x1</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>128x1</td>
</tr>
<tr>
<td>Streaming + download FACHs (kbps)</td>
<td>-</td>
<td>128x1</td>
<td>-</td>
</tr>
<tr>
<td>FACH Bit rate (kbps)</td>
<td>256</td>
<td>128</td>
<td>256</td>
</tr>
<tr>
<td></td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>256x1</td>
<td>64x1</td>
<td>256x1</td>
</tr>
<tr>
<td></td>
<td>32x2</td>
<td>32x1</td>
<td>128x1</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>64x1</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>64x2</td>
</tr>
</tbody>
</table>

As seen from Table 4-7, the worst case transmit power occurs when the services with different bit rates and QoS demand (i.e. streaming or download) are carried by the same S-CCPCH. For example, in the worst case channel multiplexing scenario, there are three FACHs carrying both streaming and download services with differences in bit rates (e.g. 256kbps vs. 64kbps). On the other hand, POA-based mapping can group services with similar mean rate and QoS rank into one S-CCPCH. Almost all FACHs in this scenario carry a unique type of service, only one 128kbps FACH carries both streaming and download services (i.e. 32kbps streaming & 32/64 kbps download) and in this case the POA algorithm fills the residual capacity left by streaming with download services with similar rate.

![Figure 4-17. Total physical channel transmit power under different traffic mixes.](image-url)

Figure 4-17 shows the simulation results of total transmit power under different physical channel capacities but with the same traffic mix in Scenario 1B, 2B, 3B and 4B. It shows that the power conserving ratios for physical channel capacity are 8.6%, 7.6%, 6.0%, and 13.9% for these scenarios, respectively.
Table 4-8. Mapping derived for scenario 2B: (a) POA-based scenario; (b) Worst-case scenario.

<table>
<thead>
<tr>
<th></th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S-CCPCH Bit rate(kbps)</strong></td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td><strong>Streaming FACHs(kbps)</strong></td>
<td>128x1</td>
<td>64x1</td>
<td>-</td>
</tr>
<tr>
<td><strong>Download FACHs (kbps)</strong></td>
<td>-</td>
<td>64x1</td>
<td>128x1</td>
</tr>
<tr>
<td><strong>Streaming + download FACHs (kbps)</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>FACH Bit rate(kbps)</strong></td>
<td>128</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td><strong>Streaming MTCHs(kbps)</strong></td>
<td>128x1</td>
<td>32x2</td>
<td>-</td>
</tr>
<tr>
<td><strong>Download MTCHs (kbps)</strong></td>
<td>-</td>
<td>-</td>
<td>32x2</td>
</tr>
</tbody>
</table>

(a) POA-based scenario

(b) Worst-case scenario

Table 4-8 shows the mapping derived for the POA-based and worst-case non-POA-based methods in the scenario 2B. Although this is a much simpler version of channel multiplexing, the POA algorithm still brings together services with similar mean rate and QoS rank. Unlike the mixed services carried by FACHs in the worst case scenario, all FACHs are assigned with unique type of service by using the POA algorithm.

4.6 Summary

The SDMB system implements a satellite based BC/MC layer over the unicast terrestrial 3G mobile telecommunication system infrastructure aimed at the efficient delivery of the interactive
MBMS services to a widely dispersed user group. The adoption of UMTS air interface in the SDMB system achieves maximum commonalities with terrestrial mobile network radio interfaces, and this leads to significant reduction on additional complexities for both mobile terminals and access networks.

After analysing the problems in the existing RRA strategy, in order to seek improvements in the system transmission capacity and optimize the transmit power usage, we have proposed a novel RRA strategy for the SDMB system. A new 2-level channel multiplexing scheme, which performs channel multiplexing at both transport and physical channels, is the essence of this new RRA strategy. After a feasibility study of this new approach, the overview of the proposed scheme was highlighted. Then the procedure of this new scheme was described step-by-step. Notably, two new optimization algorithms, namely optimum estimation algorithm and power-oriented adaptation algorithm, were developed to optimize the proposed 2-level channel multiplexing scheme. The purpose of the optimum estimation algorithm is to minimize the required numbers of FACHs and MTCHs, and reduce the residual capacities on the transport channels. At the same time, the power-oriented adaptation algorithm effectively minimizes the total transmission power.

Performance evaluation of the proposed approach was carried out for a wide range of scenarios comprising different traffic mixes and physical channel capacities. The results show that, comparing with the existing single-level channel multiplexing scheme, the new approach requires fewer transport channels, has better usage of the residual capacities on the transport channels, can improve the total transmission capacities on the logical channels, and reduces the TFCS sizes. The performance improvement on the power consumption by using the power-oriented optimization process has also been demonstrated. The 2-level channel multiplexing scheme features higher computation complexity than its single-level counterpart. However, given relatively small number of MTCHs, the performance gain by using the proposed scheme will far outweigh the increase in the complexity.
Chapter 5

5 Proportional Differentiation Based Packet Scheduling

5.1 Introduction

5.1.1 Overview

The SDMB system is unidirectional and provides point-to-multipoint services. There is no power control mechanism and no channel-state information available. The design of a packet scheduling scheme, which is a key element of RRM functionalities implemented at the SDMB access layer, is thus challenging. The purpose of this chapter is to develop more efficient packet scheduling algorithms based on the proportional differentiation concept, and evaluate the performance of the novel algorithms in the SDMB RRM framework.

Much research [106]-[108] has been devoted to develop efficient scheduling schemes for effective QoS provisioning, both in wired and wireless systems. For instance, one of the interesting topics is delay differentiated scheduling, where waiting time and queuing delay are considered in prioritizing and scheduling, such as in waiting time priority (WTP) and proportional delay differentiation (PDD) schemes proposed in [89] for terrestrial networks. In [106], scheduling algorithms for general wireless systems have been discussed. Nevertheless, as already stated above, those schemes cannot be directly adopted in the SDMB system due to its specific characteristics. Previous studies [13] have systematically addressed the RRM problems in the SDMB system via traditional packet scheduling schemes, namely MLPQ and WFQ. Comprehensive simulations of RRM frameworks over S-UMTS networks have been performed within IST SATIN [109] [110] [111] and MAESTRO projects [35]. However, both of these feature major weaknesses in the provisioning of QoS differentiated multimedia services with respect to efficiency and fairness.

To overcome these inherent deficiencies in MLPQ and WFQ, the challenge to the design of packet scheduling algorithms is to optimally utilise resources and efficiently schedule traffic whilst guaranteeing the prescribed QoS demands for each user. In this chapter, we proposed two
novel packet scheduling schemes, namely buffer-length related queue (BLRQ), delay differentiation queue (DDQ), which are more applicable for multimedia contents delivery in SDMB.

This chapter is organised as follows. In Section 5.2, BLRQ scheduling is presented. We then proceed in section 5.3 with the description of the DDQ algorithm. In Section 5.4, we present the performance evaluation for both BLRQ and DDQ in comparison to existing packet scheduling algorithms. We summarise our proposal and conclude this chapter in Section 5.5.

5.1.2 Design of packet scheduling algorithm

The role of the packet scheduler in SDMB is not that dominant in determining the system throughput as in the T-UMTS case. Nevertheless, the scheduler is still responsible for two important tasks that are executed with a period equal to the TTI of the radio bearers [112].

- Time-multiplexing of flows with different QoS requirements into fixed physical channels, in a way that can satisfy these requirements.

- Adjusting the transmit power of the physical channel carrying the data flows on the basis of the required reception quality of the service (in terms of the target BLER) under the constraint that the total available power for all the physical channels within a beam is fixed.

The packet scheduling strategy can in general be described into two steps, as in Figure 5-1. These two steps effectively constitute the scheduling discipline of the packet scheduler; they differ from each other and define the capability in each to integrate the service QoS requirements with an efficient system resource utilisation. The details of the packet scheduler are described below.

Service prioritization

In SDMB, each service is mapped one-to-one onto an MTCH, a logical channel, which is then mapped onto the FACH transport channel. At the physical level, the S-CCPCHs can carry one or more FACH(s). The incoming service requests are ordered according to a priority criterion. In selecting the respective criteria, the service attributes are considered, which are normally mapped onto the traffic handling priorities, as defined by the UMTS QoS classes. Note that the prioritization can be more or less dynamic; in a more dynamic prioritization, the relative priority of the different channels may change in each resource allocation interval (this is normally one TTI), depending for example on the maximum delay tolerated by a service or the number of packets buffered, as compared to a more static approach.
Chapter 5. Proportional Differentiation Based Packet Scheduling

- **Service prioritisation**
  
  Prioritizing each MTCH based on its instant priority

  The MTCHs are sorted and their IDs are queued according to the respective assigned priority

- **Resource allocation**
  
  Get the MTCH ID from MTCH ID queue according to their priority

  Select TBS for the selected MTCHs

  - **Check power constraints satisfied?**
    - False
      - MTCH i is not allocated
      - Mark MTCH ID
    - True
      - MTCH i is allocated
      - Reduce TFCS

  - **All MTCHs in MTCH ID queue have been checked?**
    - No
    - Exit and go for next TTI, TFC is decided
    - Yes

  Figure 5-1: Packet scheduling procedure.

As stated above, the service prioritization can be performed more dynamically. Rather than performing service prioritization in a static way, a more efficient packet scheduling algorithm performs service prioritization dynamically depending on the waiting time/queuing delay experienced by packets in each MTCH/FACH at the beginning of each TTI. Resource is then allocated to respective physical channels (i.e. S-CCPCH) according to the priority assigned to each MTCH/FACH flow as long as their power and load condition can be satisfied.

Within this framework, novel scheduling schemes are proposed. Firstly, BLRQ is introduced which introduces the buffer status into the scheduling. Furthermore, DDQ scheduling scheme is also introduced. DDQ is developed based on the Hybrid Proportional Delay (HPD) scheduling scheme [89], which is widely used in differentiated services networks. It assumes that there are QoS ratios between different QoS priority classes. It is worth noting that the packet scheduling algorithms retain the assumption of one-to-one mapping from logical channels (MTCHs) to transport channels (FACHs). The various proposed packet scheduling algorithms are evaluated in the following sections.
Chapter 5. Proportional Differentiation Based Packet Scheduling

Resource allocation

Once all the services to be transmitted are prioritized, the next step is the resource allocation which consists of bit rate and transmit power assignments within the specific resource allocation interval (i.e., TTI). The data rate assignment consists of the selection of the TFCs, which directly determine the per FACH transport block size, namely how much data from each transport channel mapped to the physical channel will be forwarded to the physical layer in each TTI. For each active physical channel (S-CCPCH), the exact TFC is selected from the TFCS, which is passed during the admission of a new service and its mapping onto a specific bearer. This TFC selection step is of paramount importance since the capacity allocated to each service is strongly related to the QoS perceived by the end users, and therefore, the selection of the TFC has to take into consideration constraints in terms of service requirements (e.g., minimum guaranteed rate, maximum tolerated delay) as well as system-level constraints (system load, transmit power per beam).

As shown in Figure 5-1, for all S-CCPCHs, the packet scheduler serves the FACHs according to their instantaneous priorities, which are dynamically calculated from priority criteria in the particular TTI. The FACH queues with higher priorities will be served ahead of the lower priority FACH queues in non-preemptive order.

For each FACH, the packet scheduler scans the TFCS of its corresponding physical channel to find all the different TBS sizes and then seeks to allocate the maximum TBS size to the selected FACH based on data queued at its buffer.

The packet scheduler then checks the power requirement on the basis of the BLER curve of the active service flow. If power allocation satisfies the power and load constraints, the scheduler will allocate this FACH and derive a reduced TFCS; otherwise, this service is not scheduled.

As for the power allocation, the transmit power setting for the S-CCPCH is based on the required reception quality of the active service flows mapped to the S-CCPCH, which in our case is defined in terms of the most demanding target FER among these service flows. The calculated power is only allocated as long as it is within the constraint of the total available power for all the physical channels, which is fixed within a beam. In the resource allocation step, the S-CCPCH TFC selection and power allocation happen in parallel.

5.2 Buffer length related queuing

When a finite length buffer size is assumed, it is essential to maintain a reasonable buffer status to prevent excess packet loss due to buffer overflow. In order to take account of buffer status during the packet scheduling procedure, Buffer-Length Related Queuing (BLRQ) scheme is proposed,
Chapter 5. Proportional Differentiation Based Packet Scheduling

aimed at balancing all the traffic flows with regard to their respective queue lengths. This approach is designed to reduce the probability of packet loss due to buffer overflow in the case of finite RLC buffer size. Since BLRQ can be regarded as a modified form of MLPQ, it is still a priority scheduling scheme, in which the packets in higher priority queues will be processed first. For those queues having same priority class, the queue with the longest packet queue in its buffer will be served first, instead of adopting the traditional round-robin approach.

BLRQ is essentially an enhanced version of MLPQ; the scheduler operates exactly the same for traffic flows featuring different QoS rank. The difference between them is that BLRQ will provide service differentiation between the traffic flows within the same QoS rank.

In each TTI, the packet scheduler will scan all FACH queues and schedule packets accordingly. Firstly, the queue with highest QoS rank is served ahead of those with lower QoS rank. And then, once there is more than one queue within each QoS rank, the FACH queue with the longest queue length in its buffer will be served first. The mathematical presentation can be expressed as:

$$FACH\_selected = \text{MAX}\{queue\_length(i)\}$$

(5.1)

where $FACH\_selected$ is the ID of the FACH queue with the longest queue length; $queue\_length(i)$ is the queue length for the $i^{th}$ FACH within the same QoS rank at current TTI slot.

The scheduler will allocate resource according to both the QoS rank and the buffer status of each FACH queue on a TTI-scale. In this respect, the differentiation is made available for those FACHs with same QoS rank but featuring discrepancies in their relative queue length, which can arise from the asymmetry of the network resource allocation (i.e. different number of FACHs mapped onto respective S-CCPCH), different traffic mixes under certain scenarios, traffic dynamics for each incoming traffic flow as well as the propagation and interference variations in the satellite system.

By considering the buffer status of individual queue, the BLRQ can effectively improve the queuing performance in terms of both buffer occupancy and packet drop rate. However, the BLRQ can only differentiate queues with same QoS rank according to the buffer length, there are other performance metrics should also be considered.
5.3 Delay differentiation queuing

5.3.1 Overview

In order to achieve better packet scheduling performance in terms of both efficiency and fairness, inherited from the proportional delay differentiation (PDD) scheme in the context of differentiated service networks. A DDQ scheme is proposed, offering improved performance in delay, jitter, and channel utilisation. DDQ was proposed for the delay differentiation services in a satellite environment, assuming there are QoS ratios between different traffic priority classes. For each resource allocation interval (e.g., TTI), the serving indices are obtained based on the average waiting delay for all packets currently in the queue, the average queuing delay for all the packets having left the queue, the packet arrival rate and QoS ratio. In this scheme, the instantaneous queuing delay is effectively considered for queues with the same QoS rank. Compared with WFQ and MLPQ, DDQ offers improved performance in delay, jitter, and channel utilisation. However, DDQ experiences unbalanced performance among multiple QoS attributes, namely the gain achieved in one performance attribute leads to the performance degradation in other attributes. Furthermore, multimedia services feature differentiated delay constraints and applies the delay constraints for differentiated services in an equal way may lead to poor QoS guarantee for high priority queues. Therefore the delay profile has to be considered against the respective delay constraints (i.e., maximum acceptable delay) specified by the class of service. Finally, rather than scheduling competing flows in a static manner, to provide more flexible QoS provisioning and maintain optimal resource utilisation, it is highly desired that the scheduler is capable of choosing the best scheduling policy according to diverse QoS preferences of the services and instantaneous performance dynamics.

5.3.2 Description of DDQ scheduling algorithm

The incoming service requests are first ordered according to a priority criterion. In order to select the respective criteria, the service attributes are considered, which are normally mapped onto the traffic handling priorities, as defined by the UMTS QoS classes. It is noted that the prioritization can be more or less dynamic. In a more dynamic prioritization, the relative priorities of channels may change in each resource allocation interval (i.e., one TTI), depending for example on the QoS rank or average queuing delay status of competing flows.

DDQ performs service prioritization dynamically depending on the QoS and the waiting time/queuing delay experienced by packets in each FACH. It assumes that each MBMS session
maintains a separate FACH queue and that there are QoS ratios between different QoS priority classes. In each TTI, the serving indices are calculated for each queue. These serving indices are obtained based on the average waiting delay for all the packets currently in the queue, the average queuing delay for all the packets that have left the queue prior to this TTI, and the QoS priority ratio index.

The QoS factor $\alpha$ indicates the QoS priority of the MBMS services. In the SDMB system, there are three different service classes: streaming, hot download and cold download.

The fairness factor $\delta$ indicates the fairness among the MBMS services, and is expressed by the average waiting delay for all the packets currently in the queue and the average queuing delay for all the packets that have left the queue before the current TTI.

Mathematical formulation of the DDQ can be expressed as follows.

Let $\delta_i(n)$ be the average queuing/waiting delay at current time slot $n$ for each queue $i$. This measure describes the delay status of all packets passing through the respective queue, including both the packets which are currently in the queue and those packets which have already left the queue (been served). Delay index will be calculated for each queue $i$ in each TTI as:

$$\delta_i(n) = \frac{\sum_{j=1}^{N_q} W_{i,j}^q(n) + \sum_{j=0}^{N_d} W_{i,j}^d(n)}{(N_q + N_d)}$$

(5.2)

where $\delta_i(n)$ is the fairness factor for queue $i$, $N_q$ is the number of packets that are currently in the queue, $W_{i,j}^q$ is the waiting delay for packet $j$ currently in the queue $i$, $N_d$ the number of packets that have left the queue before the current TTI, $W_{i,j}^d$ is queuing delay for packet $j$, which has left the queue $i$ before the current TTI.

Let $\alpha_i$ be the QoS priority factor for the service flow at the FACH queue $i$; the priority for queue $i$ in TTI $n$ can be defined as:

$$P_i(n) = \alpha_i \times \delta_i(n)$$

(5.3)

where $\alpha_i$ is the QoS class factor, which is essentially a time-independent parameter designated, for each queue $i$.

Consequently, the serving orders are calculated and assigned to each FACH by (5.3) at the beginning of each TTI.
With the above approach of dynamic service prioritization in mind, the dynamically changing priorities of MTCHs indicate the serving order of FACHs and S-CCPCHs for each TTI by the scheduler. It must also be noted that it is generally assumed that only services with similar characteristics and QoS requirements are multiplexed together on the same transport channel.

Once the instantaneous priorities of each MTCH are determined at the beginning of each TTI, the scheduler selects TF(C) (i.e. allocating resource) for each S-CCPCH based on the assigned priorities of MTCHs mapped onto it. This procedure repeats itself in each TTI-scale until all the S-CCPCHs are assigned.

5.4 Performance evaluation of scheduling algorithms

5.4.1 Performance evaluation of BLRQ

The scenario considered for the radio bearer mapping is given in Table 5-1.

Table 5-1. Radio bearer mapping for the evaluation of BLRQ.

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming MTCH/FACHs (kbps)</td>
<td>1x128</td>
<td>-</td>
<td>1x384</td>
</tr>
<tr>
<td>Download MTCH/FACHs (kbps)</td>
<td>2x128</td>
<td>1x256</td>
<td>-</td>
</tr>
<tr>
<td>SF/TTI (ms)</td>
<td>8/80</td>
<td>8/80</td>
<td>8/80</td>
</tr>
</tbody>
</table>

5.4.1.1 Queuing delay and delay variation evaluation

Table 5-2. Mean queuing delay for MLPQ and BLRQ (seconds)

<table>
<thead>
<tr>
<th>MTCH/FACH id</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPQ</td>
<td>0.12</td>
<td>0.37</td>
<td>0.74</td>
<td>0.18</td>
<td>2.51</td>
<td>0.08</td>
</tr>
<tr>
<td>BLRQ</td>
<td>0.14</td>
<td>0.19</td>
<td>0.37</td>
<td>0.10</td>
<td>0.35</td>
<td>0.10</td>
</tr>
</tbody>
</table>
As shown in Table 5-2, the mean queuing delay experienced by download service (FACH 2-5) is dramatically reduced, at the same time, no significant degradation is introduced for delay-sensitive streaming service (FACHs 1 and 6).

5.4.1.2 Buffer status analysis

As seen from Figure 5-2, the average buffer length is greatly reduced by using the BLRQ-based scheduling. Furthermore, all queues serving download class traffic (FACH 2-5) have similar
buffer length, which essentially affect the impact of queue length on differentiating the services featuring the same QoS rank. It is noted that there is slight difference between the Max and Mean queue length on the streaming queue length (i.e. FACHs 1 and 6), it is therefore proven that for the majority of the time the streaming service will be given more priority than download service so that their queue lengths remain stable at a mean value approaching to their maximum values...

5.4.2 Performance evaluation of DDQ

Simulation was carried out using the system-level SDMB platform (in Appendix D) for a wide range of simulation scenarios by using the software package ns2 [113], which involved different traffic mixes and different physical channel capacities. The proposed DDQ packet scheduling algorithm, compared with the MLPQ-based scheduling that are previously used in the baseline SDMB scenario.

The performance of our proposed strategy and that of previous studies are compared via performance metrics, such as delay, delay variation, packet queue length, system utilisation and fairness.

In order to examine the performance of the proposed scheme with respect to the service type and the data rate, an indicative simulation scenario is chosen as shown in Table 5-3.

<table>
<thead>
<tr>
<th>S-CCPCH ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-CCPCH Bit rate (kbps)</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACH (kbps)</td>
<td>256x1;64x1</td>
<td>256x1;128x1</td>
<td>-</td>
</tr>
<tr>
<td>Hot download FACH (kbps)</td>
<td>64x1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cold download FACH (kbps)</td>
<td>-</td>
<td>-</td>
<td>384x1</td>
</tr>
</tbody>
</table>

According to the QoS traffic classes defined in the SDMB system, with consideration to service differentiation and fairness, three QoS classes are assigned in the designated scenario, namely streaming, hot download and cold download. In addition, we choose different guaranteed data rates in order to examine the performance between users with different rates. As shown in Table 5-3, three S-CCPCHs are used for carrying homogeneous multimedia services with different classes:

- S-CCPCH 1: Has three FACHs. The hot download FACH (i.e. FACH 1) has a data rate at 64kbps, and the two streaming FACHs (i.e. FACH 2, 3) have data rates at 256kbps and 64kbps, respectively;
• S-CCPCH 2: Has two streaming FACHs (i.e. FACH 4, 5), whose data rates are 256kbps and 128kbps respectively;

• S-CCPCH 3: Has one download FACH (i.e. FACH 6), and its data rate is 384kbps;

Here we assume one-to-one mapping between MTCHs to FACHs, whilst multiplexing only occurs from transport channel to physical channel. Thereby, the FACHs transport channel to physical channel multiplexing scenario is specified in Table 5-3. In the simulation, the estimation of transmission power is derived using the Downlink Pole Equation elaborated for WCDMA system in [115].

The performance of DDQ and MLPQ are compared via simulation metrics, such as delay, delay variation and channel utilisation.

5.4.2.1 Analysis of delay and delay variation

As illustrated in Figure 5-3, by using DDQ-based packet scheduling algorithms, download multimedia services (i.e. FACH 1 and FACH 6) experience much less mean delay compared with MLPQ. It is noted that the significant reduction on delay of lower class services does not result in dramatic performance degradation on the higher class counterparts (i.e FACHs 2 to 5) beyond the guaranteed QoS rank. It is demonstrated that DDQ allows the download service to attain the highest degree of utilisation of the spare resources remaining for the streaming service for no significant degradation on the QoS target of the streaming users.

Figure 5-3: Packet mean delay at RLC buffers for different PS algorithms

As illustrated in Figure 5-3, by using DDQ-based packet scheduling algorithms, download multimedia services (i.e. FACH 1 and FACH 6) experience much less mean delay compared with MLPQ. It is noted that the significant reduction on delay of lower class services does not result in dramatic performance degradation on the higher class counterparts (i.e FACHs 2 to 5) beyond the guaranteed QoS rank. It is demonstrated that DDQ allows the download service to attain the highest degree of utilisation of the spare resources remaining for the streaming service for no significant degradation on the QoS target of the streaming users.
Figure 5-4 shows the mean jitter experienced by each individual service when employing MLPQ and DDQ packet scheduling. Obviously, the DDQ-based approach features much lower jitter for both streaming service and download service than those in MLPQ, especially for lower class and lower date rate users.

Since the unidirectional streaming service in SDMB is quite sensitive to delay-variation (jitter), this result proves that the DDQ-based packet scheduling provides a way to balance all FACH queues in order to achieve minimum delay variation for streaming services.
5.4.2.2 Analysis of channel utilisation ratio

Figure 5-5: Code channel utilisation for MLPQ and DDQ

Figure 5-5 plots the instantaneous variation of S-CCPCH physical channel utilisation on a TTI-scale for MLPQ-based and DDQ-based schedulers. Both schedulers managed to achieve throughput close to the optimum. For instance, the code channel utilisation ratios are: 97.8%, 96.2%, 85.4% respectively under MLPQ scheduling; whilst they achieve 98.4%, 96.2%, 86.4% respectively under DDQ scheduling. Therefore, DDQ manages to obtain a slight channel utilisation improvement on those code channels carrying background traffic.

To summarise, DDQ algorithm achieves the following advantages over the existing scheduling scheme:

- Achieving dynamic proportional delay-driven prioritization.
- Balancing all FACHs irrespective of their QoS class.
- Background class can achieve the highest utilisation without posing significant degradation on streaming class.
- Significant improvement on delay and delay variation performance
- Better overall system utilisation.
5.5 Summary

The packet scheduling algorithm is a key element within the RRM framework of the SDMB system. Existing algorithms have been shown to have drawbacks with regard to efficiency and fairness. Motivated by the proportional differentiation concept in general differentiated networks, novel packet scheduling algorithms are proposed for the SDMB system. The algorithms take into account the impact of important performance factors reflecting service QoS demands in order to provide traffic differentiation and optimize the overall system performance.

As the simplest adaptation of the proportional differentiation model, the BLRQ scheme considers the scheduling problem in accordance with the instantaneous buffer status of each FACH queue. Simulation and evaluation results show that the proposed packet scheduling schemes achieve better performance compared with the existing schemes. Average buffer length is greatly reduced by using the BLRQ-based scheduling.

More importantly, as the single delay metric that can be available and controllable at the satellite hub in the unidirectional satellite system, queuing delay is envisaged as one of the most crucial performance parameters for the packet scheduling. By adopting the DDQ-based scheduling, all FACH queues are balanced irrespective of their QoS class. The background class can achieve the highest utilisation without posing significant degradation on the streaming class. Compared with the MLPQ-based and WFQ-based scheduling scheme, dramatic improvements are achieved in delay, delay variation, and physical channel utilisation performance. It is noticed that performance enhancements of the proposed scheduling algorithms heavily depend on the selected specific traffic mix scenarios and incoming traffic dynamics. Our simulation demonstrated that the proportional differentiation based scheduling variations perform better in a scenario with relatively complex traffic mix and highly dynamic traffic statistics.
Chapter 6

6 Packet Scheduling with Cross-layer Design

6.1 Introduction

In this chapter, we investigate the cross-layer concept in the packet scheduling scheme. Novel packet scheduling algorithms are proposed using the cross-layer approach. The proposed algorithms jointly consider several key factors which span from the application layer to the physical layer, aiming at simultaneously guaranteeing diverse QoS whilst utilising radio resource efficiently under the system’s power/resource constraints. Simulation results demonstrate that the proposed cross-layer scheme achieves significantly better performance than existing schemes in queuing delay/jitter and channel utilisation.

Firstly, two new service prioritization algorithms, namely Combined Delay and Rate Differentiation (CDRD) and Cross-layer Joint Priority Queuing (CJPQ), are proposed. The former can be regarded as an enhanced version of DDQ, considering the data rate and delay constraint attribute; and the latter considers comprehensive performance criteria across multiple protocol layers to further optimize overall QoS guarantees. Both of the above algorithms are independent of the resource allocation procedure which operates after the service prioritization procedure.

Furthermore, we consider a priority-oriented resource allocation at the RRC layer for multi-sessions with diverse QoS requirements, where each session employs the proposed dynamic rate matching (DRM) at the physical layer. The proposed DRM-based resource allocation scheme considers physical layer data rate information into packet scheduling in order to achieve improved channel utilisation efficiency.

The combined packet scheduling scheme outperforms the existing scheme in various aspects. Firstly, the service prioritization procedure takes into account the session’s traffic priorities and QoS requirements in terms of maximum tolerable queuing delay and required data rate. Secondly, queuing dynamics are encapsulated into service prioritization to track the instantaneous variations at the RLC layer. Furthermore, the instantaneous data rate at the transport channel (TrCH) in the physical layer is proposed as another important criterion to minimize the unnecessary discontinuous transmission (DTX) and thereby optimize the resource utilisation.
Chapter 6. Packet Scheduling with Cross-layer Design

The chapter starts with a brief review of cross-layer concepts, emphasizing on their adaptation into RRM algorithms. The chapter proceeds with a detailed description of the novel CDRD packet scheduling algorithm and discussion of its performance enhancements and trade-offs. The CJPQ is then presented as a comprehensive QoS-based packet scheduling scheme, which tries to satisfy the overall QoS guarantee in terms of queuing delay, buffer occupancy, packet loss, throughput, data rate, etc. One of the most important aspects of the cross-layer design is the involvement with the DRM, which was initially proposed in [14] and thereafter employed as a contributing element to the resource allocation part of the proposed overall cross-layer packet scheduling framework.

6.2 Combined delay and rate differentiation queuing

6.2.1 Motivation

An efficient packet scheduling algorithm in SDMB is not only required to satisfy the QoS requirements of different services, but also has to optimize the transmission power setting of each physical channel on the basis of the required reception QoS level and under the constraint that the total available transmission power within a satellite beam is fixed.

To overcome inherent drawbacks in MLPQ and WFQ, and further improves the performance of the DDQ scheme, in this section, we propose a novel packet scheduling scheme, namely CDRD packet scheduling. This scheme is distinct from most existing scheduling algorithms in that: 1) it guarantees the prescribed QoS requirements, 2) it accounts for queuing behaviours at link layer (i.e., queuing delay and data rate), while at the same time, 3) it preserves the physical layer power/resource constraints. To the best of the authors' knowledge, no evaluation of such scheme is written on current RRM framework in SDMB, and this is the first time this concept is investigated. Although we herein apply the scheme to the SDMB system, it can be applied adaptively to any WCDMA-based BC/MC network.

6.2.2 Description of CDRD scheduling algorithm

The framework of the proposed CDRD scheduling scheme for heterogeneous multimedia services in SDMB system is illustrated as Figure 6-1. The accepted sessions can comprise different sessions with different QoS requirements. Each session is assumed to retain an individual queue in the FACH transport channel buffers. The types of user service supported within the SDMB system are streaming and download, which correspond to UMTS QoS class streaming and background respectively [37]. Herein the traffic class considered can be divided into two main
categories depending on delay constraint. The first category is delay-sensitive streaming class. The second category is the delay-insensitive download class, which can be further sub-categorised into two sub-classes according to its sensitivity to delay, namely "hot download" and "cold download". Packets in FACH buffers are first prioritized by a "Service Prioritization" module with their respective QoS metrics as criteria. Consequently, the "Dynamic Adaptive Packet Scheduler" module deals with all the queues in FACH buffers according to their instantaneous priorities instead of their inherent QoS class. Nonetheless, since packets belonging to the download class have no explicit delay constraint, the scheduler will restrict these download service so as to preserve QoS requirements for delay-constraint classes if the available resources is scarce. On the contrary, as long as there is spare resource remaining in the streaming queues, the scheduler will adaptively enable download services to fill the gap left by the streaming services.

The selected queues will be passed to the "Resource Allocation" module and be allocated with the required resources, which consist of bit rate and transmission power assignments. The scheduled packets will be delivered to the S-CCPCH in the form of TB [11]. For each active physical channel, the exact TFC, which consists of multiple TBSs at each TrCH will be selected (i.e. the amount of data from each transport channel mapped to the physical channel) from the TFCS. It is noted that a separate TFCS is provided for each S-CCPCH.

The admission control module has tight interaction with the packet scheduler. In order to allow the packet scheduler to handle with satisfactory QoS provisioning for ongoing sessions, the admission control module limits the traffic load so as to preserve the required QoS/GoS requirements. This chapter focuses on the packet scheduling technique and the design of admission control is not considered herein.
As stated above, the packet scheduling strategy can be described in the following two steps: service prioritization and resource allocation. For CDRD, only the former is changed with the following algorithms, whilst the latter step remains the same.

Based on the incoming service requests, these different requests are ordered according to priority criteria. In selecting the respective criteria, the service attributes are considered in a joint judgment function (JJF) to provide a dynamic intelligent scheduling task whilst considering several essential QoS factors that have crucial impact on system performance.

The joint judgement function used by the proposed CDRD is developed based on the DDQ scheme [116]. The difference between them is that DDQ only focuses on delay differentiation and does not consider other QoS factors, but CDRD takes into account several other key performance parameters (i.e. required data rate and maximum acceptable delay for the specified service) for assuring satisfaction of service QoS requirements. In brief, CDRD aims to balance all service flows in order to achieve high-grade QoS satisfaction and improve overall system performance.

In the CDRD scheme, each of the QoS parameters is represented as a contributing factor in the JJF. For the service flow at FACH queue \( i \) at current time slot (i.e. TTI for UMTS) \( n \), the JJF is defined as follows:

\[
P_i(n) = \alpha_i \times \delta_i(n) \times \lambda_i(n) \times \gamma_i(n)
\]  

\((6.1)\)
where $P_i(n)$ is the priority index for each queue $i$ at current time slot $n$. $n$ is the sequence number of the TTI at current time.

$\alpha_i$ is QoS class factor, which is essentially a time-independent parameter designated for queue $i$. Since CDRD assumes that there are QoS ratios between different QoS priority classes, the value of $\alpha_i$ represents the relative QoS priority class of the corresponding service. This factor is set according to the type of service carried by queue $i$, i.e. streaming service has higher priority than hot download service, while hot download service has higher priority than cold download service. In this way, QoS class affects the priority criterion (i.e. $P_i(n)$) by $\alpha_i$ and priority differentiation within the same QoS class is provided by other parameters in (6.1).

$\delta_i(n)$ is the delay serving index at current time slot $n$ for queue $i$. This measure describes the delay status of all packets passing through the respective queue, including both the packets which are currently in queue $i$ and the packets which have already been served by queue $i$ up to current TTI $n$. $\delta_i(n)$ is calculated for each queue $i$ in each TTI as follows:

$$\delta_i(n) = \frac{\sum_{j=1}^{N_q} W_{i,j}^q(n) + \sum_{k=1}^{N_d} W_{i,k}^d(n)}{N_q + N_d}$$

where $W_{i,j}^q(n)$ is the waiting delay for the $j^{th}$ packet currently in the queue $i$; $N_q$ is the number of packets that are currently in the queue $i$; $W_{i,k}^d(n)$ is the queuing delay for the $k^{th}$ packet which has been served by queue $i$ before the current TTI (i.e. current time slot $n$); $N_d$ is the number of packets that have been served and left the queue before the current TTI $n$.

$\lambda_i(n)$ represents the data rate factor for queue $i$ at current time slot $n$. It is based on the ratio of the service data rate required against the average transmitted data rate. The average transmitted data rate $\overline{\lambda}_i(n)$ for queue $i$ at time slot $n$ can be expressed as follows:

$$\overline{\lambda}_i(n) = \sum_{k=1}^{N_d} \frac{S_{i,k}}{(n-1) \times T_{mi}}$$

where $S_{i,k}$ is the packet size for $k^{th}$ packet in queue $i$; $N_d$ is the number of packets that have left the queue prior to this TTI; $T_{mi}$ is the value of TTI (i.e. 0.08 seconds in our case).

Therefore, the data rate factor $\lambda_i(n)$ is defined as follows:
where $\lambda_{req}$ is the required/guaranteed data rate specified by the service QoS level. If the average transmitted data rate served by the scheduler is smaller than the required data rate targeted by the specific service, $\lambda_i(n)$ is larger than 1, thus the priority index for this queue is increased for better chance of being served; otherwise, it is smaller than 1, and the priority index is decreased for this over-satisfied queue. This factor is used to fine-tune the priority and leads the transmitted data rate to approach the guaranteed data rate.

$\gamma_i(n)$ is the delay constraint factor for queue $i$ at current time slot $n$, depending on the maximum queuing delay tolerated by a service. This factor can be expressed as follows:

$$
\gamma_i(n) = \begin{cases} 
2, & \forall n : \frac{\sum_{j=1}^{N_i} W_{ij}(n)}{N_q} \geq W_i^{\text{threshold}} \\
1, & \forall n : \frac{\sum_{j=1}^{N_i} W_{ij}(n)}{N_q} < W_i^{\text{threshold}} 
\end{cases}
$$

(6.5)

where $W_{ij}(n)$ is the waiting delay for the $j^{th}$ packet currently in the queue $i$; $N_q$ is the number of packets that are currently in queue $i$; $W_i^{\text{threshold}}$ is the delay threshold for the service queue $i$.

If the average queuing delay for queue $i$ is larger than its delay threshold, the delay constraint factor $\gamma_i(n)$ is set to 2, which doubles the priority of this queue for improved chance to be processed; otherwise, it is set to 1. It is noted that delay threshold can be chosen as an adjustable parameter, which depends on the maximum tolerable delay of the corresponding service. $\gamma_i(n)$ is only in effect when the average queuing delay beyond the designated delay threshold, which provides a more efficient action to be taken for better QoS provisioning amongst differentiated traffic flows.

In each TTI, the scheduler will sort the FACH queues according to their priority index calculated from the JJF in descending order. The FACH queues with higher priorities will be served ahead of their lower priority counterparts.
6.3 Cross-layer Joint Priority Queuing packet scheduling

6.3.1 Overview of CJPQ Scheme

As illustrated in Figure 6-2, the novel packet scheduling algorithm can accommodate multi-session multimedia content delivery with diverse QoS requirements. The accepted ongoing sessions are first configured onto a radio bearer by RRA before going through the queuing buffer. Then the packet scheduler prioritizes all the queues in the buffer according to the QoS demands and queuing behaviours of the particular session. Prioritized queues are then allocated with required resource and power subject to system's resource and power constraints.
Figure 6-3 illustrates the layer/sublayer interactions of the proposed cross-layer packet scheduling scheme in the SDMB GW.

Figure 6-3 illustrates the layer/sublayer interactions of the proposed cross-layer packet scheduling scheme. The RRM is mainly handled at the data link layer, which can be further divided into RRC, RLC and MAC sublayers. As seen from the proposed scheme, the cross-layer/sublayer correspondence is set up from both top-down and bottom-up directions to the packet scheduler at the MAC sublayer of the data link layer. Firstly, the MBMS sessions' prescribed QoS demands are passed to the RRA at the RRC sublayer of the data link layer at the beginning of each admitted session. During the radio bearer configuration, the RRA abstracts the prescribed QoS demands of admitted sessions and passes them to the packet scheduler as one set of priority criteria. The queuing dynamics in the RLC buffer are monitored and passed to the packet scheduler as another set of priority criteria.

By utilising the cross-layer correspondence through the layered protocol stacks, the proposed scheme exploits multimedia QoS requirements and seeks to provide better system performance by dynamically adapting to the queuing behaviours of each competing flow. To efficiently schedule wireless resources (such as bandwidth and power) and satisfy diverse QoS guarantees, the QoS demand at both application layer and transport layer as well as the instantaneous queuing
behaviours induced by heterogeneous multimedia traffics are used in the design of the proposed scheduling scheme.

To differentiate and schedule the multiplexed sessions with diverse QoS and queuing status, a joint priority function (JPF) is defined and applied for each admitted session and updated periodically in each TTI depending on several cross-layer factors as shown in Table 6-1:

<table>
<thead>
<tr>
<th>Layers</th>
<th>Performance metrics</th>
<th>Target performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Layer</td>
<td>Service type (i.e. QoS rank)</td>
<td></td>
</tr>
<tr>
<td>Transport Layer</td>
<td>Average packet loss rate</td>
<td>Acceptable packet loss rate</td>
</tr>
<tr>
<td></td>
<td>Average throughput</td>
<td>Target throughput</td>
</tr>
<tr>
<td>RLC Layer</td>
<td>Average queuing delay</td>
<td>Queuing delay bound</td>
</tr>
<tr>
<td></td>
<td>Instantaneous queue length</td>
<td>Buffer threshold</td>
</tr>
<tr>
<td>Physical Layer</td>
<td>Average data rate</td>
<td>Required data rate</td>
</tr>
</tbody>
</table>

The contributing factors can be classified into two main sets, according to their frequency of variation. The first set of factors is set at the beginning of each session and kept constant during the session transmission. These factors include service type, required data rate, delay or packet loss constraints, etc. The second set of factors, which determine the average performance behaviours of the competing sessions achieved prior to the current TTI during the transmission, are reset at the beginning of each TTI and kept constant within a TTI. These factors include the average history queuing delay, instantaneous buffer length, etc. For multimedia delivery, the QoS rank of the specific session is the decisive factor in the JPF function, whilst other factors depend on the satisfaction of the above performance metrics by comparing the instantaneous performance metrics with their corresponding targets.

In the duration of each TTI, the JPF value is evaluated and assigned to each competing session. The session with the highest priority evaluated from the JPF function is scheduled first compared with their lower priority counterparts.

6.3.2 Description of CJPQ scheduling algorithm

In order to simultaneously guarantee the QoS criteria and adapt to the dynamic variations on queuing behaviours, we define a joint priority $\theta(n)$ in (6.6) for competing transport channel $j$ at
the time slot $n$, where $n \in \mathbb{N} := \{1, 2, \ldots, N\}$, $N$ is the total number of TTI during the transmission of session $j$.

$\theta_j(n) = \alpha_j \cdot \bar{\tau}_j(n) \cdot \lambda_j(n) \cdot T_j(n) \cdot \Lambda_j(n) \cdot \Gamma_j(n) \cdot \Sigma_j(n) \cdot H_j(n)$ \hspace{1cm} (6.6)

where $\alpha_j$ is the application prescribed QoS rank for the $j^{th}$ session, providing the service classification, whilst the fairness in the proposed algorithm is provided with the following profiles. $\bar{\tau}_j(n)$ is the average queuing delay until the $n^{th}$ TTI, $\lambda_j(n)$ is the average queue length achieved until the $n^{th}$ TTI. Parameter $T$, $\Lambda$, $\Gamma$, $\Sigma$, $H$ are the performance coefficients for queuing delay, buffer occupancy, data rate, packet loss rate, throughput, respectively, which denote the current status of the queuing behaviours for each queue at current TTI.

### 6.3.2.1 Queuing delay

Let $\gamma^*$ denote the maximum queuing delay tolerable for the $j^{th}$ FACH queue specified by the session QoS requirements. We associate with each FACH queue $j$ a queuing delay criterion $T_j(n)$ as a coefficient in the JPF given by:

$T_j(n) = \begin{cases} 1 & \text{if } \bar{\tau}_j(n) \leq \gamma_j^* \\ \frac{\bar{\tau}_j(n)}{\gamma_j^*} & \text{if } \bar{\tau}_j(n) > \gamma_j^* \end{cases}$ \hspace{1cm} (6.7)

$\bar{\tau}_j(n)$ denotes the average queuing delay of the $j^{th}$ FACH achieved up to TTI $n$, which is defined as:

$\bar{\tau}_j(n) = \sum_{m=1}^{n} \tau_j(m)$ \hspace{1cm} (6.8)

where $\tau_j(m)$ represents the instantaneous queuing delay for all the packets served by the $j^{th}$ FACH at TTI $m$, which is given by:

$\tau_j(m) = \frac{\sum_{k=0}^{\Delta_j} \tau_{j,k}^\Delta(m) + \sum_{k=0}^{\Lambda_j} \tau_{j,k}^\Lambda(m)}{N_j^\Delta(n) + N_j^\Lambda(n)}$ \hspace{1cm} (6.9)

where $\Delta := \{1, 2, \ldots, N_j^\Delta(n)\}$, $N_j^\Delta(n)$ is the number of packets that are currently in the queue, $\tau_{j,k}^\Delta(m)$ is the packet queuing delay for packet $k$ currently in the queue $j$, $\Theta := \{1, 2, \ldots, N_j^\Theta(n)\}$, $N_j^\Theta(n)$ the number of packets that have left the queue prior to this TTI, $\tau_{j,k}^\Theta(m)$ is packet queuing delay for packet $k$, which has left the queue $j$ prior to this TTI. The value of $N_j^\Delta(n) + N_j^\Theta(n)$ equals to the total number of packet arrived in the FACH queue.
It is worth noting that the most important and decisive delay profile that impacts the overall achieved QoS is the end-to-end delay in the network. However, due to the unidirectional nature of the SDMB system and its long propagation delay, this end-to-end delay is not available at the satellite hub (Satellite hub), therefore the queuing delay experienced in the buffer is the only controllable delay element pertinent to the RRM functionality and is employed as a limited version of delay criterion in this chapter. A delay tolerance threshold is assigned to each admitted session, representing the maximum acceptable queuing delay for each session. If the queuing delay exceeds the delay threshold, the session is given higher priority for better chance of being processed.

6.3.2.2 Buffer occupancy

Herein we assume a finite size buffer in the satellite hub. It is vital, especially for loss-sensitive service, to maintain the queue length at a reasonably safe level to prevent the system having excessive packet loss due to buffer overflow. Let $\Lambda_j^*$ denote the buffer length for the $j^{th}$ FACH queue specified by application. We associate with each FACH queue $j$ a buffer occupancy criterion $A_j(n)$ as a coefficient in the JPF given by:

$$
A_j(n) = \begin{cases} 
1 & \text{if } \lambda_j(n) \leq \lambda_j \cdot \sigma_j \\
\frac{\lambda_j(n)}{\lambda_j \cdot \sigma_j} & \text{if } \lambda_j(n) > \lambda_j \cdot \sigma_j
\end{cases}
$$

(6.10)

where $\sigma_j$ is the buffer occupancy threshold, providing a safe bound for the $j^{th}$ FACH queuing buffer, $\lambda_j(n)$ denotes the instantaneous queue length of the $j^{th}$ FACH at current TTI.

6.3.2.3 Data rate

Let $\gamma_j^*$ denote the guaranteed data rate for the $j^{th}$ FACH queue, a data rate profile $\Gamma_j(n)$ of the $j^{th}$ FACH queue is defined as:

$$
\Gamma_j(n) = \begin{cases} 
1 & \text{if } \gamma_j(n) \leq \gamma_j^* \\
\frac{\gamma_j(n)}{\gamma_j^*} & \text{if } \gamma_j(n) > \gamma_j^*
\end{cases}
$$

(6.11)

where $\gamma_j(n)$ denotes the mean data rate of $j^{th}$ FACH up to TTI $n$, which is determined as:

$$
\gamma_j(n) = \sum_{k=1}^{N_j(n)} S_{j,k} / n \cdot T_{mi}
$$

(6.12)
where $S_{jk}$ represents packet size for $k^{th}$ packet in queue $j$. $T_m$ is the value of TTI, e.g. 80ms in our simulation. $N_j^i(n)$ the number of packets that have left the $j^{th}$ queue before TTI $n$.

### 6.3.2.4 Packet loss rate

Let $\xi_j^*$ denote the target packet loss rate (PLR) due to buffer overflow for the $j^{th}$ FACH queue at the satellite hub RLC buffer. We associate with each FACH queue $j$ a PLR criterion $\Xi_j(n)$ as a coefficient in the JPF given by:

$$\Xi_j(n) = \begin{cases} 
1 & \text{if } \bar{\xi}_j(n) \leq \xi_j^* \\
\frac{\bar{\xi}_j(n)}{\xi_j^*} & \text{if } \bar{\xi}_j(n) > \xi_j^* 
\end{cases} \quad (6.13)$$

$\bar{\xi}_j(n)$ denotes the average PLR of the $j^{th}$ FACH achieved up to TTI $n$, which is defined as:

$$\bar{\xi}_j(n) = \frac{N_j^d(n)}{N_j^q(n) + N_j^i(n)} \quad (6.14)$$

where $N_j^d(n)$ represents the total number of packets that are dropped due to buffer overflow for the $j^{th}$ FACH up to TTI $n$.

Similar to the queuing delay, the PLR metric available at the satellite hub is also confined to the packet loss rate due to buffer overflow, although the packet loss in the propagation path is the most crucial factor impacting the QoS performance. Nevertheless, the packet scheduler is unable to control that information and the overflow loss is the single metric that can be monitored and controlled by the RRM entity.

### 6.3.2.5 Throughput

Let $\eta_j^*$ denote the target throughput for the $j^{th}$ FACH queue at the satellite hub RLC buffer. We associate with each FACH queue $j$ a throughput criterion $H_j(n)$ as a coefficient in the JPF given by:

$$H_j(n) = \begin{cases} 
1 & \text{if } \bar{\eta}_j(n) \leq \eta_j^* \\
\frac{\bar{\eta}_j(n)}{\eta_j^*} & \text{if } \bar{\eta}_j(n) > \eta_j^* 
\end{cases} \quad (6.15)$$

$\bar{\eta}_j(n)$ denotes the average throughput of $j^{th}$ FACH achieved up to TTI $n$, which is defined as:
where $\bar{\eta}_j(n) = \sum_{m=1}^{n} \eta_j(m)$ represents the instantaneous throughput for the $j^{th}$ FACH at TTI $m$.

The throughput is defined as the ratio between the total bits arrived in a specific queue until current timing (i.e. $m^{*}TTI$) and the total bits that has been successfully scheduled and delivered to the physical channel for radio frame transmission.

In each TTI, the multiplexed FACH queues are sorted according to the instantaneous value of $\eta_j(n)$, queues with the higher priority will be served first in a non-preemptive manner.

### 6.4 Dynamic resource allocation

#### 6.4.1 Background

As shown in Figure 6-4, time multiplexing from transport channels to physical channels is performed independently from other physical layer functions, such as CRC, turbo coding and interleaving. However, the performance of rate matching, which is responsible for matching input data rate to the allocated data rate by performing puncturing or repeating the bits, is highly relevant to the transport channel multiplexing.

It is noted that extensive research [117]-[119] has been performed on the upper layer Turbo coder, and interleaving, but limited research has been carried out with respect to the rate matching technique. In traditional static rate matching (SRM) techniques, the allocated data rate is based on the maximum data rate supportable for each physical channel. This strategy can only influence long term resource allocation, whilst the short term physical layer data rate variation can waste
system capacity. Since the rate matching functionality is performed at the physical layer in accordance with other physical layer procedures, cross-layer interactions between physical layer and MAC layer can lead to beneficial impact on resource utilisation/allocation.

The SRM scheme has been used within the baseline SDMB RRM. The objective of the SRM is to minimize the number of DTX bits required for the maximum data rate supported by the TFCS for a given physical channel. The rate matching ratio is calculated according to the maximum data rate at the beginning of each session. However, when the instantaneous data rate is lower than the maximum data rate, the rate matching ratio is held constant, which results in unnecessary puncturing and inefficient resource/power utilisation in the physical layer.

6.4.2 Dynamic rate matching

In order to minimize the DTX insertion in the downlink SRM, a novel rate matching technology, namely DRM, was proposed for delivering a highly rate-variable MBMS service in the SDMB system [14]. DRA is developed based on the novel DRM technique and performs the resource allocation task at the MAC sublayer according to the cross-layer data rate information from the physical layer. This new technique is more resource-friendly in that the radio resource is allocated dynamically based on the instantaneous data rate for each application class in a short-term TTI-scale.

The objective of downlink DRM is to minimize the number of DTX bits required for the chosen TFC at a given TTI according to the available physical layer resources. Rather than per-session- and maximum data rate-based rate matching calculation in SRM, in the DRM case, the rate matching ratio is calculated in each TTI based on the instantaneous data rate of each TrCH. Therefore, the DRM employs a variable rate matching ratio to prevent unnecessary DTX insertion.

The scenario undertaken is applicable for the TrCHs featuring identical TTI-scale. In each TTI, the following phases are performed in DRM:

- Phase 1 (TFC reordering) - All the TFCs within the TFCS list are re-organised according to their corresponding total data rate (i.e. based on the TFC size).
- Phase 2 (Rate matching ratio calculation) - The rate matching ratio is calculated based on the instantaneous data rate (i.e. the TBS size) of each TrCH for each different TFCs allowed for a given physical channel.
- Phase 3 (Bit matching) - According to the selected TFC, a tentative value of repetition/puncturing bits value is calculated for each TrCH.

Finally, the rate matching module performs the tentative value corrections and a rate matching pattern is generated.
In order to apply the DRM technique, modifications are identified and applied at both the sending and receiving side:

- New rate matching interval on a TTI-scale, during which the rate matching ratio remains constant.
- Before the transmission starts, rate matching ratios need to be calculated for all possible TFCs.

By reducing the DTX insertion, the total transmit power can be minimized whilst the power consumption is optimized. Furthermore, in some cases, DRM can facilitate higher bit repetition, which has essential impact on improving the BER performance. Finally, the transmit power saved can be utilised by other channels for improved transmission capacity; alternatively, it can also be used for improving the packet loss rate of allocated channels. The main limitation of DRM lies in that more processing and memory are required for the rate matching calculation and storage.

6.4.3 DRM-based resource allocation

The task of resource allocation is to select the required bit rate and transmit power for all physical channels within the specific resource allocation interval (i.e. one TTI) according to the QoS guarantees and physical channel constraints. The existing SRM technique at the physical layer has been traditionally designed separately from higher layers, the rate matching and power allocation are separately optimized. Based on the DRM technique at the physical layer, we develop a cross-layer design, which optimizes the resource utilisation, to improve the overall system efficiency, when combined with resource allocation at the network layer.

Unlike SRM, where the rate matching ratio is fixed during the session transmission, in DRM, until all the sessions are scheduled, the precise rate matching ratio for each TrCH will not be known and the DRM only knows the possible TFC candidates (i.e. TFC subset). The power requirements are evaluated based on every possible TFCs within the TFC subset under the maximum power constraints. In each TTI, a TFC is selected for a given physical channel, the DRM will then calculate the number of bits required to be repeated/punctured.

The proposed DRA algorithm, as shown in Figure 6-5, relies on the DRM technique to evaluate the required transmit power for respective TrCH according to their instantaneous data rate requirements. In this case, the physical layer instantaneous supportable data rate is effectively the TBS size allowable for the total transmit power estimation. Along with the priority decided by the service prioritization function, the instantaneous data rate information is fed into the resource allocation module for power estimation, the derived TFC is thereby priority-based, DRM-
associated, and data-rate-confined, which is then allocated to the corresponding physical channel. In this way, physical layer resources are effectively considered within the packet scheduling.

Figure 6-5. Procedures of DRM-based resource allocation algorithm.

The procedure for the proposed power allocation algorithm based on DRM is described in the following.

For each FACH, the packet scheduler scans the TFCS of the physical channel to find the different TBS sizes that could be used, namely serve the whole or part of the queued data at the FACH buffer. A sorted list of all candidate TBS sizes, in decreasing order, is created. The scheduler firstly seeks to allocate the maximum TBS size to the first FACH.

As shown in Figure 6-6, the scheduler first checks the rate matching ratio based on (6.17). It calculates the RM ratios for all the subsets of TFCs from the full TFC set, and these values are stored against each TFC set. Then the scheduler checks which subset the chosen TFC belongs to, and based on this information the rate matching ratio obtained from the stored data, and a tentative value is determined according to the selected TFC. Following this, the scheduler performs tentative value correction and then rate matching patterns are generated.
Calculate the transmit power for the considered physical channel

\[
RF_{p,q} = \frac{N_{Data}}{\sum_{k=1}^{\omega} \left( (l_k + l_{CRC}) \times CR + l_{tailbits} \right) \times N_k} 
\]

\[\text{for } 1 \leq q \leq \omega \]  

(6.17)

where, \( p \) is the SCCPCH ID, \( q \) is the TFC ID, and \( \omega \) is the TFC subset that its TBS that has been calculated. \( N_{Data} \) is the allowed data. \( l_k \) is the TB length of FACH k in this TFC. \( N_k \) is the number of TB with allocated TBS. \( CR \) is the coding rate.

According to the calculated RM ratio values, the scheduler determines the required \( E_b/N_0 \) value according to each session BLER requirement. The scheduler then checks whether the selected TBS size for the new session satisfies the total transmit power criteria.

- If this is not satisfied, it will check whether all the possible TBS sizes have been checked for total transmit power criteria. For the next TBS size (less than the previous one), the above procedure is performed. If none of the TBS sizes satisfy the power criteria, the scheduler assigns TBS to zero.
- If the power criteria is met, based on each session and RM combinations, the transmit power for each session is calculated separately according to (6.18) and the highest power requirement assigned as the physical channel transmit power.

Figure 6-6. Flowchart for the DRA algorithm.
\[ P_t(n) = \frac{P_n \times L \times \rho \times SF}{R_s \times CR \times RF} \] (6.18)

where \( P_t(n) \) is the required transmit power at time slot \( n \), \( P_n \) is the thermal noise, \( \rho \) is the \( E_b/N_0 \) requirement, \( L \) is path loss, and \( RF \) is the rate matching ratio, \( R_s \) is the modulation scheme, and \( SF \) is the spread factor.

These procedures are repeated recursively until all the FACHs mapped to each S-CCPCH are assigned.

In conclusion, the DRA algorithm offers two main advantages over the existing SRM-based scheme: 1) it allows better DTX minimization, and 2) requires less transmit power when the instantaneous data rates are less than the maximum data rate. However, the DRA involves with inevitable more processing and memory compared to SRM-based resource allocation.

### 6.5 Performance evaluation

#### 6.5.1 Simulation scenarios

In order to demonstrate the performance enhancement of the proposed CDRD and CJPQ packet scheduling schemes, a system-level simulator implementing the SDMB system has been developed as described in Appendix D and a wide range of simulation scenarios have been identified. In addition, we choose different guaranteed data rates in order to examine the performance between users with different rate requirements. An indicative scenario is selected, as in Table 6-2, to demonstrate the simulation outcomes, where three S-CCPCHs are used for carrying heterogeneous multimedia services with different QoS classes:

- S-CCPCH 1 carries streaming FACHs 2 and 3 with data rate of 256kbps and 64kbps respectively and hot download FACH 1 with data rate of 64kbps;
- S-CCPCH 2 carries streaming FACHs 4 and 5 with data rate of 256kbps and 128kbps respectively;
- S-CCPCH 3 carries cold download FACH 6 with data rate of 384kbps.
Table 6-2: Channel multiplexing configuration

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming (kbps)</td>
<td>256x1;64x1</td>
<td>256x1;128x1</td>
<td>-</td>
</tr>
<tr>
<td>Hot Download (kbps)</td>
<td>64x1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cold Download (kbps)</td>
<td>-</td>
<td>-</td>
<td>384x1</td>
</tr>
</tbody>
</table>

Link level simulation results provide the $E_b/N_0$ v.s BLER requirements of each TrCH. The radio propagation channel model features either classical Gaussian characteristics for satellite-associated path, or modified Rayleigh multipath fading channel for UE-associated path with the consideration of both Doppler Effect and propagation impairments.

In the SDMB system, the queuing delay tolerance threshold is assumed to be 20-100ms for video streaming, and 200-2000ms for push-and-store service. Accordingly, various queuing delay threshold values are applied and examined for the specific scenario, showing the performance variation against tuning the delay threshold parameter.

The performance of our proposed scheme and that of previous studies are evaluated and compared via simulations. Since the performance of WFQ-based scheduling algorithm is worse than that of MLPQ-based algorithm in terms of both delay and delay variation, only MLPQ is used as reference in this chapter. The following main parameters, which have significant impact upon the overall system performance, will be evaluated and discussed in this section.

- Queuing delay, delay variation evaluation
- PDF/CDF delay distribution
- S-CCPCH and FACH utilisation evaluation

6.5.2 Performance evaluation for DDQ with DRA

We first examine the performance impact of DRM, in comparison to SDM, on the DDQ-based scheduling algorithm. Since the selection of rate matching techniques feature major differences between the proposed scheduling algorithm and that of the previous research, we identify the scheduling algorithm scenarios as combinations of scheduling algorithms and rate matching techniques:

- Weighted Fair Queue scheduling algorithm with Static Rate Matching technique: WFQ-SRM;
Multi-level Priority Queue scheduling algorithm with Static Rate Matching technique: MLPQ-SRM;

Delay Differentiation Queue scheduling algorithm with Static Rate Matching technique: DDQ-SRM;

Delay Differentiation Queue scheduling algorithm with Dynamic Rate Matching technique: DDQ-DRM. This is our proposed CL-DDQ scheduling scheme.

The following key parameters, which have significant impact upon the overall system performance, have been examined in our evaluation:

- Queuing delay;
- Queuing delay variation;
- Physical channel throughput/utilisation.

### 6.5.2.1 Queuing delay evaluation

The mean queuing delay experienced by packets of each FACH queue in RLC buffers is an important parameter in evaluating the performance of packet scheduling algorithms. Figure 6-7 show the queuing delay of different scheduling schemes: MLPQ with SRM, WFQ with SRM, DDQ with SRM and DDQ with DRM.

As illustrated in Figure 6-7, compared with MLPQ-SRM and DDQ-SRM, WFQ-SRM has much lower delay for the high data rate services (i.e. FACH 6 at 384kbps), but has much longer delay.
for the relative low data rate services (i.e. FACHs 1, 3 and 5). However, it is noted that FACH 6 carries the download services and FACH 3 and 5 carry the streaming services. This means the predominance of high data rate down load services causes long delays on lower data rate streaming services since the priority between services classes is almost ignored by WFQ-SRM according to its data-rate-based nature.

The DDQ-SRM has slightly better performance compared with MLPQ-SRM, as shown in Figure 6-7. However, when the DRM is applied, download multimedia services (i.e. FACHs 1 and 6) experience much less mean delay in DDQ-DRM compared with the MLPQ-based scheme, whilst the mean delay experienced by streaming services features similar performance.

Numerically, hot and cold download classes have a reduction of 51.3% and 74.0% on their mean queuing delay, respectively. It is noted that, by using DDQ-DRM, the significant reduction on delay of the lower class service (i.e. download service) has been achieved. This indicates that, unlike the SRM scheme, DDQ is more workable for the DRM scheme.

As stated above, both DDQ and DRM achieve better queuing delay than previous approaches. When the performance of DDQ with DRM is compared with those of MLPQ and WFQ with SRM, significant performance enhancement are seen.

6.5.2.2 Queuing delay variation

Figure 6-8 shows the mean queuing jitter experienced by each individual FACH for three different packet scheduling schemes. As shown in Figure 6-8, by using SRM, compared with MLPQ and DDQ, the WFQ has much lower delay jitter for the high data rate services (i.e. FACH 6 at 384kbps), but has much longer delay jitter for the relative low data rate services (i.e. FACH 1, 3 and 5). It is noted that the unidirectional streaming service in the SDMB system is quite sensitive to delay-variation (jitter). The delay variation of the flows should be limited in order to preserve the time variation between packets of the stream. Although WFQ assigns higher priority to high data rate services, there is no differentiation made with respect of QoS rank, therefore WFQ makes delay-tolerant high data rate download services have higher priority than the delay-sensitive lower data rate streaming services.
Chapter 6. Packet Scheduling with Cross-layer Design

Queuing Jitter at RLC Buffer

<table>
<thead>
<tr>
<th>Scheme</th>
<th>FACH 1 d-64k</th>
<th>FACH 2 s-256k</th>
<th>FACH 3 s-64k</th>
<th>FACH 4 s-256k</th>
<th>FACH 5 s-128k</th>
<th>FACH 6 d-384k</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFQ-SRM</td>
<td>9.3</td>
<td>2.2</td>
<td>8.7</td>
<td>1.8</td>
<td>5.0</td>
<td>0.0</td>
</tr>
<tr>
<td>MLPQ-SRM</td>
<td>4.3</td>
<td>1.4</td>
<td>1.3</td>
<td>1.2</td>
<td>2.3</td>
<td>19.5</td>
</tr>
<tr>
<td>DDQ-SRM</td>
<td>3.8</td>
<td>1.3</td>
<td>1.1</td>
<td>1.1</td>
<td>2.0</td>
<td>18.7</td>
</tr>
<tr>
<td>DDQ-DRM</td>
<td>1.1</td>
<td>0.8</td>
<td>1.0</td>
<td>0.7</td>
<td>0.6</td>
<td>4.9</td>
</tr>
</tbody>
</table>

Figure 6-8. Queuing jitter of different scheduling schemes

As shown in Figure 6-8, DDQ-SRM has better performance for delay jitter than both MLPQ and WFQ with SRM. If the DRM is used, DDQ-DRM features much lower jitter for both streaming and download services. Typically, the average jitter reduction for download services (i.e. FACH 1, 6) is as much as 73.0%, whilst the average jitter reduction for the streaming service (i.e. FACH 2 to 5) is 38.2%. Therefore, the proposed cross-layer packet scheduling scheme achieves better queuing jitter than existing schemes.

6.5.2.3 Physical channel throughput/utilisation evaluation

<table>
<thead>
<tr>
<th>Scheme</th>
<th>S-CCPCH 1</th>
<th>S-CCPCH 2</th>
<th>S-CCPCH 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFQ-SRM</td>
<td>359.0</td>
<td>350.3</td>
<td>328.1</td>
</tr>
<tr>
<td>MLPQ-SRM</td>
<td>365.2</td>
<td>361.9</td>
<td>328.1</td>
</tr>
<tr>
<td>DDQ-SRM</td>
<td>366.4</td>
<td>363.4</td>
<td>328.1</td>
</tr>
<tr>
<td>DDQ-DRM</td>
<td>378.2</td>
<td>378.9</td>
<td>328.1</td>
</tr>
</tbody>
</table>

Figure 6-9. Throughput of different packet scheduling schemes

Figure 6-9 plots the mean S-CCPCH throughput for WFQ-SRM, MLPQ-SRM, DDQ-SRM and DDQ-DRM, respectively. DDQ-DRM achieves much higher throughput than for all the other schemes. It reaches 378.2kbps on SCCPCH 1, and 378.9kbps on SCCPCH 2 and 328.1 kbps on
SCCPCH 3, which are equivalent to 98.4%, 98.6% and 85.4% of the S-CCPCH capacity, respectively. Its nearest candidate, DDQ-SRM, reaches 366.4kbps on SCCPCH 1, and 363.4kbps on SCCPCH 2 and 328.1 kbps on SCCPCH 3, which are equivalent to 95.5%, 94.7% and 85.4% channel utilisation, respectively. It is noted that the performance on SCCPCH 3 are the same since there is a power limit on the spot beam, and the power is allocated from SCCPCH 1 to SCCPCH 3. Hence, DDQ-DRM manages to obtain channel utilisation improvement on those physical channels carrying background traffic. From the results, it can be inferred that the proposed DDQ-DRM scheme not only improves the delay performance among different QoS classes, but also has essential impact on increasing the throughput and physical channel utilisation.

As a summary, the proposed cross-layer scheduling algorithm has the following advantages over the existing packet scheduling scheme:

- Achieves dynamic proportional delay-driven prioritization;
- Balances all FACHs irrespective of their QoS class;
- Has significant improvement on delay and delay variation performance;
- Achieves more efficient resource allocation;
- Achieves higher throughput and channel utilisation.

6.5.3 Performance evaluation for CDRD with DRA

Since both service prioritization and resource allocation procedures of the proposed scheme feature major differences to the previous scheme, we identify the packet scheduling schemes as combinations of service prioritization algorithms (i.e. MLPQ and CDRD) and rate matching techniques (i.e. SRM and DRM) -based resource allocation algorithms. We compare the performance gain obtained by applying the proposed cross-layer optimization on the following essential performance metrics:

- Queuing delay evaluation
- Queuing delay/delay variation(jitter)
- CDF queuing delay
- Delay threshold variation
- Channel throughput/utilisation analysis
- Instantaneous data rate
- Average channel utilisation
Figure 6-10. Queueing delay CDF distribution for streaming performance evaluation at RLC layer.
6.5.3.1 Queuing delay evaluation

We evaluate the performance of the CDRD service prioritization by investigating the mean queuing delay experienced by packets in each FACH queue at the RLC buffer. As illustrated in Figure 6-9, download multimedia services (i.e. FACHs 1 and 6) experience much less mean queuing delay in CDRD than MLPQ, whilst the mean delays experienced by streaming services (i.e. FACHs 2 to 5) feature similar performance between the algorithms. Numerically, hot download and cold download enjoy an average reduction of 32.6% and 23.7% on their mean queuing delay respectively, whilst the maximum increase in the mean queuing delay of streaming service is 7.6%. It can be inferred that the significant reduction on the delay of the lower QoS class does not pose dramatic performance degradation on its higher QoS class counterpart, i.e. the QoS of streaming application can be guaranteed. It also implies that the CDRD enables the download service to efficiently utilise the spare resources of the streaming service without causing significant degradation on the QoS satisfaction of the streaming services. The cumulative distribution function (CDF) of queuing delays for the respective streaming FACHs in Figure 6-9 further verifies that CDRD achieved better queuing performance than MLPQ for all the streaming FACHs.

Figure 6-11 compares the mean queuing jitter experienced by each FACH between MLPQ and CDRD. Obviously, the latter achieves much lower jitter for both streaming and download services. Typically, the average jitter reductions for download and streaming services are 45.5% and 29.1% respectively. It is worth noticing that the unidirectional streaming service in SDMB is quite sensitive to delay variation (i.e. jitter), thereby the delay variation of the flows should be limited in order to preserve the time variation between packets in the stream [43]. The results in Figure 6-11 show that the proposed CDRD scheme provides better delay variation in RLC queues irrespective of their traffic priorities.
Here the probability density function (PDF) statistic of the queuing delay is used to describe statistical distribution density of queuing delay for each FACH. Figure 6-12 - Figure 6-13 shows the PDF curves of queuing delay for each FACH by using MLPQ and CDRD. It is noted that CDRD has better delay distribution characteristics than MLPQ. Especially, it achieves both lower average delay and delay variation for download service (i.e. FACHs 1 and 6).
Unlike those for the download service, the PDF curves of the queuing delays of streaming FACHs (i.e. FACHs 2 to 5) are indicated at the bottom-left corner of Figure 6-13.

In fact the explanation of PDF and CDF queuing delay performance becomes straightforward if the conclusions of mean delay/jitter are recalled.

![Mean queuing delay for FACH queues under different delay thresholds](image)

**Figure 6-14. Queuing delay performance evaluation under different delay thresholds.**

As explained earlier, the queuing delay threshold can be tuned as an adjustable parameter indicating the queuing delay tolerance for multimedia traffics. Observed from Figure 6-14, tuning the delay thresholds effectively influences the queuing delay performance of the FACH queues. The average queuing delay of particular FACH queues is affected directly by its delay threshold, namely more stringent delay thresholds lead to better performance of the corresponding queue, whilst causing longer delays for the other service classes.

### 6.5.3.2 Physical channel utilisation evaluation

Channel utilisation performance is presented in this section, considering utilisation ratios on both code channels (S-CCPCHs) and transport channels (FACHs). Figure 6-15 displays average code channel utilisation status when adopting MLPQ and CDRD scheduling algorithms for the given scenario.
Figure 6-15: Mean code channel utilisation for MLPQ and CDRD scheduling

Noticeable improvement can be seen for those code channels which are comprised with transport channel carrying background class services (i.e. S-CCPCHs 1 and 3). From the results, it can be inferred that the proposed algorithm not only improves the delay performance among different QoS classes, but also has essential impact on increasing the overall physical channel utilisation.
Figure 6-16: Instantaneous code channel utilisation for MLPQ (a) and CDRD (b) scheduling.

Figure 6-16 plots the instantaneous channel utilisation status of each S-CCPCH on a TTI-scale for both MLPQ-based and CDRD-based scheduler, respectively. Both schedulers managed to achieve throughput close to the maximum. For instance, the code channel utilisation ratios are: 97.8%, 96.2% and 85.4% for respective S-CCPCHs under MLPQ scheduling; whilst they achieve 98.4%, 96.2% and 86.4% respectively when CDRD scheduling is adopted. Therefore, CDRD manages to obtain slight channel utilisation improvement on those code channels carrying background traffic (i.e. S-CCPCHs 1 and 3).

The explanation of the above statement can be seen from Figure 6-17, where the comparison of the mean channel utilisation ratio of FACHs is displayed. Contrary to the MLPQ scheduling, for the CDRD scheduling scheme, it appears to achieve higher utilisation score, especially for background class traffic (i.e. download FACHs 1 and 6). These results also confirm those of Figure 6-15, explaining the main improvement on code channel utilisation is virtually caused by the higher channel utilisation ratio achieved by CDRD scheduling on those FACHs carrying background traffic.
The QoS demand expressed by the required data rate is considered as another factor effectively driving the session's instantaneous achieved data rate towards to the required data rate. This can significantly influence the overall system performance for the highly rate-variable multimedia transmissions. Comparing the Figure 6-18 (a) and (b), the CDRD allows better fairness in terms of data rate provisioning among multiplexed flows. By providing better QoS guarantees and considering the physical channel capabilities, the proposed scheme can increase the total mean transmission rate at the physical channels by 12.8%. Moreover, the data rate distribution amongst all the competing flows becomes more even, which means that better fairness is provided with the CDRD service prioritization than the strict prioritized MLPQ scheduling scheme. Figure 6-18 (c) shows the instantaneous physical channel utilisation achieved for different scheduling and rate matching schemes during a sample simulation period (i.e. around 125 TTIs). The total channel utilisation achieved by CDRD enjoys noticeable improvement through the session transmission. Numerically, the average channel utilisation reaches around 83.8% in CDRD-based scheduling, compared with 78.2% for the MLPQ-based scheduling. On the other hand, with the SRM applied, MLPQ only achieved an average of 62.4% utilisation because of its frequent unnecessary puncturing and repetition.
Chapter 6. Packet Scheduling with Cross-layer Design

Figure 6-18. Performance evaluation on channel throughput/utilisation achieved at the physical layer: (a) Instantaneous data rate of FACH with MLPQ scheduling, (b) Instantaneous data rate of FACH with CDRD scheduling, (c) Instantaneous physical channel utilisation for MLPQ and CDRD scheduling using SRM and DRM techniques.
6.5.4 Performance evaluation for CJPQ with DRA

6.5.4.1 Delay performance

In Figure 6-19, the queuing delay performance of CJPQ is compared to those of MLPQ and DDQ for all the allocated real-time streaming and download sessions. Results show that although DDQ achieves much better queuing delay performance than MLPQ via selecting the packet scheduling policy based on instantaneous queuing delay performance, the mean queuing delay on streaming service performs slightly worse than the MLPQ case. For example, cold download FACH 6 experiences less queuing delay from 10.62 seconds to 6.79 seconds, whilst the queuing delay of streaming FACH 5 is increased from 0.87 seconds to 1.06 seconds.

Rather than achieve lower download delay by sacrificing streaming delay performance in the DDQ case, the proposed CJPQ manages to deliver download sessions with even lower queuing delay whilst maintaining a similar or better performance on streaming sessions. Numerically, compared with DDQ, the CJPQ is capable of reducing the mean queuing delay for both streaming and download services by 17.1% and 34.9%, respectively. In conclusion, the CJPQ achieves improved queuing delay performance on all competing sessions, irrespective of their QoS rank.
6.5.4.2 Channel utilisation performance

Figure 6-20. Comparison on physical channel utilisation with different scheduling schemes.

By providing delay differentiation, DDQ provides better channel utilisation than MLPQ in that the resource is effectively and dynamically allocated to all multiplexed sessions. Therefore, download service have a better chance of being processed. Performance from Figure 6-20 shows that, when DDQ is used instead of MLPQ, there is an improvement of 7% on channel utilisation of download service (i.e. S-CCPCH 3), yet the channel utilisation on those S-CCPCHs carrying streaming services (S-CCPCH 1, 2) has a slight reduction.

One of the most desired features of the proposed CJPQ scheduling algorithm is the higher score achieved in terms of physical channel resource utilisation. Thanks to the cross-layer interaction in CJPQ, the channel utilisation is balanced between different S-CCPCHs, which provide improved fairness amongst competing sessions on utilising the radio resources. For example, in Figure 6-20 comparing CJPQ with DDQ, the channel utilisation for S-CCPCH carrying cold download service has been increased by 11%, whilst there is a less than 2% reduction on the channel utilisation for the other two S-CCPCHs carrying higher priority traffics, which will lead to significant performance gain achievable in resource/bandwidth utilisation.
6.5.4.3 Packet loss rate performance

As aforementioned, the packet loss rate herein refers to the packet loss probability due to buffer overflow. By taking account of the packet loss rate criterion for respective traffic type, the CJPQ scheduler is able to maintain a desired packet loss rate for corresponding sessions whilst performing efficient and fair scheduling task.

From Figure 6-21, when DDQ is applied and compared with MLPQ, it can be shown that there are reductions on packet loss rate performance of download FACHs which are essentially caused by the less delay achievable for low QoS class queues (i.e. lower queue length). In the mean time, the streaming FACH (e.g. FACH 2) suffers from slightly higher packet loss rate caused by the longer queuing delay in DDQ. When CJPQ is employed, by adaptively taking account of the packet loss rate performance, it can be seen that the packet loss rate is maintained below predefined performance targets, especially for those FACHs carrying lower priority services (FACHs 1 and 6). It is shown that CJPQ not only provides better performance on queuing delay for all competing sessions, but also is capable of providing desired packet loss rate performance against session’s QoS criterion.

By taking into account the impact of multiple essential QoS aspects across protocol stacks, the proposed CJPQ packet scheduling scheme not only satisfies higher layer diverse QoS demands, but also is capable of adapting the queuing dynamics of multiplexed sessions, and thereby offering balanced priority and fairness among different competing flows. The scheme has been evaluated for the SDMB system via simulations, and the results show that the proposed packet scheduling scheme achieves better performance for queuing delay, channel utilisation and packet loss compared with those of the existing schemes.
6.6 Summary

In this chapter, we have proposed a cross-layer packet scheduling scheme for multimedia delivery in the SDMB system. This scheme not only takes into account the impact of key performance factors reflecting service QoS demands and queuing dynamics, but also utilises a DRM technique at the physical layer in order to maximise spectrum efficiency and resource utilisation. Simulation has been conducted over extensive scenarios for various performance metrics and results show that the proposed packet scheduling scheme achieves better performance than the existing schemes for delay, jitter and channel utilisation. Discussions on the impact of a tuneable threshold parameter further demonstrate the flexibility and scalability achievable in this scheme. Simulation outcomes given in this chapter demonstrate that the certain extent of the performance enhancements varies from one traffic mix scenario to another, depending also heavily on the incoming traffic dynamics. However, from the results shown in selected traffic scenarios, it appears that the proposed cross-layer packet scheduling schemes is capable of gaining more performance improvement in complex traffic mix and variable rate scenarios than in relatively simple and constant rate scenarios. The reason for these trends lies in the dynamic prioritization nature of the proposed schemes.
Chapter 7

7 Adaptive Multi-dimensional QoS-based Packet Scheduling

7.1 Introduction

The explosive advances in the mobile multimedia broadcasting have created opportunities and challenges for the mobile and broadcast industries to deliver multimedia services to mass-market in a spectrum-efficient and cost-effective way. The high-speed and high-quality multimedia communications is becoming increasingly sophisticated, entailing diverse QoS requirements to be supported for various multimedia applications including voice, data as well as real-time video streaming.

Given the unique nature of the SDMB system, the design of an efficient packet scheduling scheme satisfying diverse QoS requirements is an especially challenging task. Numerous studies on packet scheduling schemes have been proposed in the literature for wireless networks in order to achieve high QoS satisfactions. For example, one popular subject within this context is to exploit the channel quality of fast-varying wireless links for more efficient packet scheduling [83]; however, given the unidirectional nature and long propagation delay, the SDMB system is unable to track real-time CSI from the mobile terminal side, which makes the channel-state dependent scheduling not feasible. Future multimedia applications feature increasingly diverse range of capabilities and QoS requirements, which should be considered in the design of the packet scheduling scheme so that these requirements can be well guaranteed.

To provide better QoS guarantee whilst achieving more efficient resource utilisation, in this chapter, we propose a novel adaptive multidimensional QoS-based (AMQ) packet scheduling scheme for provisioning heterogeneous multimedia services. By taking into account essential aspects of QoS provisioning whilst preserving the system power/resource constraints, the AMQ packet scheduling scheme is capable of satisfying diverse QoS requirements and adaptively optimizing the resource utilisation for satellite multimedia broadcasting. The proposed scheme is implemented by two separate and cooperative algorithms: adaptive service prioritization (ASP) algorithm and adaptive resource allocation (ARA) algorithm. By taking into account multiple
essential performance attributes, the former is capable of prioritizing contending sessions based on their QoS preferences and traffic dynamics, whilst the latter performs the resource allocation, in a dynamic and adaptive manner, according to the current QoS satisfaction degree of each session.

Compared with most existing packet scheduling algorithms used in satellite communication systems, e.g. [13], [116], the proposed packet scheduling scheme is distinct in that it is capable of: 1) satisfying multiple essential QoS requirements, 2) adaptively tracking the queuing dynamics induced by heterogeneous traffics, 3) dynamically adapting itself to the most appropriate scheduling policy according to service QoS preferences and instantaneous performance variations, and 4) intelligently allocating the radio resources to contending sessions based on their degree of instantaneous QoS satisfaction. The proposed AMQ scheme is evaluated in a unidirectional geostationary satellite broadcast system (i.e., SDMB) through extensive analytical and simulation studies; nevertheless, the proposed methodology can also be applied adaptively to other WCDMA-based BC/MC networks.

### 7.2 Motivation for the AMQ scheduling scheme

The motivation and objective of this research is two-fold. The first is to improve the efficiency and QoS performance in the design of service prioritization algorithm. In doing so, we take stock of existing work in this area, highlight their potential problems in terms of both efficiency and fairness, and propose a feasible approach for a more efficient and QoS-based prioritization scheme. The second is to resolve the inefficiency in the design of resource allocation algorithm. We do so by employing an adaptive resource sharing mechanism that is capable of allocating and sharing resources, in a more dynamic and adaptive manner, according to the instantaneous QoS satisfactions of respective sessions.

In order to achieve better packet scheduling performance in terms of both efficiency and fairness, inherited from the proportional delay differentiation (PDD) scheme in the context of differentiated service networks, DDQ was proposed in our early work [116], offering improved performance for delay, jitter, and channel utilisation. However, DDQ experiences unbalanced performance among multiple QoS attributes, namely the gain achieved in one attribute leads to the performance degradation in other attributes. Furthermore, multimedia services feature differentiated delay constraints, applying the delay constraints for differentiated services in an equal way may lead to poor QoS guarantee for high priority queue, therefore the delay profile has to be considered against the respective delay constraints (i.e. maximum acceptable delay threshold) specified by the class of service. Finally, rather than scheduling competing flows in a static manner, to provide more flexible QoS provisioning and maintain optimal resource utilisation, it is highly desired that
the scheduler is capable of choosing the best scheduling policy according to different QoS preferences of the services and instantaneous performance dynamics.

For these reasons, we propose a novel scheme, namely ASP algorithm, at the MAC layer that considers multiple performance criteria across layers in order to adopt the most appropriate packet scheduling policy in response to diverse QoS demands and traffic dynamics. By taking into account the session's traffic priorities, QoS requirements at both application layer and transport layer, and the queuing dynamics induced by heterogeneous traffic at the RLC layer, the proposed ASP can satisfy multiple essential QoS requirements and provide efficient resource utilisation. Moreover, we exploit the desired flexible feature of the ASP in dynamically adapting itself to the most appropriate scheduling policy according to service QoS preferences and instantaneous performance variations.

The traditional resource allocation procedure operates based on the existing SRM technique, where the allocated data rate is based on the maximum data rate supportable for each physical channel. This strategy can only influence long term resource allocation, whilst the short term physical layer data rate variations can dramatically waste system capacity. Since the rate matching functionality is performed at the physical layer in accordance with other physical layer procedures, cross-layer interactions between physical layer and MAC layer are proven to be capable of obtaining performance gain in resource utilisation. A novel DRM scheme has been proposed in [14]. The proposed DRM relies on instantaneous data rate instead of maximum data rate used in the SRM. The rate matching ratio is calculated for every TTI and corresponds to the instantaneous data rate of each TrCH. Based on the novel DRM technique, resource allocation is desired to be performed in conjunction with DRM for higher utilisation efficiency. Therefore, a dynamic DRA scheme was proposed in our previous work [116]. This new resource allocation algorithm uses DRM to select the required transmission power for all physical channels according to their instantaneous data rate requirements. Therefore, it offers two main advantages: 1) it allows better DTX minimization, and 2) it requires less power when the instantaneous data rate is lower than the maximum data rate. However, compared with SRM technique, DRM technique involves more processing and memory.

Previously, resource allocation operated separately with the service prioritization procedure, which provides the serving orders upon scheduling the contending traffic sessions on a TTI-by-TTI basis. Based on the instantaneous supportable data rate derived from the DRM functions, the resources are allocated to the selected FACH queue in a strict-priority based manner, i.e., the tentative TF size is checked and assigned, from the maximum supported TF size to zero, in the highest priority FACH queue prior to the lower priority FACH queues. In this case, the high priority queues are always allocated with resources ahead of their low priority counterparts, TFs
in lower priority queues are only checked when all the TFs in higher priority queue cannot be granted. In this scheme, high priority queues always obtain a high degree of QoS satisfaction, whilst the lower priority queues can only be allocated with resource at the expense of higher priority queues, which leads to inferior performance in terms of both delay and throughput.

To tackle this challenge, we propose a more adaptive and dynamic resource allocation algorithm, namely adaptive resource allocation (ARA) algorithm. The introduction of this scheme will allow low priority queues to be allocated with more bandwidth by moderately utilising the resources which should be assigned to those higher priority queues with enough QoS satisfaction. It is noted that to maintain the QoS satisfaction for high priority queues above their required level, only high priority queues with adequate QoS satisfaction performance at the particular resource allocation interval is eligible for sharing their resources with other lower priority queues with unsatisfied QoS performance. For each resource allocation interval, queues with either high-priority unsatisfied QoS or low-priority satisfied QoS are excluded from the adaptive resource sharing mechanism of the ARA algorithm. To this end, the proposed ARA scheme enables the maximum possible resource sharing between diverse QoS traffic classes at the minimum expense of the performance degradations on high QoS traffic classes.

### 7.3 Adaptive multidimensional QoS-based packet scheduling

#### 7.3.1 Overview

The proposed AMQ scheme takes into account several key performance criteria simultaneously in order to assure comprehensive QoS satisfaction. On one hand, rather than differentiating the competing sessions with respect to their inherent traffic priorities (i.e. service types), the AMQ scheme considers the application prescribed QoS requirements as a combination of multiple attributes. On the other hand, the queuing dynamics of the competing flows at the RLC layer are monitored and evaluated in response to the fast-varying traffic dynamics. The proposed AMQ mechanism operates in the MAC sub-layer of the data link layer within the RRM scheme in SDMB.

As shown in Figure 7-1, the admitted ongoing sessions can comprise multiple MBMS sessions with diverse QoS demands. In SDMB, each session is assumed to retain an individual FACH queue in the RLC queuing buffer. Packets in the FACH queues are prioritized in decreasing order, with input parameters abstracted from both the RRA at the beginning of each session as well as the RLC queuing buffer per-TTI scale. The parameter abstractions are then subject to two formulated mechanisms: service classification and queue differentiation. The former is the QoS
classification of competing service flows, depending on their QoS requirements, performed once during the phase of service establishment or re-negotiation. Whilst the latter keeps track of the queuing dynamics for competing flows during each session’s transmission, which are passed to the packet scheduler on a TTI-by-TTI basis.

To consider both QoS criteria and queuing behaviours, we introduce an adaptive priority function (APF) for handling the contributing parameters from the aforementioned two modules. The parameters involved can be effectively sub-categorised into two main streams: Static priority factor (SPF) and dynamic priority factor (DPF).

- SPF refers to a set of prescribed QoS demands of each service class which is kept constant during the service session transmission. The parameters included in the SPF list are the QoS guarantees expressed in terms of application prescribed QoS rank, required data rate, upper bound on queuing delay/buffer occupancy, and target PLR and throughput.

- DPF refers to the performance criteria which keeps track of queuing status of each queue dynamically and updates on a TTI scale. The DPF parameters represent the dynamic queuing behaviour in terms of queuing delay, queue length, PLR and throughput.

Upon receiving the SPF and DPF parameters on either a per-session or a per-TTI scale, the APF function carries out the ranking and priority derivation process and comes up with a quantified priority associated with each FACH queue for the current TTI. The FACH queue with the highest priority traffic flow is served ahead of the other competing flows. The objective of the AMQ scheme can be identified as: to provide the high-level diverse QoS satisfaction among
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

heterogeneous multimedia services, subject to the system resource and power constraints. The prioritized queues are then passed to resource allocation and are allocated the required resources.

7.3.2 Adaptive service prioritization

To perform the adaptive service prioritization, we define the APF function $\vartheta_j(n)$ for the $j^{th}$ TrCH at the $n^{th}$ TTI as:

$$\vartheta_j(n) = \alpha_j \cdot T_j(n) \cdot A_j(n) \cdot \Gamma_j(n) \cdot \Xi_j(n) \cdot H_j(n), \quad j = 1, \ldots, J; \quad n = 1, \ldots, N. \quad (7.1)$$

where $\alpha_j$ is the application prescribed QoS rank for the $j^{th}$ session, $J$ is the total number of competing FACH queues, $N$ is the total number of TTI during the session transmission. For each of the $n^{th}$ TTI, the instantaneous queuing behaviours in the $j^{th}$ FACH queue can be characterised by a multi-dimensional vector $(T_j(n), A_j(n), \Gamma_j(n), \Xi_j(n), H_j(n))$, which denotes the performance coefficients for queuing delay, buffer occupancy, data rate, packet loss rate and throughput, reflecting the current distances between instantaneous performance and its desired performance.

The first profile involved $\alpha_j$, namely the QoS profile, is essentially a time-independent parameter designated for each queue during the session’s transmission, reflecting the comparative traffic priority level of the service carried by the $j^{th}$ FACH queue. The higher $\alpha_j$ the higher priority of the session is. It is worth noting that the QoS profile is the premier criterion in the APF, which means that for the majority of the time, the high QoS session will be served ahead of their low QoS counterparts. However, this is not necessarily true once one or more performance criteria are degraded to such an extremely severe condition that the scheduler must take immediate action to prevent the corresponding queue suffering undesired performance loss (e.g., buffer overflow, exceptional long queuing delay).

Due to the unidirectional nature of SDMB, the end-to-end delay in the network is not obtainable at the satellite hub. Queuing delay experienced in the RLC buffer is thereby employed in defining the delay-related metric in this chapter. We define the mean queuing delay for the $j^{th}$ FACH queue until the $n^{th}$ TTI as:

$$\tau_j(n) = \frac{\sum_{k \in \Delta_j(n)} \tau_j^k(n) + \sum_{k \in \Theta_j(n)} \tau_j^k(n)}{N_j^1(n) + N_j^4(n)}, \quad j = 1, \ldots, J; \quad n = 1, \ldots, N. \quad (7.2)$$

where $N_j^1(n)$ the number of packets that have left the $j^{th}$ queue before the $n^{th}$ TTI, $N_j^4(n)$ is the number of packets that are queuing in the $j^{th}$ FACH buffer at the $n^{th}$ TTI, $\Delta_j(n) := \{1, 2, \ldots, N_j^1(n)\}$,
$\Theta(n) := \{N_j^1(n) + 1, N_j^2(n) + 2, ..., N_j^q(n) + N_j^q(n)\}$, \(\tau_{j,k}(n)\) is the current queuing delay for the \(k^{th}\) packet arrived in the \(j^{th}\) FACH queue, which is defined as:

$$\tau_{j,k}(n) = \begin{cases} T_{j,k}(n) - T_{j,k}(k) & \text{if } k \leq N_j^1(n) \\ n \cdot T_m - T_{j,k}(k) & \text{if } k > N_j^q(n) \end{cases}$$

where \(n \cdot T_m\) represents current timing, \(T_{j,k}(k)\) and \(T_{j,k}(k)\) denote the arrival time and leaving time for the \(k^{th}\) packet in the \(j^{th}\) queue, respectively.

In SDMB, a queuing delay threshold is assigned to each admitted ongoing session, representing the maximum acceptable queuing delay for the corresponding service class. Let \(\tau^*_j\) denote the maximum acceptable queuing delay for the \(j^{th}\) FACH queue specified by session’s QoS requirements. We associate with each FACH queue \(j\) a queuing delay profile \(T_j(n)\), which is given by:

$$T_j(n) = \begin{cases} 1 & \text{if } \tau_j(n) \leq \tau^*_j \\ \frac{\tau_j(n)}{\tau^*_j} & \text{if } \tau_j(n) > \tau^*_j \end{cases}$$

This attribute is derived based on the maximum queuing delay tolerated by the corresponding service. It proportionally adjusts itself in response to the difference between the mean queuing delay and its delay threshold. It is only effective when the mean queuing delay is beyond the designated delay threshold. It is noted that the delay threshold can be regarded as a tuneable parameter to balance the overall system performance.

Once the finite length buffer at the satellite hub is employed, it is vital, especially for loss-sensitive service, to maintain the queue length at a safe level to prevent the system from excessive packet loss due to buffer overflow. Let \(\lambda^*_j\) denote the maximum buffer length for the \(j^{th}\) FACH queue. A buffer occupancy profile \(\Lambda_j(n)\) is defined for the \(j^{th}\) FACH queue, which is given by:

$$\Lambda_j(n) = \begin{cases} 1 & \text{if } \lambda_j(n) \leq \lambda^*_j \cdot \sigma_j \\ \frac{\lambda_j(n)}{\lambda^*_j \cdot \sigma_j} & \text{if } \lambda_j(n) > \lambda^*_j \cdot \sigma_j \end{cases}$$

where \(\sigma_j\) is the buffer occupancy threshold, providing a safe bound for the buffer length, \(\lambda_j(n)\) denotes the instantaneous queue length of the \(j^{th}\) FACH at current TTI.
The data rate profile is calculated as the ratio of the service required/guaranteed data rate against the mean data rate at current time. The instantaneous priority of each queue is affected proportionally to the difference between the mean transmitted data rate and the required data rate of each queue. Let \( y^* \) denote the guaranteed data rate for the \( j^\text{th} \) FACH queue, the data rate profile of the \( j^\text{th} \) FACH queue is then defined as:

\[
\Gamma_j(n) = \begin{cases} 
1 & \text{if } \frac{\overline{y}_j(n)}{\gamma_j^*} \leq y^*_j \\
\frac{\overline{y}_j(n)}{\gamma_j^*} & \text{if } \frac{\overline{y}_j(n)}{\gamma_j^*} > y^*_j
\end{cases} \quad j = 1, \ldots, J, \quad n = 1, \ldots, N.
\] (7.6)

where \( \overline{y}_j(n) \) denotes the mean data rate of \( j^\text{th} \) FACH achieved up to TTI \( n \), which is determined by:

\[
\overline{y}_j(n) = \sum_{k=1}^{N_{j,n}} S_{j,k} / T_m \quad j = 1, \ldots, J, \quad n = 1, \ldots, N.
\] (7.7)

where \( S_{j,k} \) represents packet size for the \( k^\text{th} \) packet in the \( j^\text{th} \) FACH queue.

Similar to the queuing delay profile, the packet loss available at the satellite hub is also confined to the packet loss due to buffer overflow, although the packet loss in the propagation path is the most crucial factors impacting the QoS performance. Nevertheless, the packet loss due to buffer overflow is the single metric that can be monitored and controlled by the packet scheduling entity. Let \( \xi_j^* \) denote the acceptable PLR due to buffer overflow for the \( j^\text{th} \) FACH queue, \( \delta_j \) is the PLR threshold for the \( j^\text{th} \) FACH queue. The packet loss rate profile \( \Xi_j(n) \) is defined as:

\[
\Xi_j(n) = \begin{cases} 
1 & \text{if } \frac{\overline{\xi}_j(n)}{\xi_j^* \cdot \delta_j} \leq \xi_j^* \cdot \delta_j \\
\frac{\overline{\xi}_j(n)}{\xi_j^* \cdot \delta_j} & \text{if } \frac{\overline{\xi}_j(n)}{\xi_j^* \cdot \delta_j} > \xi_j^* \cdot \delta_j
\end{cases} \quad j = 1, \ldots, J, \quad n = 1, \ldots, N.
\] (7.8)

where \( \overline{\xi}_j(n) \) denotes the mean PLR of the \( j^\text{th} \) FACH achieved until the \( n^\text{th} \) TTI, which is defined as:

\[
\overline{\xi}_j(n) = \frac{N_j^f(n)}{N_j^s(n) + N_j^f(n)} \quad j = 1, \ldots, J, \quad n = 1, \ldots, N.
\] (7.9)

where \( N_j^f(n) \) represents the total number of packets that are dropped due to buffer overflow for the \( j^\text{th} \) FACH up to TTI \( n \).

---

Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

We specify the throughput as the ratio between the total bits arrived in a specific queue up to current time to the total bits that has been successfully scheduled and delivered to the physical channel for radio frame transmission. Let \( \eta_j^* \) denote the target throughput for the \( j^{th} \) FACH queue, the throughput profile \( H_j(n) \) for the \( j^{th} \) FACH queue is given by:

\[
H_j(n) = \begin{cases} 
1 & \text{if } \frac{\eta_j(n)}{\eta_j^* \cdot \varphi_j} \leq \eta_j^* \\
\frac{\eta_j(n)}{\eta_j^* \cdot \varphi_j} & \text{if } \frac{\eta_j(n)}{\eta_j^* \cdot \varphi_j} > \eta_j^* 
\end{cases} 
\]

(7.10)

where \( \varphi_j \) is the throughput ratio threshold of the \( j^{th} \) FACH, \( \eta_j(n) \) denotes the mean throughput of the \( j^{th} \) FACH that has been achieved so far, which is defined as:

\[
\eta_j(n) = \frac{B_j^e(n)}{B_j^a(n)} 
\]

(7.11)

where \( B_j^e(n) \) represents the total number of bits that are successfully scheduled for transmission for the \( j^{th} \) FACH until current TTI, \( B_j^a(n) \) represents the total number of bits that have arrived in the \( j^{th} \) FACH so far.

### 7.3.3 Adaptive resource allocation

Upon receiving the serving orders for all competing flows on a TTI-by-TTI basis, the packet scheduler performs priority-based resource allocation, which is operated by assigning each FACH queue with the most appropriate TF size subject to the system resource/power constraints, namely how many bits can be granted for radio frame transmission. The assignment of TFs for all competing FACH queues are performed with the prior knowledge of a reduced TFCS list, which is derived at the physical layer according to instantaneous data rate of the corresponding session prior to each session’s transmission.

Each TF element in the reduced TFCS list corresponds to a two-dimensional vector: priority of each queue and power requirement expressed by \( E_b/N_0 \). The first factor is determined by the serving order evaluated for FACH queues, and remains the same for each TBS in the TFS of the queue, i.e., one column in TFCS list. The second factor is determined by the BLER requirement for each session and the chosen TBS size, a larger TBS size corresponds to a higher data rate and transmit power requirements.
The assignment of TFs for all the FACH queues during the resource allocation intervals can be performed in either a strict-priority based or a flexible-priority based manner. In the traditional strict-priority based approach, the assignment of TFs starts from the FACH queues with the highest priority to the FACH queues with the lowest priority. The power requirement of each tentative TF, from the maximum supported TF size to each next lowest TF size, are evaluated and compared with the system power constraints until all candidate TFs are exhausted, i.e., the TF size equals zero. Based on the strict-priority based nature, the lower FACH queues can only be assigned when all candidate TFs in higher priority queues are exhausted. Apparently, the scheme may lead to over-satisfied performance for high priority queues and causes a severe starvation problem for those queues with lower priority, especially in the scenario of heterogeneous multimedia service provisioning. Therefore, regardless of their instantaneous QoS satisfaction performance, high priority queues may occupy the majority of bandwidth whilst the low priority queue can only survive at the expense of the high priority queues.

To tackle this challenge, it is highly desired that the assignment of TFs can be performed more adaptively taking account of the instantaneous performance demands. We propose an innovative approach, namely adaptive resource allocation (ARA), which is capable of allocating the resource based on the current performance and QoS satisfactions of respective FACH queues.

The proposed methodology can be summarised as: the resource can be shared between high priority queues with over-satisfied QoS performance and those low priority queues with under-satisfied QoS performance, under the constraints that the QoS demands of high priority queues are guaranteed to be met, i.e. the sharing mechanism will not apply to those high priority queues with under-satisfied QoS performance. The amount of the resource to be shared is proportional to the QoS satisfaction factor of the high priority queues, which is derived on a TTI-scale; the better QoS satisfaction the high priority queue has, the more resources that could be shared with other demanding low priority queues.

In each TTI, based on the current QoS performance and service QoS guarantees, the QoS satisfaction factors are evaluated from the “QoS satisfaction estimation” and passed to the ARA. To measure the current QoS performance for each queue against its QoS demands, we define a QoS satisfaction factor $\Omega_j(n)$ as in (7.12), considering multiple QoS attributes at current TTI for the $j^{th}$ FACH queue, thereby reflecting the current QoS satisfaction status of each queue.

$$\Omega_j(n) = \frac{\tau_j(n)}{\tau_j(n)} \cdot \frac{\lambda_j(n)}{\lambda_j(n)} \cdot \frac{\gamma_j(n)}{\gamma_j(n)} \cdot \frac{\xi_j(n)}{\xi_j(n)} \cdot \frac{\eta_j(n)}{\eta_j(n)}$$

(7.12)
It is noted that, in this chapter, the evaluation applies the average performance metrics from the beginning of a session up to the current TTI, against their QoS constraints. However, this evaluation can be adjusted on a short-term basis, e.g., 10 TTIs, the window size determines the effectiveness and sensitivity of the QoS satisfaction estimation.

The value of the $Q_j(n)$ represents the QoS satisfaction measurement, which is used in the judgement of ARA algorithm as follow:

- $Q_j(n) > 1$: the FACH queue is over-satisfied at current time slot $n$. This means the queue may be less aggressive for acquiring resources, which can be shared by other demanding queues. The amount of resources can be shared are proportional to the value of $Q_j(n)$.

- $Q_j(n) = 1$: QoS satisfaction is exactly satisfied for this FACH queue at current time slot $n$. The queue needs to be assigned with resources at current rate; otherwise, its QoS may not be guaranteed for the following time slots.

- $0 < Q_j(n) < 1$: the FACH queue is currently experiencing under-satisfied QoS, the queue has to be assigned with more resources as the value becomes smaller.

Figure 7-2. Proposed adaptive resource allocation procedures.
The proposed adaptive resource allocation procedures are described in Figure 7-2, and can be summarised as follows:

- Based on the physical layer DRM function, a reduced TFCS list is generated for each FACH queue according to the required $E_b/N_0$ values.

- According to the priority of each FACH queue derived from the APF, the TFS column in the TFCS list are resorted in decreasing order, i.e. the first column of TFS has the highest priority.

- For the first TFS column, the most appropriate power requirements are checked for that FACH queue according to system power constraints, from the highest TF size to zero. The QoS satisfaction are checked as:
  - If the QoS for this queue is unsatisfied, assign the selected TF size to the FACH queue, and thereby notify the packet scheduler with the remaining power. Then the scheduler checks the next adjacent queue until all FACH queues are exhausted.
  - If the QoS for this queue is satisfied, the scheduler determines the available shared resource according to the satisfaction degree of the corresponding session and assigns the corresponding resource for the FACH queue.

- Once all the queues are checked and assigned with appropriate TF size, the TFC of this TTI is determined for radio frame transmission. The selection of TFC is performed periodically in each resource allocation interval, i.e., on a TTI-scale.

![Figure 7-3. An example of TFC selection mechanism under different resource allocation schemes.](image-url)
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

An example of the TFC selection process employing traditional resource allocation and adaptive resource allocation is illustrated in Figure 7-3. With the traditional resource allocation approach, the tentative TF size selection procedure operates in accordance with the system power constraints from the highest priority queue to the lowest priority queue (i.e. FACH 1 to 3), whilst the TF sizes in each queue are checked from the maximum allowable TF size (i.e. corresponds to the required data rate of the FACH channel) to zero TF size. Following this rule, the TF size is finally selected and assigned to each FACH with the maximum possible physical resource until all the system power is exhausted. It is noted that the advantage of ARA lies in the capability of sharing resource in the QoS-satisfied queue with other QoS-unsatisfied queues. For instance, in Figure 7-3, once ARA is applied, having the knowledge of the QoS-satisfied status in FACH 2, the scheduler has assigned a portion of the maximum possible TF size, i.e., 5120 bits from 10240 bits, leaving the remaining 5120 bits to other lower priority QoS-unsatisfied queues (i.e. FACH 3). This intelligent adaptive allocation approach is desirable in that it not only improves QoS performance of low priority queues by assigning more resource to those QoS-unsatisfied queues, but also minimizes the detrimental impact of adaptive resource sharing on the performance of high priority queues to a maximum possible level. The resource adaptive sharing mechanism will be choked once the QoS performance of the high priority queues cannot be guaranteed.

In such a dynamic and adaptive way, the service prioritization and resource allocation cooperate closely with each other towards the QoS targets: the former will keep track of the current QoS performance and assign higher priority for those starved queues with poor QoS performance, whilst the latter will monitor those queues with over-satisfied performance and capture their resources for improving the performance of other demanding queues. By employing both schemes, the packet scheduler not only takes care of those queues with unacceptable QoS performance, but also is capable of prevent those queues with over-satisfied performance from occupying excess resources.

7.3.4 Flexibility and scalability analysis

In the above context, we assume that all the contributing profiles behave and influence the APF in an equal way during the session transmission. However, fixed settings upon all performance criteria may not work well in provisioning multimedia data with different QoS demands and fast-varying traffic dynamics. The performance gain achieved in one profile may sacrifice other profiles, which may be even more important for the specific service. To offer more flexibility and enhance the system performance, the proposed AMQ algorithm provides a tuning mechanism over essential performance profiles to further optimize the scheduling performance. By observing the QoS preferences specified by the service and the behaviours of queuing dynamics, the AMQ
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

scheme dynamically adjusts the following "tuning knobs" on a TTI-scale: 1) queuing delay threshold ($\sigma_j$), 2) PLR threshold ($\delta_j$), and 3) throughput threshold ($\varphi_j$). By selecting an appropriate combination of the above threshold parameters for each FACH queue, the serving orders of competing flows can be effectively managed. According to the sensitivity preferences of differentiated QoS traffic classes, through giving flexible importance to different profiles in terms of delay, PLR and throughput, it is therefore possible to adaptively select the best possible scheduling policy to allow for different treatments of diverse QoS demands and maintain optimal resource utilisation. For example, the $\sigma_j$ is preferred to be set higher for delay-tolerant PLR-sensitive service, whilst preserving a target $\delta_j$, $\varphi_j$. Some applications have stringent constraints on the throughput rather than PLR, thus the scheduler should apply a higher $\varphi_j$ whilst releasing the constraints on $\sigma_j$, $\delta_j$.

From the implementation point of view, the proposed AMQ scheme introduces extra computational complexity due to its nonlinear (with loop iterations for selection sort operations) and nondeterministic (with unpredictable variables) nature. In order to examine the scalability of the proposed AMQ scheme, the Big O notation is employed for determining the involved computational complexity [99]. It is assumed that there are $n$ sessions to be transmitted to UEs in a number of multicast groups, located within multiple sectors of a satellite beam. We consider the computational complexity for AMQ algorithm during one TTI period, with all the tuneable thresholds already assigned for the current TTI. Derived from the worst case scenario, where the processing time is the most expensive among all possible scenarios, with the input size of $n$ (i.e., total number of TrCHs), the involved computational time complexity (running time) required for MLPQ and DDQ are derived as $O(n)$ and $O(n^2)$ respectively, whilst the AMQ scheme requires an overall computational complexity of $O(n^2)$, featuring typical quadratic statistics.

7.4 Performance Analysis

7.4.1 Simulation scenario

The AMQ packet scheduling scheme is physically implemented in the satellite hub (Node-B) within our developed system-level ns2 simulator described in Appendix D, employing the SDMB function. The performance of our proposed scheme was evaluated via simulations over a wide

---

17 It is noted that the adaptive tuning upon performance thresholds can be specified according to some predefined rules. For example, the threshold can be changed adaptively based on the knowledge of the corresponding performance during the previous time period (e.g. the last 10 or 100 TTIs), the updated threshold values are determined by the distance between the current performance and their target thresholds.
variety of traffic mix scenarios. In these scenarios, we consider individual MBMS session with
diverse QoS profiles in terms of service type, data rate, and QoS constraints for broadcast
transmission, each of which is carried by a single FACH queue. Multiple S-CCPCHs are used for
carrying heterogeneous multimedia services and the considered radio bearer mapping scenarios
are given in Table 7-1 to Table 7-3.

Table 7-1. Radio bearer mapping configuration (kb/s) for Scenario I

<table>
<thead>
<tr>
<th>S-CCPCH id</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-CCPCH bit rate</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>FACH id</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Streaming</td>
<td>128</td>
<td>64</td>
</tr>
<tr>
<td>Hot Download</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cold Download</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7-2. Radio bearer mapping configuration (kb/s) for Scenario II

<table>
<thead>
<tr>
<th>S-CCPCH id</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-CCPCH bit rate</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>FACH id</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Streaming</td>
<td>256</td>
<td>128</td>
<td>-</td>
</tr>
<tr>
<td>Hot Download</td>
<td>-</td>
<td>-</td>
<td>256</td>
</tr>
<tr>
<td>Cold Download</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7-3. Radio bearer mapping configuration (kb/s) for Scenario III

<table>
<thead>
<tr>
<th>S-CCPCH id</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-CCPCH bit rate</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>FACH id</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Streaming</td>
<td>-</td>
<td>256</td>
<td>64</td>
</tr>
<tr>
<td>Hot Download</td>
<td>64</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cold Download</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

The traffic mix scenarios consist of different mixtures of service types, i.e. streaming, hot download and cold download, reflecting diverse QoS guarantees of heterogeneous multimedia services.

- Scenario I is formed by 5 FACHs and 2 S-CCPCHs carrying homogenous multimedia service, i.e. streaming, where no QoS-differentiation is envisaged in this scenario, herein the parameter of interest is the different data rates of these streaming services. Therefore, the QoS thresholds (e.g. delay threshold and PLR threshold) settings are same for all the multiplexed FACH queues.

- Scenario II considers two additional lower priority traffic classes: hot download and cold download. However, the scenario is defined as only a single service type is carried by each S-CCPCH. In this case, the adaptive priority function in the packet scheduler plays the key role in differentiating the diverse QoS demands with the consideration of multiple performance criteria.

- Scenario III describes a more complex scenario; where heterogeneous traffic types are carried by arbitrate S-CCPCHs. The task of the packet scheduler not only includes the differentiation of the session within a single S-CCPCH, but also embraces the traffic differentiation between FACHs which are carried by different S-CCPCHs.

We compare the performance of the proposed AMQ packet scheduling with those of MLPQ and DDQ, and analyse several main parameters, which have significant impact upon the overall system performance.

7.4.2 Queuing delay evaluation

![Mean queuing delay for scenario 1](a)
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

(b) CDF delay of FACHs for scenario 1 using MLPQ

(c) CDF delay of FACHs for scenario 1 using DDQ
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

To examine the queuing delay performance for Scenario I, where only streaming sessions are envisaged, we set the simulation parameters to the heavy traffic load scenario, i.e. all queues are non-empty during session transmission time. Figure 7-4 provides an indication of the performance achievement by the AMQ scheme. Since the queuing delay for streaming services features slight differentiation, we demonstrate the performance gain by analysing the CDF of queuing delay, which represents the overall queuing delay performance for each queue over simulation time under various scheduling plans. The overall steepness of the CDF slopes for contending sessions indicates proportionally the scalability and applicability of the scheduling mechanism. The performance shows that, for homogenous traffic mix, the MLPQ does not provide any differentiation for competing flows, i.e. queues are served in round robin manner, which leads to the queuing delays being determined mainly by their packet arrival rates. For example, when MLPQ is applied, FACH 4 carried with 256kbps streaming experiences much longer queuing delay than other low rate queues. From the CDF distribution curves for MLPQ in Figure 7-4 (b), the CDF curves for the higher rate sessions (i.e. FACHs 1 and 4) feature significantly less steep
slopes than the slopes of other sessions, in that they have the highest data rate amongst all sessions within the respective S-CCPCHs, and thereby suffer from much more incoming traffic. However, when MLPQ is employed in this case, the round robin fashion offers arbitrarily fairness amongst all sessions, giving poor queuing performance for those high data rate sessions. With the involvement of delay differentiation, the DDQ is capable of balancing the queuing delay performance of all queues and mitigating the unfavourable performance caused by different arrival rates. From the CDF distributions in Figure 7-4 (c), compared with Figure 7-4 (b), the steepness of the slope are increased for higher data rate queues (i.e. FACHs 1 and 4) while the sacrifices are the slight decreased slope steepness for lower rate session (i.e. FACHs 2 and 3). However, when multiple QoS aspects are taken account of, notable performance improvements are shown in Figure 7-4 (d) compared with Figure 7-4 (b) and (c), the AMQ can provide even better queuing delay performance. Numerically, the CDF delay of AMQ achieves similar steepness of slope for all queues with the upper bound of 3.2 seconds, whilst the curves in MLPQ and DDQ case appear to be a more variable (i.e. steeper for lower rate session and smoother for higher rate sessions) with upper bounds of 28 and 10 seconds, which means the queuing delays are distributed over longer intervals.

![Mean queuing delay for scenario II](image)

<table>
<thead>
<tr>
<th></th>
<th>FACH 1</th>
<th>FACH 2</th>
<th>FACH 3</th>
<th>FACH 4</th>
<th>FACH 5</th>
<th>FACH 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPQ</td>
<td>1.17</td>
<td>0.30</td>
<td>2.56</td>
<td>0.31</td>
<td>14.89</td>
<td>5.81</td>
</tr>
<tr>
<td>DDQ</td>
<td>0.68</td>
<td>0.69</td>
<td>2.51</td>
<td>2.41</td>
<td>3.58</td>
<td>3.25</td>
</tr>
<tr>
<td>AMQ</td>
<td>0.51</td>
<td>0.52</td>
<td>0.83</td>
<td>0.91</td>
<td>1.88</td>
<td>1.35</td>
</tr>
</tbody>
</table>

(a)
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

PDF delay of FACHs for scenario II using MLPQ

Packet Delay (seconds)

PDF delay of FACHs for scenario II using DDQ

Packet Delay (seconds)
The effectiveness of AMQ in terms of the QoS differentiation is analysed in Figure 7-5, where Scenario II is selected as a representative of traffic mixes with diverse QoS ranks. The PDF of queuing delay is analysed to compare the performance of different scheduling schemes. From Figure 7-5 (a), MLPQ achieves the best delay performance for streaming services and the worst delay performance for download services because of its strict-priority based scheduling mechanism. DDQ, with its delay-differentiation nature, provides improved performance for low class services by sacrificing the performance of high class sessions. By adaptively utilising the resources in an intelligent manner, the AMQ has managed to maintain even lower queuing delays for all sessions in that the available resources are efficiently utilised and shared between the contending flows. From the PDF distributions in Figure 7-5 (b)-(d), it is further shown that the AMQ spreads the queuing delay probability distributions over a fairly small time interval with an upper delay bound of 2.25 seconds and thereby improves the overall applicability and stability of the system.
Chapter 7. Adaptive Multi-dimensional QoS-based Packet Scheduling

Figure 7-6. Queuing delay performance at the RLC buffer in the satellite hub under different scheduling schemes for Scenario III: (a) mean queuing delay, (b) mean queuing jitter for different scheduling schemes

The performance of queuing delay and its probability distribution are analysed in the above context. Subsequently, we investigate the queuing delay variation (i.e. queuing jitter) performance under Scenario III, which provides the most complex traffic mix scenario. In Figure 7-6 (a), the queuing delay performance for AMQ is compared to those of MLPQ and DDQ for Scenario III, where the streaming and download sessions are multiplexed in a combined manner for different S-CCPCHs. Rather than achieving lower download queuing delay by sacrificing streaming delay performance in the DDQ case, the AMQ scheme manages to deliver download sessions with lower queuing delay whilst maintaining similar performance in streaming sessions. Specifically, compared with MLPQ scheduling, DDQ reduces the mean queuing delay of download services by 28.5% with a considerable increasing of 68.9% on the mean queuing delay of streaming services, whilst AMQ is capable of reducing the queuing delay of download services by 35.6% with a 16.4% increase of the queuing delay of the streaming service. As seen from Figure 7-6 (b), although the download services (i.e. FACHs 1 and 6) suffer from higher jitter for the AMQ case
compared with the DDQ case, a considerable performance gain on the queuing jitter, i.e. increased by 35.8% and 28.1% compared with MLPQ and DDQ respectively, is achieved for all streaming services, which makes it an attractive solution for real-time jitter-sensitive streaming service.

Analysis of the queuing delay performance in the above context demonstrates that the proposed AMQ packet scheduling is capable of offering favourable performance improvements under various traffic mix scenarios. It is indicated that the impact of AMQ on the queuing delay performance can be identified as: 1) for homogenous QoS traffic class, the AMQ provides the best queuing delay performance, compared with those of existing schemes, with a good fairness score achieved in terms of queuing delay distribution, 2) for heterogeneous QoS traffic class, unlike DDQ, which improves performance of low class services at the expense of the performance of high class sessions, the AMQ can effectively offer reasonably lower queuing delays for all sessions regardless of their QoS rank. 3) AMQ has managed to achieve better performance on the queuing jitter for streaming service, which makes it an attractive approach for delivering real-time jitter-sensitive multimedia content.

7.4.3 Delay threshold flexibility

To examine the adaptive feature of the proposed AMQ scheme, variable delay thresholds are applied to illustrate performance improvements, 4 types (i.e. Types A to D) has been identified in order to specify different delay threshold preferences for respective QoS classes, giving flexible importance to streaming service and hot/cold download service. Type A is the base delay...
threshold setting which is applied by default for our simulation study and performance variations are presented in comparison to its performance in Type A. On one hand, Type B gives more priority to streaming service via assigning it with a more stringent delay threshold value. In fact, this setting essentially reduces the relative priorities of other QoS classes and degrades their delay performance. On the other hand, releasing the delay threshold of lower QoS class can be regarded as another effective approach to differentiate the delay performance amongst heterogeneous traffic flows. For example, Type C increases the importance of both streaming and cold download services via increasing the delay threshold of hot download services.

It is noted that only the streaming class is envisaged in Scenario I, i.e. no differentiation is made on the delay thresholds amongst competing sessions. To this end, we evaluate the mean queuing delay performance for Scenarios II and III via tuning the delay threshold parameters for their QoS classes. Figure 7-7 presents the range of performance gain obtained by tuning the delay threshold values for respective QoS service classes amongst competing flows.

As shown in Figure 7-7 (a) Type B, streaming service is given more priority via assigning a more stringent delay threshold from 0.1 seconds to 0.02 seconds, therefore the mean queuing delay experienced for streaming service (FACHs 1 and 2) is greatly reduced by 55% and 60%, at the expense of higher delay suffered in download services (FACHs 3 to 6). When we compare Type C with Type A, by increasing the delay threshold of hot download service (FACHs 3 and 4), it is appeared that the delay performance of both streaming and cold download are improved.

Figure 7-7 (b) shows the delay threshold performance evaluated for Scenario III. In comparison to Type A, cold download FACH 6 suffers from worse delay (i.e. from 8.7 seconds to 10.6 seconds) in Type D when its delay threshold is increased from 0.8 second to 2.0 second, but this leads to the performance gain achieved on both streaming and hot download FACHs. The results show that the AMQ is capable of balancing the delay performance amongst all sessions, namely, the delay can be reduced for specific QoS class by either tightening the delay threshold of the corresponding QoS class or loosening the delay threshold of other QoS classes.

AMQ has adaptive tuning characteristics, in terms of not only the delay threshold, but also the PLR and throughput threshold. Via flexible considerations to different profiles according to the sensitivity preferences of differentiated QoS classes and their instantaneous performance, the AMQ packet scheduling is shown to be able to adaptively select the most appropriate scheduling policy which allows different treatments to diverse QoS demands whilst maintaining optimal resource utilisation.
7.4.4 Physical channel utilisation

Table 7-4. Mean physical channel utilisation for different PS schemes (%)

<table>
<thead>
<tr>
<th>Scenario I</th>
<th>S-CCPCH id</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPQ</td>
<td>89.4</td>
<td>94.3</td>
<td></td>
</tr>
<tr>
<td>DDQ</td>
<td>91.8</td>
<td>92.7</td>
<td></td>
</tr>
<tr>
<td>AMQ</td>
<td>96.7</td>
<td>95.9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario II</th>
<th>S-CCPCH id</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPQ</td>
<td>88.2</td>
<td>68.1</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>DDQ</td>
<td>81.6</td>
<td>71.5</td>
<td>51.6</td>
<td></td>
</tr>
<tr>
<td>AMQ</td>
<td>85.8</td>
<td>83.5</td>
<td>76.7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario III</th>
<th>S-CCPCH id</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLPQ</td>
<td>89.1</td>
<td>83.8</td>
<td>62.3</td>
<td></td>
</tr>
<tr>
<td>DDQ</td>
<td>85.6</td>
<td>82.5</td>
<td>79.9</td>
<td></td>
</tr>
<tr>
<td>AMQ</td>
<td>93.3</td>
<td>90.6</td>
<td>89.5</td>
<td></td>
</tr>
</tbody>
</table>

The impact of channel utilisation with the AMQ scheme is studied in this section. It is worth mentioning that the performance of the proposed AMQ scheme on the physical channel utilisation is affected heavily by the incoming traffic dynamics, radio bearer configurations, and maximum transmit power settings. With its QoS-differentiation nature, AMQ is capable of offering better resource utilisation as the complexity of QoS class mixture increases, and this affect is more noticeable for those physical channels carrying lower QoS traffic classes. From the results shown in Table 7-4, by adaptively re-utilising wasted resources among diverse QoS class services, AMQ manages to offer better physical channel utilisation compared with those of the existing schemes.

In Scenario I, where traffic mix is limited to single streaming traffic class, i.e., no QoS-differentiation is envisaged for the scheduling plan, MLPQ can be regarded as a simple round robin discipline; therefore the discrepancy featured in the physical channel utilisation is mainly due to the incoming traffic dynamics. DDQ is able to provide primary differentiation in terms of its queuing delay behaviour and the impact of traffic dynamics on the channel utilisation is largely eliminated, i.e. smaller difference on S-CCPCH channel utilisation. The proposed AMQ scheme, with its adaptive resource sharing capability, achieves better overall resource utilisation; the mean physical channel utilisation for all S-CCPCHs reaches 96.3%, compared with 91.9% and 92.2% in MLPQ and DDQ, respectively.

From the results obtained for Scenario II, where the single QoS traffic classes are carried within a single S-CCPCH, we can see that by adaptively utilising the resource in an intelligent way, the
AMQ achieves higher resource utilisation than other schemes in that it either utilises the wasted resources or re-utilises/shares the resources in a more efficient way. The overall channel utilisation has been improved by 48.2% and 20.6% from MLPQ and DDQ respectively. The performance enhancement on channel utilisation can also be seen in Scenario III, where the traffic mix is the most complex. AMQ outperforms MLPQ and DDQ with a significant saving on radio resources and thereby greatly improves the physical channel utilisation. As the complexity of traffic mix increases, AMQ perform better on all the contending sessions irrespective of their QoS classes, for example, AMQ achieves channel utilisation at 93.3%, 90.6% and 89.5% in Scenario III, all of which are higher than those of other schemes. In conclusion, the results in Table 7-4 show that AMQ is capable of achieving a high utilisation ratio for all QoS class sessions, compared with the existing MLPQ and DDQ schemes.

7.4.5 Fairness analysis

In this section, we examine the fairness score achieved by the proposed AMQ scheme, hereby the fairness is referred to the overall achieved performance and QoS guarantees among diverse QoS traffic classes. Hence, we present the fairness performance in the most complex traffic mix scenario, i.e. Scenario III, which groups a wide range of traffic priority classes with different QoS demands.
Figure 7-8 describes the comparison of PDF for the instantaneous allocated TFC in Scenario III, namely the instantaneous allocated bits for each FACH that are to be transmitted over physical channels. As shown in the figure, the download services (FACHs 1 and 6) transmit more capacity in AMQ, since the probability is increased towards high TFC size, whilst the TFC applied for streaming services is more evenly shared amongst all available TFC sizes. For example in FACH 6, which carries cold download service with 384 kbps and corresponds to a mean TFC size of 30720 bits. In MLPQ, with its lowest priority assigned, FACH 6 carries no traffic (i.e. TFC size equals to zero) for 64 percentage of time, whilst the maximum TFC size (i.e. 30720 bits) is achieved in this queue for only 28 percentage of time. By applying delay differentiation in the DDQ case, the maximum TFC size is achieved over 75 percentage of time with the channel idle duration of 12%. The proposed AMQ scheme, by giving flexible importance to the performance
metrics and resource sharing capability, is able to transmit maximum TFC size for up to 90 percentage of the time, whilst the transmission idle time is drastically reduced to less than 5%. The performance results indicates that the AMQ balances all the FACHs by flexibly and adaptively choosing the TF size for FACHs rather than the strict-priority based MLPQ and delay-differentiated DDQ cases, and thereby improves the overall throughput.

Herein the main parameter of interest so far considered is the throughput ratio, which is obtained by dividing the total bits successfully scheduled with the total bits arrived at the scheduler. The variance of individual throughput ratios is considered as another important metric representing the fairness of a packet scheduling algorithm in the time domain; lower variance means a fairer scheduling scheme. In Figure 7-9, the performance of the proposed scheme is compared to MLPQ and DDQ in terms of fairness, where the AMQ achieves the lowest variance values with the fastest convergence, which means that it can provide UEs with better throughput equality in a shorter time. Numerically, as time elapses, the throughput variances of MLPQ and DDQ scheduling suffer from frequent and high max-min variations of 0.05 and 0.03 seconds. However, AMQ is capable of reducing variance of throughput ratio to 0.1 seconds within 700 TTIs (i.e. 56 seconds) and maintaining the throughput variance below 0.015 seconds with relatively low max-min variation, i.e. below 0.01 seconds.
7.5 Summary

To provide more efficient QoS provisioning and resource utilisation whilst preserving the system power/load constraints, a novel AMQ packet scheduling scheme is proposed for the SDMB system in this work. By taking into account the impact of multiple essential performance aspects simultaneously, the proposed AMQ scheme not only satisfies diverse QoS demands, but also is capable of adopting the best possible scheduling policy according to traffic priorities, QoS preferences and queuing dynamics. Moreover, the instantaneous QoS satisfactions of competing sessions are effectively considered in the resource allocation procedure, which leads to essential performance improvements in terms of both channel utilisation and throughput.

The proposed packet scheduling scheme employs an adaptive mechanism to both service prioritization and resource allocation procedures, giving flexible and controllable importance to competing sessions with diverse demands according to the distance between the instantaneous performance and their prescribed QoS guarantees, which, under the heterogeneous multimedia traffic scenario, offers desired flexibility and scalability characters. The proposed scheme was implemented in the satellite hub for the SDMB system, its performance was evaluated via simulation studies over extensive traffic mixes for indicative performance metrics. The results show that, compared with the existing schemes, the AMQ packet scheduling scheme is capable of achieving considerable performance gain on queuing delay/jitter, throughput, channel utilisation and fairness.
Chapter 8

8 Conclusions and Future Work

8.1 Thesis conclusions

This thesis focuses on the research and development of efficient RRM schemes in the SDMB system. The SDMB system defines a unidirectional multimedia delivery system via satellite broadcasting. The unique nature of SDMB poses challenges for the design of a feasible and efficient RRM scheme. By investigating these challenges and problems with existing RRM schemes, we propose and develop new advanced RRM schemes for higher resource utilisation and better QoS guarantees.

Our first essential contribution focuses on the optimization techniques for RRA, where we develop a novel 2-level channel multiplexing scheme. Two optimization algorithms, namely OEA and POA, are proposed to optimize the channel utilisation and minimize the power consumption. The performance of the 2-level channel multiplexing is evaluated through extensive analytical and simulation studies. Comparing with the existing single-level channel multiplexing, the new approach achieves significant performance gain in various aspects such as channel utilisation, power consumption and transmission capacity.

A significant part of this thesis evolves around one of the most important resource management functions in the SDMB, namely the packet scheduling. The performance of the scheduling is greatly improved when the proportional differentiation concept are introduced into packet scheduling. This innovative concept is implemented and examined via the proposal of BLRQ and DDQ, which consider the buffer occupancy and the queuing delay in their scheduling decisions. These schemes are capable of differentiating traffic flows with the same QoS class on the basis of diverse performance metrics, e.g., buffer length, queuing delay and QoS guarantee, and thereby facilitate low QoS class traffic queues with unfavourable performance. Numerically, the average reduction on the mean queue length of streaming services achieved by BLRQ reaches 90.9%; whilst the DDQ reduces the mean queuing delay for download services by 24.4% and maintains similar performance for streaming services.

The cross-layer approach for packet scheduling is examined by utilising important information across protocol layers. Simulation results show that the cross-layer packet scheduling scheme
enjoys a performance gain on queuing delays of download services of 50%-70%, whilst maintaining similar delay performance for streaming services.

Finally, we investigate an AMQ packet scheduling scheme via prioritizing queues according to multiple essential performance metrics. To take into account the traffic QoS preferences and instantaneous queuing dynamics, an adaptation concept is introduced for both service prioritization and resource allocation. The adaptive sharing mechanism is capable of giving flexible adaptive importance to different performance measures, based on the service QoS requirement and instantaneous performance variations. From extensive simulation studies, the propose AMQ scheme can greatly improve the performance of the baseline scheme on multiple performance metrics whilst offering good scalability and flexibility features.

![Queuing delay for different scheduling schemes](image)

**Figure 8-1. Comparison of queuing delay for different scheduling schemes for Scenario III.**

Fig. 8-1 compares the queuing delay performance for different scheduling scheme under Scenario II, we found that each stage of our proposed schemes provides performance improvements against previous one. And AMQ achieves the best performance amongst all the schemes, e.g. the mean queuing delay in AMQ is reduced by an average of 26.8% for all FACHs when compared with CJPQ.

By investigating respective RRM functions via different aspects, various efficient RRM schemes are proposed and evaluated in this research. It is shown that the overall performance of the SDMB system is improved at each stage of these optimization techniques and the performance gains achievable for these schemes can be variable in different traffic mix scenarios and flow dynamics.

From the implementation point of view, the proposed scheduling schemes will introduce more computational complexity due to increased dimensions and functions. However, in employing the Big O Notiation for examining their long-term complexities, the increased parameters are involved with either addition or sequence equations, which add extra linear elements (i.e. $O(n)$) in
the overall complexity function, with \( n \) approaching the infinity, the most demanding element \( O(n^2) \) will be the final overall complexity. When compared with MLPQ, which has a complexity of \( O(n) \), the proposed DDQ, CDRD, CJPQ and AMQ have the same long-term complexity of \( O(n^2) \). Therefore, given relatively small number of multiplexed service flows in the SDMB system, the AMQ is suggested as the most efficient scheduling scheme with acceptable involved complexity.

8.2 Future work

8.2.1 Cross-layer RRM investigating TCP-MAC interactions.

Due to the inherent nature of wireless transmission, satellite communications suffer from strong variations of the received signal power caused by shadowing and multipath fading. Shadowing of the satellite signal is due to obstacles in the propagation path (buildings, trees, bridges, etc). Whereas for multipath, the fading occurs because the satellite signal is received not only via the direct LOS path but also being reflected from objects in the surrounding area. The difference in propagation distances for the multipath signals may add destructively and lead to deep fades. Unlike its terrestrial counterpart, the design of the RRM scheme in the satellite environment cannot rely on better utilisation of the instantaneous information reflecting frequent channel variations, since its long propagation delay for a GEO satellite makes it impossible to utilise the CSI from lower layer. Therefore, we suggest a cross-layer approach for SDMB to utilise information from higher layers of protocol stack, e.g., application layer and transport layer.

One promising solution of cross-layer RRM in SDMB could be a TCP-driven MAC scheme. The transport layer is in charge of establishing end-to-end network connections and maintaining target transmission quality and reliability. For example, TCP will deem large delays or packet losses as a signal reflecting the wireless channel congestion status and thereby adjust its mechanism accordingly. However, in satellite communication system, large delays or packet loss event occur more frequently than terrestrial case, therefore, appropriate mechanism has to be designed to avoid TCP misunderstanding these indicative signals. MAC protocols play a fundamental role in guaranteeing good performance to higher-layer functions, by managing the arbitration of radio access. In fact, decisions made within the satellite RRM in MAC can significantly impact the end-to-end performance of TCP flows over a satellite network. Some research effort has been devoted to this subject nowadays, as in [51], [120], [121], [122]. As indicated in [120], increasing MAC level retransmissions can effectively avoid TCP retransmission and improve the power
8.2.2 SDMB with return links supporting interactive applications

Another research challenge foreseen is that of supporting interactive applications over satellite broadcasting network, which can be regarded as a promising solution in delivering future advanced multimedia applications. Previously, the return link from mobile users to the S-RAN was not envisaged in the baseline SDMB system and the S-RAN has to perform the resource allocation without knowledge of CSI for different users. With the growing demand for supporting advanced multimedia applications, it is highly desired that the return link can be exploited in future SDMB systems for providing two-way interactive transmissions. As a follow on from this direction, research can be conducted to investigate both the system infrastructure and the algorithm optimization for an efficient delivery of interactive multimedia content to mobile users via SDMB with return links, on either terrestrial or satellite components.

It is noted that the schemes considered in this thesis are designed for the unidirectional SDMB system; they may not perform at an optimum level when applied to bi-directional systems. Other optimization techniques can be carried out within this scope. For example, instantaneous CSI estimated at the mobile devices could be exploited in conjunction with the resource management scheme, aimed at offering improved performance on delay/jitter, packet loss as well as throughput and channel utilisation. The proposed satellite communication infrastructure could provide two-way high-speed high-quality transmissions, supporting a variety of multimedia applications, such as interactive TV/video broadcasting, video/telephone conferences, disaster recovery and emergency broadcasting. This innovative concept of providing interactive services in advanced SDMB system will have major impact on the satellite broadcasting industry.

Given the absence of a return link in the current SDMB architecture, reliable transport is provided mainly by FEC function, therefore only UDP-based applications can be supported in the baseline SDMB system. When the return link via the terrestrial/satellite network infrastructure is adopted in future designs of the SDMB system, reliable transport protocol(TCP/RMTP) based applications, such as FTP (File Transfer Protocol), HTTP (HyperText Transfer Protocol - WWW), TELNET (TELeteyp NETwork), SMTP (Simple Mail Transfer Protocol) and etc., are expected to be supported in SDMB. In this way, supporting interactive applications in SDMB via feasible RRM strategies deserves further investigation.
Publications

Chapters in books


Journal articles


Publications

Conference proceedings


Other publications


Bibliography

[26] EU IST FP5 VIRTUOUS project website: http://www.ebanet.it/virtuous.htm
[27] EU IST FP5 FUTURE project website: http://www.ebanet.it/future.htm
[28] EU IST FP5 MOBILITY project website: http://www.rose.es/mobility/
[33] ESA ARTES 5 Advanced Satellite UMTS testbed (ATB) project description: http://telecom.esa.int/telecom/www/object/index.cfm?fobjectid=617
[36] EU IST FP6 SatNEx project website: http://www.satnex.org/


[54] Li-Chun Wang, Yi-Cheng Chen and Ming-Chi Chen, “A cross-layer cost-function based rate adaptation mechanism for the WCDMA system with multi-class services by transport format selection”, Wireless Communications and Mobile Computing (Wiley Inter Science), Vol. 5, No 4, pp. 397 406, July 2005.


[58] 3GPP 25.211 V6.2.0: “Physical channels and mapping of transport channels onto physical channels (FDD)”, September 2004.


[101] ETSI TS 101 851-1, v1.2.1 Satellite Earth Stations and Systems (SES); Satellite Component of UMTS/IMT2000; G-family; Part 1: Physical channels and mapping of transport channel into physical channels (S-UMTS-A 25.211).


[113] ns website, http://www.isi.edu/nsnam


Appendix A: Streaming Service Characterisation

The first step of the simulation is to estimate the streaming service characterisation. As we have described in the Chapter 4, the characterisation for each individual streaming service can be calculated from the 3-tuple – bit rate, system arrival rate and the service duration of the streaming service. The following section shows the evaluation results for the scenarios defined in the last chapter. The evaluation results in this section will be used as the input parameters for the estimation of required MTCHs and FACHs.

In order to evaluate the performance of 2-level channel multiplexing, simulation has been carried out for a wide range of different traffic mixes and physical channel capacities. The S-CCPCH capacity configurations selected are: 3 x 384 kbps; 1 x 384 kbps; 3 x 128 kbps; 1 x 384kbps + 3 x128 kbps.

For each physical channel configuration, different traffic mixes of x% streaming service and (1-x)% download service are considered: 80%-20%; 50% - 50%; 20% - 80%.

A.1 Scenario 1: 3 S-CCPCH with 384kbps each

A.1.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
<th>Normalized load per application type (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>2.75</td>
<td>7.333</td>
<td>0.637</td>
<td>2.225</td>
</tr>
<tr>
<td></td>
<td>64</td>
<td></td>
<td>3.43</td>
<td>18.29</td>
<td>1.588</td>
<td></td>
</tr>
<tr>
<td>Video broadcast</td>
<td>64</td>
<td>300</td>
<td>1.67</td>
<td>8.907</td>
<td>0.773</td>
<td>2.940</td>
</tr>
<tr>
<td></td>
<td>128</td>
<td></td>
<td>0.12</td>
<td>1.28</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td></td>
<td>256</td>
<td></td>
<td>1.11</td>
<td>23.68</td>
<td>2.056</td>
<td></td>
</tr>
</tbody>
</table>

Table A-1: Streaming service characterisation for 80% streaming -20% download given 3x384kbps S-CCPCH

A.1.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
<th>Normalized load per application type (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>0.21</td>
<td>0.56</td>
<td>0.0486</td>
<td>0.1551</td>
</tr>
<tr>
<td></td>
<td>64</td>
<td></td>
<td>0.23</td>
<td>1.2257</td>
<td>0.1056</td>
<td></td>
</tr>
</tbody>
</table>
### Appendix A: Streaming Service Characterisation

#### A.1.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
<th>Normalized load per application type (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>0.08</td>
<td>0.2133</td>
<td>0.0185</td>
<td>0.0185</td>
</tr>
<tr>
<td>Video broadcast</td>
<td>64</td>
<td>300</td>
<td>0.0011</td>
<td>0.0059</td>
<td>0.0005</td>
<td>0.1116</td>
</tr>
</tbody>
</table>

#### A.2 Scenario 2: 3 S-CCPCH with 128kbps each

#### A.2.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
<th>Normalized load per application type (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>1.25</td>
<td>3.1111</td>
</tr>
<tr>
<td>Video broadcast</td>
<td>64</td>
<td>300</td>
<td>1.34</td>
<td>7.1467</td>
<td>1.8611</td>
<td></td>
</tr>
</tbody>
</table>

#### A.2.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load(kbps)</th>
<th>Normalized load per application (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>1.25</td>
</tr>
<tr>
<td>Audio broadcast</td>
<td>128</td>
<td>300</td>
<td>0.12</td>
<td>1.28</td>
<td>0.33</td>
</tr>
</tbody>
</table>

#### Table A-2: Streaming service characterisation for 50% streaming -50% download given 3x384kbps

#### Table A-3: Streaming service characterisation for 80% streaming -20% download given 3x384kbps

#### Table A-4: Streaming service characterisation for 80% streaming -20% download given 3x128 kbps

#### Table A-5: Streaming service characterisation for 50% streaming -50% download given 3x128 kbps

### Appendix A: Streaming Service Characterisation

A.2.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>0.625</td>
</tr>
</tbody>
</table>

Table A-6: Streaming service characterisation for 20% streaming -80% download given 3x128 kbps S-CCPCH

A.3 Scenario 3: 1 S-CCPCH with 384kbps

A.3.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>1.25</td>
</tr>
<tr>
<td>Audio broadcast</td>
<td>64</td>
<td>300</td>
<td>1.34</td>
<td>7.1467</td>
<td>1.8611</td>
</tr>
<tr>
<td>Video broadcast</td>
<td>64</td>
<td>300</td>
<td>0.28</td>
<td>1.4933</td>
<td>0.3889</td>
</tr>
<tr>
<td>Video broadcast</td>
<td>128</td>
<td>300</td>
<td>0.12</td>
<td>1.28</td>
<td>0.3333</td>
</tr>
</tbody>
</table>

Table A-7: Streaming service characterisation for 80% streaming -20% download given 1x384 kbps S-CCPCH

A.3.2 50% streaming-50% download given 1x384kbps S-CCPCH

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>1.25</td>
</tr>
<tr>
<td>Audio broadcast</td>
<td>128</td>
<td>300</td>
<td>0.12</td>
<td>1.28</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table A-8: Streaming service characterisation for 50% streaming -50% download given 1x384 kbps S-CCPCH

A.3.3 20% streaming-80% download given 1x384kbps S-CCPCH

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>0.625</td>
</tr>
</tbody>
</table>

Table A-9: Streaming service characterisation for 20% streaming -80% download given 1x384 kbps S-CCPCH
A.4 Scenario 4: 1S-CCPCH of 384kbps and 3 S-CCPCH with 128kbps each

A.4.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
<th>Normalized load per application type (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>2.71</td>
<td>7.2267</td>
<td>0.9410</td>
<td>1.1007</td>
</tr>
<tr>
<td>Audio broadcast</td>
<td>64</td>
<td>300</td>
<td>0.23</td>
<td>1.2267</td>
<td>0.1597</td>
<td></td>
</tr>
<tr>
<td>Video broadcast</td>
<td>64</td>
<td>300</td>
<td>0.0462</td>
<td>0.2464</td>
<td>0.0321</td>
<td>0.0741</td>
</tr>
<tr>
<td>Video broadcast</td>
<td>128</td>
<td>300</td>
<td>0.019</td>
<td>0.2027</td>
<td>0.0264</td>
<td></td>
</tr>
<tr>
<td>Video broadcast</td>
<td>256</td>
<td>300</td>
<td>0.0056</td>
<td>0.1195</td>
<td>0.0156</td>
<td></td>
</tr>
</tbody>
</table>

Table A-10: Streaming service characterisation for 80% streaming-20% download given 1x384kbps + 3x128 kbps S-CCPCH

A.4.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load (kbps)</th>
<th>Normalized load per application (percent)</th>
<th>Normalized load per application type (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>1.8</td>
<td>4.8</td>
<td>0.625</td>
<td>2.25</td>
</tr>
<tr>
<td>Audio broadcast</td>
<td>64</td>
<td>300</td>
<td>2.34</td>
<td>12.48</td>
<td>1.625</td>
<td></td>
</tr>
<tr>
<td>Video broadcast</td>
<td>64</td>
<td>300</td>
<td>1.28</td>
<td>6.83</td>
<td>0.889</td>
<td>1.056</td>
</tr>
<tr>
<td>Video broadcast</td>
<td>128</td>
<td>300</td>
<td>0.12</td>
<td>1.28</td>
<td>0.167</td>
<td></td>
</tr>
</tbody>
</table>

Table A-11: Streaming service characterisation for 50% streaming-50% download given 1x384kbps + 3x128 kbps S-CCPCH

A.4.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>Application type</th>
<th>Guaranteed bit rate (kbps)</th>
<th>Mean service duration (sec)</th>
<th>System-level arrival rate (times per hour)</th>
<th>Load(kbps)</th>
<th>Normalized load per application (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio broadcast</td>
<td>32</td>
<td>300</td>
<td>0.08</td>
<td>0.2133</td>
<td>0.0278</td>
</tr>
<tr>
<td>Audio broadcast</td>
<td>128</td>
<td>300</td>
<td>0.09</td>
<td>0.96</td>
<td>0.1250</td>
</tr>
</tbody>
</table>

Table A-12: Streaming service characterisation for 20% streaming-80% download given 1x384kbps + 3x128 kbps S-CCPCH
Appendix B: Radio Bearer Configuration

B.1 Required number of streaming MTCHs and FACHs

In this step of the simulation, the RRA invokes well-known results of classical queuing theory to estimate the required number of MTCHs that can guarantee the target service blocking probability. The value of the system arrival rate is finally accepted when the number of MTCHs estimated from the RRA satisfies the target blocking probability and tallies with the number of MTCHs determined earlier. The required number of streaming MTCHs and FACHs varies for a given service group/characterisation and target QoS requirements. QoS requirement is implemented by means of limiting the maximum allowable blocking probability in the algorithm. By employing the 2-stage bin-packing with optimum estimation algorithm described in Chapter 4, the required FACHs can be estimated, and which will be mapped down to S-CCPCHs.

B.1.1 Scenario 1: 3 S-CCPCH with 384kbps each

B.1.1.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Table B-1: Required streaming MTCHs and FACHs for 80% streaming-20% download given 3x384 kbps S-CCPCH

B.1.1.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>2</td>
</tr>
</tbody>
</table>

Table B-2: Required streaming MTCHs and FACHs for 50% streaming-50% download given 3x384 kbps S-CCPCH
Appendix B: Radio Bearer Configuration

B.1.1.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-3: Required streaming MTCHs and FACHs for 20% streaming-80% download given 3x384 kbps S-CCPCH

B.1.2 Scenario 2: 3 S-CCPCH with 128kbps each

B.1.2.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-4: Required streaming MTCHs and FACHs for 80% streaming-20% download given 3x128 kbps S-CCPCH

B.1.2.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-5: Required streaming MTCHs and FACHs for 50% streaming-50% download given 3x128 kbps S-CCPCH

B.1.2.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-6: Required streaming MTCHs and FACHs for 20% streaming-80% download given 3x128 kbps S-CCPCH
Appendix B: Radio Bearer Configuration

B.1.3 Scenario 3: 1S-CCPCH with 384kbps

B.1.3.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-7: Required streaming MTCHs and FACHs for 80% streaming – 20% download given 1x384 kbps S-CCPCH

B.1.3.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-8: Required streaming MTCHs and FACHs for 50% streaming – 50% download given 1x384 kbps S-CCPCH

B.1.3.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-9: Required streaming MTCHs and FACHs for 20% streaming – 80% download given 1x384 kbps S-CCPCH

B.1.4 Scenario 4: 1S-CCPCH of 384kbps and 3 S-CCPCH with 128kbps each

B.1.4.1 80% streaming-20% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table B-10: Required streaming MTCHs and FACHs for 20% streaming-80% download given 1x384kbps+3x128kbps S-CCPCH
Appendix B: Radio Bearer Configuration

B.1.4.2 50% streaming-50% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table B-11: Required streaming MTCHs and FACHs for 50% streaming-50% download given 1x384kbps+3x128kbps S-CCPCH

B.1.4.3 20% streaming-80% download

<table>
<thead>
<tr>
<th>MTCH rate</th>
<th>32kbps</th>
<th>64kbps</th>
<th>128kbps</th>
<th>256kbps</th>
</tr>
</thead>
<tbody>
<tr>
<td># of streaming MTCHs</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>FACH rate</td>
<td>32kbps</td>
<td>64kbps</td>
<td>128kbps</td>
<td>256kbps</td>
</tr>
<tr>
<td># of streaming FACHs</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

Table B-12: Required streaming MTCHs and FACHs for 20% streaming-80% download given 1x384kbps+3x128kbps S-CCPCH

B.2 Channel mapping configuration derived

B.2.1 Scenario 1: 3 S-CCPCH with 384 kbps each

B.2.1.1 80% streaming – 20% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs(kbps)</td>
<td>256x1; 128x1</td>
<td>256x1; 32x1</td>
<td>256x1</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>384</td>
<td>288</td>
<td>256</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-13: Mapping derived for 80% streaming – 20% download given 3x384 kbps S-CCPCH
As can be seen in Table B-13 after the mapping for the streaming services has been performed, there exist two types of residual capacity: the first is the residual capacity of 96 kbps on FACH 4, which cannot be utilised further and hence appears as a waste of FACH capacity; the other is the residual capacity on S-CCPCH 3, which can be assigned to download FACHs for carrying download services. For instance, the 128 kbps residual capacity on S-CCPCH 3 is allocated to one download FACH, which in turn is assigned equally in capacity to 2 MTCHs of 64 kbps each, so as to accommodate two download applications. The overall channel mapping configuration for this scenario is clearly illustrated in Figure B-2, which also depicts both the residual capacity on the FACH (which goes unutilised) and on the S-CCPCH (which is utilised to carry download services).

![Figure B-2: Channel mapping structure utilising FF+BF bin-packing algorithm combination for 80% streaming–20% download traffic mix.](image)

### B.2.1.2 50% streaming – 50% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>256x1; 128x1</td>
<td>256x1</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>384</td>
<td>256</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>128x1</td>
<td>256x1; 128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>-</td>
<td>128</td>
<td>384</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>256</td>
<td>64</td>
<td>64</td>
<td>256</td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>256x1</td>
<td>32x2</td>
<td>-</td>
<td>128x1</td>
<td>64x2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>256</td>
<td>64</td>
<td>-</td>
<td>256</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>32x2</td>
<td>-</td>
<td>128x1</td>
<td>128x2</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>64</td>
<td>-</td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-14: Mapping derived for 50% streaming – 50% download given 3x384 kbps S-CCPCH
Table B-14 shows the channel mapping configuration derived for 50% streaming - 50% download given 3x384 kbps S-CCPCH. From the results, it can be found that all the FACHs are fully occupied by both streaming MTCHs (on FACHs 1, 2, 4) and download MTCHs (on FACHs 3, 5, 6, 7); hence zero residual capacity in FACHs is achieved. The total transmitted streaming capacity \((256\text{kbps} + 64\text{kbps} + 256\text{kbps} = 576\text{kbps})\) is equal to the total transmitted download capacity \((64\text{kbps} + 128\text{kbps} + 256\text{kbps} + 128\text{kbps} = 576\text{kbps})\), which corresponds to the specific 50% streaming - 50% download traffic mix.

### B.2.1.3 20% streaming - 80% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>384</td>
<td>384</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>(128\times2)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>256</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>(128\times1)</td>
<td>(256\times1)</td>
<td>(256\times1)</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>128</td>
<td>384</td>
<td>384</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>128</td>
<td>128</td>
<td>256</td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>(128\times1)</td>
<td>(64\times1)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>128</td>
<td>96</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>(32\times1)</td>
<td>(128\times2)</td>
<td>(64\times2)</td>
<td>(128\times2)</td>
<td>(64\times2)</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>32</td>
<td>256</td>
<td>128</td>
<td>256</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-15: Mapping derived for 20% streaming - 80% download given 3x384 kbps S-CCPCH

### B.2.2 Scenario 2: 3 S-CCPCH with 128 kbps each

In this section, we have a simpler scenario than scenario 1 in that the S-CCPCH physical channel capacity is reduced to 3 S-CCPCH with 128 kbps each. Various traffic mixes (80% streaming-20% download; 50% streaming-50% download; 20% streaming- 80% download) are considered and evaluated for this Scenario.
### B.2.2.1 80% streaming – 20% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>128x1</td>
<td>128x1</td>
<td>64x1</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>128</td>
<td>128</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>64x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>128</td>
<td>64</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>128x1</td>
<td>64x2</td>
<td>32x2</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>128</td>
<td>128</td>
<td>64</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>32x2</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>64</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table B-16: Mapping derived for 80% streaming – 20% download given 3x128 kbps S-CCPCH**

### B.2.2.2 50% streaming – 50% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>128x1</td>
<td>64x1</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>128</td>
<td>64</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>64x1</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>-</td>
<td>64</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>64</td>
<td>64</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>128x1</td>
<td>32x2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>128</td>
<td>64</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>32x2</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>64</td>
<td>128</td>
</tr>
</tbody>
</table>

**Table B-17: Mapping derived for 50% streaming – 50% download given 3x128 kbps S-CCPCH**
Table B-17 shows a simpler channel mapping configuration compared to Table B-13, whereby the same total rate is allocated to the download and streaming services, given a traffic mix of 50%-50%. There is one S-CCPCH of rate 128 kbps each for streaming and download service, while the other S-CCPCH of 128 kbps is assigned to carry 64kbps of streaming FACH and 64kbps of download FACH. The results show that all the FACHs are fully occupied by streaming MTCHs (on FACHs 1, 2) and download MTCHs (on FACHs 3, 4) with zero residual capacity carried on them.

B.2.2.3 20% streaming – 80% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs(kbps)</td>
<td>64x1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>64</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>64x1</td>
<td>128x1</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>64</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>64</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs(kbps)</td>
<td>32x2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>64</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>64x2</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-18: Mapping derived for 20% streaming – 80% download given 3x128 kbps S-CCPCH

B.2.3 Scenario 3: 1 S-CCPCH of 384 kbps:

In this section, S-CCPCH physical channel capacity of 1 S-CCPCH with 384kbps is consider. Various traffic mixes (80% streaming-20% download; 50% streaming-50% download; 20% streaming- 80% download) are considered and evaluated for this scenario.

B.2.3.1 80% streaming – 20% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs(kbps)</td>
<td>128x2</td>
</tr>
<tr>
<td></td>
<td>64x1</td>
</tr>
</tbody>
</table>
Appendix B: Radio Bearer Configuration

### Streaming FACHs sum (kbps)
- 320

### Download FACHs (kbps)
- 64x1

### Download FACHs sum (kbps)
- 64

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>128</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>128x1</td>
<td>64x2</td>
<td>32x2</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>128</td>
<td>128</td>
<td>64</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>32x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>64</td>
</tr>
</tbody>
</table>

**Table B-19: Mapping derived for 80% streaming – 20% download given 1x384 kbps S-CCPCH**

#### B.2.3.2 50% streaming – 50% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>384</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>128x2</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>256</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>128x1</td>
<td>32x2</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>128</td>
<td>64</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>64x1</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>64</td>
<td>128</td>
</tr>
</tbody>
</table>

**Table B-20: Mapping derived for 50% streaming – 50% download given 1x384 kbps S-CCPCH**

#### B.2.3.3 20% streaming – 80% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>384</td>
</tr>
</tbody>
</table>

Appendix B: Radio Bearer Configuration

### Streaming FACHs

<table>
<thead>
<tr>
<th>(kbps)</th>
<th>64x1</th>
</tr>
</thead>
</table>

### Streaming FACHs sum (kbps)

| 64 |

### Download FACHs (kbps)

| 320x1 |

### Download FACHs sum (kbps)

| 320 |

#### Table B-21: Mapping derived for 20% streaming – 80% download given 1x384 kbps S-CCPCH

#### B.2.4 Scenario 4: 1 S-CCPCH of 384 kbps and 3 S-CCPCH with 128 kbps each

In this section, we have a more complex scenario in that the S-CCPCH physical channel capacity is combined as 1 S-CCPCH with 384kbps and 3 S-CCPCH with 128kbps each. Various traffic mixes (80% streaming-20% download; 50% streaming-50% download; 20% streaming- 80% download) are considered and evaluated for this Scenario.

#### B.2.4.1 80% streaming – 20% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>384</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>256x1</td>
<td>64x2</td>
<td>128x1</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>384</td>
<td>128</td>
<td>128</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate(kbps)</td>
<td>256</td>
<td>128</td>
<td>64</td>
<td>64</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>
### Appendix B: Radio Bearer Configuration

#### Streaming MTCHs (kbps)

<table>
<thead>
<tr>
<th></th>
<th>256x1</th>
<th>128x1</th>
<th>32x2</th>
<th>32x1</th>
<th>64x2</th>
<th>-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>256</td>
<td>128</td>
<td>64</td>
<td>32</td>
<td>128</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>32x1</td>
<td>-</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>32</td>
<td>-</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-22: Mapping derived for 80% streaming - 20% download given 1x384+3x128 kbps S-CCPCH

#### B.2.4.2 50% streaming - 50% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>384</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>256x1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>384</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>-</td>
<td>128x1</td>
<td>128x1</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>-</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>256</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>128x1</td>
<td>64x1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>256</td>
<td>128</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>-</td>
<td>64x2</td>
<td>64x2</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>-</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-23: Mapping derived for 50% streaming - 50% download given 1x384+3x128 kbps S-CCPCH

#### B.2.4.3 20% streaming - 80% download

<table>
<thead>
<tr>
<th>S-CCPCH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>384</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming FACHs (kbps)</td>
<td>128x2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming FACHs sum (kbps)</td>
<td>256</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download FACHs (kbps)</td>
<td>128x1</td>
<td>128x1</td>
<td>128x1</td>
<td>128x1</td>
</tr>
<tr>
<td>Download FACHs sum (kbps)</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>
### Appendix B: Radio Bearer Configuration

<table>
<thead>
<tr>
<th>FACH</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate (kbps)</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Streaming MTCHs (kbps)</td>
<td>128x1</td>
<td>32x1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Streaming MTCHs sum (kbps)</td>
<td>128</td>
<td>32</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Download MTCHs (kbps)</td>
<td>-</td>
<td>48x2</td>
<td>64x2</td>
<td>64x2</td>
<td>64x2</td>
</tr>
<tr>
<td>Download MTCHs sum (kbps)</td>
<td>-</td>
<td>96</td>
<td>128</td>
<td>128</td>
<td>128</td>
</tr>
</tbody>
</table>

Table B-24: Mapping derived for 20% streaming – 80% download given 1x384+3x128 kbps S-CCPCH
Appendix C: Transport Channel Multiplexing and Transport Format Combination

In the radio access network (RAN), RLC and MAC layers are designed to accommodate the multiplexed incoming sessions. The session transmission starts after the establishment of multiple bearers at the same time. In UMTS, the UEs are designed so as to support the running of multiple applications at the same time. These applications are carried by multiple logical channels, with diverse assigned priority. As shown in the Figure C-1, Logical channels are responsible for transmitting the data traffic to MAC and eventually be multiplexed to transport channels, which is
transmitting the data traffic to physical channels. Each transport channel defines a specific format for the transmission of the data traffic. The available resources mapped onto the logical channels are scheduled by the network according to some specific criteria, e.g., the traffic priority and buffer occupancies. The scheduling procedure is essentially fulfilled via TFC selection, some general approach in selecting TFC for MAC scheduling in WCDMA are discussed in [123].

During the transport channel multiplexing, the MAC selects the adequate TFC for each transport channel from the TFCS specified by the RRC. Upon selection of the TF, MAC selects the TF that can transmit high-priority data on the basis of the traffic priority and RLC buffer status. MAC multiplexes and de-multiplexes logical channels carried on the same transport channel using the MAC header and measures the traffic volume for each transport channel and reports it to RRC.

The data format in the exchange across protocol layers/sublayers is MAC Protocol Data Unit (PDU) - includes the MAC header and MAC Service Data Unit (SDU), where the former includes TCTF, UE-id type, UE-id, C/T, and the latter serves as RLC-PDU.

- **TCTF** provides an identification of which logical channel type (e.g. CCCH) is carried in the payload.
- **UE-id type** indicates whether the following UE ID is a U-RNTI (32 bits) or a C-RNTI (16 bits), both of which are explained in [124].
- **UE id** will be either a U-RNTI or a C-RNTI as identified by UE ID type.
- **C/T** field identifies the logical channel number (1-15) where there are multiple logical channels in either DCH or FACH. It is also the main field used for channel multiplexing from a number of radio bearers to the transport channels.

During the phase of **TFC** selection, the following terms relating to data mapping by MAC can be identified and described as:

- **Transport Block** – TB is a basic unit exchanged between MAC layer and physical layer. A TB corresponds to a RLC PDU, and is the unit added with Cyclic Redundancy Check (CRC) in Layer 1.
- **Transport Block Set** – is defined as a set of TBs exchanged between physical layer and MAC layer at the same time instance using the same transport channel.
- **Transport Block Size** – refers to the length of the TB, represented in bits.
- **Transport Block Set Size** - refers to the length of the TB set, represented in bits.
• Transmission Time Interval (TTI) – is defined as the interval of the time at which TB sets arrive between layers, and is equal to the time interval at which TB sets are forwarded by Layer 1 on the radio interface. In SDMB, the typical TTI values are 10, 20, 40, and 80 ms.

• Transport Format (TF) – refers to the format in which the TB is supplied at every TTI on a transport channel. It consists of:
  • Dynamic part: TB size and TBS size;
  • Semi-static part: TTI, error-correction and CRC.

• Transport Format Set (TFS) – it defined as a set of TFs used in the transport channel. Within the TFS, the Semi-static parts of the TFs are same, and Dynamic parts of the TFs are changing every TTI.

• Transport Format Combination (TFC) – is defined as a combination of TFs in transport channels that are forwarded to L 1 in each TTI interval.

• Transport Format Combination Set (TFCS) – a set of TFCs carried on a single physical channel.

• Transport Format Indicator (TFI) – the identifier of a TF assigned to every TB set forwarded to physical layer from MAC layer, indicating which TF is being used in TFS.

• Transport Format Combination Indicator (TFCI) – the identifier of a TFC generated from the TFI by physical layer and transmitted over the radio interface.

Transport channel multiplexing and TFC selection are important procedures in WCDMA networks. A TFC is to be selected by the MAC layer considering the status of the logical channels and the provided radio resources of the transport channels. By selecting appropriate TFC for each TTI, the packet scheduling function is essentially performed.
Appendix D: Description of the SDMB system-level simulator

To evaluate the performance of the advanced radio resource allocation and packet scheduling algorithms, we use a system-level simulator that has been developed using the software package ns2. Ns2 is a widely used network simulation environment in mobile communication field, both academic and industrial. It operates on the interoperation between C++ code module and OTcl code module, offering an excellent platform for study comprehensive aspects of complex communication systems. Taking advantage of its available built-in code blocks in various aspects of wireless and satellite related functions, we developed additional code modules implementing SDMB specific features.

The first attempt of this system-level simulator was developed within the context of the IST SATIN project, implementing the basic functions for S-UMTS network such as: traffic models/trace files, critical MAC/RLC/RRC functions, i.e., packet scheduler, queuing buffer, channel multiplexing, segmentation/concatenation and etc. The system-level simulator was thereafter greatly enhanced within the context of IST MODIS project, focusing on user mobility analysis, channel propagation models, antenna characteristics, transport/application FEC, outer coding, and link-level look-up tables. Within the context of IST MAESTRO projects, the system-level simulator is further enriched with power scheduling model, data repetition technique, interleaving to the transport-level FEC and etc.

The main contribution of this research to the system-level simulator can be summarised as:

- A complete two-level channel multiplexing framework is implemented for both logical channels and transport channel at RLC and MAC layer, in addition to the existing single-level channel multiplexing framework.

- New optimization algorithm models are developed for optimizing the novel two-level multiplexing scheme in term of both channel utilisation and power efficiency.

- The functional model of TFCS derivation is developed so as to support the novel two-level channel multiplexing framework.

- Packet scheduler model is significantly upgraded from the previous MLPQ and WFQ scheduling algorithms. New interfaces between the algorithm model and the system functional models are defined so as to adapt the different advanced packet scheduling algorithms.
• Instantaneous buffer status is effectively tracked by the packet scheduler via the
development of BLRQ algorithm code model.

• Tracking mechanisms are established to effectively monitor the queuing delay performance
and report to the DDQ packet scheduling algorithm code model.

• Data rate performance is considered in combination with DDQ via the development of
CDRD algorithm code model.

• Cross-layer correspondences have been built across the protocol stacks within the
simulator, from application layer, transport layer to physical layer. Newly defined priority
function model are introduced to the simulator as an external agent, abstracting useable
information metrics from respective layers and handling the calculation and derivation
procedures for all these metrics.

• Adaptive service prioritization function and adaptive resource allocation function are
developed for multiple performance metrics via an adaptive adjustment mechanism
functional model, which leaks to the priority function model and updates itself in response
to the instantaneous performance variations.

• Comprehensive simulation scenarios are added to the simulator reflecting possible radio
bearer configurations.

To provide a better view of the system-level simulator with respect to the interaction and protocol
stack distribution, the major involved functional code models of the system-level simulator can be
described as in Figure D-1:
Appendix D: Description of the SDMB system-level simulator

As illustrated in Figure D-1, the system-level simulator evolves around four main parts based on the layer in which the respective functional model is positioned. In SDMB, three types of QoS classes are supported, namely: 1) streaming, 2) hot download, and 3) cold download. The streaming traffic model applies publicly available trace files [114] for video streaming traffics. Traffic characteristics associated with hot- and cold- download services -or, push-and-store services- follow the ns2 Pareto distribution\(^\text{18}\), with different traffic priority assigned. In addition, we choose different guaranteed data rates in order to examine the performance between users with different data rate requirements.

The packet scheduler is physically implemented in the Node-B employing the SDMB function (i.e. satellite hub), handling the per-TTI scale time-multiplexing task. It is coupled with the admission control and load control mechanisms to take account of the power constraint of the SDMB system.

\(^{18}\) The Pareto On/Off Traffic Generator is a traffic generator embodied in the ns2 OTcl class Application/Traffic/Pareto, which generates packets at a fixed rate (as specified by the application requirements) during on periods, whilst no packets are sent during off periods. Packets sent from a Pareto distribution generator are assumed to have constant packet size. Due to these characteristics, a Pareto distribution is often considered as one of the best solutions for simulating the background or download service traffics. In our simulation, we set the Pareto shape parameter as 1.5 as it is defaulted in ns2, however, different shape values can be set for various traffic dynamics or estimated based on some prior knowledge [126].
Appendix D: Description of the SDMB system-level simulator

Packets generated by the Traffic Engine are passed to the queuing buffers at satellite gateway for segmentation and scheduling. The functionalities of the segmentation/queuing block are to segment/reassemble the packets, initiate the corresponding RLC functions.

At the beginning of each admitted session starts, the RRA will derive the appropriate TFCS and pass it to the packet scheduler for its short-term resource allocation. The packet scheduler prioritizes the FACH queue and sorts them in descending order of the calculated priorities based on the specific metrics and rules. Prior to the prioritization procedure, the packet scheduler abstract information from the queuing model for instantaneous buffer behaviours and check the power constraints from the $E_b/N_0$ v.s. BLER look-up tables.

The link budget simulation results provide the $E_b/N_0$ v.s. BLER look-up curves of each FACH. The radio propagation channel model features either classical Ricean characteristics for satellite-associated path, or Rayleigh multipath fading channel for UE-associated path with the consideration of both Doppler Effect and propagation impairments. The maximum SDU size is 1500 bytes, TTI equals to 0.08 seconds and Turbo coder and QPSK is applied. The simulation period is set as 1000s or 12500 TTIs. Key simulation parameters are listed in Table D-1.

<table>
<thead>
<tr>
<th>Simulation Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of operation (GHz)</td>
<td>2.5</td>
</tr>
<tr>
<td>Chip rate (Mchip/s)</td>
<td>3.84</td>
</tr>
<tr>
<td>Spreading Factor</td>
<td>8</td>
</tr>
<tr>
<td>TTI (ms)</td>
<td>0.08</td>
</tr>
<tr>
<td>Modulation</td>
<td>QPSK</td>
</tr>
<tr>
<td>Coding</td>
<td>Turbo Code</td>
</tr>
<tr>
<td>Code Rate</td>
<td>1/3</td>
</tr>
<tr>
<td>Pareto shape factor</td>
<td>1.5</td>
</tr>
<tr>
<td>Maximum bit rate (kbit/s)</td>
<td>384</td>
</tr>
<tr>
<td>Packet size (bytes)</td>
<td>1280</td>
</tr>
<tr>
<td>Terrestrial channel model</td>
<td>Rayleigh</td>
</tr>
<tr>
<td>Satellite channel model</td>
<td>Ricean</td>
</tr>
</tbody>
</table>

Table D-1. System simulation parameters