This document discusses possible radio resource management, connection admission control, and handoff schemes for a single HAP multiple cell scenario that is intended to deliver up to 120Mbps per cell to fixed and high-speed mobile users. The radio resource management schemes are developed taking into account the unique properties of a cellular HAP, particularly the overlapping cells that occur as a result of their antenna beam generation. The schemes developed have shown that it is possible to exploit this overlap to increase the overall capacity of the system, while at the same time ensuring uniform qualities of service over the coverage area. A guard channel connection admission control algorithm supporting batch handoffs has been developed for the CAPANINA scenario with simplified calculation of near optimal channel guards. It is designed to keep the dropping probability of connections below a predefined threshold on handoff from one cell to another, while achieving a high maximum channel utilisation. Handoff schemes are also developed to overcome the relatively loose station-keeping characteristics, reducing the need for mechanical stabilisation. These schemes have been evaluated using several HAP mobility models and shown to be robust. They can be incorporated into the radio resource management schemes. The final part of the work examines the exploitation of advanced buffering techniques applied on on-board buffers in a high-speed train scenario as a way of mitigating against link outage for streaming applications when trains travel through tunnels. It is shown that a scheme that provides higher capacity in the initial stages of a download (when the user buffer levels are low) can significantly reduce outage probability compared with the no buffer case.

**Keyword list:** High Altitude Platforms, Resource Allocation, Connection Admission Control, Handoff, Buffering
Document History

<table>
<thead>
<tr>
<th>Date</th>
<th>Revision</th>
<th>Comment</th>
<th>Author / Editor</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P01</td>
<td>First Complete Draft</td>
<td>Konstantinos Katzis, Linjun Dong, David Grace, Paul Mitchell, Luong Dinh Dung</td>
<td>UOY, BUTE</td>
</tr>
<tr>
<td></td>
<td>P02</td>
<td>Addition of Intro, Conclusions and Summary</td>
<td>DG</td>
<td>UOY</td>
</tr>
<tr>
<td></td>
<td>P03</td>
<td>Addition of BUTE’s work and other edits</td>
<td>PDM</td>
<td>UOY</td>
</tr>
<tr>
<td>14/08/06</td>
<td>01</td>
<td>Issued Version</td>
<td>Konstantinos Katzis</td>
<td>UoY</td>
</tr>
</tbody>
</table>

Document Approval (CEC Deliverables only)

<table>
<thead>
<tr>
<th>Date of approval</th>
<th>Revision</th>
<th>Role of approver</th>
<th>Approver</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/08/06</td>
<td>P06</td>
<td>Editor (Internal reviewer)</td>
<td>Mihael Mohorcic</td>
<td>JSI</td>
</tr>
<tr>
<td>14/08/06</td>
<td>01</td>
<td>On behalf of Scientific Board</td>
<td>David Grace</td>
<td>UoY</td>
</tr>
</tbody>
</table>
EXECUTIVE SUMMARY

This document forms a deliverable from CAPANINA WP2.4 on resource allocation and handoff techniques for High Altitude Platform (HAP) systems. Existing techniques for terrestrial and satellite systems are not directly applicable in a HAP system. New techniques have been derived based on the HAP architecture.

Previous work has shown that cell overlap can be exploited to assist with resource allocation. This can be achieved by taking advantage of the physical architecture of a HAP communication system and allowing cells to symmetrically overlap each other, as detailed in chapter 2.

Chapter 3 builds on the methodology identified in chapter 2 and investigates a range of resource allocation schemes designed to improve Quality of Service in terms of blocking probability, data rate and fairness. In particular, the Area Based Fixed Channel Allocation (ABFCA), the Regional Based Fixed Channel Allocation (RBFCA), the Uniform Fixed Channel Allocation (UFCA) and its improved version UFCA-II schemes were investigated and compared with the basic fixed channel allocation (FCA) scheme. It is shown how it is possible to achieve overlap between adjacent cells and how the system capacity increases. Furthermore, it is shown, how it is possible to ensure uniform and fair quality of service (QoS) to all users irrespectively of their position.

A suitable approach to connection admission control for the CAPANINA scenario is described in chapter 4. The proposed strategy is based on the widely accepted approach of utilising guard channels to prevent the acceptance of new connections when the available resources become scarce. Unique characteristics of the CAPANINA scenario make the design of a suitable scheme particularly challenging, notably the necessity to handle batch handoffs when a train travels across a cell boundary, and the requirement to support multiple traffic classes. A suitable algorithm is detailed which keeps the dropping probability on handoff from one cell to another below a predefined threshold, whilst maintaining high maximum resource utilisation efficiency.

One of the design challenges that have been addressed in a High Altitude Platform architecture is the relatively loose station keeping, resulting in the movement of the cellular coverage on the ground, requiring mechanical stabilisation or connection handoff. Handoff schemes and mobility models are detailed in chapter 5, where an investigation of the impact of the aerial platform movements in a HAP telecommunication system is carried out. Results have shown that handoff technique reduces the need for mechanical stabilisation. It has also been shown that handoff mechanism is necessary to ensure continuity of the connections being affected by the platform movements. Furthermore, by employing guard channels it was possible to ensure lower dropping levels but at the expense of higher blocking levels. The immediate handoff scheme based on area based fixed channel allocation (ABFCA) exploiting cell overlap model has shown that by allowing overlap, the blocking probability has decreased from 6.5% to 2% compared with the non-overlapping case whilst the dropping probability was also much lower than in the no overlap case. It was also shown that the dropping probability has been significantly reduced (when compared with the no overlap case) since the users that required handoff were located in the areas of overlap. Thus, they experienced higher trunking efficiency and as a result the handoff was successful.

Chapter 6 explores the potential of on-board buffering to mitigate outages due to tunnels in broadband service delivery to trains, for non-real time services. A range of schemes have been investigated whereby information is downloaded to the users at a faster rate than required by the application, with the additional data stored in a buffer on-board the train. This information is then read from the buffer during periods of outage (when the line of sight connection is lost) to keep the applications running until a connection to the HAP is regained. The alternative techniques differ in the way in which capacity is allocated to the users. It has been shown that simple capacity assignment strategies such as fixed and demand assignment can provide some mitigation against outages caused by tunnels, but are not ideal solutions since the allocation mechanism does not take account of different user requirements. Significant improvements in performance can be obtained by taking individual user buffer levels into account and assigning more capacity to users with lower buffer levels (those that have started their
applications more recently). The proportional buffer and exponential buffer schemes offer the best performance, reducing the outage probability to a very low level compared with the no buffer circumstance.

This work has demonstrated the potential of the HAP architecture when providing broadband communication services. A number of radio resource management, connection admission control, and handoff schemes have been developed for a single HAP multiple cell scenario intended to deliver up to 120Mbps per cell to fixed and high-speed mobile users. The results demonstrated that the HAP architecture presents an opportunity to offer broadband services whilst efficiently utilise the spectrum provided. In addition, it has shown that the station keeping of the platform can be addressed by employing handoff techniques, thus minimising the complexity and reducing the weight of the payload. Furthermore, practical challenges such as providing broadband services on high-speed vehicles such as trains undergoing a consecutive number of tunnels can be addressed by employing an on-board buffering as an outage mitigation strategy.
TABLE OF CONTENTS

1. INTRODUCTION ..............................................................................................................14
   1.1 References .............................................................................................................. 16

2. RADIO RESOURCE MANAGEMENT TECHNIQUES USING CELL OVERLAP ............17
   2.1 Introduction............................................................................................................. 17
   2.2 Exploiting cell overlap in HAP system ................................................................. 18
   2.3 Multi-Cell System ................................................................................................. 23
   2.4 References ............................................................................................................. 27

3. RESOURCE ALLOCATION SCHEMES PROVIDING FAIR SERVICES WHILST EXPLOITING CELL OVERLAP .......................................................................................28
   3.1 Introduction............................................................................................................. 28
   3.2 Background assumptions ....................................................................................... 28
   3.3 Fixed Channel Allocation with no Cell Overlap .................................................... 29
       3.3.1 Description...................................................................................................... 29
       3.3.2 Performance.................................................................................................... 29
   3.4 Area Based Fixed Channel Allocation Scheme (ABFCA) ........................................ 30
       3.4.1 Description...................................................................................................... 30
       3.4.2 Performance.................................................................................................... 30
   3.5 Region Based Fixed Channel Allocation Scheme (RBFC) ......................................... 31
       3.5.1 Description...................................................................................................... 31
       3.5.2 Performance.................................................................................................... 31
   3.6 Uniform Fixed Channel Allocation Scheme (UFCA) .............................................. 33
       3.6.1 Description...................................................................................................... 33
       3.6.2 Performance.................................................................................................... 34
   3.7 Uniform Fixed Channel Allocation - II (UFCA - II) .............................................. 36
       3.7.1 Description...................................................................................................... 36
       3.7.2 Performance.................................................................................................... 36
       3.7.2.1 Cell overlap based on minimum received power threshold ...................... 37
       3.7.2.2 Effects of Interference on UFCA scheme performance ......................... 39
       3.7.2.3 UFCA-II Algorithm.................................................................................... 40
       3.7.2.4 Simulation Model without Overlap......................................................... 42
       3.7.2.5 Simulation Model with Overlap........................................................... 42
       3.7.2.6 Applying RAF........................................................................................ 43
   3.8 Summary and Conclusions ...................................................................................... 44
   3.9 References ............................................................................................................. 46

4. GUARD CHANNEL CONNECTION ADMISSION CONTROL ..................................47
   4.1 Introduction............................................................................................................. 47
   4.2 Problem Formulation .......................................................................................... 48
5. HANDOFF TECHNIQUES AND STATION KEEPING ........................................57
   5.1 Introduction ..............................................................................................................57
   5.2 Handoff Techniques ...............................................................................................58
   5.2.1 What is handoff? ..................................................................................................58
   5.2.2 Why handoff is necessary? ................................................................................58
   5.2.3 Quantifying Handoff ..........................................................................................59
      5.2.3.1 Handoff Probability .......................................................................................59
      5.2.3.2 Dropping Probability ...................................................................................59
   5.2.4 Terrestrial Handoff Schemes .............................................................................60
   5.2.5 Satellite Handoff Schemes ...............................................................................60
   5.3 High Altitude Platform Mobility Models ...............................................................62
      5.3.1 Drift with respect to the x, y and z axis .........................................................63
      5.3.2 Rotation with respect to the x, y and z axis .....................................................67
      5.3.3 Random Walk ..................................................................................................68
      5.3.4 Reflection ........................................................................................................69
   5.4 Steerable Antennas and Handoff ...........................................................................69
      5.4.1 Aperture Antenna Steering Solutions .............................................................70
   5.5 Handoff Simulation Model .......................................................................................72
   5.6 Handoff Simulation ..................................................................................................76
      5.6.1 Drift Movement Results ..................................................................................77
      5.6.2 Pitch Movements ..............................................................................................81
      5.6.3 Reflection, Rotation and Random Walk Movement Results .........................83
   5.7 Summary and Conclusions ......................................................................................85
   5.8 References ................................................................................................................87

6. ON-BOARD BUFFERING TECHNIQUES ................................................................89
   6.1 Introduction ..............................................................................................................89
   6.2 Fundamental On-Board Buffering Techniques ......................................................89
      6.2.1 Single Tunnel Scenario ....................................................................................89
      6.2.2 Statistical Distribution for UK Tunnel Lengths ..............................................90
      6.2.3 System Performance .........................................................................................92
      6.2.4 Meeting Quality of Service Requirements ....................................................94
   6.3 Advanced Buffer Techniques for QoS Provisioning in HAPs .................................95
      6.3.1 Tunnel Outage Simulation Scenario ...............................................................95
      6.3.2 Buffer-less download scenario ........................................................................96
      6.3.3 Buffering Schemes with Equal Sharing of Resources ......................................97
         6.3.3.1 Fixed Assignment Scheme .......................................................................97
         6.3.3.2 Demand Assignment Scheme .................................................................98
         6.3.3.3 Comparison of Fixed Assignment and Demand Assignment Schemes .......98
      6.3.4 Buffering Schemes with Adaptive Buffer Level ............................................100
6.3.4.1 Step Buffer Scheme .......................................................................................................................... 100
6.3.4.2 Proportional Buffer Scheme ............................................................................................................. 101
6.3.4.3 Comparison of Step Buffer and Proportional Buffer Schemes .......................................................... 103
6.3.4.4 Exponential Buffer Scheme .............................................................................................................. 104
6.3.5 Final Comparison of Schemes ............................................................................................................. 105
6.4 Summary and Conclusions .................................................................................................................... 107
6.5 References ............................................................................................................................................ 108

7. SUMMARY AND CONCLUSIONS........................................................................................................... 109
LIST OF FIGURES

Figure 2-1 High Altitude Platform - Concept of Operation (size not to scale) ........................................... 17
Figure 2-2 Illustration of a typical 7-cell (beam) scenario............................................................................ 18
Figure 2-3 Cells with no overlap.................................................................................................................. 19
Figure 2-4 Cells considering overlap......................................................................................................... 19
Figure 2-5 Illustration of Regions – 3 cell example .................................................................................... 20
Figure 2-6 Illustration of Regions - 7 cell examples..................................................................................... 21
Figure 2-7 Illustration of Areas - 7 cell examples......................................................................................... 21
Figure 2-8 Percentages of Overlapping Area A, B and C with respect to the normalised radius $R_i$ - Circular shape footprints are considered for a 3 cell case scenario.................................................. 23
Figure 2-9 Circular margin of a 121 cells coverage area.............................................................................. 23
Figure 2-10 Inner cell consists of 6B's, 6C's and one region A ................................................................. 24
Figure 2-11 Inner cell consists of 6B’s, 6C’s and one area A .................................................................... 24
Figure 2-12 Outer cell overlaps with 5 cells............................................................................................... 25
Figure 3-1 Fixed Channel Allocation with - no Cell Overlap ...................................................................... 30
Figure 3-2 Area Based FCA Scheme (ABFCA). ......................................................................................... 31
Figure 3-3 Partitioning areas into regions .................................................................................................... 32
Figure 3-4 Region Based FCA scheme (RBFCA). Overlap has improved system fairness. ......................... 33
Figure 3-5 UFCA – RAF defines probability of acceptance in area B and C .............................................. 34
Figure 3-6 Uniform FCA scheme (UFCA). The blocking probability has been reduced ............................... 35
Figure 3-7 Uniform FCA scheme (UFCA) blocking probability in centre cell for different degrees of overlap radius ......................................................................................................................... 36
Figure 3-8 CNIR levels of No Overlap Model and Overlap Model .............................................................. 40
Figure 3-9 UFCA-II Channel allocation flowchart ...................................................................................... 41
Figure 3-10 CNIR and Channel usage plots without cell overlap ............................................................... 42
Figure 3-11 CNIR and Channel usage plots with cell overlap ...................................................................... 43
Figure 3-12 CNIR and Channel usage plots with RAF .............................................................................. 44
Figure 4-1 General operating scenario ....................................................................................................... 47
Figure 4-2 One dimensional Markov chain after transformation ............................................................... 51
Figure 4-3 Link utilisation versus G1 and G2 ............................................................................................... 55
Figure 5-1 Handoff occurs when user moves from one cell to another ..................................................... 59
Figure 5-2 Handoff requests are queued while users move from point (a) to point (b) .............................. 60
Figure 5-3 Two Level Handoff Scheme – Case 1 [5.21] ............................................................................ 61
Figure 5-4 Two Level Handoff Scheme – Case 2 and 3 [5.21] .................................................................. 61
Figure 5-5 HAP position cylinder for 99% and 99.9% of time as defined by HeliNet [5.23] ...................... 63
Figure 5-6 Example of y-axis drift ............................................................................................................. 64
Figure 5-7 Graphical representation of equation (4.3) ............................................................................... 64
Figure 5-8 Example of z-axis drift ............................................................................................................. 65
Figure 5-9 Antenna mask example ............................................................................................................. 66
Figure 5-10 Drift on the z-axis scenario ..................................................................................................... 66
Figure 5-11 Example of y-axis rotation (pitch) ............................................................................................ 67
Figure 5-12 Example of z-axis rotation (yaw) ............................................................................................. 68
Figure 5-13 Example of random walk HAP movement .............................................................................. 68
Figure 5-14 Example of reflection HAP movement .................................................................................... 69
Figure 5-15 Steerable antennas can be utilised to correct drift movements and therefore maintain constant coverage area ................................................................. 70
Figure 5-16 Steerable antennas can be utilised to correct drift movements with respect to the z-axis ............................................................................................. 71
Figure 5-17 Steering mechanism for correcting HAP drift on the z-axis .................................................................................................................................105
Figure 5-18 Steerable antennas can be utilised to correct pitch and roll movements and therefore maintain constant coverage area ........................................ 72
Figure 5-19 Users are generated within a circle of 9km and served by 19 cells of 3.15km radius each ....................................................................................... 73
Figure 5-20 Immediate Handoff Initiation Algorithm ................................................................. 74
Figure 5-21 Identify and Remove Handoff Users .................................................................... 75
Figure 5-22 Add Handoff Users ............................................................................................. 76
Figure 5-23 Handoff and dropping probability for drift movement ........................................ 77
Figure 5-24 Moving footprints cause connection dropping .................................................. 78
Figure 5-25 Cell boundaries with no overlap ........................................................................ 78
Figure 5-26 Cell boundaries with overlap ............................................................................... 79
Figure 5-27 Handoff and dropping probability with respect to the distance from centre of coverage area ...................................................................................... 80
Figure 5-28 Blocking probability .......................................................................................... 80
Figure 5-29 Pitch HAP movement in three stages ................................................................. 81
Figure 5-30 Handoff and dropping probability for pitch HAP movement ................................. 82
Figure 5-31 Handoff and Dropping probability with respect to the distance from centre of coverage area ...................................................................................... 83
Figure 5-32 Blocking Probability .......................................................................................... 83
Figure 5-33 Handoff and dropping probability for a HAP system as a function of ground speed when subject to different types of motion ....................................... 84
Figure 5-34 Handoff and dropping probability with respect to the distance from the centre of the coverage area ................................................................. 85
Figure 6-1 Scenario of communication from a HAP to a mobile train ...................................... 90
Figure 6-2 CDF of Tunnel Lengths ....................................................................................... 91
Figure 6-3 Probability of outage for different applications, as a function of the amount of information stored in the buffer on entry to the tunnel ................................. 91
Figure 6-4 The model considering the distance from tunnel .................................................... 92
Figure 6-5 Probability of outage as a function of distance from a tunnel (mpeg4 application) .. 93
Figure 6-6 Probability of outage as a function of distance from a tunnel (AAC application) ..... 93
Figure 6-7 Download rate requirement as a function of minimum distance between tunnels ... 94
Figure 6-8 Scenario of communication from a HAP to mobile commuters on a train ............... 95
Figure 6-9 Outage probability as a function of distance between tunnels (buffer-less download scenario) ................................................................................. 97
Figure 6-10 Probability of outage as a function of distance between tunnels .......................... 98
Figure 6-11 CDF for outage time in different schemes .......................................................... 99
Figure 6-12 System performances with different thresholds ................................................. 101
Figure 6-13 The system performance for different initial download rates ................................. 102
Figure 6-14 Probability of outage as a function of distance between tunnels (proportional buffer scheme) .......................................................................... 103
Figure 6-15: Probability of outage as a function of overall download rate (comparing step buffer scheme with proportional buffer scheme) ....................................... 104
Figure 6-16 Outage probability with different initial download rate ......................................... 105
Figure 6-17 Probability of outage as a function of distance between tunnels in different schemes.

Figure 6-18 Probability of outage as a function of overall download rate for the proportional and exponential schemes.
LIST OF TABLES

Table 2-1 Limits of Overlap for the communication model .......................................................... 20
Table 2-2 List of Areas for various Overlapping Cases ................................................................. 25
Table 2-3 Definitions of regions .................................................................................................... 25
Table 3-1 Simulation Parameters for FCA schemes ................................................................ 29
Table 3-2 Channel Allocated per region in centre cell using RBFCA ........................................ 33
Table 3-3 $a$, $\beta$ and $OT_{\text{Optimum}}$ parameters used for overlap radius $R=1.25$ .......... 35
Table 3-4 Default Parameter values used to assess performance .............................................. 38
Table 3-5 Modulation and Coding figures used to determine capacity ..................................... 40
Table 3-6 RAF parameters ........................................................................................................... 43
Table 3-7 Blocking levels generated from each scheme ............................................................... 44
Table 4-1 Transformation rules for connections from class $j$ to class $i$ ................................. 51
Table 4-2 Parameters for the two traffic classes ......................................................................... 54
Table 4-3 Comparison of results obtained by simulation and approximation method ............ 54
Table 5-1 Handoff Scenario Simulation Parameters ................................................................. 73
Table 6-1 System simulation parameters ..................................................................................... 96
Table 6-2 Comparison of the fixed assignment and demand assignment schemes ............... 99
Table 6-3 Comparison of the outage probability of the step buffer scheme with the proportional buffer scheme .................................................................................................................. 104
### LIST OF ACRONYMS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC</td>
<td>Advanced Audio Coding</td>
</tr>
<tr>
<td>ABFCA</td>
<td>Area Based Fixed Channel Allocation</td>
</tr>
<tr>
<td>BE</td>
<td>Best Effort service</td>
</tr>
<tr>
<td>BFWA</td>
<td>Broadband Fixed Wireless Access</td>
</tr>
<tr>
<td>BpC&lt;sub&gt;thres&lt;/sub&gt;</td>
<td>Bits per connection per frame</td>
</tr>
<tr>
<td>BPSK</td>
<td>Binary Phase Shift Keying</td>
</tr>
<tr>
<td>BS</td>
<td>Base Station</td>
</tr>
<tr>
<td>BW&lt;sub&gt;RF&lt;/sub&gt;</td>
<td>RF Bandwidth</td>
</tr>
<tr>
<td>CA</td>
<td>Coverage Area</td>
</tr>
<tr>
<td>CAC</td>
<td>Connection Admission Control</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CIR</td>
<td>Channel to interference ratio</td>
</tr>
<tr>
<td>CNIR</td>
<td>Carrier to Noise plus Interference Ratio</td>
</tr>
<tr>
<td>CNR</td>
<td>Channel to Noise Ratio</td>
</tr>
<tr>
<td>CPE</td>
<td>Customer Premises Equipment</td>
</tr>
<tr>
<td>DAS</td>
<td>Demand Assignment Scheme</td>
</tr>
<tr>
<td>DBT</td>
<td>Distance Between Tunnels</td>
</tr>
<tr>
<td>DH</td>
<td>Directed Handoff</td>
</tr>
<tr>
<td>Dmax</td>
<td>Maximum Directivity</td>
</tr>
<tr>
<td>DR</td>
<td>Directed Retry</td>
</tr>
<tr>
<td>Eff-CNIR</td>
<td>Effective CNIR</td>
</tr>
<tr>
<td>f</td>
<td>Carrier Frequency</td>
</tr>
<tr>
<td>FAS</td>
<td>Fixed Assignment Scheme</td>
</tr>
<tr>
<td>FCA</td>
<td>Fixed Channel Allocation</td>
</tr>
<tr>
<td>FWA</td>
<td>Fixed Wireless Access</td>
</tr>
<tr>
<td>GC</td>
<td>Guard Channel</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>Gref</td>
<td>Reference Gain</td>
</tr>
<tr>
<td>h</td>
<td>Platform Altitude</td>
</tr>
<tr>
<td>HAP</td>
<td>High Altitude Platform</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HQ</td>
<td>Handoff Queuing</td>
</tr>
<tr>
<td>K</td>
<td>Cluster Size</td>
</tr>
<tr>
<td>LEO</td>
<td>Low Earth Orbit</td>
</tr>
<tr>
<td>LOS</td>
<td>Line Of Sight</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>Na</td>
<td>Number of active users</td>
</tr>
<tr>
<td>NCQ</td>
<td>New Call Queuing</td>
</tr>
<tr>
<td>nrtPS</td>
<td>Non-Real-Time Polling Service</td>
</tr>
<tr>
<td>NRX</td>
<td>Noise Received Level</td>
</tr>
<tr>
<td>ODR</td>
<td>Overall Download Rate</td>
</tr>
<tr>
<td>OT</td>
<td>Offered Traffic</td>
</tr>
<tr>
<td>P(Block)</td>
<td>Blocking Probability</td>
</tr>
<tr>
<td>PD</td>
<td>Dropping Probability</td>
</tr>
<tr>
<td>PH</td>
<td>Handoff Probability</td>
</tr>
<tr>
<td>PTX</td>
<td>Transmit Power</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>R</td>
<td>Radius of Overlapping Cell</td>
</tr>
<tr>
<td>RAF</td>
<td>Random Acceptance Factor</td>
</tr>
<tr>
<td>RBFCA</td>
<td>Regional Based Fixed Channel Allocation</td>
</tr>
<tr>
<td>Rd</td>
<td>Download Rate</td>
</tr>
<tr>
<td>Re</td>
<td>Radius of the circle which encloses the hexagonal cell (external radius)</td>
</tr>
<tr>
<td>Ri</td>
<td>Radius of the circle enclosed inside the hexagonal cell (internal radius)</td>
</tr>
<tr>
<td>rtPS</td>
<td>Real-Time Polling Service</td>
</tr>
<tr>
<td>SHOT</td>
<td>Selective Handover for Traffic balance</td>
</tr>
<tr>
<td>SPP</td>
<td>Sub Platform Point</td>
</tr>
<tr>
<td>TTA</td>
<td>Time To Arrive</td>
</tr>
<tr>
<td>UAV</td>
<td>Unmanned Aerial Vehicle</td>
</tr>
<tr>
<td>UGS</td>
<td>Unsolicited Grant Service</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Network</td>
</tr>
</tbody>
</table>
1. Introduction

High Altitude Platform (HAP) systems have the promise to deliver high capacity broadband communications to a regional coverage area. One way of delivering the promised high capacities is to use multiple cells. This brings together a number of challenges, including development of appropriate network infrastructure, design of the payload, appropriate radio resource management strategies, as well as dealing with the unique properties of the HAP system. Here we focus on the latter two points, the radio resource management and the impact of movement of the HAP on radio resource management.

The studies here are based around the IEEE 802.16 [1.1] standard given that it is seen as the most promising standard [1.2] on which to base HAP broadband communications. It is likely that this may require some modest modifications to cope with HAP mobility and centralised cell structure.

In general a cell can be considered as a localised coverage area in which an assignment of the radio spectrum is used. Other cells then reuse the spectrum such that the interference between cells using the same spectrum is sufficiently low to allow suitable demodulation of the signal at the receivers. In the case of terrestrial systems it is the physical separation distance between cells that is important and the signal losses on the propagation path between the cells. In the case of satellite and HAP systems, the angular separation of cells originating from the HAP or satellite, and the characteristics of the antenna beam, producing the cells on the craft, are of most importance. This type of spectrum reuse allows different information to be sent in each of the cells, increasing the information capacity of the whole system. Therefore, the information capacity of the system is increased, and the number of users that can benefit is dependent on the total information rate and capacity requirement per user.

HAP cellular architectures have been examined in [1.3], specifically addressing the antenna beam characteristics required to produce an efficient cellular structure on the ground, and the effect of antenna sidelobe levels on channel reuse plans. Cells can be regularly spaced, as their area and location are substantially unaffected by geography and terrain, and since they all originate from the same HAP this centralisation can be additionally exploited by the resource management strategy. Such strategies have already been explored for HAP use [1.4]. Spectrum efficiency can be further enhanced by exploiting the antenna directionality of the user and employing multiple HAPs to serve the same coverage area [1.5].

Another difference between HAPs and terrestrial and satellite systems are their relatively loose station-keeping characteristics. When a HAP is used with a cellular structure this imperfect station-keeping results in movement of the footprints on the ground, unless the payload is stabilised. For non-stabilised or partially stabilised payloads there is a requirement to handover communications from one antenna footprint (cell) to another. Performing handoffs at bit rates of up to 120Mbps, coupled with the stringent quality of service and latency requirements demanded by multimedia traffic, make conventional type handoff schemes less than suitable. One advantage of the HAP architecture is that the cell transmitters are centralized, unlike with terrestrial broadband deployments, allowing aspects of the resource management and handoff to be managed collectively. The requirements for geographical locations requiring handover for different modulation strategies when subject to lateral drift in the HAP with corresponding payload stabilisation strategies have been investigated in [1.6], the effect on UMTS of platform instability [1.7] and platform motion on IMT-2000 soft handover [1.8], with earlier work carried out by El-Jabu and Steele [1.9], while Foo et al have modelled conventional 3G handoff for mobile users assuming stationary HAPs. Albagory [1.10] has also examined handoff for HAPs, developing a ring shaped cell which will cope better with circling platforms.

This centralized HAP architecture is similar to that seen by multi-spot beam satellites, but in the case of geostationary satellites, the cells do not move appreciably on the ground, and in the case of low-earth orbit (LEO) satellites, movement of beams is highly predictable, allowing handoff requirements between beams and satellites to be predicted well in advance, ensuring that traffic streams remain practically contiguous [1.11], [1.12]. However, the proposed data rates for HAPs are significantly higher.

The purpose of this document is to analyse in detail HAP cellular radio resource management and handoff issues, illustrating the potential capacity gains and performance constraints. The sister deliverable to this document will examine the impact of a multiple HAP architecture as a way of enhancing the available capacity. The work is focussed into three broad areas, radio resource
management schemes incorporating the benefits of cell overlap and new connection admission control policies, handoff, and link outage mitigation.

Firstly, Chapter 2 discusses the unique properties of the HAP cellular structure, specifically examining the degree of cell overlap that is available for use with the radio resource management scheme. Overlap occurs due to the way in which the cells are generated from the antenna beams as opposed to distance-based reuse in a terrestrial system. This overlap can be used to improve flexibility of the radio resource management scheme and improve handoff thereby enhancing the available capacity of the HAP system.

In Chapter 3 several different radio resource management schemes are examined that exploit the overlap as a way of improving capacity and/or quality of service. A critical issue is to try and maintain fairness across the coverage area when the amount of resource available varies depending the degree of cell overlap. Chapter 4 describes a connection admission control algorithm for the CAPANINA high-speed train scenario, which employs guard channels for different traffic classes to prevent the acceptance of new connections when resource usage becomes high, leaving the remaining channels available for handoff requests. An approximation method is proposed for the calculation of guard channels to keep the resource reservation as high as possible while guaranteeing a minimum bandwidth for each connection and keeping the dropping probability of hand-off connections under a predefined threshold.

Chapter 5 then examines the handoff schemes specifically designed for the HAP architecture, looking in particular at the impact of different HAP mobility models on the overall performance of the system. The handoff techniques proposed ensure continuity of the connections being affected by the platform movements and therefore reduce the need for mechanical stabilisation. As a result, the proposed handoff schemes are saving on valuable payload weight, volume, and power.

The final main chapter, Chapter 6, deals with mitigating the impact of link outage in a high-speed train scenario by using buffers at the train end of the link. Such a strategy is intended to improve the perceived end-user quality of service for non-real time audio and video situations.

A comprehensive summary and conclusions is presented finally in Chapter 7.
1.1 References


2.1 Introduction

High Altitude Platforms can achieve a high capacity by using a large number of wireless transceivers, each using a directional antenna to create cells on the ground. These transceivers are co-located on the platform, and they offer line-of-sight communication to a geographic service area of approximately 60km diameter [2.1] assuming the minimum elevation angle above 30 degrees.

Figure 2-1 High Altitude Platform - Concept of Operation (size not to scale)

Figure 2-1 depicts an example of the concept of operation of a High Altitude Platform communication system. The platform is allowed to move within a positional cylinder, whose size has been defined in HeliNet project [2.1] which is more relaxed than the proposed ITU boundaries [2.2] but is chosen to be conservative. This is to guarantee that the platform will consistently deliver broadband communications over the nominal coverage area irrespectively whether this is an airship or an airplane. The coverage area is subdivided into cells based on the number of antennas collocated on the platform. Each cell is assigned a group of channels. It has been shown in [2.3], [2.4] and [2.5] that with the nature of the HAP architecture, it is possible to achieve overlap between adjacent cells and thus maximise the system capacity whilst ensuring uniform and high quality of service (QoS) to all users. The cell overlap occurs because of the way the power decreases away from the boresight of the antenna.

The cells formed on the ground are assumed to be circular and of equal size; this can be readily achieved by a careful design of the antenna beam profiles [2.6]. Interference between cells is largely due to the gain profile and sidelobe levels of the antennas used. It has been shown that the antenna gain profiles at the cell edges can create useful cell overlap [2.3].

Although cell overlap does occur, it has not always been used to redirect traffic from one cell to another. One example is the case of fixed wireless terrestrial systems where the user antenna requires to be redirected when moving between cells, a comparatively slow process, and during this time the user will be out of contact with any base station. In addition, this will require the system to be highly centralised, due to the need to exchange data about the channels available within each cell at high speed [2.7]. Nevertheless the concept of overlapping cells has been investigated in the past [2.8] [2.9] for mobile terrestrial systems with schemes such as directed retry (DR), directed handoff (DH) and a variety of selective handover for traffic balance (SHOT).
schemes. From work [2.8] it has been shown that with the DR, an increase in the overlapping between cells leads to an increase in the quality of service (QoS) provided by the system. Furthermore, the DH scheme proved to have good sensitivity properties with respect to variation in the spatial profile of the system. From work [2.9] it has been shown that the SHOT schemes improve traffic handling capacity and enhance resource utilisation. The more cell overlapping the more the traffic carried provided that the interference is acceptable.

### 2.2 Exploiting cell overlap in HAP system

Most of the practical problems that the terrestrial systems face are not applicable in the case of HAPs. This is because of the nature of the system: all transceivers are co-located on the platform, and the platform provides a line-of-sight communication link with the stations on the ground. This means that there will be fewer obstacles between the users and the platform (therefore no shadowing and multipath – when operating in the mm-wave band), and the cell overlap can effectively be applied in all cells of the system. The HAP itself can therefore keep track of all channels in use within its coverage area by making use of a centralised architecture. The footprints are positioned to cover a nominal coverage area. Their size must be large enough not to leave any parts of the coverage area unattended. An example can be seen in Figure 2-2.

![Figure 2-2 Illustration of a typical 7-cell (beam) scenario](image)

Overlap is generated due to the characteristics of the mainlobe of the antenna of each cell. The size of the overlapping area is determined by setting a minimum received power limit and Channel to Interference Ratio (CIR) threshold. This is determined from the link budget and Channel to Interference Ratio (CIR) values. Calculation of the link budget is related to the power roll-off of the antenna directivity pattern. The fact that the cells overlap with each other can be proved beneficial for the performance of the system since this allows a more flexible allocation of channels to the cells.

There are two ways that a user can select which base station to connect to. The first one is based on their distance from a station and the second one is based on a minimum received power threshold. In either case, cells are set to overlap with each other as shown in Figure 2-2 to ensure full coverage. However, for the case where users connect to the closest base station the cell overlap is effectively ignored since despite the extent of the cell coverage, users will still connect to their closest base station. To take advantage of the cell overlap users must be allowed to choose to connect to any base station that happens to be within range (i.e. have the received power higher than the threshold or minimum CIR required).
Figure 2-3 illustrates the concept of the cell radius. Here, $R_i$ defines the radius of the circle enclosed inside the hexagonal cell (internal circle) and $R_e$ the radius of the circle which encloses the hexagonal cell (external circle). The relationship between $R_e$ and $R_i$ is $R_e = \frac{2R_i}{\sqrt{3}}$.

Now we define a cell radius for a general overlapping scenario (see Figure 2-4). The minimum value of the radius of the overlapping cell is equal to the external radius of the cell $R_e$ in order to avoid leaving any areas without service (if $R$ was less than $R_e$). The maximum possible value ($R$) taken here must always be less than 1.5$R_e$ radius, in order to limit the maximum number of overlapping cells to three. The aim is to limit overlap to only three cells instead of four or even more and to prevent the system from becoming overly complicated by numerous overlapping cells, which would result in co-channel interference. In theory, four or even more cells can overlap if the cell radius is increased sufficiently, assuming the co-channel interference and received power levels remain acceptable.
The radius of overlapping cell $R$ will be normalised either with respect to the external circle radius $R_e$ or to the internal circle radius $R_i$ (depending on the simulation).

### Table 2-1 Limits of Overlap for the communication model

<table>
<thead>
<tr>
<th>Radius of Overlapping Cell</th>
<th>Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{Min}$</td>
<td>Equal to $R_e$. This is the minimum value of radius so that all the coverage area will be served.</td>
</tr>
<tr>
<td>$R_{Max}$</td>
<td>Less than $1.5R_e$, in order to avoid overlapping between 4 or more cells.</td>
</tr>
</tbody>
</table>

![Figure 2-5 Illustration of Regions – 3 cell example](image)

Whenever two or more adjacent cells overlap, they form a set of individual regions, which can be categorised into three types (A, B or C) according to the numbers of cells that overlap at any given time. They can therefore be assigned a channel from one, two or three cells respectively. If we sum up the regions in one cell according to the three types of degrees of overlap, we then have what we call areas.

To give a more complete picture of what a region is and what an area is let us assume that seven cells overlap each other like the case in Figure 2-6.
In this case the centre cell is partitioned into six regions of (overlap) type B (B1 to B6), six regions of type C (C1 to C6) and one region of type A, provided that the radius of overlapping cell R is less than $1.5R_e$.

If the regions (of the same type) illustrated in Figure 2-6 are grouped together, they will form three types of areas (again of type A, B and C). Figure 2-7 illustrates these three types of areas. In the seven-cell overlap example, the size of Area A equals to the size of region A, the size of area B equals to the sum of regions B1 to B6 and the size of area C equals to the sum of regions C1 to C6.
The significance of having regions and areas is that it is possible to perform the channel allocation based on either region or area, which will be shown later in this report. If for example the channel allocation is based on regions (assuming a fixed channel allocation scheme and a uniform offered traffic), then each region will be assigned a nominal number of channels. The number of channels depends on the size of the region and the assigned channels can only be used in that specific region. If on the other hand the channel allocation is based on areas (i.e. a sum of regions), then the channels’ allocation should be performed according to the size of the area. Therefore, the total number of channels allocated to an area will be greater than the number of channels allocated to a region. These channels can be assigned anywhere within the area, which means that they can geographically be assigned to any of the regions consisting the area.

In a general fixed channel allocation scheme (FCA), areas can be assigned different numbers of channels $C_A$, $C_B$, and $C_C$. This allows a channel allocation scheme to be implemented while minimising the blocking probability $P(\text{Block})$. The task is to decide how to assign different numbers of channels to the respective regions, areas or cells so that the blocking probability will be the same in every part of the cell.

The mathematical derivation which calculates the area as a function of the varying radius of the three overlapping circles (see Figure 2-5) is fully presented in [2.10] and [2.11]. This task requires finding the percentage that these areas cover with respect to the total area of the cell. Here, it was assumed that all three circular cells are of the same size and as a result the extended overlapping circles will also be of the same area. However, if it is required to have more than three cells overlapping the same area then the frequency reuse factor has to be greater than 3. This is because co-channel interference can occur when the size of the overlapping circles extends more than the half of the maximum frequency reuse distance $(D/2)$ in order to include a fourth circle into the overlapping area. Although more than three overlapping cells is a feasible case scenario, it will not be attempted to be analysed at this point due to its complexity.

While varying the radius of the overlapping circles (normalised to internal circle radius $R_i$), the percentages of area A, B and C with respect to the size of the overlapping circle were plotted. Figure 2-8 depicts these variations for the case with only three cells in the system. The overlapping area is extended up to 10-times the internal radius $(R)$ of the cell. From the plot in Figure 2-8, it can be seen that when $R = R = 2R_i$, area A (non overlapping area) is significantly larger than areas B and C. Area B however increases almost linearly up to the point where radius $R$ is 1.73 times the internal radius $R_i$. Also, area C remains zero up to the point where $R = 2/\sqrt{3}$. This is however outside the range set previously $(R_{\text{max}} < \sqrt{3}R_i)$ and it will not be considered in this study. Nevertheless, it has been a good starting point understanding the concept of cell overlap. Knowledge of the behaviour of overlapping areas can be useful for the channel assignment strategy, and it is shown later in this section that the channel assignment will depend on these areas according to which frequency reuse factor K the system is operating at.
2.3 Multi-Cell System

When the cells overlapped each other symmetrically, then this formed regions A, B and C as seen in Figure 2-5. The three-cell approach presented above however is not the most common case seen in a multi-cell system. Hence the three-cell approach was further analysed to apply to a multi-cell system. The following picture depicts a 121-cell HAP communication system.

The coverage area is defined as the area enclosed within the bold circle. The cells marked with the number of neighbours represent the cases that need to be considered when calculating the
overlap areas. The case of 6 neighbours however will be mostly considered when describing the channel allocation schemes, since it is the most common and generally applicable case.

Most cells inside the coverage area have 6 neighbouring cells, which overlap with each other. According to the definition of region and areas presented before, this means that each cell (with 6 neighbouring cells) consists of 6\(B_{\text{Small}}\) (= area B) and 6\(C\) (= area C) regions plus one region A (= area A) as shown in Figure 2-10, provided that the radius of overlapping cell R is less than \(\sqrt{3}R_i\). Notice that there are no regions formed by 4 or more overlapping cells.

![Figure 2-10 Inner cell consists of 6B’s, 6C’s and one region A](image)

Varying the radius of the cells \((R_i \leq R < \sqrt{3}R_i)\), it can be observed that the three types of areas (sum of all regions) vary differently. Area A, B and C plots are depicted in the following figure as a function of the overlapping radius R normalised to the internal cell radius \(R_i\).

![Figure 2-11 Inner cell consists of 6B’s, 6C’s and one area A](image)
In the cases where there are either 3, 4 or 5 neighbours, there are two types of regions B formed. Region B\textsubscript{Small} and B\textsubscript{Large}. Figure 2-12 illustrates the case where an outer ring cell overlaps with 5 cells. From this, it can be seen than in this case, area B consists of both types of Region B\textsubscript{Small} and B\textsubscript{Large}.

**Figure 2-12 Outer cell overlaps with 5 cells**

Table 2-2 and Table 2-3 represent all the cases of overlap, with exact derivations of each area in terms of the overlap cell radius \( R \), the internal and external normal size cell radius \( R_i \) and \( R_e \) respectively and the parameter \( \theta \) [2.10]. These overlap cases are the cases where the cell of interest is located in the outer ring of the coverage area and its number of neighbours can be 3, 4, 5 or 6 depending on this position.

**Table 2-2 List of Areas for various Overlapping Cases**

<table>
<thead>
<tr>
<th>Category</th>
<th>Area A</th>
<th>Area B</th>
<th>Area C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Cell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 - Neighbours</td>
<td>Cell _Area – (Area B + Area C)</td>
<td>6 Region B\textsubscript{Small}</td>
<td>6 Region C</td>
</tr>
<tr>
<td>External Cell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 - Neighbours</td>
<td>Cell _Area – (Area B+ Area C)</td>
<td>2 Region B\textsubscript{Large} + 3 Region B\textsubscript{Small}</td>
<td>4 Region C</td>
</tr>
<tr>
<td>4 - Neighbours</td>
<td>Cell _Area – (Area B+ Area C)</td>
<td>2 Region B\textsubscript{Large} + 2 Region B\textsubscript{Small}</td>
<td>3 Region C</td>
</tr>
<tr>
<td>3 Neighbours</td>
<td>Cell _Area – (Area B+ Area C)</td>
<td>2 Region B\textsubscript{Large} + 1 Region B\textsubscript{Small}</td>
<td>2 Region C</td>
</tr>
</tbody>
</table>

**Table 2-3 Definitions of regions**

\[
\text{Region A} = 2\pi R^2 - \text{(Area B + Area C)}
\]

\[
\text{Region B}_{\text{Small}} = \left[ R^2 \cdot a \cos \left( \frac{R_i}{R} \right) - R_i \left( \sqrt{R^2 - R_i^2} \right) \right] - \frac{2R^2}{2} \left[ \sqrt{3} \cdot (1 - \cos(\theta_c)) + 3 \cdot (\theta_c - \sin(\theta_c)) \right]
\]

\[
\text{Region B}_{\text{Large}} = \left[ R^2 \cdot a \cos \left( \frac{R_i}{R} \right) - R_i \left( \sqrt{R^2 - R_i^2} \right) \right] - \frac{R^2}{2} \left[ \sqrt{3} \cdot (1 - \cos(\theta_c)) + 3 \cdot (\theta_c - \sin(\theta_c)) \right]
\]

\[
\text{Region C} = \frac{R^2}{2} \left[ \sqrt{3} \cdot (1 - \cos(\theta_c)) + 3 \cdot (\theta_c - \sin(\theta_c)) \right]
\]

where

\[
R_e = \frac{2R_i}{\sqrt{3}} \quad \text{or} \quad R_i = \frac{3R_e}{2} \quad \text{and} \quad \theta_c = \arccos \left( 1 - \frac{3 \cdot \left( \sqrt{R^2 - R_i^2} - \frac{R_i}{\sqrt{3}} \right)^2}{2R^2} \right)
\]
The investigation above was useful at a later stage when the traffic was considered. It was required though, to categorise each cell based on its number of neighbours in order to obtain an accurate approximation of the total size of areas A, B and C. Since some of the cells of the outer ring (external cells) fell outside the coverage area (see Figure 2-9), they were not taken into account for these calculations. In fact we have considered the case of the 6 neighbouring cells which is the most common case to be found in a large system such as the example of 121-cell system proposed in HeliNet [2.1].
2.4 References


3. Resource Allocation Schemes Providing Fair Services Whilst Exploiting Cell Overlap

3.1 Introduction

In Chapter 2 we have looked at the concept of cell overlap and how it is being defined. In this Chapter we will be looking at a number of Fixed Channel Allocation (FCA) schemes, the benefits of cell overlap and how it can be ensured fair access to the actual resources. The chapter is organised as follows. First we describe the background assumptions regarding the cellular coverage area and traffic load taken into account in performance evaluation of different channel allocation schemes. Next the basic FCA scheme is presented, which does not make use of cell overlap, followed by the Area Based Fixed Channel allocation scheme (ABFCA) employing cell overlap. ABFCA proves that by exploiting cell overlap, the system performance improves but results show that the Quality of Service (QoS) is not uniform across the cells. Then we introduce the Regional Based Fixed Channel Allocation Scheme (RFCA), which has been implemented in order to unify the blocking probability levels between the regions. The performance of RFCA, however, was worse than the case of ABFCA scheme with no overlap even though more channels per cell were assigned. Next we present the Uniform Fixed Channel Allocation (UFCA) scheme, which is based on the Random Acceptance Factor (RAF) and is intended to achieve both uniform blocking across the cell as well as better performance than the simple FCA scheme which does not employ cell overlap. Finally, the UFCA-II is presented, which is an enhanced version of UFCA scheme so that apart from ensuring uniform blocking level it also ensures uniform data rates across the cell area.

3.2 Background assumptions

To ensure realistic results within a reasonable simulation time, 37 circular cells were used instead of 121. 37 cells were sufficient to eliminate the problem of edge-effects [3.1]. Statistics were only collected from the centre cell and this is because the majority of the cells in a larger scale cellular scenario have six neighbours (e.g. centre cell). Each of these 37 cells is assigned a different group of 30 channels within its coverage area and therefore co-channel interference is not present. We have also assumed that there is direct line-of-sight communication between the user and the HAP. 100,000 users have been assumed with 500,000 conversations to be made. The arrival process is a Poisson distribution and the length of the phone calls has a negative exponential distribution. The users are uniformly distributed within the coverage area, and the offered traffic (OT) is quantified in terms of Erlangs per square unit area. The total OT in the 37-cell coverage area ranges from 8 - 10 Erlang per square unit times the actual size of the coverage area (one unit of length is taken to be the external radius of the cells, R).

The total size of the coverage area (CA) for 37 cells can be expressed with Equation (3.3) based on Region \(B(\text{No Area C})\) defined in Equation (3.1) and Region \(C\) given in Equation (3.2).

\[
\text{Region } B(\text{No Area C}) = 2 \cdot \left( R^2 \cdot a \cos \left( \frac{R_i}{R} \right) - R_i \cdot \left( \sqrt{R^2 - R_i^2} \right) \right) \tag{3.1}
\]

\[
\text{Area } C = \text{Region } C = R^2 \cdot \left[ \sqrt{3} \cdot (1 - \cos(\theta)) + 3 \cdot (\theta - \sin(\theta)) \right] \tag{3.2}
\]

\[
CA = 37\pi R^2 - (90 \cdot \text{Region } B(\text{No Area C}) - 54 \cdot \text{Region } C) \tag{3.3}
\]
Assuming that \( R_e \) is equal to 1 unit, then \( R_i \) is equal to \( \sqrt{\frac{3}{2}} \) units, and the total coverage area is equal to 99.9 square units. This value remains fixed for any overlap radius despite the fact the cell radius might change. This is to ensure constant OT within the coverage area. Simulation parameters used for comparison of different channel allocation schemes are summarised in Table 3-1.

### Table 3-1 Simulation Parameters for FCA schemes

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Number of cells (c)</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Number of channels per cell</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Offered traffic</td>
<td>8-10</td>
<td>erlang/sq. unit area</td>
</tr>
<tr>
<td>4</td>
<td>Number of users</td>
<td>100,000</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Number of connection requests</td>
<td>500,000</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Coverage area</td>
<td>99.9</td>
<td>sq.units</td>
</tr>
<tr>
<td>7</td>
<td>Average connection duration</td>
<td>6</td>
<td>Minutes</td>
</tr>
<tr>
<td>8</td>
<td>Connection request arrival process</td>
<td>Poisson</td>
<td></td>
</tr>
</tbody>
</table>

Furthermore, the following Traffic Model Assumptions were used in the simulations that follow:

1. Call arrivals are described by a Poisson process with rate \( \lambda \) while the call duration is exponentially distributed with mean \( 1/\mu \).
2. Blocked calls are cleared and do not return.
3. Users connect to the base station with the most available channels within range. They can choose channel(s) from one of up to the three base stations depending on the area in which they are located.
4. Both HAP and users are considered to be stationary.
5. Handoff or Call Dropping is not implemented.

### 3.3 Fixed Channel Allocation with no Cell Overlap

In this study we use the basic fixed channel allocation (FCA) as a reference scheme in order to compare the performance of the newly implemented channel allocation schemes that exploit cell overlap, so we first describe the FCA scheme.

#### 3.3.1 Description

The FCA scheme was implemented such that each cell was allocated a fixed set of channels and only allowed users to be allocated a channel from the nearest cell. No overlap was considered and hence no choices were made by the users about which cell should allocate them a channel.

#### 3.3.2 Performance

Figure 3-1 depicts how the standard FCA scheme where users are connecting to the closest cell performs against offered traffic (OT). It clearly shows that the system can maintain a 4%
blocking probability for an OT less than 9.5 Erlang per square unit. As mentioned before, the users in this scheme could only use channels from the cell they are located in.

![Figure 3-1 Fixed Channel Allocation with - no Cell Overlap](image)

3.4 Area Based Fixed Channel Allocation Scheme (ABFCA)

3.4.1 Description

The ABFCA scheme is based on the standard FCA scheme presented above. As before, each cell has a fixed number of channels, and can allocate them to any user within its coverage area. However, the cell radius (R) is now equal to 1.25 times the initial cell radius (R₀). This means that users positioned within radius R of the centre of any cell can connect to this cell. The users will first search for the number of cells they can connect to (up to 3 cells) and then they will pick a channel from the cell having the highest number of available channels. Thus, the user located in area C will pick a channel from one out of three cells within the range that has most channels available.

3.4.2 Performance

In this simulation the radius of the cells defining the degree of overlap is set to R=1.25. Notice that although the overlap radius has been increased, the OT within the total coverage area remains the same (as in the previous case with no overlap). Figure 3-2 depicts the blocking levels for areas A, B and C in the centre cell for the case where we consider cell overlap, in contrast to the standard FCA scheme. From the plot, it can be seen that overall (considering the total cell blocking levels), the overlap scheme performs better than the one with no overlap (previous scheme); however the blocking levels for the users positioned in area A, B and C vary considerably. Users in area A experience worse blocking levels than the users in area B and C. For areas B and C, where the users have the option to choose from more than one cell, the blocking levels are much lower than the case with no overlap.
From the results above we can verify that the areas formed by the overlap of two or more cells enjoy the advantage of higher trunking efficiency, and as a result, the blocking levels are much lower than the case where no overlap is considered. Therefore, the overall blocking levels of the ABFCA scheme are much lower than the no overlap case. Nevertheless, it is not permissible to favour any area (in this case B and C) for the sake of maintaining the lowest possible overall blocking levels in the system. The ABFCA scheme, however, indicates that utilising cell overlap can significantly reduce the blocking levels. In addition, there is great potential to improve the performance of ABFCA by careful analysis of the channel distribution based on the performance of the individual areas. More specifically, in this scheme it is useful to reduce the blocking levels in area A, by increasing slightly the blocking levels in area B and C. The aim is to try to do this by keeping the total cell blocking level below the non-overlap case while increasing the fairness of the scheme.

### 3.5 Region Based Fixed Channel Allocation Scheme (RBFCF)

#### 3.5.1 Description

The approach proposed at the end of section 3.4.2 has been used in design and implementation of Region Based Fixed Channel Allocation Scheme (RBFCF). Here, we are trying to control the channel allocation on a region-by-region basis. The total coverage area was partitioned into a number of small regions and instead of assigning a fixed number of channels within each cell we now have a size-based optimisation mechanism. This mechanism performs a numerical analysis in order to find the optimum set of channels that must be allocated in each of the regions.

The main concept when implementing this scheme was to ensure that there would be fairness in the service, in terms of uniform blocking probability. This model has been based on the ABFCA scheme with the only difference that the coverage area was partitioned into a number of small regions and instead of assigning a fixed number of channels within each cell, we had a size-based optimisation mechanism to assign the channels into the small regions. This mechanism was developed in order to find the right combination of channels to give the smallest blocking probability while at the same time ensuring uniform blocking probability all over the cell.
This required a number of iterations during which the blocking as well as the discrepancy between the areas were monitored in order to identify the optimum number of channels to be assigned in these regions. The number of channels allocated was a function of the size of the individual regions formed due to the overlap and the number of cells forming these regions. Channels assigned to a region can only be used within this region, and inter-cell channel borrowing (i.e. channels being used to another part of the cell apart from the one it is designated) is not permitted.

This optimisation technique does not perform well when cells are allocated a small number of channels, since having a small number of channels overall, a very small number will be allocated to each region. As a result, there might not be enough channels to be allocated to all regions, or some of the regions might have more channels than others, causing large non-uniformities in terms of blocking probability within the coverage area. This goes back to the fact that the channels allocated to the regions must be integers. It has been shown [3.2] that it is not possible to allocate the exact number of channels based on the size of the regions. It is therefore necessary to use the floor-function.

Partitioning areas into regions has been essential in order to examine the blocking in each region rather than in each area.

3.5.2 Performance

In this simulation, some of the basic parameters used in Table 3-1 were changed. The total number of cells was set to be 7, the number of channels per cell was increased to 50 channels, the total OT in the coverage area was from 6-12 Erlang per square unit area times the actual size of the coverage area which was equal to 19.817 units (for 7 cells). The radius of the cells defining the degree of overlap was set again to R=1.25.

The following table lists the number of channels allocated per region based on the optimisation technique. These channels refer to the centre cell illustrated in Figure 3-3 as this is the cell where the statistics are gathered. The rest of the areas in the surrounding cells are not of any interest.
Table 3-2 Channel Allocated per region in centre cell using RBFCA

<table>
<thead>
<tr>
<th>Region Number</th>
<th>0</th>
<th>7-12</th>
<th>19-24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region Type</td>
<td>A</td>
<td>B_{small}</td>
<td>C</td>
</tr>
<tr>
<td>Channels per Region</td>
<td>14</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

The reason for reducing the total number of cells to 7 and not keeping it to 37 is to reduce the complexity. This was feasible because the edge-effect problem was not an issue in this scenario. This is primarily because the channels were not allowed to be moved between regions and also because statistics were gathered from the centre cell.

Figure 3-4 depicts how the levels of blocking in different regions have become more fair than in the case of ABFCA scheme. However, the blocking levels in all regions have also increased beyond the standard FCA scheme's levels (that has been simulated again based on the new parameters). This comes as a result of partitioning the coverage area into smaller regions. Each cell now consists of different size regions and therefore the group of channels assigned to this cell is also partitioned into different smaller groups. As a result the trunking efficiency is significantly reduced.

![Figure 3-4 Region Based FCA scheme (RBFCA). Overlap has improved system fairness.](image)

Furthermore it can be seen in Figure 3-4 that the blocking levels are not completely uniform across the cell especially at higher OT levels. This is because it is difficult to accurately divide the channels based on the offered traffic of every region. In order to improve the uniformity, a greater number of channels should be used, however their availability is limited by the spectrum allocation.

### 3.6 Uniform Fixed Channel Allocation Scheme (UFCA)

RBFCAl has shown that it is possible to introduce fairness to the system but this comes at the expense of low trunking efficiency and therefore higher blocking levels. ABFCA has shown the potential in terms of its trunking efficiency and the reduced blocking probability across the cell area at the expense of non-uniform blocking across the cell area. UFCA is intended for the first
time to exploit both the fairness given by the RBFCA and also the improved trunking efficiency achieved by ABFCA.

### 3.6.1 Description

Here, each cell has a fixed number of channels, which can be allocated to any user within its coverage area. In this scheme, certain restrictions are imposed in order to prevent a proportion of the channels from being allocated to the overlap areas, to allow them to remain available to areas with no overlap (area A). The parameters chosen for this scheme take into account the performance of the ABFCA scheme shown in Figure 3-2. The aim of this scheme is to improve the results illustrated in Figure 3-2, so that the blocking in all regions becomes uniform as well as lower than the standard FCA model. From Figure 3-2 it is apparent that more channels have to be allocated for the users in area A, and less channels for the users in areas B and C.

One way of doing this without directly shifting channels from one area to the other, which requires partitioning the cell into small regions (e.g. RBFCA in Figure 3-4), is by blocking a proportion of users in area B and C even though there are channels available. The channels saved from area B and C can then be used in area A. As a result, the blocking of area A will be decreased and the blocking in area B and C will increase. The proposed model saves the last channel in area B and the last two channels in area C from being automatically used. Instead, these channels are only used on a random basis, i.e. when a random number generated is greater than a random acceptance factor (RAF). The RAF thus comes into effect every time a user in either area B or C requests a channel and there is only one (or two in the case of area C) left available for that area. After experimentation, it was found that this combination is proven to perform better than the other combinations tried since it was impossible to ensure uniform blocking in each region. Figure 3-5 illustrates diagrammatically how the UFCA scheme operates.

![Diagram of UFCA scheme with RAF](image)

**Figure 3-5 UFCA – RAF defines probability of acceptance in area B and C**

The optimum value for RAF was found, after a numerical investigation, to be a function of the level of offered traffic per unit area, and is given by:

$$\text{RAF} = \alpha + \beta \cdot \ln \left( \frac{\text{OT}_{\text{optimum}}}{\text{OT}_{\text{varying}}} \right)$$  \hspace{1cm} (3.4)

where $\alpha$ defines the probability that the last channel will be saved by blocking the user in the area of reference (area B or C) and $\beta$ is a scaling factor for the RAF to optimise it for a range of
offered traffic. OT\textsubscript{Optimum} is the chosen value of offered traffic we wish to optimise the system for. Initially, \( \beta \) is set to zero in order to find the best probability factor \( a \) by requiring the blocking levels in all three areas to be as close as possible to each other. For example, in Figure 3-2, it can be clearly seen that the blocking of area A increases approximately exponentially. We therefore need to increase individually the blocking in area B and area C using different RAF scaling factors and as a result decrease the blocking in area A. This technique enables us to control the number of channels being allocated into certain areas without partitioning the group of channels of a cell into smaller groups, and retaining the maximum number of channels to be available for localised “hot-spots” in traffic demand.

3.6.2 Performance

As before, the radius of the cells defining the degree of overlap is set to \( R=1.25 \). The parameters \( a, \beta \) and the OT\textsubscript{Optimum} used in this simulation are given in Table 3-3.

<table>
<thead>
<tr>
<th>Area Type</th>
<th>( a )</th>
<th>( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.405</td>
<td>1.7</td>
</tr>
<tr>
<td>C</td>
<td>0.045</td>
<td>1.7</td>
</tr>
<tr>
<td>OT\textsubscript{Optimum}</td>
<td>8.8</td>
<td></td>
</tr>
</tbody>
</table>

From the results shown in Figure 3-6 we can clearly see that the total cell blocking has been reduced when compared to the non-overlapping FCA scheme. Furthermore, the blocking levels in all regions are approximately the same.

In comparison with the other schemes, this model has improved the QoS by reducing the blocking probability in all areas, and has also achieved uniform blocking levels across the cell. This absolute level of improvement has been achieved with even fewer channels than the RBFCA. More specifically, the blocking probability has dropped by approximately 0.025 at OT\textsubscript{Optimum} and the OT supported increased by 10.2% than in the standard case with no overlap. This improvement does not require any prior knowledge of the interference environment.
The significance of the RAF is that the control mechanism of channel allocation has now changed from a limited number of channels that was before (e.g. RBFC) to potentially unlimited granularity in terms of time. If for example there are 100 channels per cell there could be a granularity of +/-1 channel (e.g. RBFC). With the RAF assuming an unlimited time we have an unlimited granularity and although we are restricting the number of users coming into the system, we have maintained maximum trunking efficiency. RAF avoids the assignment of discrete integer number of channels to each area or region because RAF is applied over a very large number of connection requests. RAF performs the optimisation based on time which is potentially infinite rather than a restricted number of channels. It therefore controls the QoS over time rather than actually restricting channels indefinitely.

Up to now, we have assumed a cell radius of $R=1.25$. Figure 3-7 however shows that this technique is also applicable for other overlap radii. The simulation was performed for a fixed OT of 8.8 Erlangs per square unit area and a range or radius $R$ ranging from 1-1.5.

![Figure 3-7 Uniform FCA scheme (UFCA) blocking probability in centre cell for different degrees of overlap radius.](image)

As the values of $a$ and $\beta$ for areas B and C are optimised for $R=1.25$, the optimisation clearly does not work for lower overlap radius as the blocking probabilities in area A, B and C are significantly different; showing that the values of $a$ and $\beta$ are functions of the amount of overlap. Although the blocking probability decreases as the cell overlap increases, a radius of greater than $R=1.25$ assumes that users can cope with the increased level of interference.

### 3.7 Uniform Fixed Channel Allocation - II (UFCA - II)

#### 3.7.1 Description

In section 3.6, it has been shown that cell overlap can be exploited to improve blocking levels and enhance system capacity for a scheme that uses fixed channel allocation. The areas that are served by more than one cell benefit from the higher trunking efficiency and as a result they will have much lower blocking than the areas served by one cell. In order to improve uniformity and reduce the blocking probability within a cell, a technique was developed called Uniform Fixed Channel Allocation (UFCA). As shown before, this technique has made it possible to control the number of channels being allocated to certain areas without partitioning the channels allocated in a cell (and thereby reducing the trunking efficiency) while still ensuring uniform
blocking levels within the coverage area. Channel assignment strategies such as in \[3.3\] exploited cell overlap, but because they did not take into account the different blocking rates in different areas they were inherently unfair. The UFCA scheme \[3.4\] took into account the different blocking rates in each area to make the system fair but it did not exploit CNIR effectively.

Based on previous work \[3.4\] and \[3.3\], a new channel allocation scheme, in fact an improved UFCA scheme, has been developed which takes into account the effect of interference on channel allocation based on cell overlap. What UFCA-II effectively does is to take into account the carrier to noise plus interference ratio (CNIR) values, and uses a group of multilevel modulation schemes and multiple channels to achieve better QoS and fairness by exploiting the differences in CNIR that exist between the regions. More specifically, it chooses to pick up a number of channels from the base station that is within range and that has the most available channels. The number of channels required depends on the minimum bits per connection BpC threshold (normalised to the frame rate) defined to ensure fair quality of service in all regions and the modulation scheme that can be supported in each channel.

3.7.2 Performance

Before moving into describing the operation and results of the UFCA-II scheme, it is important to explain in more detail how the cell overlap model was implemented and what sort of additions / alterations had to be made in comparison to the previous model used for the UFCA scheme.

3.7.2.1 Cell overlap based on minimum received power threshold

As mentioned before, cell overlap occurs because of the way the power decreases away from the boresight of the antenna. For UFCA-II, the size of the overlapping area is now determined by setting a minimum received power threshold that is determined from the link budget and is related to the power roll-off with angle from the antenna gain profile \[3.5\]. This minimum received power threshold is a function of the transmit power of the base stations located on the HAP. So, in effect the radius of cells is set by the fixed transmit power of the base stations at the platform. This is different from the previous cell overlap model used for UFCA where cell overlap was entirely based on the actual distance from the centre of the cells.

To elaborate further, all the calculations for defining the boundaries of the cells in terms of radius \( R \) were made based on a number of assumptions. The first assumption made was that all calculations were made in the absence of interference. Furthermore, the HAP base stations transmit power should be fixed and the value has been calculated using the following equation:

\[
P_{TX} = \frac{P_{RX} \cdot d^2}{G_{TX} \cdot G_{Ref}} \tag{3.5}
\]

where, \( P_{TX} \) is the base station transmit power, \( d \) is the distance between the user and the HAP, \( G_{TX} \) is the gain of the HAP antennas and \( G_{Ref} \) is the reference gain, which can be calculated based on the user antenna gain \( G_{RX} \) and the wavelength of the frequency of interest.

\[
G_{Ref} = \frac{G_{RX}}{\left(\frac{4\pi}{\lambda}\right)^2} \tag{3.6}
\]

where \( \lambda \) is the wavelength of the signal carrier.

The peak directivity \( D_{max} \) is approximated by using the flat sidelobe model as been described in \[3.6\]:

---

Page 37 of 110
\[ D_{\text{max}} = \frac{32 \ln(2)}{\left[ 2 \cdot \arccos \left( \frac{n_T}{\sqrt{2}} \right) \right]^2 + \left[ 2 \cdot \arccos \left( \frac{n_F}{\sqrt{2}} \right) \right]^2} \tag{3.7} \]

where \( n_T \) and \( n_F \) are the indices for optimising directivity at the cell edges [3.5]. Both are functions of the antenna 3dB beamwidths.

The directivity \( D \) seen at a point where a user is positioned can be calculated as:

\[ D = D_{\text{max}} \cdot \left\{ \cos(\theta_{\text{user}} \cdot \cos(\phi_{\text{user}})) \right\}^{n_T} \cdot \left\{ \cos(\theta_{\text{user}} \cdot \sin(\phi_{\text{user}})) \right\}^{n_F} \tag{3.8} \]

where \( \theta_{\text{user}} \) and \( F_{\text{user}} \) are the elevation and azimuth angles for a user on the ground relative to the boresight of the base station of interest.

Equation (3.9) can therefore be applied to calculate the power received at this position which is a distance \( R \) away from where we see the maximum directivity. Thus, the power level required \( (P_{\text{RX}}) \) from a user to be able to connect to the base station of interest is defined as follows:

\[ P_{\text{RX}} = D_x \cdot \frac{G_{\text{Ref}}}{d^2} \tag{3.9} \]

Assuming that a user is \( R \) units away from the centre of the cell, then let \( \theta_x \) and \( F_x \) be the two angles that can be used to calculate the radius of the cell. As a result, the directivity at this position will become \( D_x \).

Applying equation (3.8) to equation (3.9) and then to equation (3.5) we can calculate the value of the transmit power required from the base station to ensure coverage of a cell of radius \( R \). The following table lists some of the link budget parameters we have used for the simulation:

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Noise received level ( (N_{\text{RX}}) )</td>
<td>-133.9</td>
<td>dBm</td>
</tr>
<tr>
<td>2</td>
<td>Carrier frequency ( (f) )</td>
<td>28</td>
<td>GHz</td>
</tr>
<tr>
<td>3</td>
<td>Reference gain ( (G_{\text{Ref}}) )</td>
<td>-82.4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Platform height ( (h) )</td>
<td>22</td>
<td>km</td>
</tr>
<tr>
<td>5</td>
<td>Transmitter power ( (P_{\text{TX}}) )</td>
<td>-26.6</td>
<td>dBm</td>
</tr>
<tr>
<td>6</td>
<td>Number of cells ( (c) )</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Number of channels per cell</td>
<td>240</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Cluster size ( (K) )</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>External cell radius ( (R_e) )</td>
<td>1</td>
<td>km</td>
</tr>
<tr>
<td>10</td>
<td>Overlap radius ( (R) )</td>
<td>1.25</td>
<td>km</td>
</tr>
<tr>
<td>11</td>
<td>Offered traffic ( (OT) )</td>
<td>1</td>
<td>Erl/km^2</td>
</tr>
<tr>
<td>12</td>
<td>RX antenna gain ( (G_{\text{RX}}) )</td>
<td>38</td>
<td>dB</td>
</tr>
</tbody>
</table>
In order to obtain realistic results and to allow the multi-level channel assignment of the users to operate, an increased number of channels per cell has been used. These results cannot be directly compared with the previous schemes. UFCA-II however exploits the previous findings to improve the QoS whilst maintaining the blocking fairness.

Figure 2-4 in Chapter 2 depicts a hexagonal layout of cells, and marked on one of the cells are the circle which encloses a hexagon, it has a radius of $R_e$. The value of the overlap radius $R$ varies within limits defined from the minimum power received threshold. The minimum value of the radius of the overlapping cell is chosen to be equal to the original radius of the cell $R_e$ in order to avoid leaving any areas without service. As mentioned previously, the maximum value is taken to be equal to $1.5R_e$ radius, in order to limit the maximum number of overlapping cells to three: but theoretically four or even more can overlap if the cell radius is increased sufficiently, assuming the co-channel interference and received power is acceptable.

As before, the UFCA-II work has been based on a 6 neighbouring cell model. This is the most common case in a large-scale cellular communication system. The overlap radius $R$ was set to 1.25km whereas the initial external radius of the cells $R_e$ was set to 1km (Figure 2-4 in Chapter 2). For an overlapping radius of 1.25km, a fixed transmit power from the HAP of 0.07mW was required. This value was calculated based on the CNIR levels of Table 3-5.

In simulations, the major difference considered in traffic model assumptions as compared with the UFCA scheme is that user’s selection of base stations depends on their minimum received power threshold as opposed to the distance that UFCA employed. Also, new users are blocked if they cause degrading of CNIR levels of existing users.

### 3.7.2.2 Effects of Interference on UFCA scheme performance

It has been shown that the UFCA model [3.4] provides fair channel distribution among different cell areas while ensuring minimum blocking levels. However the data rate was determined by the area with the worst Channel to Noise Ratio (CNR) levels by setting a minimum data rate threshold and having a fixed modulation scheme that could be supported in all areas. This would ensure uniform data rates as well as uniform blocking levels. However, it did not exploit the differences in the CNIR levels between the areas. Introducing interference to the new cell overlap model and using CNIR levels as metric for determining which of the supported modulation schemes from Table 3-5 to use, it was possible to exploit these differences.

We have first looked at the CNIR plots for the case with no overlap and for the case where we allow overlap of $R=1.25km$. The following plot illustrates the CNIR levels of the centre cell for cluster size of 7.
From Figure 3-8 it can be seen that by allowing overlap of $R=1.25\text{km}$, the users at the edges of the cell (i.e. the users located in the overlap areas) experience lower CNIR levels than the users in the non-overlapping areas (area A). As a result users in these areas cannot connect to high rate modulation schemes that can give higher bit-rates. Although the blocking probability in the areas of overlap is much lower than in area A (and the case of no overlap), connections in these areas will be of much lower data rates. The modulation schemes assumed for this simulation are listed in Table 3-5.

### Table 3-5 Modulation and Coding figures used to determine capacity

<table>
<thead>
<tr>
<th>Mod. Scheme</th>
<th>BPSK</th>
<th>QPSK</th>
<th>8PSK</th>
<th>32-QAM</th>
<th>64-QAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bits / Symbol</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>CNIR(dB)</td>
<td>2.3</td>
<td>10</td>
<td>15.5</td>
<td>17</td>
<td>25</td>
</tr>
<tr>
<td>Eff-CNIR(dB)</td>
<td>2.3</td>
<td>15</td>
<td>20.5</td>
<td>22</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 3-5 lists the modulation schemes that have been used along with the standard CNIR value and the effective CNIR value that users must fulfil to be able to use the respective scheme. For this work it has been assumed that the interference is Gaussian.

### 3.7.2.3 UFCA-II Algorithm

As mentioned before, the initial UFCA scheme ensured uniform blocking levels between the regions but it has been shown that users in different regions can support different data rates. To overcome this problem the following algorithm graphically presented in Figure 3-9 has been devised.
Stage 1: To solve the problem, more than one channel is allocated to a user in order to improve the total number of bits per connection per frame (BpC) and therefore increase its data rate to an acceptable level. A user will keep requesting more channels in order to satisfy the minimum acceptable level BpC_{thres} of 15 bits/C/F. (Table 3-4) Channels can support different number of bits per symbol depending on the chosen modulation scheme. The most extreme case will be that of when all available channels can support only 1 bits per symbol (bps) and therefore more channels will be required for one user to satisfy the BpC_{thres}.

However, better modulation schemes can be supported in the areas with the least interference and therefore fewer channels will be required per user. According to Figure 3-8, users in area A are more likely to request fewer channels than users in area B or area C to reach the minimum bits per symbol threshold.

Stage 2: Users select the BS within range that has the most available channels. The channel selection is based on which channel is first found to be available in the chosen BS with an acceptable CNIR level for BPSK.

The extra channels a user will request should come from the same base station that the initial channel has been assigned from. It could be possible to allow users to be allocated channels from any base station within range (for users in area B and C) but for simplicity it was preferred to allow channel assignment from one cell only.

However, new users are blocked if they disturb any currently active users by degrading their bit rate level below their current value. In order to prevent this happening too often, a safeguard CNIR margin called Eff-CNIR (Table 3-5) of 5dB has been set when deciding which modulation scheme to use. This value was chosen upon investigation which gave the lowest blocking probability levels. This is based on the concept of hysteresis [3.7] where there is an acceptance Eff-CNIR threshold and an interruption CNIR threshold. By introducing this margin, users are less susceptible to interference and can afford to lose some of their CNIR levels before they actually reach the interruption CNIR levels, thus allowing the new users to use the same channel in a different cell without being blocked or without degrading the service of existing users.
More specifically, new users can connect to the system if their CNIR levels are above the Eff-CNIR levels and the current users CNIR levels do not drop below the interruption CNIR levels. As a result, for this work users connect to the scheme that can be supported from their CNIR level assuming that their CNIR is 5dB less than what it really is. This margin is applied when choosing a modulation scheme from Table 3-5 higher than BPSK, so that users will not be completely blocked as far as their CNIR can support BPSK. In the case that not enough channels have been obtained to sum up and satisfy the $B_{pC_{thres}}$ threshold, all channels that have been reserved will be released and the user will be blocked.

In the following we report simulation results describing the performance of Uniform Fixed Channel Allocation Scheme II (UFCA-II). All tests investigated the effect of CNIR on capacity and blocking levels. The first set of results presented is for the case where there is no overlap. Then we introduce overlap but without applying the Random Acceptance Factor (RAF) technique. Finally, the RAF technique is applied.

3.7.2.4 Simulation Model without Overlap

Users in this model have to connect to the closest BS, so areas B and C do not exist. The cell radius is set to 1km. More information regarding the parameters used for this simulation can be found in Table 3-4.

![Figure 3-10 CNIR and Channel usage plots without cell overlap](image)

Figure 3-10 depicts the cumulative distribution function (cdf) of CNIR and the cdf of the channels in use for the case of no overlap. These results were used to compare the performance of the no overlap case with the overlap case.

3.7.2.5 Simulation Model with Overlap

The radius of each cell now is set to 1.25km. Users can connect to up to 3 BSs depending on their location. This simulation was also based on the parameters used in Table 3-4.

From Figure 3-11 it can be seen that we have a multilevel usage of channels in each area. This depends primarily on the CNIR levels which for the overlap areas B and C are worse than for area A. As a result, users in these regions have to use lower modulation schemes and therefore there is higher demand for channels to satisfy the minimum bits per connection threshold level ($B_{pC_{thres}}$). Taking as an example the channel usage in area C, we can see that most of the time
users require on average more channels than users in area A or B. Although the blocking levels in area A, B and C differs (Table 3-7), we have managed to guarantee uniform bit rate for all users.

3.7.2.6 Applying RAF

After ensuring uniform bps service, it is important to ensure uniform blocking service within the coverage area. To do so, the RAF technique is being used. Repeating the same technique as in [3.4], we have numerically calculated a new set of $a$ and $\beta$ parameters listed in Table 3-6. After a numerical investigation, it was found that the RAF becomes active when there are only 5 channels left in each of the base stations that are within range.

<table>
<thead>
<tr>
<th>RAF Parameters</th>
<th>$a$</th>
<th>$\beta$</th>
<th>$\text{OT}_{\text{saving}}$</th>
<th>$\text{OT}_{\text{optimum}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.6</td>
<td>0.2</td>
<td>6.5 erlang/km$^2$</td>
<td>6 erlang/km$^2$</td>
</tr>
</tbody>
</table>

Figure 3-12 depicts the results for the RAF model. The CNIR statistics vary very little when compared with the previous scheme. However, applying the RAF we can provide more equal blocking probabilities in areas A, B and C. Applying the RAF the overall level of blocking is approximately 2.5% compared with the no overlap case with 4.2% as shown Table 3-7. Clearly, we have improved the QoS significantly at no expense of the capacity of the system.
Note that overlap areas B and C experience lower CNIRs than area A. As a result, more channels are required to be assigned per connection in these areas in order to guarantee the $B_{pC_{thres}}$ level set. However, users in these areas get increased flexibility that effectively reduces the number of channels a cell may have to assign them. As a result, users in area A will on average have a greater number of channels available.

### 3.8 Summary and Conclusions

This section of the report, has been dedicated to the exploitation of cell overlap and the investigation of various channel allocation techniques for improving the Quality of Service in terms of blocking probability, data rate and fairness. These techniques are ABFCA, RBFCA, UFCA and UFCA-II, and their performance has been compared to that of the basic FCA scheme without exploiting of cell overlap.

Comparing the ABFCA scheme with the no overlap case showed that the system performance in terms of blocking probability has improved since the overlap areas benefit from the increased channel availability. This however introduces a non-uniform blocking between the areas and therefore across the coverage area that is not fair for the users. To address this problem, RBFCA has been developed. This is an optimisation scheme that allocates channels directly to regions in order to unify the blocking across the cell. From the results, it has been shown that allocating channels into regions improves fairness across the cell but this comes at the cost of...
much higher blocking levels. Furthermore, it has been shown that trunking efficiency becomes an issue when allocating channels to regions and as a result the optimisation technique cannot guarantee uniform blocking.

The results obtained with the UFCA scheme have shown that cell overlap can be exploited to improve the performance for a scheme that uses fixed channel allocation (FCA). The areas served by more than one cell (overlap areas) benefit from the higher trunking efficiency and as a result they have much lower blocking than the areas served by one cell. To improve uniformity and reduce the blocking probability within a cell, a technique has been developed called Random Acceptance Factor (RAF) where certain restrictions have been imposed in order to prevent a proportion of the channels from being allocated to the overlap areas. As a result, these channels are re-directed in areas with no overlap (area A). This technique enables to control the number of channels being allocated into certain areas without partitioning the coverage area into smaller regions. It will therefore be of particular benefit in situations where there is non-uniform distribution of traffic and will also aid handoff procedures.

The UFCA model [3.4] has been perfectly fair in terms of the channel distribution while ensuring minimum blocking levels. However the data rate was determined by the area with the worst Channel to Noise Ratio (CNR) levels by setting a minimum data rate threshold and having a fixed modulation scheme that could be supported in all areas. This ensured uniform data rates as well as uniform blocking levels. However, it did not exploit the differences in the CNR levels between the areas.

The improved UFCA-II scheme looks at different ways of selecting a base station (BS) and assigning a channel or more to a user. The UFCA-II model exploits both cell overlap and CNIR and delivers uniform QoS. This has been achieved by introducing a bit per connection threshold level $B_{\text{thres}}$ that all user connections must satisfy. This allowed the CNIR to be better exploited in the channels. This threshold in conjunction with the RAF has ensured uniform blocking across the areas as well as equal data rates across the user connections. For the parameters chosen it has been shown that the blocking probability can be reduced from 4% in no overlap case using the basic FCA scheme to 2.6% by exploiting the UFCA-II scheme with RAF technique.
3.9 References


4. Guard Channel Connection Admission Control

4.1 Introduction

The general operating scenario considered for the development of the Connection Admission Control algorithms (CAC) is shown in Figure 4-1. Platform interconnection via inter-platform links provides extended system coverage with significantly reduced terrestrial infrastructure. To support communication between adjacent platforms without any ground network elements each HAP payload includes a switching device and one or more inter-platform link terminals. Depending on the link budget analysis we can choose between optical and radio frequency terminals. In this scenario ground stations are used mainly as gateways to other public and/or private networks, while providing also a backup interconnection between platforms in the case of IPL failure.

![Figure 4-1 General operating scenario](image)

In the general operating scenario, fixed users and mobile users select the most suitable platform (possibly the one with strongest signal level) and initiate the authentication and admission procedure in order to gain access to the network. Arrivals or departures of fixed users can be considered as single events while the arrivals or departures of handoffs occur in batches since the mobile users travel in vehicles and are connected to HAP via on-board collective terminals.

As in other mobile networks, connection admission control in HAP networks should give higher priority for connections with a handoff request. The aim of CAC algorithms should be to keep the dropping probability under a predefined threshold while minimising the blocking probability and maximising the bandwidth reservation at the same time. Dropping probability is the probability that a request for handoff is rejected, while blocking probability is the probability that a new connection is blocked.

The guard channel policy adopted here is widely used in CAC schemes for mobile networks with handoff requests. Guard channels are defined for each traffic class to prevent the acceptance of new connections from a particular class, leaving the remaining channels available for handoff
requests when the resource usage becomes high. Mobile systems with multiple class traffic make the process of finding optimal guard channels very complicated and CPU-time-consuming due to multi-dimensional Markov chains [4.1]. Additionally, the consideration of batch handoffs makes the development of effective CAC algorithms for HAP networks more challenging.

The use of the guard channel CAC algorithm for the general HAP operating scenario is investigated in this chapter. An approximation method is proposed for the calculation of the guard channels to keep the resource reservation as high as possible while guaranteeing a certain degree of Quality of Service (QoS), i.e. guaranteeing minimum bandwidth for each connection while keeping the dropping probability of handoff connections under a predefined threshold.

4.2 Problem Formulation

Modern data communication networks should be designed and dimensioned to serve real-time and non real-time traffic classes. For instance, the IEEE 802.16 standard, which has been identified in CAPANINA as the most suitable candidate for the provision of broadband communications services [4.2], specifies four different QoS classes for the uplink:

- **UGS** Unsolicited Grant Service - For real-time service flows which generate fixed size packets on a periodic basis, e.g. VoIP without silence suppression.
- **rtPS** Real-Time Polling Service - For real-time service flows which generate variable size packets on a periodic basis, e.g. MPEG video.
- **nrtPS** Non-Real-Time Polling Service - For non-real-time service flows with guaranteed minimum rate.
- **BE** Best Effort service – For service flows without any QoS guarantee.

Real-time traffic classes are sensitive to delay and/or delay variation, while non-real-time connections are not sensitive on delay factors. CAC and bandwidth allocation support the implementation of QoS at the flow level and the packet scheduling is one of the tools for QoS implementation at the packet level. At the flow level, the delay factors are not usually considered directly, or they are considered with certain constraints and assumptions about the underlying packet scheduling algorithms. At this level, the bandwidth requirements are the primary factors under consideration. Connections may require a fixed amount of bandwidth, e.g. for voice without silence suppression, or require a minimum bandwidth guarantee, e.g. for video streaming and non real-time traffic. Best effort traffic can be considered as a special non-real-time class without the minimum bandwidth guarantee.

In HAP networks there are three types of user connectivity that can be offered: First type is when user terminals are connected to the fixed type of Customer Premises Equipment (CPE), which performs the processing of signals from/to an antenna. Second type is when mobile users in “public” transport means such as buses, trains, ships etc. access the network via a collective on-board terminal interfacing between the HAP on one side and moving LAN on the other. Third type is when both fixed and mobile users are equipped with terminal for direct access to the HAP. For a mobile wireless system with the arrival of new connections and handoff requests, and with the support for QoS, the most widely used CAC strategy is the guard channel scheme, proposed in the mid 80s [4.3]. However, the development of a fast calculation method for optimal guard channels for each traffic class is still an open problem. In addition, the presence of mobile users as collective travellers on vehicles makes the requirements for the HAP scenario significantly different from other mobile systems in the sense of batch handoff. That is, if a bus or a train moves across the boundary of cells, then all users located in the vehicle have to perform handoff at the same time. The main objective of this work was to find a solution to this problem, which requires a small amount of CPU-time and addresses the specific problem of batch arrivals of handoff requests.

The following model applies to a general HAP network. Consider a system with different traffic classes, with \( b_i \) as the minimum bandwidth guarantee for class \( i \) (\( i = 1, 2, \ldots, N \)). With \( G_i \) we denote the guard channel for class \( i \). At the CAC decision time instant \( t \), we have \( n_i(t) \)
connections of class \( i \) admitted into the network. Given that the link capacity is \( C \), the guard channel admission policy is as follows:

A connection request arrives from class \( i \) at time \( t \).

If the incoming request is a new connection then

\[
\text{if } b_i < G_i - \sum_j n_j(t) b_j \text{ then}
\]

The new connection is admitted.

else

The new connection is blocked.
end if

else

\[
\text{if } b_i < C - \sum_j n_j(t) b_j \text{ then}
\]

The handoff request is admitted.

else

The handoff request is dropped.
end if

end if

The sum \( B(t) = \sum_j n_j(t) b_j \) denotes the total bandwidth reservation at time \( t \), so \( B(t) \) can be considered as a random variable. Let \( d_i \) be the dropping probability of class \( i \), then we have to find the optimal \( G = \{G_i, G_2, ..., G_N\} \) to maximise the expected bandwidth consumption while guaranteeing that dropping probabilities remain below predefined thresholds:

\[
P: \quad \max \mathbb{E}(B),
\]

subject to \( \frac{d_i}{G_i} \leq C, i=1, 2, ..., N \)

We assume that the arrival of individual new connection and batch handoff requests are Poisson processes. We denote by \( \alpha_i \) the arrival rate of new connections of class \( i \) and by \( \alpha^H \) the arrival rate of handoff batches (a batch can contain multiple connections of different classes). The expected value of the number of connections of class \( i \) inside a batch is \( m \). Furthermore, we assume that the holding time of connections of class \( i \) is exponentially distributed with expected value of \( 1/\mu_i \). We have the same assumption for the cell residence time for handoff batches with expected value of \( 1/\mu^H \).

If we separately consider the admission of connections within a batch, then the order of connections in the admission process is crucial. If a given connection is processed earlier than another connection, its chance to be admitted is obviously higher than that of the later one. That is, the order of admission consideration affects the dropping probability of each traffic class. Similarly, if we consider the admission of connections in a batch together, then in case there is not enough available bandwidth for the whole batch, the dropping policy of calls to decrease the bandwidth requirements of the batch also affects the dropping probability of each traffic class.

We assume a random order for connection admission processing. Furthermore, we only consider the common dropping probability of all connections instead of the separate dropping probability of each traffic class.
With this approach and the consideration of the common dropping probability for all traffic classes, the optimisation problem $P$ is rewritten as:

\[
P^*:
\begin{align*}
  \text{maximize} & \quad E(B), \\
  \text{subject to} & \quad P^d < d, \\
  & \quad \text{over } G_i: 0 < G_i \leq C, \; i = 1, 2, \ldots, N
\end{align*}
\]

For traditional 2G and 3G cellular mobile systems where most of the handoffs do not occur in batch mode, the optimal vector of $G$ can be obtained by writing multi-dimensional Markov chain. However, solving of such Markov chains is very complicated and heavily CPU-time consuming. Furthermore, the overall space of $G$ is too large to execute the optimization numerically. The introduction of batch handoffs makes the situation even more complicated: there are transitions not only between adjacent states in the multi-dimensional Markov chain, but also between non-adjacent states. This calls for some approximation method that allows simplified calculation of the guard channels, such as proposed in the following.

### 4.3 An Approximation Solution

In order to have simple method to approximate optimal solution for $G$ we use a transformation approach. A similar approach called the splitting approach has already been proposed in [4.4] to split the multi-dimensional Markov chain into one-dimensional Markov chains. However, the splitting solution in [4.4] does not take the multiplexing nature of aggregate traffic in a multi-class environment into account.

The transformation approach consists of $N$ steps: the guard channel $G_i$ is approximated in step $i$ ($0 < i < N$) as follows. All connections of class $j \neq i$ are replaced by connections of class $i$ with minimum bandwidth of $b_i$, instead of $b_j$, holding time of $1/\mu_i$ instead of $1/\mu_j$. By this replacement, there is only one traffic class in the system and the multi-dimensional Markov chain becomes one-dimensional. The transformed arrival rate $\alpha_{ji}$ of new connections of class $j$ is calculated as:

\[
b_j \frac{\alpha_{ji}}{\mu_j} = b_i \frac{\alpha_{ji}}{\mu_i}
\]

Therefore,

\[
\alpha_{ji} = \frac{b_i}{b_j} \frac{\mu_i}{\mu_j} \alpha_j
\]

The arrival rate of $\alpha^H_{mj}$ of handoff requests is transformed to:
In equation (4.3) $\frac{b_j}{\mu_j + \mu^H} a_j^H m_j$ represents the termination rate of handoff requests. The termination time of a handoff request is the minimum of holding time and cell residence time. The transformation of parameters is summarised in Table 4-1.

**Table 4-1 Transformation rules for connections from class $j$ to class $i$**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum bandwidth requirement</td>
<td>$b_j$</td>
<td>$b_i$</td>
</tr>
<tr>
<td>Holding time</td>
<td>$1/\mu_j$</td>
<td>$1/\mu_i$</td>
</tr>
<tr>
<td>Arrival rate for new calls</td>
<td>$a_j$</td>
<td>$\alpha_j = \frac{b_j}{b_i} \cdot \frac{\mu_i}{\mu_j} a_j^H$</td>
</tr>
<tr>
<td>Arrival rate for handoff calls</td>
<td>$a_j^H m_j$</td>
<td>$\alpha_j^H = \frac{b_j}{b_i} \cdot \frac{\mu_i}{\mu_j + \mu^H} a_j^H m_j$</td>
</tr>
</tbody>
</table>

The one-dimensional Markov chain after transformation for step $i$ is shown in Figure 4-2, where the aggregate arrival rate of new connections is:

$$\alpha_{TSN} = \sum_{j=1}^{N} \alpha_{ji}$$  \hspace{1cm} (4.4)$$

and the aggregate arrival rate of handoff requests is:

$$\alpha_{TSN}^H = \sum_{j=1}^{N} \alpha_{ji}^H$$  \hspace{1cm} (4.5)$$

$H_i$ is the guard channel measured by the number of connections ($G_i = b_i H_i$).
Let \( p_i(j, H) \) denote the steady-state probability that there are \( j \) calls in the system of step \( i \) with the guard channel of \( H \):

\[
p_i(j, H) = \begin{cases} 
\frac{\alpha_i^{TSN} + \alpha_i^{TSH}}{\mu_i} \frac{(\alpha_i^{TSN} + \alpha_i^{TSH})^j}{j!} p_i(0, H) & j \leq H \\
\frac{\alpha_i^{TSN} + \alpha_i^{TSH}}{\mu_i} (\alpha_i^{TSH})^{j-H} \frac{(\alpha_i^{TSN} + \alpha_i^{TSH})^H}{H!} p_i(0, H) & H < j \leq \lfloor C/b_i \rfloor 
\end{cases}
\]

(4.6)

The larger the \( H \), the higher the expected bandwidth reservation and the same time the higher the dropping probability. Thus we need to find the largest \( H \) that still keeps the dropping probability smaller than the threshold \( d \).

\[
H_i = \max H : P_i^d(H) < d
\]

(4.7)

We suppose that the aggregate bandwidth requirement of a batch is exponentially distributed with the expected value:

\[
S = \sum_{k=1}^{N} b_k m_k
\]

(4.8)

Given that there are \( j \) connections in the system, the expected dropping probability is:

\[
p_i^d(j) = P(\text{connection is dropped} \mid \text{in the system are} \ j \ \text{connections}) = \\
\int_{r_i(j)}^{\infty} \frac{1}{S} e^{-\frac{x}{S}} (x - r_i(j)) dx = e^{-\frac{r_i(j)}{S}}
\]

(4.9)

where \( r_i(j) = C - jb_i \)

(4.10)

The dropping probability can be calculated as:
\[ P_i^d(H) = \sum_{j=1}^{C/H} p_i(j, H) P_i^d(j) \] (4.11)

Note that \( P_i^d(j) \) is independent of \( H \), which makes the calculation of \( H \) simpler. Furthermore, to fasten the process of finding optimal \( H \), we can use the binary search function \( \text{BinarySearchGuard} \) described in the following.

**Require**: \( \text{begin} = \text{end} \)

\[ \text{if } \text{begin} == \text{end} \text{ then} \]
\[ \quad \text{if } P_i^d(H) \leq d \text{ then} \]
\[ \quad \quad \text{return } \text{begin} \]
\[ \quad \text{else} \]
\[ \quad \quad \text{return } \text{notfound} \]
\[ \text{end if} \]
\[ \text{else} \]
\[ \quad H \leftarrow \left\lfloor \frac{\text{begin} + \text{end}}{2} \right\rfloor \]
\[ \text{if } P_i^d(H) < d \text{ then} \]
\[ \quad \text{return } \text{BinarySearchGuard}(H, \text{end}) \]
\[ \text{else if } P_i^d(H) == d \text{ then} \]
\[ \quad \text{return } H \]
\[ \text{else} \]
\[ \quad \text{return } \text{BinarySearchGuard}(\text{begin}, H) \]
\[ \text{end if} \]
\[ \text{end if} \]

### 4.4 Simulation

The approach proposed in the previous section is an approximation method to calculate the optimal guard channel. The aim of simulation experiments here is to investigate the accuracy of the approximation. A system with two traffic classes is considered: a real-time class and a non-real-time class. Simulation parameters including the arrival rate of new connections, holding time, minimum bandwidth guarantee, and the batch size of each class are specified in Table 4-2. Since handoff requests for both classes arrive in common batches, we assume both traffic classes have the same cell residence time and batch arrival rate. The number of handoff
requests of each class in the batch is assumed to be geometrically distributed with the expected value of $m_1=6$ for the real-time class and $m_2=4$ for the non-real-time class.

### Table 4-2 Parameters for the two traffic classes

<table>
<thead>
<tr>
<th></th>
<th>Real-time class</th>
<th>Non-real-time class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrival rate of new connections ($s^{-1}$)</td>
<td>$\alpha_1 = 9$</td>
<td>$\alpha_2 = 6$</td>
</tr>
<tr>
<td>Minimum bandwidth (Kbps)</td>
<td>$b_1 = 64$</td>
<td>$b_2 = 96$</td>
</tr>
<tr>
<td>Holding time(s)</td>
<td>$1/\mu_1 = 10$</td>
<td>$1/\mu_2 = 20$</td>
</tr>
<tr>
<td>Batch size</td>
<td>$m_1 = 6$</td>
<td>$m_2 = 4$</td>
</tr>
<tr>
<td>Cell residence time (s)</td>
<td>$1/\mu'' = 60$</td>
<td></td>
</tr>
<tr>
<td>Arrival rate of batches ($s^{-1}$)</td>
<td>$\alpha'' = 0.3$</td>
<td></td>
</tr>
</tbody>
</table>

The resulting link utilisation is shown in Figure 4-3. The dropping threshold is selected from $\{[0.01,0.09]:0.01\}$ (this notation means the dropping threshold starts from 0.01, then increases in value in steps of 0.01 until it reaches 0.09). In simulation, to approximate the optimal $G$ for each dropping threshold, $G_1$ and $G_2$ are selected from $\{[0.7,0.95]:0.05\}$ and the dropping probability and the reserved bandwidth are calculated for each combination of $G=(G_1,G_2)$. The optimal $G$ is that one that gives maximum reserved bandwidth while its dropping probability is under the dropping threshold. Table 4-3 gives the optimal $G$ calculated by the simulation and the proposed approximation method. It has to be noted though, that simulation results also represent approximate solution due to the estimation error within the increasing step (0.05 in this case). If we consider the values obtained by simulation are the actual values of the optimal guard channel, the comparison shows that most of the estimates obtained by the approximation method have the absolute error within 0.05 and the remaining part just slightly exceeds 0.05.

### Table 4-3 Comparison of results obtained by simulation and approximation method

<table>
<thead>
<tr>
<th>Dropping threshold</th>
<th>$G_1$</th>
<th>$G_2$</th>
<th>Simulation</th>
<th>$G_1$</th>
<th>$G_2$</th>
<th>Approx. method</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.09</td>
<td>0.9</td>
<td>0.95</td>
<td>0.909091</td>
<td>0.907473</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>0.9</td>
<td>0.95</td>
<td>0.903743</td>
<td>0.900356</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.07</td>
<td>0.9</td>
<td>0.95</td>
<td>0.898396</td>
<td>0.893238</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.06</td>
<td>0.95</td>
<td>0.9</td>
<td>0.893048</td>
<td>0.88968</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>0.95</td>
<td>0.9</td>
<td>0.882353</td>
<td>0.879004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>0.95</td>
<td>0.9</td>
<td>0.871658</td>
<td>0.868327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>0.9</td>
<td>0.9</td>
<td>0.869363</td>
<td>0.857651</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.02</td>
<td>0.9</td>
<td>0.9</td>
<td>0.839572</td>
<td>0.839858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.85</td>
<td>0.9</td>
<td>0.807487</td>
<td>0.807829</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.5 Summary and Conclusions

In this chapter guard channel connection admission control algorithm for the HAP networks has been investigated. A model was constructed for the general operating scenario taking into account the batch handoff requests. An approximation method was proposed for the simple calculation of the vector of optimal guard channels for multiple traffic classes. Simulation results suggest that the proposed approximation method is fairly accurate.

This work is a part of QoS framework for HAP network containing admission control at the connection level and packet scheduling [4.5] at the packet level.
4.6 References


5. Handoff Techniques and Station Keeping

5.1 Introduction

Despite the advantages that the HAP architecture has to offer [5.1], there are also several disadvantages that need to be investigated. One of these disadvantages is the relatively loose station keeping characteristics. Although the stratosphere is a layer of relatively mild turbulence [5.2], the platform will inevitably encounter sudden wind gusts. As a result, the platform could move in any direction. Various antenna steering mechanism techniques have been proposed that can be employed to practically compensate for the HAP movement. These mechanisms can be used on both Customer Premises Equipment (CPE) as well as on the platform itself. The antenna steering mechanisms [5.3] [5.4] on the platform have several constraints associated with them. For example, they must be powered up adding to the power demands on the platform, while they also add extra weight and cost to the payload. Furthermore, even if they are employed it is impossible to guarantee a stationary position for the HAP during its service hours and so the patterns of received power and interference on the ground (footprints) will still be moving.

Handoff is a common technique employed by all cellular systems both terrestrial [5.5] and satellite [5.6], which has been proven vital both for ensuring uninterrupted connections and increasing system capacity (using directed handoff). For this work we have focused on its first use: to maintain uninterrupted connections when one of the ends of the communication link is moving and as a result the user must be transferred from one cell to another (i.e. experience handoff). For terrestrial systems, it is common that the user moves with the Base Station (BS) being fixed. In this section we are assuming that the platform in a HAP system will be moving and the users will be fixed (the case of Broadband Fixed Wireless Access (BFWA)). Taking into account the movements of the HAP, it is possible by employing a combination of handoff techniques and a steering mechanism, to avoid interruptions on the link between the user and the platform. To do so, either the Customer Premises Equipment (CPE) should be able to keep track of the platform and / or the HAP itself should employ an antenna steering mechanism to maintain a constant coverage. As mentioned before, the latter may not be feasible due to the power available onboard and the payload weight constraints imposed by the platform manufacturer. Furthermore, it is important to eliminate unnecessary handoffs / signalling and therefore unnecessary power consumption that could be vital for the continuous operation of the HAP [5.7].

There are several handoff techniques proposed in the literature depending on the type of the channel allocation scheme used. These can be classified as fixed [5.8], flexible [5.9] and dynamic [5.10] channel assignment strategies, all initially designed for terrestrial systems. These strategies are used both for new connection and handoff requests. For the case of handoff in a terrestrial communication system, either the user or the base station can make a decision about whether a handoff is required. Both user and base station will monitor the quality of the channel.

For this work we have focused on a Fixed Channel Allocation (FCA) based scheme called Area Based FCA (ABFCA) [5.11] to exploit cell overlap, and have investigated the effect of the various movements of an aerial platform in order to quantify the effectiveness of this scheme. The reason for using ABFCA instead of the more advanced UFCA scheme is because of being simple to understand and simple to implement. The aim here was to quantify the impact of station keeping onto the system and not trying to unify the levels of blocking probability. From the results, the handoff performance is improved when employing a guard channel based scheme. In this scheme a number of channels is reserved explicitly for users requiring handoff [5.12].

This report starts with examining basic handoff techniques and shows the different forms of handoff techniques used in the terrestrial and Low Earth Orbiting (LEO) satellite communication
systems. Next, a number of HAP mobility models are presented to highlight the difference between HAP movements and the terrestrial user movements on one side and the LEO satellite movements on the other. Then the impact that HAP movements have on handoff is presented and it is shown how the handoff can be controlled using mechanical steering correction mechanism.

An alternative technique to the mechanical steering correction mechanism is then examined, where the handoff simulation model is used to demonstrate the effects of limited mechanical stabilisation for rotation only. Finally, simulation results and conclusions are presented at the end of this chapter.

5.2 Handoff Techniques

5.2.1 What is handoff?

Handoff is defined as the change of the radio channel used by a wireless terminal. The new channel can be either assigned from the same base station (intra-cell handoff) or from a different base station (inter-cell handoff). In the case where a handoff is unsuccessful and the user is forced to terminate the connection, this is called dropping. Handoff procedure can be divided into three phases [5.13]:

1. Handoff Decision / Detection

   A decision has to be made when exactly to initiate and perform a handoff. Also which cell to handoff to? This decision can be made by the user's equipment or by the base station(s).

2. Handover Resource Assignment

   This phase has to manage the channels in order to ensure that there will be enough channels to minimise dropping probability.

3. Handoff Execution

   This phase of a handoff procedure includes the protocols for reliable exchange of handover data. This is the signalling procedure needed to inform the handoff connection and base station about the new resource allocation.

This chapter focuses on the second phase of the handoff procedure. Initial investigations of handoff models designed for terrestrial or satellite purposes have been vital for the development of a HAP-based handoff scheme, hence an overview of handoff techniques is presented in this section.

5.2.2 Why handoff is necessary?

Handoff is a procedure employed by cellular systems for ensuring uninterrupted connections and increasing system capacity (Directed Handoff). In the first case, a handoff is usually initiated when the radio link drops below a minimum threshold level, and in the second case it occurs when the system is rearranging its channels due to congestion in one area. A typical example for the first case is when a user with an active connection moves across the cell boundary to an adjacent cell (Figure 5-1). In this case the connection has to be handed to the neighbouring cell in order to prevent the connection being dropped. However, if the new cell does not have adequate channels to support the handoff, then the connection is dropped.
For the second case where the system is rearranging its channels due to congestion, a number of handoffs take place from congested cells to less congested cells in order to maintain acceptable blocking levels or to further increase the system capacity in a given area [5.14]. An example of a handoff scheme associated with this concept is the Directed Handoff (DH) scheme [5.15], [5.16] and [5.17]. DH effectively redirects existing calls from one cell to a neighbouring cell to further increase the system capacity. DH requires cell overlap to operate as it redirects existing calls in the overlap region from one cell to a neighbouring cell.

5.2.3 Quantifying Handoff

Handoff performance can be measured in terms of how many handoffs take place in a region or coverage area and how many of these are successful or unsuccessful. To measure the effect of the users or platform movements and assess the performance of the handoff in terms of these two aspects, we consider the Handoff Probability \( P_H \) and Dropping Probability \( P_D \) separately.

5.2.3.1 Handoff Probability

In this work Handoff Probability is defined as a measure indicating how often users experience a successful handoff irrespectively whether the connection was eventually dropped.

\[
P_H = \frac{N_H}{N - N_B}
\]  

(5.1)

where \( N_H \) is the total number of users that perform handoff, \( N_B \) is the number of user that have been blocked and \( N \) is the total number of users. So a handoff is required if for example a user moves across the boundaries of a cell (in a terrestrial system) or in the case of a HAP communication system the HAP moves away from the centre of the coverage area causing the same effect on the user.

5.2.3.2 Dropping Probability

Dropping probability represents the levels of unsuccessful handoffs in a communication system. Here:

\[
P_D = \frac{N_D}{N - N_B}
\]  

(5.2)

where \( N_D \) is the number of the users being dropped and \( N_B \) is the number of users being blocked. \( N \) is the total number of users in the system.
5.2.4 Terrestrial Handoff Schemes

In order to minimise the chances of users being dropped during a handoff, a number of channels per cell can be dedicated just for the handoff. These Guard Channel (GC) based schemes [5.18] and [5.12] provide better performance in terms of dropping probability at the expense of a reduction in the total admitted traffic and an increase in the blocking probability of new calls. Therefore, there is the risk of inefficient spectrum utilisation.

Another well-known handoff scheme is the Handoff Queuing (HQ) scheme. The main characteristic of this scheme is that none of the new calls is allocated a channel before the handoff requests in the queue are served. More specifically, a handoff request is added to the queue (list of handoffs) when a user has moved to a point where the received power signal from a neighbouring cell is higher than the one the user is currently connected to (point a in Figure 5-2). When the user’s received power signal drops below the minimum received power threshold (point b in Figure 5-2) and no channel has been found, then the connection is terminated (i.e. dropped).

![Figure 5-2 Handoff requests are queued while users move from point (a) to point (b)](image)

Therefore, a connection can be queued for a maximum period of time that the user will spend in the region formed by the two boundaries point a to point b. Handoff users are served on a first come first served basis [5.19] or on a basis of their speed and position in the handoff region [5.20]. The HQ scheme can reduce the dropping probability at the expense of increased connection blocking probability and a decrease in the ratio of carried-to-admitted traffic [5.20].

Another type of handoff is the New Call Queuing (NCQ) scheme. This is based on the fact that dropping existing users is worse than blocking new users and therefore it is better to queue new calls rather than queue existing calls (like the HQ scheme mentioned before). Results from [5.18] showed that the dropping probability decreases at a greater rate than the new call queuing blocking probability while the NCQ scheme can support much higher traffic than the guard channel based schemes.

On a similar operational basis, the two-handoff-level algorithm [5.21] employs two handoff request levels as well as a minimum power received threshold level. The way this scheme works is that when the users’ received power levels drop below the first handoff level, a handoff request is initiated. At this stage a handoff will take place only if the signal from the new cell is greater than the first handoff level of the new cell (case 1 – Figure 5-3).
Figure 5-3 Two Level Handoff Scheme – Case 1 [5.21]

In the case where the user reaches the second handoff threshold level, the call will be handed off with no conditions assuming that the user is within the range of the second handoff level of the new cell (case 2 – Figure 5-4). In the case where there is a gap (no coverage, case 3, Figure 5-4), or no available channels in the new cell, the call will continue until its signal strength drops below the minimum threshold level. In this case the call will eventually be dropped.

Figure 5-4 Two Level Handoff Scheme – Case 2 and 3 [5.21]

The failure of a call to be handed off happens when the required signal quality is below a required value for more than a given time interval (e.g. 5secs - [5.21]). This time interval depends
explicitly on the type of traffic the system supports. If for example the system supports voice calls, then the time interval will be a real-time delay sensitive service that can tolerate only short interruptions. On the other hand, for a non-real-time data service such as internet-based application services, interruption for a longer period of time might be passed unnoticed since only a small number of packets could attempt to be transmitted during that period.

5.2.5 Satellite Handoff Schemes

Satellite handoff has been largely used in the Low Earth Orbit (LEO) satellites as they revolve around the globe in a fixed orbit. Each revolution takes between approximately 90 minutes to few hours, and a group of LEO satellites is employed such that there is always a satellite in a line-of-sight. So, fixed users on the ground might experience inter-cell handoff or intra-cell handoff meaning that a handoff will be required on a cell-to-cell basis or on a satellite-to-satellite basis.

Handoff schemes for LEO satellites were designed based on the constant speed repetitive revolutions that LEO satellites perform [5.6]. Handoff design has been simplified as LEO satellite movements are predictable. It is therefore possible to predict the number of cells serving one area at any time.

5.3 High Altitude Platform Mobility Models

Unlike satellites, characterised by predictable movement in predefined orbits, HAPs may move in any direction at a varying speed and so both the HAP and the users have to cope with these movements. We have looked at all 6 degrees of freedom that a flying object such as HAP (plane or airship) can be subjected to. The movements examined are horizontal displacement with respect to the x, y and z-axis as well as yaw, pitch and roll. In addition the effect of HAP movements has been investigated assuming that the payload is stabilised against rotation but not against drift. In practice the HAP can perform any (or a combination) of the six degrees of freedom as far as it remains within certain boundaries. According to ITU [5.22] a platform should be stationed within a circle of radius 400m with height variations of ±700m, while HeliNet project specifications [5.23] required a platform to be located within a cylinder with a height of 3km and a radius of 4km for 99.9% of the time, and within a cylinder with a height of 1km and a radius of 2.5km for 99% of the time (see Figure 5-5).

For this work we have assumed that the platform could either be an airplane or an airship that should be able to serve the nominal coverage area at all times. Therefore, the most appropriate model was the location cylinder model proposed in HeliNet. The radius and the height of the cylinder depicted in Figure 5-5 are defined based on the assumption that the footprint must completely cover the cell area occupied by the users at any time.

A HAP communication system will be designed to serve a number of cells on the ground. The number of cells will depend on the type of the terrain that is to be served. For rural areas it is expected that fewer cells of large size will be employed whereas for urban areas more beams of smaller size will be employed. The user antenna is expected to be of high gain (e.g. 39.4dBi [5.23]). This sort of antenna will be highly directional and therefore it will be of a particularly small beamwidth. As a result, any movement of the HAP will have a direct impact on the user. The impact will be more significant if the user employs a fixed antenna. This will be a much cheaper solution for the user at the expense of long outages due to the HAP movements. For instance, when the HAP experiences drift and the customer premises equipment (CPE) antenna is fixed, the antenna will be effectively pointing at the sky rather than at the platform. The highly directional CPE antenna becomes blind as it is unable to communicate with the HAP.

The simplest form of steering to be applied on the platform is an individual antenna correction mechanism that adjusts each antenna such that the boresight points directly to the position on the ground the cell centre should be pointing at. This mechanism however, requires a large number of complex and heavy mechanical gimbals equal to the number of cells required to cover
the nominal coverage area. It is therefore not practically feasible. Based on this principle, a four actuator steering concept [5.3] has been proposed. This is effectively a mechanical mechanism that has been designed to provide correction to a group of antennas independently from another correction that might need to be applied. The four-actuator mechanism is presented in more detail in section 5.4.

For this work we have assumed highly directional steerable antennas for the users as well as a stabilised payload against rotation but not against drift.

![Diagram of HAP position cylinder for 99% and 99.9% of time as defined by HeliNet [5.23]]

Although a steering correction mechanism can ensure longer service availability it also adds extra weight to the platform, and also requires electrical power to operate. This implies an additional cost.

The six degrees of freedom that describe all possible movements that an aerial platform may perform, can be defined in terms of vectors $x$, $y$, $z$, $x_a$, $y_a$, $z_a$ where the first three denote movement in a specific direction and the latter three denote rotations with the subscript denoting the axis of rotation. The six degrees of freedom can be grouped as drift and rotational-based movements. The drift-based movements are the $x$, $y$ and $z$-axis drift movements that a platform might perform. The rotational-based movements are the $x$, $y$ and $z$-axis rotational movements; these are also known as pitch, roll and yaw respectively. Simulation of all six degrees of freedom has not been necessary since some of the movements were considered to have the same effect on the system. These are the $x$ and $y$-axis drift as well as the $x$ and $y$-axis rotation (roll and pitch). Thus, we have examined the effect of the drift and rotation based on the $x$ and $z$-axis only. In addition two more complex types of movements called random walk and reflection have been investigated. These two types contain both drift and rotation movements that an aerial platform might perform.

### 5.3.1 Drift with respect to the $x$, $y$ and $z$ axis

$x$ or $y$-axis drift of the platform have similar effect on the ground. This is because the movement is similar and because $x$ and $y$-axis are lying on the same plane. Results can therefore represent both cases. Here we have investigated the effect of HAP movements assuming that the payload is stabilised against rotation but not against drift. The HAP is allowed to drift up to a distance of 2.35km away from the centre of the coverage area. This is based on the cellular structure of 19 cells used for this work and it is to ensure that all users are located within the coverage area.
It is also assumed that the HAP will start moving from the centre of the coverage area to one end of the position cylinder and then it will move back to the other end. The HAP does not change its orientation during this movement, to keep each type of movement isolated. We also assume that the current position vector $\mathbf{r}_t$ is dependent on the vector $(\mathbf{r}_{t-\Delta t})$ generated $\Delta t$ seconds earlier.

Mathematically this can be represented as [5.24]:

$$
\mathbf{r}_t = \mathbf{r}_{t-\Delta t} + v \cdot \Delta t \cdot \hat{a}_t
$$

where $v$ is the HAP velocity and $\hat{a}_t = \frac{a_t - a_{t-1}}{|a_t - a_{t-1}|}$. $a_t$ represents the vector corresponding to a destination point $A_F$ and $a_{t-1}$ corresponds to the vector of the point of the previous point update. $A_F$ in this example represents the edge of the cylinder that the HAP will move towards. An example of this is shown in Figure 5-7.

![Figure 5-6 Example of y-axis drift](image)

![Figure 5-7 Graphical representation of equation (4.3)](image)
For this mobility model a starting point (0, 0, 17)km is assumed. Then the HAP is set to move to point (0, 2.35km, 17km) and then backwards to (0, -2.35km, 17km). There were no changes in height.

z-axis drift (see Figure 5-8) causes expansion and contraction of the size of every cell as well as the whole coverage area. Assuming that there is no steering mechanism on the HAP (apart from the rotation stabilisation mechanism mentioned before) and depending on the chosen cellular structure, the HAP is allowed to drift up to a distance of 1500m [5.23] away from its initial point to ensure that all users remain located within the coverage area. In both cases where the HAP moves upwards or downwards, the footprints must be able to serve the nominal coverage area.

More specifically, the footprint must cover the cell area where this is the area on the ground where the loss to one particular HAP antenna is minimum when the HAP is at the centre of its coverage and it is not moving (i.e. initial position). In this case, the size of the footprint will change whilst the HAP moves upwards or downwards. However, the effect depends on the antenna mask employed. In the case where the HAP moves downwards (point A – Figure 5-9), the footprint could get smaller provided that the antenna profile looked like the one in Figure 5-9 B. On the other hand, if the antenna profile used was the one in Figure 5-9 A, then the effective size of the footprint might become bigger since the platform is closer to the ground. Although users in this case might be at a much greater angle from the boresight (e.g. user A example in Figure 5-10), the $R^2$ is smaller and this might put the user within the footprint.
However, if Figure 5-9 B profile is used, then the footprint will shrink and the user A illustrated in Figure 5-10 will not be within range.

**Figure 5-9 Antenna mask example**

**Figure 5-10 Drift on the z-axis scenario**
On the other hand, when the HAP moves upwards things are different. In this case, when employing the antenna profile B of Figure 5-9, the footprint range will increase (see Figure 5-10). Thus the users that were located outside the footprint when the HAP was at its initial state will now have a smaller angle with respect to the boresight. As a result it is possible that they could actually connect to this antenna. If however Figure 5-9 A profile is used, then when the HAP moves upwards, the power density on the ground is expected to be lower (as the antenna is less directional) and as a result the footprint range will be reduced. It is therefore difficult to describe the effect using an equation since it is essential to know the antenna mask used.

### 5.3.2 Rotation with respect to the x, y and z axis

Pitch and Roll have similar effect on the ground. This is again because the rotation is performed on two axes both lying on the same plane. In both cases the HAP rotates ±\( \theta \) degrees such that there are no users left without coverage. For this work we consider that \( \theta \) varies between:

\[
- \arctan\left(\frac{R_{\text{print}} - R_{\text{cell}}}{h}\right) \leq x, y \leq \arctan\left(\frac{R_{\text{print}} - R_{\text{cell}}}{h}\right)
\]

\[
- \arctan\left(\frac{R_{\text{print}} - R_{\text{cell}}}{h}\right) \leq y, z \leq \arctan\left(\frac{R_{\text{print}} - R_{\text{cell}}}{h}\right)
\]

![Figure 5-11 Example of y-axis rotation (pitch)](image)

For the case of the rotation with respect to the z-axis i.e. yaw, we consider that the HAP rotates at a constant speed anticlockwise (viewing it from top-view) as shown in the figure below.
Planes such as Unmanned Aerial Vehicles (UAVs) are expected to be in a constant rotational type of orbit with respect to centre of the coverage area in order to maintain service over the coverage area.

### 5.3.3 Random Walk

To examine a more realistic type of movement, drift and yaw have been combined together to give what we call random walk. In this case the HAP moves at a constant speed in a random direction. The HAP must employ some sort of propulsion mechanism to keep it within the cylindrical boundaries as defined in HeliNet [5.25].

Figure 5-13 illustrates an example of the route that a HAP takes when assuming random walk and the movement can be described in terms of horizontal and vertical components. The direction changes after a fixed time period \( t \) during which the HAP also experiences drift. The greater the speed \( v_h \) of the platform the longer the distance it will travel within the time period \( t \). After the HAP starts from position A, it is assumed that the HAP continues on the same direction (vector) over the update interval \( \Delta t \), with velocity \( v_h \). The new horizontal component of position is therefore anywhere on a circle of radius \( r \) based on the HeliNet cylinder [5.23] at a randomly distributed angle \( \theta \). The decent/ascent rate is assumed to be uniformly distributed between \( +/-v_v \). Therefore the new locus of the new position \( A' \) is an ellipsoid centred on point A.
Resource Allocation and Handoff Techniques for High Altitude Platforms

(assuming it does not hit the edge of the position envelope). Mathematically this can be represented as [5.24]:

\[ \mathbf{r}_t = \mathbf{r}_{t-\Delta t} + \Delta \left[ v_h \cos(\theta) \mathbf{i} + v_h \sin(\theta) \mathbf{j} + v_i \mathbf{k} \right] \]  

(5.6)

5.3.4 Reflection

For the reflection type of movement the HAP moves from its current location to a new location at a predefined speed. The direction and the distance are randomly selected. The HAP position is maintained within the cylindrical boundaries defined in HeliNet [5.25].

![Figure 5-14 Example of reflection HAP movement](image)

Based on the same terminology used in the previous example also illustrated in Figure 5-7, a point \( A_f \) is randomly selected anywhere within the position envelope according to a random uniform distribution. This represents the destination point the HAP is heading towards at a constant speed \( v \). The intermediate positions are calculated every time interval \( \Delta t \) based on the speed of the platform. When the HAP reaches its final destination \( A_f \), the process is repeated [5.24].

\[ \mathbf{r}_t = \mathbf{r}_{t-\Delta t} + v \Delta t \mathbf{a}^t \]  

(5.7)

where \( \mathbf{a}^t = \frac{\mathbf{a}_t - \mathbf{a}_{t-1}}{\| \mathbf{a}_t - \mathbf{a}_{t-1} \|} \), with \( \mathbf{a}_t \) representing the vector corresponding to point \( A_f \) and \( \mathbf{a}_{t-1} \) corresponding to the vector of the point of the previous point update.

5.4 Steerable Antennas and Handoff

Steerable antennas can be used to cope with the movements of the HAP. They ensure constant coverage at the expense of increased weight of the payload if mechanically steerable antennas are employed. Operating at 28/31 GHz and 48GHz bands means that for the time being electrically steerable phase array antennas are not practical. Thus corrugated horn antennas [5.26] with exceptionally low sidelobes are assumed.

HAP movements have been addressed in the past such as in [5.3], [5.27] and in [5.4] where various techniques have been proposed to cope with various movements. In [5.4] a propulsion
mechanism is proposed in order to counterbalance the horizontal displacement with the ideal position of the HAP and the relevant correction required being specified using a global positioning system (GPS). In [5.27] it has been proposed that for the inclination effect, a gimbaling mechanism can be used at the bottom of the platform. This will limit the movement of the antennas with respect to the ground.

In [5.3] a steerable antenna correction mechanism was proposed, which needs to be applied on every antenna individually. However, this would require a complex mechanical system with a large number of motors and therefore it would add significant weight to the payload. It was therefore preferable that a HAP system would employ some sort of mechanically steerable mechanism but for a group of antennas instead. This mechanism is presented in more detail in this section.

### 5.4.1 Aperture Antenna Steering Solutions

In previous work [5.28], [5.26] it has been shown that HAP movements can change position or distort the shape of the individual cells. In the case where the HAP drifts from the centre of the coverage area, cells move from their intended position. Also, when the HAP experiences roll or pitch a part of the coverage area might be left with no service. Furthermore, the individual cell shape is distorted in terms of the received power as the angle of incidence on the ground of the main beam changes due to the roll or pitch. The power levels form an approximately elliptical footprint instead of the indented approximately circular one. Assuming a steerable antenna grid mechanism, the HAP antennas employed can be redirected (steered back as a group and not individually) to point towards the centre of the coverage area [5.3].

![Steerable antennas can be used to maintain constant coverage area.](image)

What will happen in this case is that the HAP antennas will point to a different location (except the centre cell) than the one they did when the HAP was located at the centre of the coverage area. In addition, all antennas will have a different elevation angle. This means that the power distribution of the ground will not be the same (compared with the initial case where the sub platform point (SPP) was at the centre of the coverage area). Nevertheless the HAP will still be servicing the intended coverage area thus correcting the drift motion. The example shown in Figure 5-15 depicts the HAP when drifting along the x and y-axis. However, HAP can also drift on the z-axis causing the footprint area to expand or shrink depending whether it goes upwards or downwards.
Figure 5-16 Steerable antennas can be utilised to correct drift movements with respect to the z-axis.

Figure 5-16 depicts a mechanism [5.3] that could potentially be used in order to cope with drift on the z-axis. The way this mechanism works is that all rings are interconnected so that when the HAP moves upwards or downwards, the cells of each ring will change their pointing angle symmetrically (except the centre cell which will keep pointing at the centre).

In the case of roll or pitch the HAP’s sub platform point (SPP) remains the same but the antennas are pointing in a different direction. The antennas can be steered back to their original position by employing the mechanism described below for the case of drift in the x or y-axis.
Figure 5-18 Steerable antennas can be utilised to correct pitch and roll movements and therefore maintain constant coverage area.

In this case, the antennas will still be pointing towards their intended positions and therefore the elevation angle with respect to their boresight will not be affected.

The steering mechanism must be intelligent enough to cope with a combination of these movements. Steerable antennas can be employed to guarantee constant coverage at all times assuming that the HAP is located within certain boundaries.

5.5 Handoff Simulation Model

To determine whether the additional expense of carrying a mechanically steerable antenna correction mechanism is necessary, it is important to investigate whether the platform could possibly operate without any major correction mechanism (by just being stabilised against rotation but not against drift). Therefore, it is necessary to examine the effect of the movements of the platform first and see whether it is possible by employing other techniques such as handoff to overcome the problem of station keeping. Since a HAP communication system is a centralised system it could be practically possible to perform all these necessary computations for the handoffs on a central ground station (backhaul) which then feeds all the data to the HAP which then bounces off to all the users. This is of course assuming that we have a ground station that connects the HAP with the global network.

Some cell overlap is inevitable because of the way the power decreases away from the boresight of the antenna, which has also been exploited to improve capacity as shown in section 3.7. This is determined from the link budget and is related to the power rolloff with angle from the antenna gain profile.

For this work we have assumed 19 cells of 3.15km radius each. Users have been uniformly randomly distributed within a circle of radius 9km (i.e. mainly within the central cluster of 7 cells) and the HAP was allowed to move within limits such that the users would not be left without service. Figure 5-19 illustrates how users are being distributed and how the footprints are positioned on the ground.
Figure 5-19 Users are generated within a circle of 9km and served by 19 cells of 3.15km radius each.

The following table lists the general system parameters for this work.

<table>
<thead>
<tr>
<th>No.</th>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Noise Received Level ($N_{RX}$)</td>
<td>-133.9</td>
<td>dBm</td>
</tr>
<tr>
<td>2</td>
<td>Frequency ($f$)</td>
<td>28</td>
<td>GHz</td>
</tr>
<tr>
<td>3</td>
<td>Reference Gain ($G_{Ref}$)</td>
<td>-90</td>
<td>dB</td>
</tr>
<tr>
<td>4</td>
<td>Platform Altitude ($h$)</td>
<td>17</td>
<td>km</td>
</tr>
<tr>
<td>5</td>
<td>Transmitter Power ($P_{TX}$)</td>
<td>-26.6</td>
<td>dBm</td>
</tr>
<tr>
<td>6</td>
<td>Number of Cells ($c$)</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Channels per Cell</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Total Offered Traffic (OT)</td>
<td>742.34</td>
<td>Erlang</td>
</tr>
<tr>
<td>9</td>
<td>Cluster Size ($K$)</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>HAP Altitude ($h$)</td>
<td>17</td>
<td>km</td>
</tr>
<tr>
<td>11</td>
<td>Cell Radius ($R$)</td>
<td>3.15</td>
<td>km</td>
</tr>
<tr>
<td>12</td>
<td>RX Antenna Gain ($G_{RX}$)</td>
<td>31.3</td>
<td>dBi</td>
</tr>
<tr>
<td>13</td>
<td>HAP Antenna Gain ($G_{TX}$)</td>
<td>39</td>
<td>dBi</td>
</tr>
</tbody>
</table>
The simulation performed was based on a Monte Carlo Simulation model. After every small time interval, the simulator checked whether the position of the HAP had changed and then initiated handoff if required. Figure 5-20 illustrates a general state flow diagram of the handoff algorithm performed.

*Figure 5-20 Immediate Handoff Initiation Algorithm*

If the position of the platform has changed, then all users that have been affected must first be identified and be removed temporarily from the system. The users must then be added back into the system and connected to the new cell. This is to eliminate the case where users are being dropped from a cell that is waiting for some of its current users to be connected to another cell. In this case the cell will have a number of channels available as soon as its handoff users release the channels they occupy. The point is to ensure that these channels are available for the new handoff users coming to the cell. The affected users are consisting a small proportion of the total number of users within a cell. Since the capacity is all located on a case by case basis, the overhead will not be significantly high. This simulation was performed when allowing users to connect to the closest base station (case 1 – No Overlap) or when allowing users to connect to any base station within range (case 2 – With Overlap).

In the case 2 where cell overlap was allowed, the users were potentially able to connect to a number of BSs depending on their received power and the minimum received power threshold that defined the boundaries of a cell. Furthermore, the users would connect to the BS that had the most available channels. Details on how to calculate the minimum received power threshold
can be found in [5.29]. Any user whose received power level $P_{RX}$ drops below the minimum received power threshold $P_{TH}$ requires an immediate handoff.

Figure 5-21 illustrates the state flow diagram for identifying and removing the users from the system. Users that have been removed will then have their state changed to “try again”. The code goes through the “trying again” users in numerical order to attempt to connect them to a new cell. The handoff will be successful or unsuccessful, depending on whether the new cell has any channels available. In the case where cells overlap with each other, the users will connect to the cell with the most available channels.

![Figure 5-21 Identify and Remove Handoff Users](image)

Figure 5-22 shows that all users lined up for a handoff must first choose which cell to connect to. As mentioned in [5.11] and [5.29] the BS is selected based on the number of channels available. However this is not restrictive and it could be that the users could connect to the cell that offers the best received power or even CNIR. Nevertheless choosing the cell with the most available channels improves fairness within the system [5.29].
In the case of no overlap, users always connect to the closest virtual base station. In this case the immediate handoff initiation algorithm illustrated in Figure 5-22 is repeated, with the difference that the users are not identified based on their received power but on their distance from the virtual base station. This is based on the assumption that users will not be connecting to the virtual base station with the lowest loss but the ones that are physically closest. Thus, to identify (and remove) the handoff users we use their distance from all virtual base stations (after the HAP has moved). Also, when adding the user back to the system, the cell is chosen based on the distance of the user from the centre of the cell.

In some of the simulations, in order to improve the dropping probability a number of channels have been reserved from being allocated to the new users. Thus, these channels are dedicated to the handoffs. More specifically 10% of the total number of channels have been reserved for handoff in order to provide sufficient flexibility for handling handoff without causing significant blocking.

### 5.6 Handoff Simulation

The simulation performed was based on the mobility models presented previously. To give a reasonable simulation time, a set of 19 circular cells was used. A cluster size \( K \) of 7 was employed and each cell had a group of 100 channels within its coverage area. We have also assumed that there is a direct line-of-sight communication between the user and the HAP. The traffic was set to be uniformly randomly distributed within the coverage area of radius \( R_{\text{user}} \). The total Offered Traffic (OT) in the 19-cell coverage area was 742 Erlangs and the mean connection duration was set to 5 minutes. When cell overlap was permitted it was set at 25%, thus the new cell radius would be 25% bigger than before. All simulation parameters are listed in Table 5-1. Statistics are gathered from the centre cell except the case where the handoff and dropping probability are presented with respect to the distance from the centre of the coverage area.
5.6.1 Drift Movement Results

For the x-axis drift movement the platform was set to move on a straight line on the x-axis from one side of the position cylinder to the other starting at the centre. The speed varied from 0 to 200 km/h.

Figure 5-23 illustrates that the handoff probability increases in all three cases, i.e. no overlap, overlap and no overlap with guard channels. We can also see that the handoff probability for the overlap case is slightly higher than the other two. This is because the increased perimeter of the cell boundaries affects more users while the HAP is moving.

Dropping occurs primarily because of the lack of channel availability (in a HAP scenario). Nevertheless, random peaks in the spatial location of active users could pose a more significant problem in a HAP communication system. This occurs when users were initially served across two cells are now served by only in one cell (Figure 5-24). It is therefore difficult to guarantee zero dropping probability as the random distribution will give rise to local user hotspots which when connections are admitted may fall across into two footprints (Figure 5-24 - Case 1). However, as a result of the platform movements after some time there may only be one single footprint (Figure 5-24 - Case 2) covering the cell (user hotspot area), which alone may have insufficient capacity to support all these users resulting in dropping.
Case 1: Two footprints serve one cell

<table>
<thead>
<tr>
<th>Footprint 1</th>
<th>Footprint 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAP MOVE</td>
<td>Cell (users hotspot)</td>
</tr>
</tbody>
</table>

Case 2: One footprint serve one cell

<table>
<thead>
<tr>
<th>Footprint 1</th>
<th>Footprint 2</th>
<th>Footprint 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell (users hotspot)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5-24 Moving footprints cause connection dropping

For the case of drift movement, the dropping probability has been significantly reduced when applying a 10% Guard Channel mechanism. Furthermore, the overlap case has a significantly lower dropping probability than the no overlap case. It is also noticeable that the dropping probability for the case with no overlap, when HAP drifts at high speeds, has decreased. This is because at 200km/h users are subjected to successive handoffs which within that time the spatial location of users does not change. For high speed the traversing across the cells happens more frequently in comparison with the activation or clear down of users. So, there might be a situation where users are actually undergoing multiple handoffs. If they can successfully undergo one handoff it means that the spatial density of the users is fine at that moment in time. In other words, after the first handoff situation the user density of active users is sufficiently low for a single cell to deal with all subsequent handoff situations. The dropping in this situation is lower in comparison with the total number of users supported.

Figure 5-25 illustrates the boundaries of the rings of footprints that were used to quantify the level of handoff with respect to the distance from the centre of the coverage area when cell
overlap was not considered. Figure 5-26 illustrates the boundaries for the case where overlap is allowed and is set to be at 25% (i.e. Overlap radius $R$ is 25% larger than the external cell radius $R_{e}$). Some of the results were based on these two diagrams. Since the users are fixed and the HAP is moving, the footprints will be moving along the x-axis (or y axis).

![Diagram](image)

**Figure 5-26 Cell boundaries with overlap**

Figure 5-27 illustrates the handoff and dropping probability with respect to the distance from the centre of the coverage area at a constant speed of 100km/h. More specifically, the handoff and dropping probability are calculated based on a number of concentric rings with thickness of 500m. Results indicate that the handoff increases as we move away from the centre of the coverage area. In addition it can be seen (for the No Overlap case) that users that are closer to the boundaries of the cells (see point A and B in Figure 5-25) experience a higher dropping probability because they are more likely to require handoff.
However, in the case where we employ cell overlap, the dropping probability has decreased significantly. Increasing the cell boundaries by 25% and allowing users to connect to any of the virtual base stations within range we have effectively increased the trunking efficiency [5.11] and the ability of the system to cope with a higher spatial density of users. Since handoff users are the users that are at the edge of a footprint, this means that these users are most certainly located within an area of overlap. Thus they can benefit from the greater channel choice. As a result, handoff users have more chances of getting a channel than before (with no overlap) and therefore less chances of being dropped. The 10% guard channel mechanism has performed a lower dropping probability but this is at the expense of much higher blocking probability (see Figure 5-28). Here it is important to say that the period of oscillation at 100km/h, i.e. the time required for the platform to move from one site of the defined boundaries (see Figure 5-6) to the other, has been approximately 10 minutes whereas the mean connection duration assumed was 5 minutes. As a result users will be subjected to handoff depending on their location and the time their connection has started.

Figure 5-27 Handoff and dropping probability with respect to the distance from centre of coverage area

Figure 5-28 Blocking probability
Examining the blocking levels presented in Figure 5-28, it can be clearly seen that the blocking probability is significantly higher in the case of the 10% guard channel scheme. The blocking has been tripled in comparison to the no overlap case. Furthermore, the blocking probability for the non-overlap case has decreased as the speed increased. This is due to the higher dropping levels experienced at higher speeds. There is therefore a better channel availability for the new incoming users. The overlap case performs better than the other two in terms of the blocking probability. More specifically the blocking probability has decreased from 6.5% to 2% compared with the non-overlapping case.

5.6.2 Pitch Movements

For the pitch movement the platform was set to perform a uniform angular variation with constant speed. The HAP is assumed to be positioned above the centre of the coverage area so that its sub platform point remains constant. The angular velocity was varied so that the speed experienced at the centre of the coverage area varied from 0 to 200 km/h. From Figure 5-8 it was assumed that the centre of the coverage could only be displaced by $d$. Thus the maximum possible value for $x_\gamma$ can be expressed from:

$$ x_{\gamma_{\max}} = \arctan \left( \frac{d}{h} \right) $$ (5.8)

The motion is set to start from idle, and then move to one side ($x_{\gamma_{\max}}$) around to the $y$-axis and then move back on the other side (again to an angle $x_{\gamma_{\max}}$). $x_{\gamma_{\max}}$ was set to be equal to 0.16 radians assuming that $d$ was 2.73km, in order for the footprint to keep providing coverage to the cell.

Figure 5-29 Pitch HAP movement in three stages

Figure 5-29 illustrates the HAP movement as has been described before. Notice that the cell area is always positioned within the coverage area of the footprint.
Figure 5-30 illustrates that the handoff probability increases with or without overlap. We can see that the handoff probability for the no overlap case is higher than the case with overlap. Also, the dropping probability for the overlap case is significantly lower than the case with no overlap. This is because users requiring handoff are located in an overlap area near the edge of the cell they are connected to. Thus they benefit from the high trunking efficiency provided in the region due to the overlap.

Figure 5-31 illustrates the handoff and dropping probability with respect to the distance from the centre of the coverage area (see Figure 5-25) at a constant speed of 100km/h. The handoff probability is found to be higher for the case with overlap. Looking at the overlap case, a peak at about 5.3km indicates that the users located in the first ring experience high handoff probability. This is because the oscillating type of movement changes the angle of incident of the footprints on the ground. As a result their shape is not circular anymore but approximately elliptical.
The blocking probability results have shown that the case with no overlap suffers from higher blocking levels. Also, the phenomenon of reduced blocking probability (see Figure 5-32) at higher speeds (no overlap case) is as a result of the increased dropping probability, which therefore allows more new users to join the system.

5.6.3 Reflection, Rotation and Random Walk Movement Results

Reflection, Rotation and Random Walk have been examined for handoff and dropping probability performance. In the case of the Rotation mobility model, the rotational movement has been quantified in terms of the two rings of cells that form the coverage area (Figure 5-25). Based on Figure 5-5, all HAP movements allowed (applicable in the cases of the Reflection and Random Walk mobility models) were up to 2.36km away from the centre of the coverage area. That is to ensure uninterrupted coverage for all users.

Figure 5-33 illustrates the handoff and dropping probability for all three mobility models in terms of speed (km/h). For the Reflection model, the handoff probability increases at a higher rate when HAP moves at low speeds and then it starts to saturate reaching a handoff probability of 70% at 200km/h. For the case of the random walk, the rate of the increase in handoff probability
Resource Allocation and Handoff Techniques for High Altitude Platforms

is approximately constant at 4% per 10km/h up to 100km/h and then drops to approximately 2% per 10km/h reaching a maximum handoff probability rate of 61% at 200km/h.

For the case of the rotation mobility model, the rotational velocity was set with respect to the centre of one of the cells in the second ring (this is the external ring in a 19-cell model). Thus the speed of the cells was proportional to the distance from the centre of the coverage area and the rotational velocity set. From the results illustrated in Figure 5-33 it can be seen that in both rings of cells the users experienced the same level of handoff probability irrespectively of their position. The rate of change of handoff probability with respect to the HAP rotational velocity is 1.8% per 10km/h.

For the case of the dropping probability it can be noted (Figure 5-33) that the results behave similarly to the handoff probability results. For the Reflection mobility model, the dropping probability increases at a higher rate when having lower speeds reaching the level of 1.5% dropping probability at maximum speed of 200km/h. The Random walk increases in a more linear manner reaching a point of 1.8% dropping probability. The results for the Rotation model show that some of the users in the first ring will be dropped whereas none of the users in the second ring will be dropped at all. This is because the number of users located in the second ring is much smaller than the users located in the first ring since the coverage area is set to be bigger than the area the users are located in.

The effect on handoff and dropping probability has also been examined as a function of the distance from the centre of the coverage area. Figure 5-34 illustrates the behaviour of all three mobility models in terms of this distance. For the case of the reflection mobility model, the further the users are located from the centre of the coverage the higher the handoff probability they will experience. This is because the HAP is set to drift from one site to the other of the HeliNet HAP location cylinder within a time interval during which users are greatly affected.

It can also be said that the handoff probability decreases slightly at the centre of the first ring. This is because some of the users are centrally located within a cell (of the first ring). Therefore, they have less chances of being handed off since they are away from the boundaries of the cells they are connected to. The handoff probability decreases as we move to the second ring to become zero at 9km, which is the distance of the most distant user. Random Walk mobility...
model results had a similar behaviour to the reflection mobility model. However the results showed lower handoff probability levels except at the edge of the user positional area (most distant users at around 9km form the centre of the coverage area). The rotation mobility model showed an even smaller handoff probability. Again, the handoff probability increases the more we move away from the centre of the coverage area.

![Handoff and dropping probability with respect to the distance from the centre of the coverage area](image)

For the dropping probability the reflection mobility model is higher than the rest. However for both reflection and random walk mobility models the dropping probability increases the more we move towards the edges of the cells. For example the edges of the cells of the centre ring in Figure 5-34. The reason the users located towards the outside of the coverage area experience lower dropping than the users located towards the centre of the coverage area is because some of the time these users are connected to the cells of the second ring that are less congested. The Rotational movement however shows a constant increase of the dropping probability as we move away from the centre of the coverage area up to the point of the most distant user.

5.7 Summary and Conclusions

In this section, an investigation of the impact of the aerial platform movements in a HAP telecommunication system has been carried out. Results have shown that a handoff technique reduces the need for mechanical stabilisation. It has also been shown that handoff mechanism is necessary to ensure continuity of the connections being affected by the platform movements. Furthermore, by employing guard channels we have ensured significantly lower dropping levels but at the expense of higher blocking levels. More specifically the dropping probability was reduced from 0.2% to almost 0% when compared with the no overlap case. The immediate handoff scheme based on ABFCA cell overlap model has shown that by allowing overlap, the blocking probability has decreased from 6.5% to 2% when compared with the 10% Guard Channel case in the x-axis drift case scenario. It was also shown that the dropping probability has been significantly reduced (when compared with no overlap case) since the users that required handoff were located in the areas of overlap. Thus, they experienced higher trunking efficiency and as a result the handoff was successful.
Additionally three complex types of motion have been investigated. These are the rotation, reflection and random walk type of movement. This effect is worse for rotation type motion, with the probability of dropping dependent on the cell dwell time. Ways to improve performance is by either reserving channels for future handoff use or controlling the density of admitted users within one area [5.24]. Hence, the limitation on performance of the HAP system will therefore depend primarily on the resource allocation process, and availability of resources on the platform.

Employing cell overlap has once more shown that the quality of service was significantly improved since both blocking and dropping probability were significantly reduced. From this it could possibly be concluded that cell overlap in a combination with UFCA-II scheme can provide a uniform low blocking and low dropping performance unlike the other schemes examined.
5.8 References


6. On-board Buffering Techniques

6.1 Introduction

This chapter investigates the potential of on-board buffering as an outage mitigation strategy in broadband service delivery to trains. Tunnels represent a major inhibitor to communications service provision in the mm-wave bands, generating periods of outage and connection loss. The problem associated with loss of Line of Sight (LOS) links on entry to tunnels is investigated, and a range of on-board buffering techniques is presented to reduce the outage probability. The fundamental principle is to download information to the train at a faster rate than it is required by the users, with excess information stored in one or more buffers on-board the train. When the user terminal loses its connection to the network, information is read from the buffer until network access is regained. Alternative strategies are considered to investigate the feasibility with which outages can be mitigated or perhaps eliminated, for non-real time service provision.

The first part of this chapter explores the basic principle of on-board buffering and identifies its potential as an outage mitigation technique. A statistical distributional fit of UK tunnel lengths is presented, and a single tunnel scenario is analysed to determine the typical buffer size requirement. A method for determining the overall download rate and buffer size needed to meet a specified Quality of Service (QoS) level in terms of outage probability is then described. The second part of the chapter presents a multiple tunnel scenario, which represents a worst-case scenario. A range of novel buffering techniques are applied to mitigate against outages caused by tunnels, differing in the strategy they employ for allocating resources to the different users.

6.2 Fundamental On-Board Buffering Techniques

Satellites are currently able to provide broadband services to trains, but High Altitude Platform (HAP) systems have the potential to offer significantly higher data rate services with improved Quality of Service (QoS). Broadband service delivery to high-speed trains has been identified as a major application of HAPs, and is being considered extensively in the CAPANINA project. This chapter introduces a promising approach to eliminate network connection loss due to tunnels, enabling non-real time services to be effectively provided to rail commuters.

6.2.1 Single Tunnel Scenario

The scenario of interest is shown in Figure 6-1. It is envisaged that HAPs will be able to deliver very high speed data rates to trains potentially traveling up to 300 km/h. Users inside the vehicle will use a Wireless Local Area Network (WLAN), communicating with an on-board WLAN base station connected to the broadband service delivery platform.
In the single tunnel scenario, we assume that only one tunnel appears during the whole journey, and we consider a single user running a streaming video application at 6kbps (Mpeg4) or 320kbps (Advanced Audio Coding – AAC). The user has accessed the service on board, and will watch the streaming video throughout the whole journey. A simple buffering technique is considered to mitigate against an outage being caused by the tunnel. When the train is in the open air, the HAP can offer normal broadband services. The services subsequently continue when the train enters the tunnel, by reading information stored in the buffer. After the train has left the tunnel, access to the HAP will be regained, and the buffer will start to refill. This technique should enable the user to get uninterrupted services (assuming that the buffer contains sufficient information). This buffer technique requires that the download rate to the train is greater than application rate in order that additional data can be stored in the buffer.

### 6.2.2 Statistical Distribution for UK Tunnel Lengths

The tunnel length is one of the decisive parameters influencing the single tunnel scenario. A database has recently been published which gives information about tunnel lengths in the UK. These tunnel lengths have been converted from yards into meters ($L_t$). A statistical distributional fit to UK tunnel lengths has been developed, since it permits realistic simulations to be carried out without access to the data stored in the database. A Cumulative Distribution Function (CDF) for actual tunnel lengths has been plotted in MATLAB, and best fits have been generated, including exponential, gamma, weibull and log normal, as shown Figure 6-2. From this figure, we can see that the lognormal distribution is the closest match to the actual tunnel length distribution, and consequently, we can employ the lognormal fit to model tunnel lengths in the UK.

The primary aim of this study is to evaluate the probability of outage under different circumstances, such as variable distance from the tunnel entrance at the start of the application, or as a function of application data rate. We define four parameters in the scenario: the tunnel length ($L_t$); the train speed ($V$); the amount of information stored in the buffer on entry to the tunnel ($S_b$); and the application data rate ($R_a$). During a typical journey, if the time that the train takes to pass through the tunnel is longer than the time taken for the user to use all the data stored in the buffer, an outage will occur. The probability of outage when the train goes into the tunnel can be described by the following equation:

$$ P(\text{outage}) = P\left( \frac{L_t}{V} > \frac{S_b}{R_a} \right) $$

(6.1)
In this scenario, we assume that the buffer can be continually filled until the train reaches the tunnel, the train-speed is 30m/s (108km/h), and the application rates considered are 64 kbps (Mpeg4) and 320 kbps (Advanced Audio Coding) respectively. The amount of information stored in the buffer on entry to the tunnel varies from zero to 12Mbits, and both the true tunnel length data for the UK and the lognormal fit are employed. Equation (6.1) is used to derive the probability of outage as shown in Figure 6-3.

Figure 6-3 Probability of outage for different applications, as a function of the amount of information stored in the buffer on entry to the tunnel

It can be seen that the probability of outage decreases as the amount of information stored in the buffer is increased, with the probability of outage dependent on the application rate. If the QoS requirement is 8% outage probability, then for an application rate of 64kbps, the buffer size should be at least 2.26 Mbits. For an application rate of 320 kbps, the buffer size should be at least 11.3 Mbits, which is easy to realise given the capacity of typical storage facilities. However, a larger amount of information would need to be stored to meet a tighter QoS constraint.
6.2.3 System Performance

Another scenario is shown in Figure 6-4, which illustrates the distance ($X$) between the train and a tunnel, when the user starts a streaming application.

![Figure 6-4 The model considering the distance from tunnel](image)

In this scenario we consider the download rate ($R_d$) from the HAP system, and the application rate ($R_a$) to provide the streaming video service to the user. The download rate ($R_d$) minus the application rate ($R_a$) represents the rate at which additional data is stored in the buffer. The amount of information stored in the buffer when the train reaches the tunnel entrance is:

$$S_{b1} = \begin{cases} \frac{(R_d - R_a)X}{V} & \text{if } \frac{(R_d - R_a)X}{V} \leq S_b \\ S_b & \text{if } \frac{(R_d - R_a)X}{V} > S_b \end{cases}$$ (6.2)

As shown in equation (6.2), if the amount of data potentially downloaded to the buffer is smaller than the buffer size ($S_b$), the buffer will continue to fill until the train enters to the tunnel. If the amount of data potentially downloaded to the buffer is larger than the buffer size ($S_b$), the buffer has become full and needs to be considered as such.

When the train enters to the tunnel, the amount of data that will be required to pass through the tunnel without interruption is given by:

$$S_{b2} = \frac{R_aL_T}{V}$$ (6.3)

If the amount of data stored in the buffer is lower than the amount required by the user to pass through the tunnel, an outage will occur. Therefore, the probability of outage when the distance from the tunnel ($X$) is varied is given by:

$$P(\text{outage}) = P(S_{b1} < S_{b2})$$ (6.4)

We consider two different application rates in this scenario, and therefore need two sets of parameters to show the system performance effectively.

For an application rate at 64kbps (Mpeg4), we consider 10Mbits and 2.5Mbits buffer capacities, and download rates of 128 kbps and 96 kbps, which represent twice and one and half times the application rate. The train speed ($V$) is set to 30 m/s. Figure 6-5 shows the resulting outage
probability as a function of distance ($X$) from the tunnel. For an application rate of 320 kbps (AAC), we consider 20Mbits and 5Mbits buffer capacities, and download rates of 640 kbps and 480 kbps, which represent twice and one and half of the application rate. Figure 6-6 show the resulting outage probability as a function of distance ($X$) from the tunnels.

![Figure 6-5 Probability of outage as a function of distance from a tunnel (mpeg4 application)](image)

![Figure 6-6 Probability of outage as a function of distance from a tunnel (AAC application)](image)

These two figures show three aspects of the system performance. Firstly, the outage probability decreases as the distance from the tunnel is increased. Obviously, a longer distance from the tunnel gives more time to fill the buffer until full. Secondly, the download rate from the HAP is a key parameter. A higher download rate produces a much lower outage probability, because the higher download rate enables more data to be stored in the buffer compared with the lower download rate in the same period. Thirdly, it can be seen from these results that there is a fundamental limit beyond which any additional distance from a tunnel has no effect, because the buffer has become full.
An important conclusion can be drawn from these two figures. Assuming the train speed and application rate are fixed, the outage probability is defined by the distance from the tunnel \((X)\), the download rate, and the buffer size. The system design therefore needs to be based primarily on these three parameters. Moreover, the method of system capacity, sharing and buffer size design determines the system performance. In the next section, it is shown how the system can be designed to meet a specified QoS requirement.

6.2.4 Meeting Quality of Service Requirements

If HAPs are to deliver broadband services to trains, appropriate Quality of Service (QoS) requirements need to be met. Multimedia applications such as video and voice have widely varying QoS requirements in terms of time delay, bandwidth and data loss. In this system, the QoS condition is of particular concern for the continuous reception of streaming data by the users. Consider the buffering approach introduced. Although the implementation can be complicated, the idea of buffering is very simple. In a HAP broadband communication system, a buffer is placed on the train to store enough data to maintain communications when the link between the train and the HAP becomes unavailable for a period of time. Given knowledge of the likely duration of such outages, the system can be designed to meet a given QoS requirement.

The single tunnel scenario enables the system parameters to be designed to meet a specified QoS requirement. It enables calculation of the required download rate and buffer size to meet a specified outage probability, given knowledge of the minimum possible distance between two successive tunnels, which represents the worst case. For example, to meet a Quality of Service (QoS) requirement of 5% outage probability with an application rate \((R_a)\) of 64 kbps, a buffer size of approximately 2.5\(\text{Mbits}\) is required (as shown in Figure 6-5). Using the following equation, we can then obtain the minimum download rate needed to pass through two successive tunnels without interruption:

\[
R_{d} = S_{d} I_{d} + R_{ul} \quad (6.5)
\]

Figure 6-7 shows the download rate requirement as a function of minimum distance between tunnels, given the specified QoS requirement.

![Figure 6-7 Download rate requirement as a function of minimum distance between tunnels](image)

The figure shows that download rate requirement decreases when the minimum distance between tunnels is increased, since the longer distance between tunnels provides longer time to fill the buffer prior to entering the next tunnel. Moreover, if the QoS requirement is 5\% outage...
probability and the minimum distance between tunnels is 1km. Mpeg4 requires at least 120kbps download rate. If we want QoS requirement of Mpeg4 to improve to 1% outage probability for the same minimum distance between tunnels, then a 240kbps download rate is required. Therefore, for the same distance between tunnels, the system needs a higher download rate to meet the higher QoS requirement as expected. For an AAC service, with a 1% outage probability QoS requirement and 1km minimum distance between tunnels, a 1.22Mbps download rate will be required. Compared with the Mpeg4 service, the AAC service requires a higher download rate to meet the same QoS requirement.

The train operator could use equation (6.5) to specify an appropriate system, given knowledge of the minimum distance between tunnels along the journey.

### 6.3 Advanced Buffer Techniques for QoS Provisioning in HAPs

#### 6.3.1 Tunnel Outage Simulation Scenario

The tunnel outage scenario is shown in Figure 6-8. Users inside the train communicate with an on-board WLAN base station, which is connected to the wider network via the high altitude platform. A multi-user scenario is considered, with a train journey consisting of many tunnels, and performance is evaluated based on a worst-case minimum distance between successive tunnels.

We consider a number of commuters on the train accessing non-real time services, such as advanced audio coding [6.1]. If the train enters a tunnel, access to the HAP is lost, and so a suitable outage mitigation technique is required. Before entering a tunnel, additional data is downloaded to the on-board buffer, which will maintain service for a short time should the train enter a tunnel, allowing the application to continue. We consider the minimum distance between tunnels to be the worst-case scenario with the distance between successive tunnels being fixed. For example, a fixed assignment scheme can be considered where the system capacity is shared between all the users. Each user receives an equal download rate: \( R_d \text{ Mbit/s} \) from the
The service needs an application rate \( (R_a) \) to run. Therefore, when the train is in the open air, each user will get \((R_d - R_a)\) of additional data. This data is stored in the buffer. Once the remaining application data has been downloaded into the buffer, the download will be finished. After the user enters one tunnel, the user will use the data in the buffer to maintain service. If the data in the buffer runs out before access to the HAP is regained, the user will suffer an outage. The simulation parameters are given in Table 6-1.

In this simulation, the length of tunnels is based on the UK tunnel database, which has recently been published [6.2]. The overall download rate is up to 120 Mbps, as proposed in CAPANINA, representing the highest burst rate from HAP to mobile vehicles or home (office) based WLANs. The minimum overall download rate is based on the number of users in the simulation, and is set to 32 Mbps (assuming 100 users, so that each user receives the required 320 kbps application rate). The last two parameters are related to the users. TTA (time to arrive) represents the time between a user being able to start and actually starting the application and is uniformly distributed in the range between 25s and 75s. The simulation needs to evaluate a certain number of download events during the train journey to obtain meaningful statistical results, hence, each user is considered to start many application events in one journey. The number of active users in a given moment thus depends on the application start times and download durations. The file size associated with each download is based on a uniform distribution in the specified range of 5Mbit to 45Mbit. This chapter focuses on developments of schemes to effectively allocate and distribute system capacity between the users; buffer capacity is not particularly important and so it is assumed that the buffer capacity is infinite. The probability of outage in the whole simulation is calculated, assuming network access is regained immediately on leaving the tunnel.

### Table 6-1 System simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train speed</td>
<td>30m/s (108km/h)</td>
</tr>
<tr>
<td>Distance between tunnels (DBT)</td>
<td>0.5km - 10km</td>
</tr>
<tr>
<td>Tunnel lengths</td>
<td>14m - 4388m</td>
</tr>
<tr>
<td>Overall download rate (ODR)</td>
<td>30Mbps - 120Mbps</td>
</tr>
<tr>
<td>Time to arrive (TTA)</td>
<td>25s - 75s</td>
</tr>
<tr>
<td>File size</td>
<td>5Mbit – 45Mbit</td>
</tr>
<tr>
<td>Application rate ((R_a))</td>
<td>320 kbps (AAC)</td>
</tr>
<tr>
<td>Scheduling period</td>
<td>1s</td>
</tr>
<tr>
<td>Number of users ((N_u))</td>
<td>100</td>
</tr>
<tr>
<td>Buffer capacity</td>
<td>Infinite</td>
</tr>
</tbody>
</table>

The aim of this investigation is to mitigate the probability of outage for a given HAP system overall download rate. We are looking for the best scheme, which can minimise the probability of outage. Five alternative schemes have been developed, and their performance is compared through simulation to identify the most promising for implementation in a real system.

### 6.3.2 Buffer-less download scenario

Before introducing the schemes utilising the buffer, we need to investigate the scenario where no buffer is used on-board the high-speed train, which thus represents a reference performance of the system. If the users finish the application in the open air, they will not suffer an outage. On the other hand, when users still have some remaining information to download on entering a
tunnel, an outage will occur. The download rate for each user equals the application rate, since there is no buffer to download additional data to, and so the users cannot make use of a higher download rate. The probability of outage as a function of minimum distance between tunnels is shown in Figure 6-9.

![Figure 6-9 Outage probability as a function of distance between tunnels (buffer-less download scenario)](image)

Figure 6-9 shows that if the distance between successive tunnels is sufficiently large and users start their applications a long way from the next tunnel, the user has a greater chance of completing a download in the open air. Hence, the outage probability decreases as the considered minimum distance between tunnels is increased. In this scenario without on-board buffer, the outage probability is not controlled by the system. It depends only on the distance between tunnels, and when the users start their applications. From the figure we can see that if the distance between tunnels is 2km, the outage probability is close to 56%, making the provision of such applications infeasible. Consequently, in order to mitigate the outage probability and improve the system performance we propose several buffering schemes based on the advanced on-board buffering approach.

### 6.3.3 Buffering Schemes with Equal Sharing of Resources

#### 6.3.3.1 Fixed Assignment Scheme

Based on resource management theory, the fixed assignment scheme is a basic scheme and the approach is to divide the overall download rate equally by the number of users. Each user gets an equal download rate, including the users that are not accessing streaming applications, but using other applications such as web browsing. This scheme has a buffer on-board the train. Each user will get an equal download rate, $R_d$, from the HAP as shown by the following equation:

$$R_d = \frac{ODR}{N_u}, \quad (6.6)$$

where $ODR$ is the overall download rate, $N_u$ is number of users, and $R_u$ is the download rate available to each users.
The broadband application needs an application data rate $R_a$ to run. Therefore, when the train is in the open air, the buffer fills up at a rate $(R_d - R_a)$. The additional data is stored in the buffer and will offer continued service for a short period of time when the Line of Sight (LOS) link is lost.

The principle of fixed assignment is to share the system capacity equally. The system does not take individual user requirements into account, and so it is not a particularly effective scheme. Since all users enter a tunnel at the same time, the users that started their applications later have a higher probability of suffering an outage, because their buffers have had less time to fill.

6.3.3.2 Demand Assignment Scheme

The demand assignment scheme takes the time varying requirements of users into account. With demand assignment, only the active users are provided with capacity, and so they receive a greater proportion of the overall system capacity compared with fixed assignment (apart from the case where all users are active). The download rate $R_d$ is given by:

$$R_d = \frac{ODR}{N_a}$$

Where, $N_a$ is the number of active users (those accessing a streaming service) and $N_a \leq N_u$. This scheme has to reassign the download resources to active users at each scheduling period.

6.3.3.3 Comparison of Fixed Assignment and Demand Assignment Schemes

The fixed assignment scheme and demand assignment are very similar, since they both share the overall download rate equally between a number of users. The main difference is that the fixed assignment scheme shares the available capacity between all the users and the demand assignment shares capacity between only the active users. The probability of outage as a function of the distance between tunnels and overall download rate is shown for these two schemes in Figure 6-10, where FAS is short for fixed assignment scheme, and DAS is short for demand assignment scheme.

![Figure 6-10 Probability of outage as a function of distance between tunnels](image-url)
From the figure, we can see when capacity is limited and the overall download rate is 32 Mbps, the demand assignment scheme achieves a much lower outage probability than the fixed assignment scheme. For the fixed assignment scheme 100 users share the 32Mbps overall download rate, but for the demand assignment only the active users (less than 100) share the 32Mbps, so the active users get a higher download rate enabling a greater amount of data to be stored in the buffer. As a result, a greater proportion of users manage to pass through the tunnels without suffering an outage. The advantage of the demand assignment scheme in allocating system capacity only to the active users becomes less significant when the overall download rate is large. From Figure 6-10 it can be seen that there is little difference between the demand assignment scheme and the fixed assignment scheme when the overall download rate is 120 Mbps. On the other hand, the advantage of the demand assignment scheme can be clearly seen when overall download rate is set to 15 Mbps; it achieves similar performance as the fixed assignment scheme with 32 Mbps overall download rate.

Table 6-2 shows the outage probability of both schemes as a function of the same distance between the tunnels and overall download rate. When the overall download rate is 60Mbps and the distance between tunnels is 5km, the fixed assignment scheme results in a 7.09% probability of outage, compared with 2.98% for the demand assignment scheme. These figures further highlight the benefits of the demand assignment scheme.

### Table 6-2 Comparison of the fixed assignment and demand assignment schemes

<table>
<thead>
<tr>
<th>Outage probability (%)</th>
<th>Overall download rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32Mbps</td>
</tr>
<tr>
<td></td>
<td>Fixed</td>
</tr>
<tr>
<td>DBT (km)</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>86.20</td>
</tr>
<tr>
<td>1.0</td>
<td>74.55</td>
</tr>
<tr>
<td>5.0</td>
<td>36.42</td>
</tr>
<tr>
<td>10.0</td>
<td>22.10</td>
</tr>
</tbody>
</table>

Figure 6-11 shows the Cumulative Distribution Function (CDF) of time from the start of an application until the entrance of the outage-causing tunnel of those users suffering outage. It is useful to see what percentage of outages occur at the first tunnel the users meet, because this indicates the effectiveness of the proposed resource allocation and buffering technique.

![Figure 6-11 CDF for outage time in different schemes](image)
In Figure 6-11, an overall download rate of 120Mbps and 1km fixed distance between tunnels are considered as typical parameters. The figure shows that the two schemes have a point of intersection. To the left of this point, the demand assignment achieves better performance than fixed assignment. In the case of the fixed assignment scheme, 80% of users suffering an outage have started reception up to 8 seconds before entering a tunnel, compared with 4 seconds for the demand assignment scheme. This means that most of the users will only suffer an outage if they commence their applications just before the tunnel. In such case there is very little time to download data into the buffer, so buffering techniques are not suitable solutions for improvement of the performance in these circumstances.

After the intersection point, the fixed assignment looks like it achieves better performance than the demand assignment scheme, but the demand assignment scheme is in fact better, because users suffer outage much further into the application with the demand assignment scheme. In the case of fixed assignment, 100% of users suffering outage commence their applications up to 31s from the outage causing tunnel entrance. In the case of the demand assignment scheme 95% of users commence their application up to 50s prior to the outage causing tunnel entrance. This means that a proportion of the users suffering outage with the demand assignment scheme have successfully passed through a tunnel, since it takes 33s to pass through the distance between tunnels. This is not the case with the fixed assignment scheme.

6.3.4 Buffering Schemes with Adaptive Buffer Level

6.3.4.1 Step Buffer Scheme

The individual user buffer levels are used to determine an appropriate download rate $R_d$ to each user in the step buffer scheme. If the buffer level is less than a certain amount, users will be given a higher download rate in an attempt to fill the buffer to a certain level. Therefore, the step buffer scheme is based on a buffer threshold $T$. At each successive scheduling period, if the user’s buffer level $B_i$ is less than $T$, the user is given a download rate $R_{d1}$, otherwise the download rate is set at $R_{d2}$, where $R_{d1} \gg R_{d2}$. Since all active users should receive a minimum of 320 kbps download rate to maintain the application rate, $R_{d2}$ is fixed at 320 kbps.

$$R_d = \begin{cases} R_{d1} & B_i < T \\ R_{d2} & B_i \geq T \end{cases}$$

(6.8)

The threshold $T$ is given by:

$$T = \frac{L_t}{V} R_d$$

(6.9)

The size of the threshold is based on the amount of information required in the buffer to pass through a tunnel of a specified length. If we consider a maximum threshold corresponding to the level required to pass through 98% of the UK tunnels, $L_t$ equals 1783m. Using equation (6.9) this maximum threshold is calculated at 18.57 Mbit. Figure 6-12 shows how the system performance varies when different thresholds are applied.

From Figure 6-12, we can select the optimal thresholds for the different system parameters (distance between tunnels and overall download rate). For example, it can be seen that a 7 Mbit threshold is the best for the case where the system provides an overall download rate of 32 Mbps, and the distance between tunnels is 5 km.

The advantage of the step buffer scheme is the fact that users without a sufficient amount of stored information receive higher download rate than those that do. The disadvantage of the scheme is the fact that it does not take individual user requirements into account. Some users
will obtain a higher download rate than they need and others will receive a lower download rate than they need due to the rigid threshold with only two distinct download rates.

![Figure 6-12 System performances with different thresholds](image)

### 6.3.4.2 Proportional Buffer Scheme

In this section, we have analysed three conventional resource management schemes. In order to mitigate the outage probability further, individual user buffer levels are taken into account in the proportional buffer scheme.

The previous schemes share the overall download rate equally between groups of users, and provide only one or two distinct download rates. To ensure a certain buffer level is a way to mitigate against outages more effectively. Users starting a download session a long way from a tunnel entrance have a longer time to fill up the buffer to a certain level than the users starting immediately prior to a tunnel. The concept of the proportional buffer scheme is to provide a higher download rate to users that have recently started their applications and who do not have a high buffer level. The less data the user has stored in the buffer, the higher the download rate the user will receive. To start with, the initial download rate to each user during the first scheduling period is set to \( R_0 \). The sum of the individual user download rates must equal the overall download rate, as shown below:

\[
\sum_{i=1}^{N_a} R_{di} = ODR \quad ; \quad N_a > 0 \tag{6.10}
\]

Where, \( ODR \) is overall download rate, and \( N_a \) is the number of active users. The other rule is to provide each user with a download rate inversely proportional to the current buffer level, as shown in equation 4.11:

\[
R_{di} \propto \frac{1}{B_i} \Rightarrow \frac{k}{B_i} \tag{6.11}
\]

From equations (6.10) and (6.11) we can calculate \( k \) and determine the overall download rate \( R_d \) that each user should receive:
Resource Allocation and Handoff Techniques for High Altitude Platforms

\[ R_{di} = \begin{cases} \frac{R_0 + k}{B_i} ; & B_i \neq 0 \\ R_0 ; & B_i = 0 \end{cases} \]  \hspace{1cm} (6.12)

\[ k = \frac{ODR - N a R_0}{\sum_{i=1}^{N_a} B_i} \]  \hspace{1cm} (6.13)

Where, \( N_a \) is the number of the active users, \( B_i \) is the buffer level of user \( i \), \( R_0 \) is the initial download rate. In the simulation, \( R_d \) needs to be recalculated in every scheduling period.

The initial download rate \( R_0 \) is essential for the proportional buffer scheme, and \( R_0 \) must be greater than or equal to the application rate. Before analysing the system performance, we need to decide what the initial download rate should be. Figure 6-13 shows the system performance with different initial download rates.

This figure shows that the system performance improves as the initial download rate approaches the application rate. This is because the user gets a lower initial download rate, causing less data to go into the buffer. Then, at the next scheduling period, the user will get a much higher download rate because the buffer level is so low. In the following section, we will employ \( R_0 = \frac{4}{3} R_a \).

Figure 6-14 shows the probability of outage as a function of distance between tunnels.
It can be seen from Figure 6-14 that the probability of outage decreases as the distance between tunnels is increased. When the overall download rate is 15 Mbit/s, which is not sufficient to provide every potentially active user with the initial download rate, a high outage probability results. Once the overall download rate is high enough to offer all active users the initial download rate, the outage probability drops significantly.

6.3.4.3 Comparison of Step Buffer and Proportional Buffer Schemes

The step buffer scheme and the proportional buffer scheme both use the user buffer levels to determine an appropriate download rate. However, the step buffer scheme separates active users into two categories, with each type receiving a specified download rate. The proportional buffer scheme considers active users individually, with each user receiving an individual download rate, inversely proportional to the buffer level.

Figure 6-15 shows the outage probability as a function of overall download rate with these two schemes together. It can be seen from Figure 6-15 that the proportional buffer scheme exhibits much better system performance. For example, when the distance between tunnels is 0.5km and the overall download rate is 32 Mbps, the outage probabilities are 65% for the step buffer scheme and 20% for the proportional buffer scheme respectively. It is clear that the proportional buffer scheme provides a much more effective outage mitigation. For example, with a limited system capacity of 60Mbps and a distance between tunnels of 1km, the probability of outage is approximately 32% with the step buffer scheme. When we use the proportional buffer scheme, the outage probability drops significantly to approximately 3%. This performance is generated by the fundamental principle of the proportional buffer scheme. It divides system capacity very fairly by considering the requirement of individual users (their buffer level), filling up the recently started user buffers quickly by giving them a higher download rate.
Table 6-3 shows some detailed performance results for the step buffer scheme and the proportional buffer scheme assuming different overall download rates and distances between tunnels.

**Table 6-3 Comparison of the outage probability of the step buffer scheme with the proportional buffer scheme**

<table>
<thead>
<tr>
<th>Outage probability (%)</th>
<th>Overall download rate</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32Mbps</td>
<td>60Mbps</td>
</tr>
<tr>
<td>Schemes DBT (km)</td>
<td>Step</td>
<td>Proportional</td>
</tr>
<tr>
<td>0.5</td>
<td>64.69</td>
<td>20.16</td>
</tr>
<tr>
<td>1.0</td>
<td>47.40</td>
<td>18.55</td>
</tr>
<tr>
<td>5.0</td>
<td>12.43</td>
<td>5.74</td>
</tr>
<tr>
<td>10.0</td>
<td>6.94</td>
<td>3.28</td>
</tr>
</tbody>
</table>

**6.3.4.4 Exponential Buffer Scheme**

The proportional buffer scheme already shows its advantages and good system performance, however we have investigated another scheme in an attempt to further reduce the outage probability. The idea of exponential buffer scheme is that the download rate varies exponentially with the buffer level, instead of linearly as in the proportional buffer scheme.

Once again, the sum of the user download rates must equal the overall download rate as follows.

\[
\sum_{i=1}^{N} R_{di} = ODR
\]  

(6.14)

The user download rate \( R_{di} \) at each scheduling period is given by:
Resource Allocation and Handoff Techniques for High Altitude Platforms

\[ R_{di} = R_0 + k \exp(-B_i), \]  

(6.15)

where \( ODR \) is the overall download rate, and \( N_a \) is the number of users.

From the previous two equations we can calculate \( k \):

\[ k = \frac{ODR - N_a R_0}{\sum_{i=1}^{N_a} \exp(-B_i)}, \]  

(6.16)

where \( N_a \) is the number of the active users, and \( R_0 \) is initial download rate.

The initial download rate plays an important role in the exponential buffer scheme, since it provides the initial data to the buffer, and determines which user will obtain a particularly high download rate in the next scheduling period. Figure 6-16 shows the outage probability with different initial download rates and system parameters. We can see that initial download rates close to the application rate provide better performance. Similar behaviour was also observed with the proportional buffer scheme as explained in Section 6.3.4.2. Consequently, for performance evaluation of the exponential buffer scheme we choose the initial download rate \( R_0 \) equal to 360 kbps.

![Figure 6-16 Outage probability with different initial download rate](image)

\[ \text{Figure 6-16 Outage probability with different initial download rate} \]

6.3.5 Final Comparison of Schemes

After introducing the alternative buffering schemes, we need to compare them and select the best to apply to a real system. The buffer-less download scenario provides the worst performance and represents the reference scheme in this study, which we are trying to improve. The fixed assignment scheme has been compared with the demand assignment scheme, and the latter scheme was proved as more effective. The step buffer scheme was then compared to the proportional buffer scheme, and the latter has been shown to be more effective. Thus in this section we only compare four schemes in detail: the buffer-less download scenario, the demand assignment scheme, the proportional buffer scheme, and the exponential buffer scheme.
Figure 6-17 shows the probability of outage as a function of distance between tunnels for the different schemes assuming overall download rate (ODR) equal to 32 Mbps.

Comparing the performance of the exponential buffer scheme to the buffer-less download scheme in Figure 6-17 highlights the benefits of on-board buffering. The probability of outage is reduced from 60% to 3%, at an overall download rate of 32 Mbps and with 2 km distance between tunnels. The proportional buffer scheme and the exponential buffer scheme perform better than the alternatives over the entire range of distances considered. In order to determine which scheme is better overall, it is important to investigate the performance of the alternatives as a function of overall download rate, since the distance between tunnels will be fixed by the specific train route.

Figure 6-18 shows the probability of outage as a function of distance between tunnels for the proportional and exponential schemes. A 1km distance between tunnels is considered as a tight constraint on achievable performance.
It can be seen that there is an intersection of the two lines in Figure 6-18. The exponential buffer scheme offers a lower outage probability in the region where system capacity is more constrained, while the proportional buffer scheme performs better at higher overall download rates. However, the proportional buffer scheme offers a consistent outage probability over the entire range of download rates considered and therefore the most effective overall.

### 6.4 Summary and Conclusions

The problem of communications outage in broadband service delivery to trains has been investigated. It has been shown that on-board buffering techniques can effectively mitigate against outages caused by loss of line of sight links on entering tunnels. A range of buffering techniques has been investigated, which differ in the strategy they employ for allocating capacity to users. A fixed assignment scheme has been considered which provides each user with an equal download rate. The advantages of a demand assigned strategy have been demonstrated, whereby the overall download rate is shared only between the active users, taking account of time-varying user requirements, which is important. Although the step buffer scheme is not particularly effective, it provides a useful concept, which is to base the user download rates on their current buffer levels, allocating more capacity to those with a lower buffer level, since they are more susceptible for link outage. Using its principle, the proportional buffer and exponential buffer schemes were developed. Results show that they can achieve a much lower outage probability compared with the buffer-less download scenario, making the provision of non-real time streaming services to trains from a HAP feasible. The proportional buffer scheme provides consistently good performance for a wide range of overall download rates and considered distances between tunnels.
6.5 References


7. Summary and Conclusions

This document has investigated several critical aspects relating to radio resource management and handoff for a cellular based single HAP scenario that is intended to deliver up to 120MBps per cell to both fixed and high speed train users. Much of the work has used the fixed user as a basis for the analysis in order to better develop the schemes and aid the interpretation of the results.

In Chapter 2 the amount of cell overlap that is available in a HAP system has been quantified. It has proved useful to categorise the coverage area into different regions, region A – no overlap, region B – overlap from two cells (potential to use twice the number of channels), and region C – overlap from 3 cells (potential to use three times the number of channels compared with a single cell). These regions are subject to differing amounts of interference, with a trade-off in the level of interference and the total numbers of channels available in each region.

Chapter 3 uses the methodology and cell overlap identified in the previous chapter and investigates various techniques for improving the Quality of Service in terms of blocking probability, data rate and fairness. These techniques are ABFCA, RBFCA, UFCA and UFCA-II.

It is shown that the ABFCA scheme delivers improved blocking probability performance with respect to basic FCA scheme since the overlap areas benefit from the increased channel availability. This is at the expense of introducing non-uniform blocking between the areas and therefore across the coverage area that is not fair for the users. RBFCA has been developed to overcome this problem, which is an optimisation scheme that allocates channels directly to regions in order to unify the blocking across the cell. It is shown that fairness is improved, although limited numbers of channels could lead to trunking efficiency problems in certain circumstances.

These trunking efficiency problems can be overcome by not specifically allocating channels to regions, but instead exploiting a Random Acceptance Factor (RAF) to statistically block connections within each region so as to achieve a uniform blocking across the coverage area, called the UFCA scheme. This ensures that there are more channels available for region A, when the system is becoming full, because in general region A has fewer channel options available. It is shown that uniformity is achieved, while at the same time increasing the number of simultaneous connections that can be supported by the system. These connections delivered a bit rate that was dependent on the CNIR levels, so while the blocking probability was uniform the data rate was not.

An improved variant of UFCA, the UFCA-II scheme, uses a RAF alongside a variable number of channels to ensure both equal blocking and data rate for users across the coverage area. This has been achieved by introducing a bit per connection threshold level $BpC_{thres}$ that all user connections must satisfy. This allowed the CNIR to be better exploited in the channels. For the parameters chosen it has been shown that the blocking probability can be reduced from 4% to 2.6% by exploiting this technique, while at the same time guaranteeing equal data rate.

A connection admission control algorithm has been presented in chapter 4, which uses guard channels to prevent the acceptance of new calls when the available resources become scarce. This approach is commonly used in mobile networks, but the necessity to handle multiple traffic classes and batch handoffs in the CAPANINA high-speed train scenario makes the design of effective techniques particularly challenging. The proposed approximation method for simplified calculation of near-optimal values of channel guards proves sufficiently accurate and efficient in supporting the strategy that keeps the dropping probability on handoff from one cell to another below a predefined threshold, whilst maintaining high maximum resource utilisation efficiency.

Chapter 1 presented an investigation of the impact of the aerial platform movements on the HAP radio resource management schemes. Results have shown that a handoff technique can be used to reduce the need for mechanical stabilisation, by maintaining the continuity of the connections when subject to different types of platform movement. Furthermore, by employing
guard channels the number of connections that drop due to fluctuation in the spatial distribution of users can be substantially reduced, without significantly affecting the blocking probability. Similarly exploitation of the overlap provided a further reduction in dropping, with a reduction from 0.2% to almost 0% when compared with the no overlap case. Overlap is used to increase the trunking efficiency and flexibility in the system, so also provides the added benefit of reducing the blocking probability.

The performance of these schemes was analysed also in the presence of three complex types of HAP motion - rotation, reflection and random walk. Rotational motion is shown to cause the most degradation in performance, with the probability of dropping dependent on the cell dwell time. Performance can be improved by either reserving channels for future handoff use or by controlling the density of admitted users within one area. Hence, the limitation on performance of the HAP system will therefore depend primarily on the resource allocation process, and availability of resources on the platform.

Employing cell overlap has once more shown that the quality of service was significantly improved since both blocking and dropping probability were significantly reduced. From this it could possibly be concluded that cell overlap in a combination with UFCA-II scheme can provide a uniform low blocking and low dropping performance unlike the other schemes examined.

The problem of communications outage in broadband service delivery to trains has been investigated in Chapter 6. It has been shown that on-board buffering techniques can effectively mitigate against outages caused by loss of line of sight links on entering tunnels, making the provision of non-real time streaming services to trains from a HAP feasible. The buffering techniques investigated differ in the way they allocate capacity to users. With the fixed assignment scheme each user receives an equal download rate. Performance can be improved if a demand assigned strategy is employed, whereby the overall download rate is shared only between the active users, taking account of time-varying user requirements. The best schemes evaluated were the proportional buffer and exponential buffer schemes that aim to assign increased capacity to users that have just started their transmissions. All the results indicate that the use of user terminal buffers can result in a much lower outage probability compared with no buffers.