Location and resource management for quality of service provisioning in wireless/mobile networks

Jun Zheng
University of Nevada, Las Vegas

Follow this and additional works at: https://digitalscholarship.unlv.edu/rtds

Repository Citation
https://digitalscholarship.unlv.edu/rtds/2625

This Dissertation is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Retrospective Theses & Dissertations by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.
NOTE TO USERS

This reproduction is the best copy available.

UMI®

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LOCATION AND RESOURCE MANAGEMENT FOR QUALITY OF SERVICE PROVISIONING IN WIRELESS/MOBILE NETWORKS

by

Jun Zheng
Bachelor of Science
Chongqing University, China
1993

Master of Science
Chongqing University, China
1996

Master of Science in Engineering
Wright State University, Dayton, Ohio
2001

A dissertation submitted in partial fulfillment of the requirements for the Doctor of Philosophy Degree in Electrical Engineering
Department of Electrical and Computer Engineering
College of Engineering

Graduate College
University of Nevada, Las Vegas
August 2005
The Dissertation prepared by

Jun Zheng

Entitled

"Location and Resource Management For Quality of Service Provisioning in Wireless/Mobile Networks"

is approved in partial fulfillment of the requirements for the degree of

PH.D. in Electrical Engineering

Examination Committee Chair

Dean of the Graduate College

Examination Committee Member

Graduate College Faculty Representative

PR/1017-52/1-00
ABSTRACT

Location and Resource Management for Quality of Service Provisioning in Wireless/Mobile Networks

by

Jun Zheng

Dr. Emma Regentova, Examination Committee Chair
Professor of Electrical and Computer Engineering
University of Nevada, Las Vegas

Wireless communication has been seen unprecedented growth in recent years. As the wireless network migrates from 2G to 2.5G and 3G, more and more high-bandwidth services have to be provided to wireless users. However, existing radio resources are limited, thus quality-of-service (QoS) provisioning is extremely important for high performance networking. In this dissertation, we focus on two problems crucial for QoS provisioning in wireless networks. They are location and resource management. Our research is aimed to develop efficient location management and resource allocation techniques to provide qualitative services in the future generations of wireless/mobile networks. First, the hybrid location update method (HLU) is proposed based on both the moving distance and the moving direction of mobile terminals. The signaling cost for location management is analyzed using a 2D Markov
walk model. The results of numerical studies for different mobility patterns show that
the HLU scheme outperforms the methods employing either moving distance or mov­
ing direction. Next, a new dynamic location management scheme with personalized
location areas is developed. It takes into account terminal’s mobility characteristics
in different locations of the service area. The location area is designed for each indi­
vidual mobile user such that the location management cost is minimized. The cost is
calculated based on a continuous-time Markov chain. Simulation results acknowledge
a lower cost of the proposed scheme compared to that of some known techniques. Our
research on the resource management considers the dynamic allocation strategy in the
integrated voice/data wireless networks. We propose two new channel de-allocation
schemes, i.e., de-allocation for data packet (DASP) and de-allocation for both voice
call and data packet (DASVP). We then combine the proposed de-allocation meth­
ods with channel re-allocation, and evaluate the performance of the schemes using an
analytic model. The results indicate the necessity of adapting to QoS requirements
on both voice call and data packet. Finally, a new QoS-based dynamic resource allo­
cation scheme is proposed which differentiates the new and handoff voice calls. The
scheme combines channel reservation, channel de-allocation/re-allocation for voice
call and packet queue to adapt to QoS requirements by adjusting the number of re­
served channels and packet queue size. The superiority of the propose scheme in
meeting the QoS requirements over existing techniques is proved by the experimental
studies.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Location management in wireless personal communication networks</td>
<td>2</td>
</tr>
<tr>
<td>1.1.1 Location Update Schemes</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2 Paging Schemes</td>
<td>7</td>
</tr>
<tr>
<td>1.2 Resource management in integrated voice/data wireless networks</td>
<td>12</td>
</tr>
<tr>
<td>1.3 Research Overview and Contribution</td>
<td>16</td>
</tr>
<tr>
<td>1.4 Dissertation Outline</td>
<td>18</td>
</tr>
<tr>
<td>CHAPTER 2 HYBRID LOCATION UPDATE SCHEME FOR WIRELESS PERSONAL COMMUNICATION NETWORKS</td>
<td>19</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>19</td>
</tr>
<tr>
<td>2.2 Hybrid Location Update Scheme</td>
<td>20</td>
</tr>
<tr>
<td>2.3 Performance Analysis of HLU Scheme</td>
<td>22</td>
</tr>
<tr>
<td>2.3.1 2D Markov Walk Model</td>
<td>23</td>
</tr>
<tr>
<td>2.3.2 Signaling Cost of HLU Scheme</td>
<td>24</td>
</tr>
<tr>
<td>2.4 Numerical Result</td>
<td>30</td>
</tr>
<tr>
<td>2.4.1 Signaling Cost vs. $D_S$ and $D_L$</td>
<td>31</td>
</tr>
<tr>
<td>2.4.2 Effect of $p_0$</td>
<td>33</td>
</tr>
<tr>
<td>2.4.3 Effect of $\delta_u$ and $\delta_p$</td>
<td>35</td>
</tr>
<tr>
<td>2.5 Conclusion</td>
<td>36</td>
</tr>
<tr>
<td>CHAPTER 3 DYNAMIC LOCATION MANAGEMENT WITH PERSONALIZED LOCATION AREA FOR WIRELESS PERSONAL COMMUNICATION NETWORKS</td>
<td>37</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>37</td>
</tr>
<tr>
<td>3.2 Dynamic Location Management with Personalized Location Area</td>
<td>38</td>
</tr>
<tr>
<td>3.3 Personalized Location Area Design</td>
<td>39</td>
</tr>
</tbody>
</table>
3.3.1 System Model ................................................................. 39
3.3.2 Markov Analysis .............................................................. 41
3.3.3 Location Management Cost .................................................. 43
3.3.4 Forming Personalized LA ..................................................... 44
3.4 Simulation Results .............................................................. 45
3.5 Conclusion ................................................................. 51

CHAPTER 4 CHANNEL DE-ALLOCATION SCHEMES FOR DYNAMIC RESOURCE ALLOCATION IN THE INTEGRATED VOICE/DATA WIRELESS NETWORKS ......................................................... 55
4.1 Introduction ............................................................................ 55
4.2 Channel De-allocation Schemes ................................................ 56
4.3 Analytic Model ........................................................................ 58
4.4 Numerical Results .............................................................. 64
4.5 Conclusion ............................................................................ 69

CHAPTER 5 PERFORMANCE EVALUATION OF DYNAMIC RESOURCE ALLOCATION WITH CHANNEL DE-ALLOCATION/RE-ALLOCATION IN THE INTEGRATED VOICE/DATA WIRELESS NETWORK ......................................................... 71
5.1 Introduction ............................................................................ 71
5.2 Dynamic Resource Allocation with Channel De-allocation/Re-allocation ........................................................................ 72
5.3 Analytic Model ........................................................................ 73
5.4 Numerical Results .............................................................. 82
5.5 Conclusion ............................................................................ 91

CHAPTER 6 A NEW DYNAMIC RESOURCE ALLOCATION SCHEME FOR QOS PROVISIONING IN THE INTEGRATED VOICE/DATA WIRELESS NETWORKS ......................................................... 92
6.1 Introduction ............................................................................ 92
6.2 QoS-based Dynamic Resource Allocation Scheme ......................... 94
6.3 Analytic Model ........................................................................ 98
6.4 Numerical Results .............................................................. 107
6.4.1 Effect of Packet Queuing .................................................. 108
6.4.2 Effect of Channel Reservation ......................................... 108
6.4.3 Effect of M ................................................................. 109
6.4.4 Performance Comparison for Different Schemes .................. 111
6.4.5 System Award .............................................................. 112
6.5 Conclusion ............................................................................ 115

CHAPTER 7 CONCLUSIONS AND FUTURE WORK ......................................................... 116

BIBLIOGRAPHY ................................................................. 120
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Location area-based location update</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Reporting cell-based location update</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Time-based location update</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Movement-based location update</td>
<td>7</td>
</tr>
<tr>
<td>1.5</td>
<td>Distance-based location update</td>
<td>8</td>
</tr>
<tr>
<td>1.6</td>
<td>Direction-based location update</td>
<td>9</td>
</tr>
<tr>
<td>1.7</td>
<td>Hexagonal cell topology</td>
<td>11</td>
</tr>
<tr>
<td>2.1</td>
<td>Cluster of hexagonal cells ($A_3$) in a personal communication wireless network. The six gray cells form a ring $R_1$.</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Illustration of HLU scheme ($D_S = 1$ and $D_L = 2$)</td>
<td>22</td>
</tr>
<tr>
<td>2.3</td>
<td>The six moving directions of MT</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Sample paths of ISS processes: (a) IID, (b) directional, (c) turning</td>
<td>25</td>
</tr>
<tr>
<td>2.5</td>
<td>The arrangement of paging areas with $D_S = 1$ and $D_L = 3$.</td>
<td>30</td>
</tr>
<tr>
<td>2.6</td>
<td>Signaling cost $C(D_S, D_L)$ for $\rho = 1, p_0 = 0.5, \delta_u = 10, \delta_p = 1, 1 \leq D_L \leq 8, 0 \leq D_S \leq D_L$.</td>
<td>31</td>
</tr>
<tr>
<td>2.7</td>
<td>(a) Optimal cost and (b) optimal distance thresholds vs. $p_0$ for $\rho = 0.1, \delta_u = 10$ and $\delta_p = 1$.</td>
<td>32</td>
</tr>
<tr>
<td>2.8</td>
<td>(a) Optimal cost and (b) optimal distance thresholds vs. $p_0$ for $\rho = 0.5, \delta_u = 10$ and $\delta_p = 1$.</td>
<td>33</td>
</tr>
<tr>
<td>2.9</td>
<td>(a) Optimal cost and (b) optimal distance thresholds vs. $p_0$ for $\rho = 1, \delta_u = 10$ and $\delta_p = 1$.</td>
<td>34</td>
</tr>
<tr>
<td>2.10</td>
<td>(a) Optimal cost and (b) Optimal distance thresholds vs. $\delta_u$ for $\rho = 0.5, p_0 = 0.5, \delta_p = 1$.</td>
<td>35</td>
</tr>
<tr>
<td>3.1</td>
<td>Example of dynamic location management scheme; (a) MT roaming, (b) Forming location areas.</td>
<td>40</td>
</tr>
<tr>
<td>3.2</td>
<td>(a) The cellular network and (b) the graph model</td>
<td>41</td>
</tr>
<tr>
<td>3.3</td>
<td>States of the CTMC</td>
<td>43</td>
</tr>
<tr>
<td>3.4</td>
<td>System topology of the small network with 25 cells</td>
<td>46</td>
</tr>
<tr>
<td>3.5</td>
<td>The movement model of small network with random generated transactional probabilities and cell residence times.</td>
<td>47</td>
</tr>
<tr>
<td>3.6</td>
<td>Performance comparison of three location management schemes for small network (25 cells).</td>
<td>48</td>
</tr>
<tr>
<td>3.7</td>
<td>Performance comparison of location management schemes for large network (100 cells).</td>
<td>49</td>
</tr>
</tbody>
</table>
3.8 Location management cost versus unit location update cost $C_u$ for small network with 25 cells with CMR = 0.3 ................................................................. 50
3.9 Location management cost versus unit location update cost $C_u$ of large network with 100 cells for CMR = 0.1 ................................................................. 51
3.10 The system topology of network with 20 cells ........................................................................ 53
3.11 The movement model of the network with 20 cells .......................................................................... 53
4.1 Voice call dropping probability $P_v$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ................................. 65
4.2 Data packet blocking probability $P_g$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ........................................ 66
4.3 Average data packet transmission time $T_g$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ................................. 67
4.4 Channel utilization $u$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ................................. 68
5.1 Voice call dropping probability $P_v$ of dynamic resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ........................................ 84
5.2 Data packet blocking probability $P_g$ of dynamic resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ........................................ 85
5.3 Average data packet transmission time $T_g$ of dynamic resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ........................................ 86
5.4 Channel utilization $u$ of dynamic resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$. ........................................ 87
5.5 System award factor $Q$ vs. weight factor $\alpha$ and data traffic load $\rho_g$. (a) DASV+RAS, (b) DASP+RAS), (c) DASVP+RAS. ........................................ 89
5.6 System award $Q$ for dynamic resource allocation schemes with $M = 4$, $\alpha_1 = 0.1$, 0.5 and 0.9. ................................. 90
6.1 System model for dynamic resource allocation scheme ........................................................................ 95
6.2 Effect of packet queue size $B$ for different data traffic load $\rho_g$. (a) $P_g$, (b) $P_{vn}$, (c) $P_{vh}$, (d) $u$. ................................. 107
6.3 Effect of channel reservation on different voice traffic load $\rho_v$. (a) $P_g$, (b) $P_{vn}$, (c) $P_{vh}$, (d) $u$. ................................. 109
6.4 Effect of $M$ for different data traffic load $\rho_g$. (a) $P_g$, (b) $P_{vn}$, (c) $P_{vh}$, (d) $u$. ................................. 110
6.5 Performance comparison for different schemes. (a) $P_g$, (b) $P_{vn}$, (c) $P_{vh}$, (d) $u$. ................................. 111
6.6 System award $Q$ vs. $B$ and $g$ for $a = 0.8$, $\beta = 0.1$ and $\gamma = 0.1$, (b) $a = $
0.1, $\beta = 0.8$ and $\gamma = 0.1$, (c) $\alpha = 0.1, \beta = 0.1$ and $\gamma = 0.8$ and (d) $\alpha = 0.4, \beta = 0.3$ and $\gamma = 0.3$. .......................................................... 113

6.7 System award comparison for different schemes. (a) $\alpha = 0.8, \beta = 0.1$ and $\gamma = 0.1$, (b) $\alpha = 0.1, \beta = 0.8$ and $\gamma = 0.1$, (c) $\alpha = 0.1, \beta = 0.1$ and $\gamma = 0.8$ and (d) $\alpha = 0.4, \beta = 0.3$ and $\gamma = 0.3$. ................................. 114
LIST OF TABLES

3.1 Static location areas designed by using Strategy Max_Gain of [20] .......... 52
3.2 Personalized LA designed for each cell as the initial cell .................... 52
3.3 Results for Personalized LA and Static LA ........................................... 52
4.1 DRA algorithm ..................................................................................... 57
4.2 Algorithm for DASV, DASP and DASVP ............................................. 59
4.3 Iterative algorithm to compute $\lambda_{vh}$ and $\pi_e$ .................................. 63
4.4 Performance comparisons of four dynamic channel allocation schemes (★: depends on traffic condition) ......................................................... 69
5.1 Algorithm for DASV+RAS, DASP+RAS and DASVP+RAS ............. 74
6.1 Algorithm of the proposed dynamic resource allocation scheme .......... 97
ACKNOWLEDGEMENTS

I would like to thank all the people who have supported me in the past three and half years of my Ph.D. study. Without their encouragement and help, it is impossible for me to complete my study and this dissertation.

First of all, I would like to give my great appreciation to my Ph.D. advisor, Dr. Emma Regentova, for her support, guidance, kindness and encouragement.

I would like to thank Dr. Shahram Latifi, Dr. Henry Selvaraj and Dr. Kazem Taghva for serving as the members of my dissertation committee.

I would also like to thank the Department of Electrical and Computer Engineering for the financial support of my study.

Finally, I will give my special gratitude for my wife Ai for her understanding and love and the best of all, Elias. I also give my great thanks to my parents, Xuelu Zheng and Xiaojun Pan, for their support, love and patience.
To my wife Ai and wonderful Elias

And

my parents, Xuelu Zheng and Xiaojun Pan
CHAPTER 1

INTRODUCTION

Wireless communication has been seen tremendous growth in the past few years. The number of wireless service subscribers is increasing at an exponential rate and will continue to increase in the near future. The voice, fax, email and paging services provided in the current generation (2G) wireless networks such as Global System for Mobile Communications (GSM) is giving way to mobile multimedia and other high-bandwidth services, such as data transfer, video conferencing, image transfer and other services envisioned in the future generation (2.5G and 3G) wireless networks such as the International Mobile Telecommunication System 2000 (IMT 2000) [17], the General Packet Radio Service (GPRS) [26], the Universal Mobile Telecommunication System (UMTS) [18]. However, wireless networks are resource-limited. Thus for the success of high-bandwidth multimedia services, the Quality of Service (QoS) issue is crucial in the future generation wireless networks.

In our study, we concentrate on two problems of a paramount impact on QoS in wireless networks. They are location management and resource management.
1.1 Location management in wireless personal communication networks

Location management comprises operations performed by the network for tracking the location of mobile terminal (MT) to guarantee delivery of incoming calls [3, 19, 34, 46, 47, 50]. Location management includes two basic tasks, i.e., location update and paging. Location update is initiated by the MT and lies in informing the network about its current location. To establish a connection upon call arrival, the network system pages the MT to determine its location and deliver the call. Both location update and paging require the expenditure of limited wireless resources via the involved signaling. Frequent signaling for location update and paging may result in degradation of QoS of other services. Thus efficient location management strategy is needed to balance the location update and paging operations and minimize the overall signaling cost.

There is considerable research carried out for developing efficient location management schemes. Two simplest ones are the always-update scheme and the never-update scheme [45]. In the always-update scheme, MT updates its location whenever it enters a new cell. Since the system is aware about the exact location of the MT, no paging operation is needed. This scheme results in high location update cost but minimizes the paging cost. On the other hand, in the never-update scheme, no location update is performed. Thus, the system has to page the whole service area to locate the MT upon the incoming call arrival. Obviously, this scheme demands on extensive paging but minimizes the location update cost. These two schemes represent two extreme
cases of location management. Practical schemes employed in most existing wireless systems combine these two simple strategies. In what follows, we review some typical location update and paging schemes employed in existing systems and those proposed for future systems.

1.1.1 Location Update Schemes

In this section we introduce various location update schemes proposed in the literature. They can be divided into two main categories, i.e., static and dynamic.

A. Static Location Update Schemes

In static schemes, location updates are performed in the predefined cells only and all MTs initiate their operations from the same set of cells. There are two main static location update strategies: location area-based and reporting cell-based.

(i) Location Area-based

The service area of the wireless network is divided into several location areas (LAs) each comprised of one or more cells (Fig. 1.1) [33, 35, 53]. The location update takes place when the MT leaves one LA and moves to another one (for example, from cell A to cell B). When an incoming call arrives for the MT, the system performs paging by sending polling messages to all the cells in the MT’s last reported LA to locate the MT. This kind of LA-based update and blanket-polling paging strategy works well for a relatively small number of MTs and it is widely adopted in current generation wireless networks such as GSM.

(ii) Reporting Cell-based
In this scheme, a subset of the cells in the wireless network is designed as the reporting cells (see Fig. 1.2). The MT updates its location only when it enters one of the reporting cells. If an incoming call arrives, the system will page the last-updated reporting cell and a subset of its neighboring bounded non-reporting cells. However, the problem of finding an optimal set of reporting cells is an NP-complete problem as shown in [4]. Hac et al. [12] proposed a heuristic method to find a near optimal solution for this problem. In [44], the authors proposed a parallel and distributed planning algorithm to find reporting cells based on the evolutionary computation and cellular automata.

B. Dynamic Location Update Schemes

In the dynamic schemes, location updates are performed based on MT's individual activity. According to the call arrival and mobility patterns, the individual MT will decide on the time and the place of location update. In what follows, we introduce four
major dynamic location update schemes, namely the time-based, movement-based, distance-based and direction-based schemes [2, 5, 15, 16, 28, 37].

(i) Time-based

In time-base schemes, MT performs location update periodically after a predefined time, T [5, 37]. This scheme is simple because each MT only needs a timer to trigger the location update. An example of the time-based scheme is shown in Fig. 1.3. If at time 0, a location update is performed at location A. Then the MT will trigger updates at locations B, C and D at times T, 2T and 3T, respectively. When a call arrives, the system will page all the possible cells that can be reached by the MT within the time elapsed from reporting in the last known cell.

(ii) Movement-based

In this scheme, the location update is performed by the MT when the number of cell boundary crossings reaches a predefined threshold M [2, 5]. Upon a call arrival,
the system pages all the cells within a distance of M-1 from the last known cell. An example of the movement-based scheme is shown in Fig. 1.4 for the same path as in Fig. 1.3. If the movement threshold is 3, the MT will perform the location updates first at location B and then at location C.

(iii) Distance-based

In the distance-based scheme, the MT performs location update when the distance from the location of the latest update exceeds a certain threshold D [5, 15, 28]. Fig. 1.5 shows the same path as in Fig. 1.3 and 1.4. If the threshold D is set to 2, the MT will perform location update at location B.

(iv) Direction-based
In this scheme, MT updates its location only when there is a change of moving direction [16]. For example, in Fig. 1.6, the MT changes the moving direction at locations B and C. The MT will perform location update at each of the indicated locations and informs the network of its new moving direction. In this way, the network can keep track of the MT's moving direction.

1.1.2 Paging Schemes

Paging is the procedure performed by the network to determine the exact location of the MT. In each polling cycle, the network first sends polling signals to all cells where the MT is likely to reside in through the downlink control channel. The MTs
Figure 1.5: Distance-based location update

residing in the coverage area listen to the paging message and only the targeted MT sends back a respond message over the uplink control channel. In each polling cycle, the paging process is terminated if the MT replies before the timeout. Otherwise, the network sends the polling signals to another group of cells in the next polling cycle. The paging delay is a crucial factor in the paging issue. It affects the QoS in the current and particularly future generation networks whose requirements in that for multimedia services constrain the allowable time for locating users. The maximum paging delay is the maximum polling cycles allowed for locating the MT. If the maximum polling cycle is one, the network has to locate the MT in a signal search operation.
As the paging procedure consumes the network resources, the paging cost is proportional to both the polling cycles and the number of cells being polled in each cycle. In this section we briefly review different paging schemes and discuss their main properties.

A. Blanket Polling

This paging scheme is used in current GSM and IS-41 networks [50]. When a call arrives, the network sends the polling signals simultaneously to all cells within MT’s current location area. Under this paging scheme, the paging delay is minimal that the MT can be found in one polling cycle. The major drawback of this scheme is that the paging cost can be very high when the number of cells in the LA is large.
B. Shortest-Distance-First [2, 15]

In this scheme, the residing LA of the called MT is partitioned into several sub-areas according to the given maximum paging delay and the distance threshold. The network pages the MT starting from the latest update location and then moving outwards in a shortest-distance-first order. The MT can be located under the paging delay constraint by grouping cells at different distances for each polling cycle. For example, in the hexagonal cell topology shown in Fig. 1.7, we assume cell 0 is cell of the latest update and the distance-based location update is used with a threshold 3. If the paging delay is not constrained, i.e. the maximum paging delay is 4, the paging sequence would be \{0, 1, 2, 3\}. That is, the network polls the cell 0 first. If there is no response from the MT after the timeout period, the network polls all the cells in the distance of one in the next polling cycle. This process will continue until the MT is located or all the cells have been polled. If we constrain the maximum paging delay to 2, the paging sequence becomes \{{0, 1}, \{2, 3\}\}. That is, the cells at distance 0 and 1 are polled first and cells at distance 2 and 3 are polled in the second polling cycle. When the maximum paging delay is one, the scheme will be the same as the blanket polling which is the worst case in terms of the paging cost. It should be noted that this paging scheme can be used in conjunction with different location update strategies such as movement-based, distance-based and direction-based schemes. The main advantage of this scheme is that it is simple and it does not require any additional information, such as location probability distribution of
C. Sequential Paging

In this scheme, the cells in the LA are grouped into so called paging areas (PAs) according to the location probability of the MT in each cell. The number of PAs equals to the maximum paging delay. The polling signals are sent by the network to the PAs according to the location probabilities. If the paging delay is not constrained, it is shown in [36] that the minimum paging cost is obtained by sequentially polling the cells in decreasing order of location probability. However, under the delay constraint, the problem of partitioning a LA into PAs is an NP-complete problem and the computation complexity is high when the LA is very large [1, 36]. Several partitioning strategies including reverse, semi-reverse and uniform grouping have been proposed in [48] in addition to the highest-probability-first (HPF) scheme of [36]. In [49], the optimal partitioning of LA into PAs under the delay constraint was studied.
and three conditions are provided to guarantee that the partitioning is optimal.

1.2 Resource management in integrated voice/data wireless networks

The increasing demand on data communication encountered due to the success of Internet applications motives the future generation wireless/mobile networks to support wideband data service in addition to the voice service. Due to the scarce resource of radio channels in wireless networks and QoS requirements of voice and data services, the resources are expected to be managed more efficiently to achieve the required performance.

Wireless/mobile networks such as GSM/GPRS networks and UMTS networks provide both data and voice services [26, 18]. Both GPRS and UMTS allow a single mobile station to transmit data using multiple channels to increase the data transmission rate. For example, in GPRS one to eight channels can be allocated to a single user [26]. In integrated voice/data wireless networks, the voice and data traffic are blocked when the wireless resource is insufficient. Thus, dynamic channel allocation is important to satisfy QoS requirements of different services. In what follows, we review several resource allocation schemes proposed in the literature for integrated voice/data wireless networks, focusing mainly on the GPRS system.

In [22], Lin et al. proposed four schemes for GPRS radio resource allocation. They are fixed resource allocation (FRA), dynamic resource allocation (DRA), fixed resource allocation with queue capability (FRAQ), and dynamic resource allocation
with queue capability (DRAQ). The FRA and FRAQ allocate the exact number of channels requested by data service, while DRA and DRAQ may allocate partial resources if the free channels are insufficient. For all four schemes, one channel is allocated for the voice call request. If there is no available channel, the voice call and data service requests are blocked immediately for FRA and DRA. However, FRAQ and DRAQ employ the voice buffer to queue the voice call request when all the channels are busy. Through analytic and simulation models, the performance of these four resource allocation schemes was evaluated. The study indicates that dynamic resource allocation has a higher GPRS packet acceptance rate compared to that of the fixed one. Among all four schemes, DRAQ which employs dynamic resource allocation with buffering for voice calls has the best performance.

By further considering the buffering mechanism for GPRS packets, Lin proposed two new channel allocation schemes: dynamic resource allocation with the voice and packet queues (DRAVP) and dynamic resource allocation with the packet and voice queues (DRAPV) [23]. These two schemes use a voice queue and a packet queue to buffer the voice calls and packet requests when there is no channel available. In DRAVP, the buffered voice call has higher priority to be served over the buffered data packet. On the other hand, in DRAPV, the buffered data packet has higher priority. The study indicates that the GPRS packet acceptance rate is increased significantly by using the packet queue while the voice call acceptance rate is affected only slightly. Also, when the packet arrival rate is small, DRAPV exhibits better
performance compared to that of DRAVP. As the packet arrival rate becomes larger, DRAVP works better in maintaining the QoS of both the voice calls and the data packet services.

In [29], the authors proposed a channel allocation scheme called dynamic reservation for GPRS systems. The scheme dynamically allocates channels to data when necessary by using the information about the queue length of the data buffer. The scheme constrains the queue length in deciding whether the channels are to be allocated to data or not. By using an analytic model, the authors compared the performance of dynamic reservation with that of the traditional voice priority scheme and R-reservation scheme which allocates a fixed number of channels to the data services. The numerical results have shown that the dynamic reservation provides an effective performance tradeoff for voice and data services.

Based on the dynamic resource allocation, Ferng and Tsai investigated the effects of priority, channel reservation, buffering, and threshold control on the buffer in the channel allocation for GPRS systems [10]. They proposed four channel allocation schemes based on various combinations of these approaches and analyzed the performance using Markov chain approach. The authors concluded that: (1) buffering reduces the blocking probability of voice calls and data packets but increases the delay; (2) priority and threshold control allows for improving the performance of handoff calls; (3) channel reservation improves the performance of handoff voice calls but degrades other services. Finally, they suggested a scheme suitable for GPRS systems
based on priorities over buffers, service priority buffering for both voice calls and data packets, and threshold control on the voice buffer.

As shown in [32], the multi-channel service employed for data service will result in higher blocking probability and longer delay than one that uses single-channel service. These effects can be avoided partially by allocating the resource with the dynamic multi-channel service. To use the multi-channel property and satisfy the demands of different services, Chen et al. [8] proposed a channel de-allocation scheme. The channel is de-allocated from the ongoing data service to the new voice call when there is no free channel in the system [8]. Experimental results indicate that channel de-allocation reduces the voice blocking probability significantly and achieves higher channel utilization at the expense of increased data blocking probability and longer data packet transmission time. Based on the channel de-allocation scheme, Shin et al. [40, 41, 42] and Chen et al. [9] proposed various dynamic channel allocation schemes by employing different combinations of channel de-allocation with buffering, priority and channel reservation. The system performances under the schemes are investigated using analytic models.

In the channel de-allocation scheme, when a voice or data call completes and releases the channels, the freed channels are left idle and wasted. To fully utilize the available resources, Zhang et al. proposed channel re-allocation [55]. The proposed method re-allocates the idling channels to the degraded data service. The study shows that by employing channel re-allocation, the system can achieve significantly lower
voice call blocking probability and data packet transmission time, higher channel utilization at the expense of slightly higher data packet blocking probability.

1.3 Research Overview and Contribution

In this dissertation, we address the problem of location management and resource management for QoS provisioning in wireless/mobile networks. Several new schemes for reducing the overall signaling cost of location management in wireless personal communication networks and allocating the resources to meet the QoS requirements of the voice and data services in the integrated voice/data wireless networks are proposed.

Among various dynamic location update schemes introduced in section 1.1.1, the distance-based and the direction-based methods work well for just certain mobility patterns. As indicated in [5], the distance-based location update (DRLU) scheme achieves good performance under the random walk model but not for directional mobility pattern. In contrast, the direction-based location update (DSLU) scheme performs well for the directional patterns. We propose a hybrid location update (HLU) scheme by taking into account both the moving distance and moving direction [56, 57, 60]. The performance of the HLU scheme is analyzed under different mobility patterns and call-to-mobility ratios based on a 2D Markov walk model. The numerical results demonstrate that the proposed scheme can achieve better performance than those based on only moving distance or moving direction.
In the HLU scheme, we consider that each MT exhibits the same mobility pattern in the whole service area. However, MT may have different mobility characteristics in different locations of the service area. By taking that into account, we propose a dynamic location management scheme with personalized location areas for each MT [59]. Once the MT leaves its current location area, a new location area is formed based on the transition probabilities (i.e. the boundary crossing probabilities) and the residence times of the MT in the cells. The continuous-time Markov chain (CTMC) is used to analyze the location management cost. A heuristic algorithm is developed to determine the personalized location area of a minimum cost. Simulation results show that the proposed scheme offers a lower signaling cost than that obtained by some known strategies.

In [8], channel de-allocation scheme for voice call (DASV) has been considered for dynamic resource allocation in integrated voice/data wireless networks. We design two new channel de-allocation schemes, i.e., de-allocation for data packet (DASP) and de-allocation for both voice call and data packet (DASVP) [58]. An analytic model with the general data channel requirement is derived to evaluate the performance in terms of the voice call blocking probability, data packet dropping probability, average data packet transmission time and channel utilization. By further employing channel re-allocation and combining it with the proposed de-allocation method, we analyze the performance of the schemes and identify a best one for satisfying QoS requirements for both voice call and data packet under different traffic conditions.
Finally, taking into account that handoff calls have higher priority than the new calls, we propose a new channel allocation scheme with channel reservation, channel de-allocation/re-allocation and packet queueing for integrated voice/data wireless networks. An analytic model with the general data channel requirement is developed to evaluate the performance of the scheme. Numerical results demonstrate that the scheme can adapt to different QoS requirements of the system by adjusting the reserved channel capacity and the size of the packet queue.

1.4 Dissertation Outline

The dissertation is organized as follows. In Chapter 2, a hybrid location update scheme for future generation wireless personal communication networks is designed and studied. Chapter 3 presents a dynamic location management scheme with personalized location areas for future generation wireless personal communication networks. In Chapter 4, we propose two new channel de-allocation schemes for integrated voice/data wireless networks. We analyze the performance of different dynamic channel allocation schemes considering channel de-allocation and re-allocation for QoS provisioning in integrated voice/data wireless networks in Chapter 5. In Chapter 6, a new dynamic channel allocation scheme with channel reservation, channel de-allocation/re-allocation and packet queueing is proposed. Finally, we summarize our work in Chapter 7 and discuss the perspectives of the future research.
CHAPTER 2

HYBRID LOCATION UPDATE SCHEME FOR WIRELESS PERSONAL COMMUNICATION NETWORKS

2.1 Introduction

To reduce the signaling cost for location management in wireless personal communication networks, various dynamic location update strategies have been proposed to track the MTs more efficiently including the distance-based [5, 15] and the direction-based [16] schemes. In the distance-based location update (DSLU) scheme, the location update happens when the MT’s distance from its last location update place exceeds a threshold value, D [5, 15]. In the direction-based location update (DRLU) scheme, MT updates its location when a change of direction occurs [16]. To locate the MT, the system only performs paging in the cells along the initial direction so that the paging cost is reduced. In contrast, DSLU scheme achieves good performance under the random walk model as studied in [5]. However, when the MT’s mobility pattern becomes more directional, the signaling cost tends to be high because of the increasing paging cost. On the other hand, the DRLU scheme performs well for the directional mobility pattern, i.e., when the MT doesn’t change its moving direction frequently.
To reduce the signaling cost of the location management for different mobility patterns, we propose a hybrid location update (HLU) scheme that considers both the moving distance and the moving direction bounds [56, 57, 60]. The rest of this chapter is organized as follows. The details of the HLU scheme are introduced in Section 2.2. In Section 2.3, signaling cost of the HLU scheme is analyzed based on a 2D Markov walk model. The numerical results are then presented in Section 2.4. Section 2.5 concludes the discussion.

2.2 Hybrid Location Update Scheme

The wireless network considered in this study has a hexagonal cellular configuration, as shown in Fig. 2.1. In the coordinate system, the x and y axes have their positive portions crossing at a 60° angle. We label each cell in the system by its coordinate (x, y). \( R_i \) is the ith ring of cells where \( R_0 = \{(0,0)\} \). \( A_j \) is the cluster of the cells from ring 0 to ring \( j \) with \( A_j = \bigcup_{i=0}^{j} R_i \). The number of cells in \( R_i \) and \( A_j \) are denoted as \( |R_i| \) and \( |A_j| \), respectively, where \( |R_i| = 6i \) if \( i > 0 \) and \( |R_0| = 1 \), \( |A_j| = 3j(j+1)+1 \).

In HLU scheme, two distance thresholds are employed, \( D_S \) and \( D_L \) where \( 0 \leq D_S \leq D_L \) and \( D_L > 0 \). The HLU scheme works as follows: First, the MT remembers the direction of its original move, i.e., initial direction \( d_j \). Denote \( d_{j-} \) as the direction opposite to \( d_j \), where \( j- = (j+3) \mod 6 \). If the distance of the MT from the origin \( D \) is not greater than \( D_S \), the MT will not update its location no matter what direction
Figure 2.1: Cluster of hexagonal cells (A₃) in a personal communication wireless network. The six gray cells form a ring R₁.

it takes. If $D_S < D \leq D_L$ but the MT moves in the cells along the line of the initial direction (along $d_j$ or $d_{j-}$), there is also no update. MT performs the location updates in three cases: (1) when MT moves from ring $D_S$ to $(D_S + 1)$ and the direction is neither $d_j$ nor $d_{j-}$; (2) change of direction happens when the MT moves within the range $(D_S, D_L]$, i.e. the direction is neither $d_j$ nor $d_{j-}$; (3) MT moves from cell $D_L d_j$ to $(D_L + 1)d_j$, i.e. the moving distance exceeds $D_L$ while MT still keeps the initial direction.

The HLU scheme is illustrated in Fig. 2.2 where $D_S = 1$ and $D_L = 2$. It can be seen that if $D_S = 0$, the HLU scheme is the same as the DRLU, and if $D_S =$
$D_L$, the HLU works as DSLU. In fact, $D_S$ and $D_L$ can be adjusted dynamically to achieve an optimal signaling cost for any given mobility pattern. The paging method for locating MTs is the selective paging with "shortest-distance-first" [15]. Here, we do not constrain the paging delay so that we can study the total cost with respect to its parameters related to location update.

2.3 Performance Analysis of HLU Scheme

In this section, we analyze the performance of the proposed HLU scheme based on a 2D Markov walk model [51]. First, we introduce the 2D Markov walk model in Section 2.3.1. Then we obtain the analytical solution of the signaling cost for the
2.3.1 2D Markov Walk Model

We first introduce the 2D Markov walk mobility model of [51]. In this model, a mobile user can move in six directions $d_i$, $i = 0, \ldots , 5$, which are the six states of the state space $\mathcal{S}$. The directions are defined in a counter clockwise manner as shown in Fig. 2.3, where $[d_0, d_1, d_2, d_3, d_4, d_5] = [(1, 0), (0, 1), (-1, 1), (-1, 0), (0, -1), (1, -1)]$. The location of the MT after $k$th cell boundary crossing $c^{(k)}$ can be described as $c^{(k)} = c^{(k-1)} + d^{(k)}$, where $d^{(k)} \in \mathcal{S}$ is the $k$th direction of the MT. The direction process $\{d^{(k)}, k = 0\}$ constitutes a discrete Markov chain over the state space $\mathcal{S}$. Denote the state transition probability matrix as $P = [p_{ij}]$, $i, j \in \mathbb{Z}_6$, $\mathbb{Z}_6 = 0, 1, \ldots , 5$. $p_{ij} = P(d^{(k+1)} = d_j|d^{(k)} = d_i)$ is the transition probability that the MT’s $(k+1)$th moving direction is $d_j$ on the condition that the $k$th moving direction is $d_i$. The steady-state distribution $\pi_j$ of the direction $d_j$ is given by $\pi_j = \sum_{i \in \mathbb{Z}_6} \pi_i p_{ij}$ and $\sum_{j \in \mathbb{Z}_6} \pi_j = 1$. If $\pi_i = 1/6$, the direction process is isotropic in terms of the steady
state distribution (ISS). For ISS process, the transition probability matrix $\mathbf{P}$ can be obtained by cyclically shifting the first row $\mathbf{p} = [p_0, p_1, p_2, p_3, p_4, p_5]$ as shown in Eq. (2.1).

$$
\mathbf{P} = 
\begin{bmatrix}
  p_0 & p_1 & p_2 & p_3 & p_4 & p_5 \\
  p_5 & p_0 & p_1 & p_2 & p_3 & p_4 \\
  p_4 & p_5 & p_0 & p_1 & p_2 & p_3 \\
  p_3 & p_4 & p_5 & p_0 & p_1 & p_2 \\
  p_2 & p_3 & p_4 & p_5 & p_0 & p_1 \\
  p_1 & p_2 & p_3 & p_4 & p_5 & p_0
\end{bmatrix}
$$

(2.1)

The matrix $\mathbf{P}$ represents a broad class of ISS processes that correspond to different mobility patterns. Some special cases are [51]: (1) independent and identically distributed (IID): $\mathbf{p} = [1/6, 1/6, 1/6, 1/6, 1/6, 1/6]$; (2) directional: $\mathbf{p} = [0.8, 0.025, 0.025, 0.1, 0.025, 0.025]$; (3) turning: $\mathbf{p} = [0, 0.25, 0.25, 0, 0.25, 0.25]$. The sample paths for these ISS processes are shown in Fig. 2.4.

2.3.2 Signaling Cost of HLU Scheme

Based on the above described model, we can obtain the analytic solution for the signaling cost. Particularly, we can find the average number of location updates per call arrival and the average number of cells paged per call arrival. Those can be found from the distribution of the number of location updates performed by the MT in a call arrival period and the distribution of the MT residing in a particular cell when a call comes.
To analyze the signaling cost, we need to consider all possible cases of MT moving into a cell along a certain direction. In the HLU scheme, there are two possibilities for MT to move into the cell of origin (0, 0), i.e. (1) MT moves into cell (0, 0) from cells in $R_1$; (2) MT updates its location. Depends on $D_s$ and $D_L$, there are six possible cases corresponding to the moving directions:

(i) When $D_s = 0$, the MT moves from cell $-d_j$ in direction $d_j$ if it was the initial direction or the MT changes the direction while roaming from cell (0,0) thus causing the location update;

(ii) When $D_s > 0$, the MT is revisiting (0, 0) from cell $-d_j \in R_1$;

(iii) MT moves from cell $D_L d_j$ to $D_{L+1} d_j$ that causes the location update;

(iv) When $0 < D_s < D_L$, the MT moves from ring $D_s$ to $(D_s + 1)$ and the initial direction is neither $d_j$ nor $d_j-1$;

(v) When $D_s = D_L$, the MT moves from ring $D_s$ to $(D_s + 1)$ that incurs the
(vi) Change of direction happens when the MT moves at a distance within \((D_s, D_L]\). For the cell \(c\) at a distance within \([1, D_s]\), MT will move to \(c\) from cell \((c - d_j)\) in direction \(d_j\) if \(c - d_j \in A_{<k>}\) with \(<k> = \min(k, D_s)\) or \((c - d_j) = -(D_s + 1)d_j\) with \(D_s < D_L\). For cell \(c\) at a distance within \([D_s + 1, D_L]\), the MT will move in direction \(d_j\) to \(c\) from cell \((c - d_j)\) if \(d_j\) is the direction of initial move.

Denote \(T_k = (c_k, j_k)\) as the state that the MT moves into the cell \(c_k\) along the direction \(d_{j_k}\), where \(j_k\) is the index of the direction of the MT’s \(k\)th cell boundary crossing and \(j_k \in Z_6\). The state process \(\{T_k, k \geq 0\}\) then forms a Markov chain. The initial state of this Markov chain is \(T_0 = (c_0, j_0)\), where \(c_0 = (0,0)\) is the cell of origin determined at the time of the last call arrival, \(j_0 \in Z_6\). Denote \(P(u_k = u, c_k = c, j_k = j)\) as \(h_k(u, c, j)\), where \(k \geq 0\) is the number of cell boundary crossing, \(u = 0, 1, \ldots, \lceil k/(D_s + 1) \rceil\) is the number of location updates, \(c\) is the cell coordination, \(j \in Z_6\). \(h_k(u, c, j)\) is the probability of the MT moves into cell \(c\) along the direction \(d_j\) in the \(k\)th cell boundary crossing and \(u\)th location update.

Denote \(I_A\) the indicator function whose value is 1 (0) if \(A\) is true (false). We can compute \(h_k(u, c, j)\) recursively using the algorithm below.

**ALGORITHM**

1. **Initialization**

\[
h_0(u, c, j) = \begin{cases} 
\pi_j & u = 0, c = (0,0), j \in Z_6, \\
0 & \text{otherwise}
\end{cases}
\]
2. **Recursion**

For \( k \geq 1, u = 0, 1, \ldots, \lfloor k/(D_S + 1) \rfloor, j \in \mathbb{Z}_6, \)

(1) For cell \( c = (0,0) \)

\[
h_k(u, (0,0), j) = I_{D_S=0} (\sum_{i=j-} p_{i,j} h_{k-1}(u, -d_j, i) + \sum_{i\neq j-} p_{i,j} h_{k-1}(u, (0,0), i)) \quad (i)
\]

\[
+ I_{D_S>0} \sum_{i\in \mathbb{Z}_6} p_{i,j} h_{k-1}(u, -d_j, i) \quad (ii)
\]

\[
+ I_{u>0, D_S<D_L} p_{j,j} h_{k-1}(u - 1, D_L d_j, j) \quad (iii)
\]

\[
+ I_{u>0, k>D_S} \left[ I_{0<D_S<D_L} \sum_{c\in \mathbb{R}_{D_S+1}, c-d_i\in \mathbb{R}_{D_S}} \sum_{i\neq j-} p_{i,j} h_{k-1}(u - 1, c - d_j, i) \right. \quad (iv)
\]

\[
+ I_{D_S=D_L} \sum_{c\in \mathbb{R}_{D_S+1}, c-d_i\in \mathbb{R}_{D_S}} \sum_{i\neq j-} p_{i,j} h_{k-1}(u - 1, c - d_j, i) \quad (v)
\]

\[
+ I_{D_S+1\leq D_L} (D_S+1) \sum_{i\neq j-} p_{i,j} h_{k-1}(u - 1, D_L d_i, i) + I_{D_S+2\leq D_L} (D_S+1) \sum_{i\neq j-} p_{i,j} h_{k-1}(u - 1, n d_i, i) \quad (vi)
\]

(2) For cell \( c \) in \( \mathbb{R}_{D_1} \) to \( \mathbb{R}_{D_S} \) and \( D_S > 0 \)

\[
h_k(u, c, j) = \sum_{c-d_j\in \mathbb{A}(u)} - R_5 \sum_{i\in \mathbb{Z}_6} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

\[
+ I_{D_S+1\leq D_S} \sum_{c-d_j=-(D_S+1)d_j, i=j-} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

(3) For cell \( c \) in \( \mathbb{R}_{D_S+1} \) to \( \mathbb{R}_{D_L} \), \( c = n d_j, D_S < D_L \)
If \( n = (D_L + 1), \ldots, D_L \),

\[
h_k(u, c, j) = \sum_{i=j}^{D_L} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

\[
+ I_{n=D_L+1} \sum_{i=j}^{D_L} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

\[
+ I_{n=0} \sum_{i=j}^{D_L} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

If \( n = -D_L + 2, \ldots, -(D_L + 1) \),

\[
h_k(u, c, j) = \sum_{i=j}^{D_L} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

If \( n = -D_L + 1 \),

\[
h_k(u, c, j) = \sum_{i=j}^{D_L} p_{i,j} h_{k-1}(u, c - d_j, i)
\]

Based on the calculated value of \( h_k(u, c, j) \), we can obtain the distribution of the MT performs location updates \( u \) times in a call arrival period, \( l(u) \), and the distribution of the MT resides in cell \( c \) when a call comes, \( r(c) \).

\[
l(u) = \sum_{k=0}^{\infty} \alpha_k \sum_{c \in \mathcal{A}_{DL}} \sum_{j \in \mathbb{Z}_6} h_k(u, c, j)
\]

\[
r(c) = \sum_{k=0}^{\infty} \alpha_k \sum_{u=0}^{\lfloor k/(D_L+1) \rfloor} \sum_{j \in \mathbb{Z}_6} h_k(u, c, j) \tag{2.2}
\]

where \( \alpha_k \) is the probability that the MT crosses cell boundary \( k \) times in a call arrival
period which is related to the call to mobility ratio (CMR) $\rho$. $\alpha_k$ can be derived as

$$\alpha_k = \begin{cases} 
1 - \frac{1}{\rho} [1 - f_m^*(\lambda_c)] & k = 0 \\
\frac{1}{\rho} [1 - f_m^*(\lambda_c)]^2 [f_m^*(\lambda_c)]^{k-1} & k > 0 
\end{cases}$$

(2.3)

where $\lambda_c$ is the incoming call rate and $\lambda_m$ is the cell moving rate [24]. $f_m^*(\lambda_c)$ is the Laplace-Stieltjes transform of the probability density of cell residence time. For exponential distributed cell residence time with mean $1/\lambda_m$ and variance $1/\lambda_m^2$, $f_m^*(s)$ is represented as

$$f_m^*(s) = \frac{\lambda_m}{\lambda_m + s}$$

(2.4)

The average number of location updates per call arrival is given by

$$U(D_S, D_L) = \sum_{u=0}^{[k/(D_s+1)]} u \cdot l(u)$$

(2.5)

The average number of cells paged per call arrival is given by

$$G(D_S, D_L) = \sum_{n=0}^{D_L} \sum_{c \in PA_n} r(c) \sum_{j=0}^{n} |PA_j|$$

(2.6)

where $PA_n$ is the $n$th paging area. Because we employ the "shortest-distance-first" as the paging technique and do not constrain the paging delay, $PA_n$ is defined as below

$$PA_n = \begin{cases} 
R_n & 0 \leq n \leq D_S \\
\{nd_j, nd_j-\} & D_S < n \leq D_L 
\end{cases}$$

(2.7)
Figure 2.5: The arrangement of paging areas with $D_S = 1$ and $D_L = 3$.

Fig. 2.5 shows the arrangement of paging areas when $D_S = 1$ and $D_L = 3$.

The total signaling cost $C$ as a function of $D_S$ and $D_L$ is then given by

$$C(D_S, D_L) = \delta_u U(D_S, D_L) + \delta_p G(D_S, D_L)$$  \hspace{1cm} (2.8)$$

where $\delta_u$ is the unit location update cost and $\delta_p$ is the per-cell paging cost.

2.4 Numerical Result

In this section, we compare the performance of the proposed HLU with the DRLU and DSLU using the analytical model presented in the previous section. In the numerical analysis, the MT's cell residence time is assumed to follow an exponential
distribution with the mean as $1/\lambda_m$. The incoming call arrival follows Poisson distribution with the rate $\lambda_c$. Thus the CMR is $\rho = \lambda_c/\lambda_m$. To study the effect of the mobility pattern on the performance of the proposed scheme, we set the mobility pattern as $\mathbf{p} = [p_0, (1 - p_0)/5, (1 - p_0)/5, (1 - p_0)/5, (1 - p_0)/5, (1 - p_0)/5]$. When $p_0$ becomes larger, the mobility pattern tends to become more directional. Thus $p_0$ indicates the directivity of the mobility pattern.

2.4.1 Signaling Cost vs. $D_S$ and $D_L$

We first investigate the effect of $D_S$ and $D_L$ on the signaling cost. Fig. 2.6 shows the total signaling cost $C(D_S, D_L)$ of the HLU scheme as a function of $D_S$ and $D_L$ for $\rho = 1, p_0 = 0.5, \delta_u = 10$ and $\delta_p = 1$. From the surface plot, we can find that
Figure 2.7: (a) Optimal cost and (b) optimal distance thresholds vs. $p_0$ for $\rho = 0.1$, $\delta_u = 10$ and $\delta_p = 1$.

the point of the minimum cost is $D_S^* = 1$ and $D_L^* = 4$. Note that the optimal cost of DRLU can be obtained also by locating the minimum point $D_R^*$ on $D_S = 0$ line. Accordingly, the optimal cost for DSLU is the minimum point $D_T^*$ on the $D_S = D_L$ line. Thus, DRLU and DSLU are two subsets of the HLU and they are only suitable for certain mobility patterns. The HLU spans a larger solution space that can adapt to different mobility patterns.
Figure 2.8: (a) Optimal cost and (b) optimal distance thresholds vs. $p_0$ for $\rho = 0.5$, $\delta_u = 10$ and $\delta_p = 1$.

2.4.2 Effect of $p_0$

Figs. 2.7-2.9 compare the total signaling costs of HLU, DRLU and DSLU as functions of $p_0$ for different CMR. $\rho$ varies as 0.1, 0.5 and 1 for Figs. 2.7, 2.8 and 2.9, respectively. The unit location update cost $\delta_u$ is 10 and the per-cell paging cost $\delta_p$ is 1. The cost of HLU is obtained under the optimal thresholds $D_S^*$ and $D_L^*$. The costs of DRLU and DBLU are calculated using the optimal thresholds $D_T^*$ ($D_L = D_S = D_T^*$) and $D_R^*$ ($D_L = D_R^*$, $D_S = 0$), respectively. The results confirm
that for larger values of \( p_0 \) the DRLU performs better than the DSLU. The latter is superior for small \( p_0 \). By adjusting \( D_L \) and \( D_S \) dynamically, HLU can achieve the performance superior to both DRLU and DSLU. Evidently, when the mobility pattern becomes more and more directional, the HLU approaches the performance of the DRLU scheme; otherwise with more random patterns, it tends to provide the performance of the DSLU. From Figs. 2.7-2.9, one can also see that the total signaling cost increases as CMR decreases.
Figure 2.10: (a) Optimal cost and (b) Optimal distance thresholds vs. $\delta_u$ for $\rho = 0.5$, $p_0 = 0.5$, $\delta_p = 1$.

2.4.3 Effect of $\delta_u$ and $\delta_p$

The effect of the unit location update cost $\delta_u$ and per-cell paging cost $\delta_p$ on the total signaling cost as function of $\delta_u$ for $\rho = 0.5$, $p_0 = 0.5$, $\delta_p = 1$ is shown in Fig 2.10(a). The distance thresholds associated with the optimal costs are shown in Fig. 2.10(b). It can be seen that the total signaling cost for each scheme increases with
\( \delta_u \). When \( \delta_u \) is small, DRLU performs better than DSLU. In that range the paging cost dominates the total cost and DRLU employs less paging than DSLU. When \( \delta_u \) becomes larger, the location update operation becomes more expensive and its cost gradually dominates the total cost. DSLU outperforms DRLU when \( \delta_u \) is larger than 7. HLU is superior to DRLU and DSLU. Fig. 2.10(b) shows that the two distance thresholds \( D_s \) and \( D_L \) for HLU can change adaptively according to the change of \( \delta_u \) to achieve an optimal cost. increases as CMR decreases.

2.5 Conclusion

In this chapter, the hybrid location update scheme is designed for locating mobile users in the future wireless personal communication networks. Both the moving distance and the moving direction are taken into account for adapting to various mobility patterns. This allows for reducing the signaling cost significantly. The numerical results demonstrate that under different mobility patterns the proposed scheme outperforms the methods solely based either on the distance or the direction metrics.
CHAPTER 3

DYNAMIC LOCATION MANAGEMENT WITH PERSONALIZED LOCATION AREA FOR WIRELESS PERSONAL COMMUNICATION NETWORKS

3.1 Introduction

In the HLU scheme proposed in Chapter 2, we assume that MT exhibits same mobility pattern in the whole service area. However, MTs may have different mobility characteristics in different locations of the service area. Thus a dynamic location management scheme is needed that takes into account each user's mobility activity in the system. This is expected to lower the overall signaling cost.

To this end, several schemes have been proposed in the literature. In [20], four strategies for grouping cells into location regions by considering movement behavior of individual MT's are discussed. The location regions are fixed for each MT but could be different for different MTs. The authors of [38] and [45] consider individual mobility patterns to create personalized location areas for each MT. Once the MT leaves a current location area, a new location area will be defined based on the transactional probabilities of crossing the cell boundaries. The new location area might contain cells belonging to the old one. However, the above mentioned schemes group cells only based on the transactional probabilities. The grouping is bounded by the size of
location areas, i.e., number of cells in the area. In performing the grouping procedure, changing of location management cost cannot be tracked.

In this chapter, we propose a dynamic location management scheme with a personalized location areas (PLAs) for each MT [59] and evaluate its performance. We use a continuous time Markov chain (CTMC) to analyze the location management cost; and then propose a heuristic algorithm to determine the personalized location area of a minimum cost.

The dynamic location management scheme with personalized location areas is described in Section 3.2. In Section 3.3, the continuous time Markov chain (CTMC) with absorbing states is introduced to derive and evaluate the location management cost. A heuristic algorithm is developed to design the personalized location area. Simulation results are presented in Section 3.4 and the performance analysis of the proposed scheme is given along with its comparison with the existing strategies. Finally, the conclusions are drawn in Section 3.5

3.2 Dynamic Location Management with Personalized Location Area

The essence of the proposed dynamic location management scheme lies in designing LAs for each MT individually and providing this information to them. Once the MT enters the wireless system, a personalized LA is found by minimizing the total location management cost based on the movement behavior of the MT in the system and system parameters. The system then sends the IDs of all cells in the designed
LA to the MT and the latter stores them in its local memory. If the MT moves into a new cell, it checks if the new cell’s ID is in the list. If it is not found, the MT sends a location update message to the system and a new personalized LA is calculated. When there is an incoming call, the system will page all the cells in the LA to identify the cell of MT’s current location and deliver the call. Fig. 3.1 shows an example of the proposed location management scheme. The system has 16 cells indicated in Fig.3.1(b) by circles. Initially the MT resides in cell 1 and the designed personalized LA for the MT includes cells 1, 2 and 5 (bounded by solid rectangle). When the MT moves into cell 5 or 2, there is no location update performed. If a call arrives when the MT is in cell 2, the system will page all three cells to find the MT and deliver the call. The location update is performed when the MT moves for example from cell 2 to the cell 6 (shown shaded). A new LA will be formed with cells 2, 6 and 11 (within dashed rectangle). Note that the new LA overlays the previous one.

3.3 Personalized Location Area Design

3.3.1 System Model

Most location management schemes assume a specific structure such as hexagonal or square to model the cellular network. This kind of topology is featured by the equal number of neighbors for each cell. However, in the real world, an arbitrary cell topology is more realistic. Therefore, to model the network we adopt an arbitrary topology without any special assumptions about the geometry and the interconnec-
Figure 3.1: Example of dynamic location management scheme; (a) MT roaming, (b) Forming location areas.

Similar to the model described in [39], the wireless network can be represented as a bounded-degree, directed graph $G = (V, E)$, where $V$ is the set of nodes representing...
Figure 3.2: (a) The cellular network and (b) the graph model

The cells and $E$ is a set of edges representing the interconnections between the cells. $|V|$ is denoted as the number of nodes in $G$. Two adjacent cells $i$ and $j$ relate by two directed edges $(i,j)$ and $(j,i)$ in the graph. For example, in Fig. 3.2, the set of five nodes $V = \{1, 2, 3, 4, 5\}$. The edge set is $E = \{(1, 2), (2, 1), (1, 3), , (5, 4)\}$.

The MT's movement in the network is modeled as a random walk. Under the random walk model, for each MT, there is a predefined probability $p_{ij}$ of moving from cell $i$ to cell $j$ with $\sum_j p_{ij} = 1$. The residence time of the MT in cell $i$ is assumed to be exponentially distributed with the mean $1/\lambda_{mi}$.

3.3.2 Markov Analysis

The behavior of the MT in a predefined LA is modeled after a continuous time Markov chain (CTMC) with absorbing states. The absorbing state denotes the state of MT of moving out of the current LA. The state space of the CTMC is $S = \{1, \ldots, k, k + 1\}$. As Fig. 3.3 indicates, states 1 to $k$ are transient states that represent the cells in the LA and state $k + 1$ is the absorbing state that represents the
neighboring cells of the LA (cells $k + 1$ to $n$). The generation matrix of the CTMC can be written as

$$Q = \begin{pmatrix} A & B \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -\lambda_{m1} \sum_{j \neq 1} p_{1j} & \lambda_{m1} p_{12} & \ldots & \lambda_{m1} p_{1k} & \lambda_{m1} \sum_{j=k+1}^{n} p_{1j} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_{mk} p_{k1} & \lambda_{mk} p_{k2} & \ldots & -\lambda_{mk} \sum_{j \neq k} p_{kj} & \lambda_{mk} \sum_{j=k+1}^{n} p_{kj} \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix}$$

(3.1)

where $A$ is a $k \times k$ matrix with grouping the transition rates in the transient states, $B$ is column vector with $B = -A e^T$ and $e = [1 \ 1 \ \ldots \ 1]$, $O$ is a $1 \times k$ zero matrix. Without loss of generality, we assume the first cell the MT enters in the LA is cell 1. Thus, the initial probability vector for this CMTC is $p_0 = [1 \ 0 \ \ldots \ 0]$.

Given $\tau_a$ is the time to reach the absorbing state from $t = 0$, the probability distribution of the time until absorption can be written as

$$F_a(t) = Pr\{\tau_a \leq t\} = 1 - p_0 e^{B^T e^T} e^T, t \geq 0$$

(3.2)

The mean time to absorption $E(\tau_a)$ is then given by by

$$E(\tau_a) = -p_0 B^{-1} e^T$$

(3.3)
Thus, the average residence time of the MT in the LA is \( \bar{t} = E(\tau_a) \).

3.3.3 Location Management Cost

The total location management cost for MT in a specific LA \( K \) is defined as

\[
C(K) = C_p \lambda_c N + C_u \Phi_u
\]

where \( N \) is the number of cells in the LA, \( \lambda_c \) is the call arrival rate for the MT, \( \Phi_u \) is the location update rate of the MT for the LA \( K \) which equals to \( 1/\bar{t} \), \( C_p \) and \( C_u \) are the per-cell paging cost and the unit location update cost, respectively. The first component of the right side of Eq. (3.4) corresponds to the paging cost and the addend is the location update cost.
3.3.4 Forming Personalized LA

A personalized LA is formed such that the total location management cost is minimized. Because of high complexity of computations for solving this problem, an iterative greedy heuristic algorithm that yields a sub-optimal solution is found as a viable alternative. The heuristic algorithm performs as follows.

Define:

LA: set of cells in the designed LA
TLA: set of cells in the temporary LA to be checked
Γ(A): the set of neighboring cells of LA A
v: LU cell of the MT

$C_{min}$: minimum signaling cost corresponding to the designed LA
$C^*$: minimum signaling cost corresponding to TLA

1. Initialize $LA = \{v\}$, $TLA = LA$ and $\Gamma(TLA)$, $C_{min} = C(LA) = C_p\lambda_c$
   + $C_u\lambda_mv$

2. Include a new cell into the LA

   $C^* = \infty$

   For cell $i$ in the $\Gamma(LA)$

   Let $TLA' = TLA \cup \{i\}$

   Calculate $\tilde{C}(TLA')$ and $C(TLA')$

   If $C(TLA') < C^*$

   $C^* = C(TLA')$, $TLA = TLA'$
If $C^* < C_{\text{min}}$

$$C_{\text{min}} = C^*, \ LA = TLA$$

End

3. If $C_{\text{min}} < C_p\lambda_c(|TLA|+1)$

Stop

Else

Find $\Gamma(TLA)$

Goto Step 2

End

Note that in step 1, $C_{\text{min}} = C(LA) = C_p\lambda_c + C_u\lambda_{mv}$ is the total signaling cost of the LA only including the LU cell $v$. In Step 3, the algorithm will terminate if $C_{\text{min}}$ is less than $C_p\lambda_c(|TLA|+1)$ which is the paging cost of grouping when one more cell is added to the temporary LA. If the condition is met, there is no need in further check because trivially $C_{\text{min}}$ will be less than the total cost incurred by adding any single cell.

3.4 Simulation Results

In this section, we present results of simulations performed to evaluate the performance of the proposed dynamic location management scheme. We assume the
incoming call follows a Poisson process with mean $\lambda_c$. The residence time of the MT in each cell is exponentially distributed with mean $1/\lambda_{mi}(i = 1, 2, \ldots, |V|)$.

In the first study, we use two networks, i.e. a small one with 25 cells and a large one with 100 cells. The system topology and the movement model of the small network are shown in Fig. 3.4 and Fig. 3.5, respectively. The system topology and the movement model of the large network are omitted here. We organize the system in a hexagonal grid structure for the simulation purpose, although arbitrary cell topology can be used. In the movement models, the labeled arrow indicates the movement direction and transaction probability of the MT. The size of the cell is proportional to the mean residence time of the MT in the cell.

Using the movement model, we first compare the proposed PLA scheme with the always-update (AU) and the distance-based location area (DBLA) schemes under different call-to-mobility ratio (CMR). The distance threshold of DBLA is set to $D =$
Figure 3.5: The movement model of small network with random generated transactional probabilities and cell residence times.

1 which gives the best result. The per-cell paging cost $C_p$ is 1 and the unit location update cost $C_u$ is 10. The incoming call rate (number of calls per hour) is $\lambda_c = 2$. The average mean residence time of $MT$ in the system is taken as $1/\lambda_m = 180, 360, 540s, 720s$ or $900s$ which corresponds to the CMR equal to 0.1, 0.2, 0.3, 0.4 and 0.5, respectively. The simulation is conducted for all the schemes using the same trace. The total number of calls generated for each simulation run is 10,000. 20 simulation runs are performed for each CMR instance. The results are obtained as the mean value of the 20 runs.

Fig. 3.6 shows the simulation result for the small network. For the low CMR range (CMR = 0.1 to 0.4), DBLA performs better than AU. When the CMR becomes larger
Figure 3.6: Performance comparison of three location management schemes for small network (25 cells).

(CMR = 0.5), AU is better than DBLA. As we know, DBLA employs a larger location area than that of AU to reduce the location update rate, and consequently to lower the location update cost. That in turn results in higher paging cost. When the CMR is low, the location update rate is high and the location update cost dominates the total signaling cost. When the CMR increases, the location update rate decreases and the paging cost contributes the total signaling cost. The proposed scheme defines the personalized LA by minimizing the location management cost following the MT’s movement behavior in the system and the system parameters. Thus, the performance of the scheme with the personalized LAs is better than those of AU and DBLA. Fig. 3.7 shows the simulation result for the large network (of 100 cells). The result
Figure 3.7: Performance comparison of location management schemes for large network (100 cells).

demonstrates that PLA outperforms AU and DBLA in the all range of CMR.

Next, we investigate the effect of unit location update cost $C_u$ to the performance of different schemes. Fig. 3.8 plots the location management cost versus $C_u$ for small network of 25 cells with CMR = 0.3 and $C_p = 1$. The simulation is conducted for 10,000 calls per run and 20 runs are applied for each case. It can be seen that the location management cost for each scheme increases with $C_u$. When $C_u$ is small ($C_u = 2$ and 4), AU outperforms DBLA because the paging cost dominates the location management cost and AU employs less paging than DBLA. As $C_u$ becomes larger, the location update operation becomes more expensive and its cost gradually dominates the location management cost. For different $C_u$, PLA has better performance than
Figure 3.8: Location management cost versus unit location update cost $C_u$ for small network with 25 cells with CMR = 0.3.

AU and DBLA because the LA is defined according to the MT's movement behavior and the system parameters. Fig. 3.9 plots the location management cost versus $C_u$ for large network with 100 cells with CMR = 0.1 and $C_p = 1$. One can observe that PLA outperforms AU and DBLA.

In the second study, we compare the performance of the proposed scheme with the static LA (SLA) scheme proposed in [20] with the same network structure and traffic parameters. The network with 20 cells and the corresponding movement model are shown in Fig. 3.10 and Fig. 3.11, respectively. The per-cell paging cost $C_p$ is 1 and the unit location update cost $C_u$ is 2. The mean residence time for each cell is set to 360s and the call arrival rate is 2 calls per hour. Thus the CMR is 0.2. Table 3.1 shows
Figure 3.9: Location management cost versus unit location update cost $C_u$ of large network with 100 cells for CMR = 0.1.

the partition of the system into the LAs by using the Strategy_Max_Gain method that gives the best result among the four strategies proposed in [20]. In Table 3.2, we give out the designed personalized LA for each cell using the heuristic algorithm. The simulation setup is the same as in the case one of the study. The results are shown in Table 3.3 proving that the proposed dynamic scheme outperforms the SLA scheme.

3.5 Conclusion

In this chapter, we have proposed a dynamic location management scheme designed for future wireless personal communication networks. In essence, a personalized location area is formed for each MT based on its mobility pattern in the system
The designed Static Location Areas

\{0, 1, 2\}
\{3, 4, 13, 14\}
\{5, 6\}
\{7, 17, 18, 19\}
\{8, 9\}
\{10, 11, 12\}
\{15, 16\}

Table 3.1: Static location areas designed by using Strategy Max_Gain of [20]

<table>
<thead>
<tr>
<th>Initial Cell</th>
<th>Designed PLA</th>
<th>Initial Cell</th>
<th>Designed PLA</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>{0, 1, 2}</td>
<td>10</td>
<td>{3, 10, 11, 12}</td>
</tr>
<tr>
<td>1</td>
<td>{0, 1, 2}</td>
<td>11</td>
<td>{3, 11, 12, 13}</td>
</tr>
<tr>
<td>2</td>
<td>{0, 1, 2}</td>
<td>12</td>
<td>{3, 11, 12, 13}</td>
</tr>
<tr>
<td>3</td>
<td>{3, 11}</td>
<td>13</td>
<td>{3, 4, 13, 14}</td>
</tr>
<tr>
<td>4</td>
<td>{3, 4, 13}</td>
<td>14</td>
<td>{4, 13, 14, 15}</td>
</tr>
<tr>
<td>5</td>
<td>{4, 5, 6, 15}</td>
<td>15</td>
<td>{5, 14, 15, 16}</td>
</tr>
<tr>
<td>6</td>
<td>{5, 6, 7, 18}</td>
<td>16</td>
<td>{5, 15, 16, 17}</td>
</tr>
<tr>
<td>7</td>
<td>{6, 7, 18, 19}</td>
<td>17</td>
<td>{16, 17, 18}</td>
</tr>
<tr>
<td>8</td>
<td>{1, 7, 8, 9}</td>
<td>18</td>
<td>{6, 7, 17, 18}</td>
</tr>
<tr>
<td>9</td>
<td>{8, 9, 10}</td>
<td>19</td>
<td>{7, 18, 19}</td>
</tr>
</tbody>
</table>

Table 3.2: Personalized LA designed for each cell as the initial cell

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Location Management Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLA</td>
<td>61,008</td>
</tr>
<tr>
<td>SLA</td>
<td>65,299</td>
</tr>
</tbody>
</table>

Table 3.3: Results for Personalized LA and Static LA
Figure 3.10: The system topology of network with 20 cells

Figure 3.11: The movement model of the network with 20 cells

and with the location management cost as the objective function to be minimized.

The implementation of the proposed scheme is as follows: The system pre-computes
the personalized LA for each cell. If the MT moves within the LA, no location update is performed. Otherwise, MT updates the location. In response to this event, the system designates a new LA and sends the IDs of the cells comprising the designed LA to the MT. The numerical results of our study acknowledge a significant decrease of the location management cost compared to that achieved by known schemes such as AU, DBLA and static LA. We are inferring that the scheme can be considered as a viable candidate for future wireless personal communication networks.
CHAPTER 4

CHANNEL DE-ALLOCATION SCHEMES FOR DYNAMIC RESOURCE ALLOCATION IN THE INTEGRATED VOICE/DATA WIRELESS NETWORKS

4.1 Introduction

In integrated voice/data wireless networks such as GSM/GPRS networks, the data services share the resource with the voice services. Since multiple channels can be allocated for the data service to increase the data transmission rate, efficient channel allocation is crucial for meeting QoS requirements on both voice and data services. Lin et al. [22] proposed fixed resource allocation (FRA) and dynamic resource allocation (DRA) for integrated voice/data wireless networks. In FRA, data packet can be served only when the number of freed channels is not less than the number of the requested channels. On the other hand, DRA can allocate partial resources for data packet, i.e., the number of channels used for transmission can be less than the number of requested channels. Experimental results indicate that DRA outperforms FRA in terms of data packet dropping probability. Based on DRA, to further reduce the voice call blocking probability, Chen et al. [8] proposed channel de-allocation scheme for voice call (DASV) by de-allocating a channel from an on-going data service to the new arrived voice call when there is no free channel available in the system. In this chapter,
we propose two new channel de-allocation schemes for dynamic resource allocation in the integrated voice/data wireless networks: de-allocation for data packet (DASP) and de-allocation for both voice call and data packet (DASVP) [58].

To evaluate the performance of the proposed de-allocation schemes and compare them with DRA and DASV, we derive an analytic model. In contrast to the models of [8, 9, 22, 23] in which for the sake of analytical simplicity the maximum number of channels for data packet is fixed to 1, 2 or 3, we derive a model with a relaxed assumption, i.e., with generalized maximum data channel requirement. An analogue of that can be found in [55]. This generalization offers more versatility for the performance analysis, i.e. the study can be carried out under different maximum data channel requirements.

The rest of this chapter is organized as follows. In Section 4.2, we introduce the details of the two new channel de-allocation schemes. The analytic model is derived in Section 4.3. Section 4.4 presents the numerical results of the performance of schemes under different traffic conditions. Finally Section 4.5 concludes this chapter.

4.2 Channel De-allocation Schemes

In this section, we will describe the two new channel de-allocation schemes (DASP and DASVP) and the two reference schemes (DRA and DASV). We consider that in the integrated voice/data wireless network the cells are homogeneous. The base station (BS) in each cell has $C$ channels shared by the voice calls and data packets.
Case 1: Voice Call Arrival

if \( C_F > 0 \)
Accept the voice call;
else
Reject the voice call;

Case 2: Data Packet Arrival

if \( C_F \geq M \) {
Accept the data packet;
Allocate \( M \) channels to the data packet;
}
elseif \( 1 \leq C_F < M \) {
Accept the data packet;
Allocate \( C_F \) channels to the data packet;
}
else
Reject the data packet;

Table 4.1: DRA algorithm

The maximum number of channels used to service a data packet is \( M \). The data packet can be transmitted using \( m \) (\( m = 1, ..., M \)) channels, which is called a type-\( m \) data call. The number of free channels in the system is denoted as \( C_F \) which equals to \( C - n_v - \sum_{m=1}^{M} m n_{g_m} \), where \( n_v \) is the number of voice calls in service, \( n_{g_m} (m = 1, 2, ..., M) \) is the number of type-\( m \) ongoing data packet transmissions.

First we describe the basic dynamic channel allocation scheme (DRA) proposed in [22]. It will be used as a reference technique in the comparison with the schemes employing channel de-allocation. In DRA, if there are free channels in the BS, a voice call will be served. For a data packet arrival, the BS will dynamically allocate channels according to the number of free channels. If there is no free channel in the BS, the voice call is blocked or the data packet is dropped. The algorithm for DRA is illustrated in Table 4.1.
For schemes employing channel de-allocation, when there is no free channel in the BS, de-allocation allows one channel from an ongoing type-\(m\) (\(m > 1\)) data call (also called degradable data call) to service the arrived voice call or data packet, depending on the scheme we use. If no degradable data call exists, the arrived voice call or data packet will be blocked. The algorithm for the three channel de-allocation schemes is shown in Table 4.2.

When channel de-allocation is applied, the channel is de-allocated from the data call with the highest number of channels, e.g. a type-\(m\) data call can be degraded if there is no type-\(q\) (\(q = m + 1, \ldots, M - 1, M\)) data call in the BS.

### 4.3 Analytic Model

We first describe assumptions used in the analytic model. We assume that the new voice calls follow a Poisson process of rate \(\lambda_{vn}\). The duration of a voice call (also called voice call holding time) is a random variable exponentially distributed with mean \(1/\mu_{ch}\). We also assume that the handoff voice call follows a Poisson process with rate \(\lambda_{vh}\). \(\lambda_{vh}\) is related to other parameters such as call arrival rate, call service time etc. and can be determined from these parameters. The total voice call arrival also follows a Poisson process with rate \(\lambda_v = \lambda_{vn} + \lambda_{vh}\). The residence time for voice call is assumed to be exponentially distributed with mean \(1/\mu_{cr}\). The channel holding time of voice call is then exponentially distributed with rate \(\mu_v = \mu_{ch} + \mu_{cr}\). For data packets, we assume the new data packets are generated according to a Poisson process with rate
### Case 1: Voice Call Arrival

If $C_F > 0$

- Accept the voice call;

Else

- If (DASV or DASVP) & (exist degradable data call)
  
  /* channel de-allocation for voice call */
  
  Accept the voice call;

  De-allocate one channel from degradable data call to the voice call;

  Else

  Reject the voice call;

### Case 2: Data Packet Arrival

If $C_F \geq M$

- Accept the data packet;

  Allocate $M$ channels to the data packet;

ElseIf ($1 \leq C_F < M$)

- Accept the data packet;

  Allocate $C_F$ channels to the data packet;

Else

- If (DASP or DASVP) & (exist degradable data call)
  
  /* channel de-allocation for data packet */
  
  Accept the data packet;

  De-allocate one channel from degradable data call to the data packet;

  Else

  Reject the data packet;

---

Table 4.2: Algorithm for DASV, DASP and DASVP

$\lambda_g$. The packet transmission time is assumed to follow an exponential distribution with mean $1/\mu_g$ when one channel is used. If $m$ channels are allocated to a data packet, the mean packet transmission time also follows an exponential distribution with mean $1/m\mu_g$. Notice that in real wireless systems the packet inter-arrival time and the packet transmission time are typically non-exponential distributed. However, this assumption is widely used in modeling for many studies [8, 9, 14, 22, 23, 40, 41, 42, 55] and it provides an useful mean value analysis [14, 23]. In this study, the
handoff of the data packet transmission is not considered, because the transmission
time of individual packet is negligible, and transmission can be completed before the
handoff procedure starts [22, 23].

We model the dynamic channel allocation scheme as a \((M + 1)\)-dimension Markov
process. A state in this process is denoted as \(s = (n_v, n_{g1}, n_{g2}, \ldots, n_{gM-1}, n_{gM})\). The
state space \(S\) of the Markov process is given by

\[
S = \left\{ s = (n_v, n_{g1}, n_{g2}, \ldots, n_{gM-1}, n_{gM}) \mid n_v + \sum_{m=1}^{M} mn_{gm} \leq C, \right. \\
0 \leq n_v \leq C, 0 \leq n_{gm} \leq \left\lfloor \frac{C}{m} \right\rfloor, m = 1, 2, \ldots, M \right\}
\]

Denote the steady state probability for state \(s\) as \(\pi_s\). For all states \(s \in S\),
\(\sum_{s \in S} \pi_s = 1\). To find the steady state probability matrix \(\Pi\), we need to obtain the
generator matrix \(Q = [q_{s \to s'}]_{s \in S, s' \in S}\), where \(q_{s \to s'}\) \((s \in S, s' \in S)\) is the transition rate
from state \(s = (n_v, n_{g1}, n_{g2}, \ldots, n_{gM-1}, n_{gM})\) to state \(s' = (n'_v, n'_{g1}, n'_{g2}, \ldots, n'_{gM-1}, n'_{gM})\).
Define \(I_A\) as the indicator function which equals to 1 (0) when the event \(A\) is true
(true). Denote \(n_{ga}\) as the first non-zero value in the sequence \((n_{gM}, n_{gM-1}, \ldots, n_{g2})\),
where \(\alpha = \max(m \mid n_{gm} > 0, m = M, M - 1, \ldots, 2)\). For \((n_{gM}, n_{gM-1}, \ldots, n_{g2}) \equiv 0, \alpha\n\)is set to -1. Thus if \(\alpha \geq 2\), there exists a degradable data call. \(M\) denotes the state
space of the data call type, \(M = (1, 2, \ldots, M - 1, M)\).

For the dynamic resource allocation scheme (DRA) without channel de-allocation
of [22], we can derive the transition rate \(q_{s \to s'}\) by considering five cases.

(1) A voice call (new call or handoff call) arrives and there are free channels in the
system ($C_F > 0$).

$$q_{s \to s'} = I_{C_F > 0} \lambda_v, n'_v = n_v + 1, n'_g = n_g(k \in M) \quad (4.2)$$

(2) A data packet arrives and the number of free channels $C_F \geq M$.

$$q_{s \to s'} = I_{C_F \geq M} \lambda_g, n'_v = n_v, n'_g_M = n_g + 1, n'_g = n_g(k \in M - \{M\}) \quad (4.3)$$

(3) A data packet arrives and $1 \leq C_F < M$.

$$q_{s \to s'} = I_{1 \leq C_F < M} \lambda_g, n'_v = n_v, n'_g_C = n_g + 1, n'_g = n_g(k \in M - \{C_F\}) \quad (4.4)$$

(4) A voice call terminates or is handoff to neighboring cell

$$q_{s \to s'} = n_v \mu_v, n'_v = n_v - 1, n'_g = n_g(k \in M) \quad (4.5)$$

(5) A type-$m$ data call completes

$$q_{s \to s'} = n_m \mu_g, n'_v = n_v, n'_g_m = n_g - 1, n'_g = n_g(k \in M - \{m\}) \quad (4.6)$$

For the dynamic resource allocation schemes with channel de-allocation, two more cases are considered below.

(6) A voice call arrives and there is no free channel ($C_F = 0$) and $\alpha \geq 2$ that means
we can de-allocate a channel from a degradable data call to serve the arrived voice call. This case supports DASV and DASVP.

\[ q_{s\rightarrow s'} = I_{C_F=0} \alpha \geq 2 \lambda_v, n'_u = n_u + 1, n'_{g_{a-1}} = n_{g_{a-1}} + 1, \]
\[ n'_{g_a} = n_{g_a} - 1, n'_{g_k} = n_{g_k} (k \in M - \{ \alpha - 1, \alpha \}) \]

(4.7)

(7) A data packet arrives and \( C_F = 0 \) and \( \alpha \geq 2 \). This case fits to DASP and DASVP.

\[ q_{s\rightarrow s'} = I_{C_F=0} \alpha \geq 2 \lambda_v, n'_u = n_u, n'_{g_1} = n_{g_1} + 1, n'_{g_{a-1}} = n_{g_{a-1}} + 1, \]
\[ n'_{g_a} = n_{g_a} - 1, n'_{g_k} = n_{g_k} (k \in M - \{ 1, \alpha - 1, \alpha \}) \]

(4.8)

The transition rate \( q_{s\rightarrow s} \) can be obtained as

\[ q_{s\rightarrow s} = - \sum_{s' \neq s, s' \in S} q_{s\rightarrow s'} \]

(4.9)

Using equations (4.2) to (4.9), we can obtain the generator matrix \( Q \) for the \((M+1)\)-dimension Markov chain. The steady-state probability matrix \( \Pi \) is obtained by solving the following linear equations using a numerical method [43]

\[
\begin{cases}
\Pi e = 1 \\
\Pi Q = 0
\end{cases}
\]

(4.10)

where \( e \) is a unit column vector.

Knowing the steady-state probability \( \pi_s \) of the Markov chain, we can represent
Step 1: Select a random initial value for $\lambda_{vh}$.
Step 2: Let $\lambda_{vh, old}$ equal to $\lambda_{vh}$.
Step 3: Get generation matrix $Q$ by using Eq. (4.2) - (4.9).
Step 4: Compute the steady-state probability matrix $\Pi$ by using eq. (4.10).
Step 5: Compute $\lambda_{vh}$ by using eq. (4.11).
Step 6: If $|\lambda_{vh} - \lambda_{vh, old}| > \delta \lambda_{vh}$ where $\delta$ is a small threshold set to $10^{-7}$,
go to Step 2. Otherwise, go to Step 7.
Step 7: The values of $\lambda_{vh}$ and $\pi_s$ converge.

<table>
<thead>
<tr>
<th>Table 4.3: Iterative algorithm to compute $\lambda_{vh}$ and $\pi_s$</th>
</tr>
</thead>
</table>

the handoff voice call arrival rate as

$$\lambda_{vh} = \sum_{s \in S} n_v \pi_s / \mu_{cr} \tag{4.11}$$

Since $\lambda_{vh}$ and $\pi_s$ are mutually related, the following iterative algorithm shown in
Table 4.3 should be applied to compute them [22, 23, 25].

The metrics used for performance comparison are the voice call blocking probability $P_v$, the data packet dropping probability $P_g$, the average data packet transmission time $T_g$ and the channel utilization $u$. Once we get the values of $\pi_s$, they can be represented as

$$P_v = \begin{cases} 
\sum_{n_v+n_{d1}=C, s \in S} \pi_s, & \text{for DASV and DASVP} \\
\sum_{n_v+\sum_{k=1}^{M} kn_{gk}}=C, s \in S \pi_s, & \text{for DASP and DRA}
\end{cases} \tag{4.12}$$

$$P_g = \begin{cases} 
\sum_{n_v+n_{d1}=C, s \in S} \pi_s, & \text{for DASV and DASVP} \\
\sum_{n_v+\sum_{k=1}^{M} kn_{gk}}=C, s \in S \pi_s, & \text{for DASP and DRA}
\end{cases} \tag{4.13}$$
4.4 Numerical Results

Based on the analytic model developed in the previous section, we evaluate the performance of the proposed channel de-allocation schemes DASP and DASVP and compare them with DRA and DASV. We normalize the traffic parameters $\lambda_{vn}$, $\lambda_g$, $\mu_{cr}$ and $\mu_g$ by $\mu_{ch}$ as in [22, 23]. For example, if the expected call holding time is $1/\mu_{ch} = 180$ s, $\lambda_{vn} = 2\mu_{ch}$ means that the expected time between two call arrivals is 360 s. The number of channels in a cell $C$ is assumed to be 7, i.e., one frequency carrier per cell.

Figures 4.1 through 4.4 show the performance metrics as functions of the maximum number of channels used to serve a data packet, $M$, with $\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$. Different traffic conditions are considered and the traffic conditions are differentiated according to voice traffic load $\rho_v = \lambda_{vn}/\mu_v$ and data traffic load $\rho_g = \lambda_g/\mu_g$. There are four different conditions: (1) light voice and data traffics ($\rho_v = 1$ and $\rho_g = 1$); (2) heavy voice traffic and light data traffic ($\rho_v = 5$ and $\rho_g = 1$); (3) light voice traffic and heavy data traffic ($\rho_v = 1$ and $\rho_g = 5$); (4) heavy voice and data traffics ($\rho_v = 5$ and $\rho_g = 5$).

Figures 4.1(a) through 4.1(d) plot the voice call blocking probability $P_v$ versus $M$.
Figure 4.1: Voice call dropping probability $P_v$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$.

under different traffic conditions. When $M$ equals to one, all channel de-allocation schemes will converge to DRA. Thus all schemes have the same $P_v$ performance when $M = 1$. This also holds for data packet dropping probability $P_g$, average data packet transmission time $T_g$ and channel utilization $u$. From Fig. 4.1(a) to Fig. 4.1(d), it can be seen that in terms of $P_v$, DASV and DASVP always outperform DASP and DRA due to channel de-allocation for voice call. For all traffic conditions, the performance
Figure 4.2: Data packet blocking probability $P_g$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$.

of $P_v$ from the worst to the best are DASP, DRA, DASVP and DASV. $P_v$ of DASP and DRA increases as $M$ increases. This is because more channels are used to service data calls as $M$ increases. For DASV and DASVP, $P_v$ almost keeps the same as $M$ changes from 2 to 7.

Figures 4.2(a) through 4.2(d) plot the data packet dropping probability $P_g$ versus $M$ under different traffic conditions. One can observe that in terms of $P_g$, DASP and
DASVP always outperform DASV and DRA due to channel de-allocation for new arriving data packet. Under all traffic conditions, the performances according to $P_g$ from the worst to the best are DASV, DRA, DASVP and DASP. As $M$ increases, $P_g$ of DASV and DRA increase since a data service needs more channels and new data packets are more likely to be dropped. For DASP, $P_g$ decreases as $M$ increases. This is because more channels can be de-allocated for new data arrival as $M$ becomes
Figure 4.4: Channel utilization $u$ of resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$.

larger.

Fig. 4.3(a) through 4.3(d) plot the average data packet transmission time $T_g$ versus $M$ under different traffic conditions. For all traffic conditions, DRA has lower $T_g$ compared to that of schemes employing channel de-allocation. This is because in these schemes the channels are de-allocated from the on-going data service to serve the new arrived voice call or data packet which results in the reduced transmission rate of data packet. For those schemes who employ channel de-allocation, the $T_g$
Table 4.4: Performance comparisons of four dynamic channel allocation schemes (★: depends on traffic condition)

<table>
<thead>
<tr>
<th></th>
<th>DRA</th>
<th>DASV</th>
<th>DASP</th>
<th>DASVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_v$</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>$P_g$</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$T_g$</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>$u$</td>
<td>4</td>
<td>★</td>
<td>★</td>
<td>1</td>
</tr>
</tbody>
</table>

performances from the worst to the best are DASVP, DASP and DASV. For all schemes, $T_g$ decreases as $M$ increases because more channels are used to serve data requests.

Figures 4.4(a) though 4.4(d) plot the channel utilization $u$ versus $M$ under different traffic conditions. Easily we can find that DASVP achieves the best channel utilization among all the schemes under all traffic conditions while DRA has the worst performance in channel utilization. When voice traffic is heavy and data traffic is light (Fig. 4.4(b)), DASV has better channel utilization than DASP. On the other hand, when data traffic is heavy and voice traffic is light (Fig. 4.4(c)), DASP outperforms DASV.

We summarize the performance of the four dynamic channel allocation schemes in Table 4.4.

4.5 Conclusion

In this chapter, we present two new channel de-allocation schemes, i.e., de-allocation for data packet (DASP) and de-allocation for both voice call and data packet (DASVP).
An analytical model with general data channel requirement is derived to evaluate the performance of the schemes. The performances of DRA, DASV, DASP and DASVP under different traffic conditions are analyzed. Compared to the schemes with channel de-allocation, DRA attains the best $T_g$ but rather bad performance for other three performance metrics. For schemes with channel de-allocation, DASV achieves the lowest $P_y$ and DASP has the lowest $P_g$. However, high $P_g$ and $P_v$ are encountered by DASV and DASP, respectively. DASVP achieves the best channel utilization and a good balance between $P_v$ and $P_g$. Among the three schemes employing channel de-allocation, only DASVP offers both lower $P_v$ and $P_g$ than DRA. In next chapter, we will study dynamic channel allocation schemes with both channel de-allocation and re-allocation under the QoS requirements of the system to voice call and data packet.
CHAPTER 5

PERFORMANCE EVALUATION OF DYNAMIC RESOURCE ALLOCATION
WITH CHANNEL DE-ALLOCATION/RE-ALLOCATION IN THE
INTEGRATED VOICE/DATA WIRELESS NETWORK

5.1 Introduction

In the previous chapter, we proposed two new channel de-allocation schemes: DASP and DASVP and analyzed their performance under different traffic conditions. For channel de-allocation schemes, when a voice or data call completes and releases channels, freed channels are left idle and wasted. To fully utilize the available resources, channel re-allocation scheme (RAS) is proposed in [55] such that idling channels are re-allocated to the degraded data services. The authors compared the performance of channel de-allocation for voice call (DASV) without RAS to that of DASV with RAS (DASV+RAS). The results show that RAS can reduce the voice call blocking probability and data packet transmission time significantly at the expense of a higher dropping probability of data packet. In this chapter, we analyze the performance of the proposed two new channel de-allocation schemes, DASP and DASVP with RAS (DASP+RAS and DASVP+RAS). Especially, we consider the QoS requirement of system to the voice call blocking probability and data packet
dropping probability which is measured using a system award factor $Q$. The higher the system award, the better is the performance of the scheme. The performances of DASP+RAS and DASVP+RAS are compared with that of DRA, DASV, DASP, DASVP and DASV+RAS, DASP+RAS and DASVP+RAS under different traffic conditions and different system QoS requirements.

The rest of this chapter is organized as follows. Section 5.2 describes the details of dynamic resource allocation with channel de-allocation and re-allocation. The analytic model is derived in Section 5.3. Section 5.4 presents the numerical results and finally Section 5.5 concludes the chapter.

5.2 Dynamic Resource Allocation with Channel De-allocation/Re-allocation

We assume the integrated voice/data wireless network is homogeneous. The total number of channels shared by voice calls and data packets in the BS of each cell is $C$. Let $M$ is the number of channels requested by a data packet for transmission. According to the number of free channels in the BS, one to $M$ channels can be used for data packet transmission and the data packet using $m(m = 1, \ldots, M)$ channels is called a type-$m$ data call. The number of free channels in the system denoted as $C_F$ equals to $C - n_v - \sum_{m=1}^{M} mn_{g_m}$, where $n_v$ is the number of voice calls in service, $n_{g_m}$ is the number of type-$m$ ongoing data packet transmissions.

For dynamic resource allocation with channel de-allocation and re-allocation, the events resulting in the state change are: voice call arrival, voice call completion
(termination or handoff), data packet arrival and data packet completion. For voice call arrival, if there are free channels in the BS, one channel is allocated. If there is no available channel in the BS, DAS will be used to service the new arrived voice call when there is a degradable on-going data service for schemes with DASV or DASVP.

Upon arrival of data packet, channels will be dynamically allocated according to the number of free channels in the BS. For a scheme with DASP or DASVP, DAS is employed to service the new arrived data packet if no channel is available in the BS and there is a degradable on-going data service. Upon channel release due to the completion of a voice call or a data packet, if there are type-\(m\) \((m < M)\) data calls (also called degraded data calls), RAS will be used to allocate freed channels to upgrade the degraded calls. RAS is performed using the "worst degraded first upgrading" policy [55]. That is, a degraded data call can be upgraded if all type-\(q\) \((q = 1, 2, ..., m - 1)\) data calls have been upgraded to type-\(M\). The algorithm for the dynamic resource allocation schemes with channel de-allocation/re-allocation (DASV+RAS, DASP+RAS and DASVP + RAS) is shown in Table 5.1.

5.3 Analytic Model

In this section, we derive an analytic model to evaluate the performance of the dynamic resource allocation schemes with channel de-allocation and re-allocation. The output performance metrics of the model are voice call blocking probability \(P_v\), data packet dropping probability \(P_g\), average data packet transmission time \(T_g\).
Case 1: Voice Call Arrival
if \( C_F > 0 \)
    Accept the voice call.
else
    if (DASV+RAS or DASVP+RAS) & (a degradable data call exists){
        /* Channel de-allocation for voice call */
        Accept the voice call;
        De-allocate one channel from degradable data call to the voice call;
    }
    else
        Reject the voice call;

Case 2: Data Packet Arrival
if \( C_F \geq M \){
    Accept the data packet;
    Allocate \( M \) channels to the data packet;
}
elseif \( 1 \leq C_F < M \){
    Accept the data packet;
    Allocate \( C_F \) channels to the data packet;
}
else
    if (DASP or DASVP) & (exist degradable data call){
        /* Channel de-allocation for data packet */
        Accept the data packet;
        De-allocate one channel from degradable data call to the data packet;
    }
    else
        Reject the data packet;

Case 3: Voice Call Completion
if (exist degraded data calls)
    /* Channel re-allocation */
    one channel is used to upgrade a degraded data call;

Case 3: Type-\( m \) Data Call Completion
if (exist degraded data calls)
    /* Channel re-allocation */
    \( m \) channels are used to upgrade degraded data calls;

Table 5.1: Algorithm for DASV+RAS, DASP+RAS and DASVP+RAS

channel utilization \( u \) and the system award \( Q \).

For the analytic model, we assume that the generation of new voice calls and
new data packets follows the Poisson process with rates \( \lambda_{un} \) and \( \lambda_g \), respectively. The voice call holding time, the cell residence time and the packet transmission time are assumed to be exponentially distributed with means \( \mu_{ch} \), \( \mu_{cr} \) and \( \mu_g \). In the model, we consider the effect of voice call mobility, i.e. the voice call can be handoff to another cell. The mobility of data packet is not considered due to the reason described in chapter 4. The handoff voice call is assumed to follow a Poisson process with rate \( \lambda_{vh} \) which is related to other parameters. The total voice call arrival rate is given by \( \lambda_v = \lambda_{un} + \lambda_{vh} \), while the mean channel holding time of a voice call is \( 1/\mu_v = 1/(\mu_{ch} + \mu_{cr}) \).

The dynamic resource allocation scheme is modeled as a \((M + 1)\)-dimension Markov process. A state in this process is denoted as \( s = (n_v, n_{g1}, n_{g2}, ..., n_{gM-1}, n_{gM}) \). The state space \( S \) of the Markov process is given by

\[
S = \left\{ s = (n_v, n_{g1}, n_{g2}, ..., n_{gM-1}, n_{gM}) \mid n_v + \sum_{m=1}^{M} mn_{gm} \leq C, 0 \leq n_v \leq C, 0 \leq n_{gm} \leq \left\lfloor \frac{C}{m} \right\rfloor, m = 1, 2, ..., M \right\}
\]

(5.1)

We denote the steady state probability for state \( s \) as \( \pi_s \). For all states \( s \in S \), \( \sum_{s \in S} \pi_s = 1 \). To find the steady state probability matrix \( \Pi \), we need to obtain the generator matrix \( Q = [q_{s \to s'}]_{s \in S, s' \in S} \), where \( q_{s \to s'} \) \((s \in S, s' \in S)\) is the transition rate from state \( s = (n_v, n_{g1}, n_{g2}, ..., n_{gM-1}, n_{gM}) \) to state \( s' = (n'_v, n'_{g1}, n'_{g2}, ..., n'_{gM-1}, n'_{gM}) \).

Given below are some definitions introduced for deriving the generator matrix \( Q \).

\( I_A \): The indicator function which equals to 1 (0) when the event \( A \) is true (false).
\(n_{g_0}\): The first non-zero value in the sequence \((n_{g M}, n_{g M - 1}, \ldots, n_{g_2})\), where \(\alpha = \max(m \mid n_{g_m} > 0, m = M, M - 1, \ldots, 2)\). For \((n_{g M}, n_{g M - 1}, \ldots, n_{g_2}) \equiv 0\), \(\alpha\) is set to -1. Thus if \(\alpha \geq 2\), there exists a degradable data call.

\(n_{g_0}\): The first non-zero value in the sequence \((n_{g_1}, n_{g_2}, \ldots, n_{g_{M-1}})\), where \(\beta = \min(k \mid n_{g_k} > 0, k = 1, 2, \ldots, M - 1)\). For \((n_{g_1}, n_{g_2}, \ldots, n_{g_{M-1}}) \equiv 0\), \(\beta\) is set to -1. Thus if \(\beta \geq 0\), there exists a degraded data call.

\(M\): The state space of the data call type, \(M = (1, 2, \ldots, M - 1, M)\).

To obtain the transition rate \(q_{s \rightarrow s'}\), we consider four cases with respect to the events of voice call arrival, voice call completion, data packet arrival and data packet completion.

**Case 1: Voice call (new call or handoff call) arrival**

(1a) There are free channels in the BS \((C_F > 0)\)

\[
q_{s \rightarrow s'} = I_{C_F > 0} \lambda_v, n_{v}' = n_v + 1, n_{g_k}' = n_{g_k} (k \in M) \tag{5.2}
\]

(1b) There is no channel available in the BS \((C_F = 0)\) and \(\alpha \geq 2\) which means that there is a degradable data call in the BS and channel de-allocation for the voice call can be applied. This case supports schemes with DASV and DASVP.

\[
q_{s \rightarrow s'} = I_{C_F = 0} I_{\alpha \geq 2} \lambda_v, n_{v}' = n_v + 1, n_{g_{\alpha-1}}' = n_{g_{\alpha-1}} + 1, n_{g_{\alpha}}' = n_{g_{\alpha}} - 1, n_{g_k}' = n_{g_k} (k \in M - \{\alpha - 1, \alpha\}) \tag{5.3}
\]
Case 2: Voice call completion (termination or handoff)

(2a) $\beta > 0$ means that there is a degraded data call in the BS and the freed channel can be re-allocated to upgrade the degraded data call.

$$q_{s \rightarrow s'} = I_{\beta > 0} n_v \mu_v, n_v' = n_v - 1, n_{g_{\beta}}' = n_{g_{\beta}} - 1,$$

$$n_{g_{\beta+1}}' = n_{g_{\beta+1}} + 1, n_{g_k}' = n_{g_k}(k \in M - \{\beta, \beta + 1\})$$

(5.4)

(2b) $\beta = 0$ means that there is no degraded data call in the BS.

$$q_{s \rightarrow s'} = I_{\beta = 0} n_v \mu_v, n_v' = n_v - 1, n_{g_k}' = n_{g_k}(k \in M)$$

(5.5)

Case 3: Data packet arrival

(3a) There are free channels in the BS and $C_F \geq M$.

$$q_{s \rightarrow s'} = I_{C_F \geq M} \lambda_g, n_v' = n_v, n_{g_M}' = n_{g_M} + 1, n_{g_k}' = n_{g_k}(k \in M - \{M\})$$

(5.6)

(3b) There are free channels in the BS and $1 \leq C_F < M$.

$$q_{s \rightarrow s'} = I_{1 \leq C_F < M} \lambda_g, n_v' = n_v, n_{g_{C_F}}' = n_{g_{C_F}} + 1, n_{g_k}' = n_{g_k}(k \in M - \{C_F\})$$

(5.7)

(3c) There is no channel available in the BS ($C_F = 0$) and $\alpha \geq 2$ which means that there is degradable data call in the BS and channel de-allocation for data packet can
be applied. This case works for schemes employing DASP and DASVP.

\[ q_{s \rightarrow s'} = I_{C_F=0 \alpha \geq 2 \lambda_g}, n_v' = n_v, n_{g_{1 \alpha}}' = n_{g_{1 \alpha}} + 1, n_{g_{a-1}}' = n_{g_{a-1}} + 1, \]
\[ n_{g_a}' = n_{g_a} - 1, n_{g_k}' = n_{g_k}, k \in M - \{1, \alpha - 1, \alpha\} \]  
(5.8)

**Case 4: Data packet completion**

(4a) Type-\( m \) data call completes.

Upon the type-\( m \) data call completion, the new state becomes

\[ s_1 = (n_v^*, n_{g_1}^*, \ldots, n_{g_m}^*, \ldots, n_{g_M}^*), \]

where \( n_v^* = n_v, n_{g_m}^* = n_{g_m} - 1, n_{g_k}^* = n_{g_k}, k \in M - \{m\} \). We introduce an index \( \theta (1 \leq \theta \leq M - 1) \) such that

\[ \sum_{k=1}^{\theta-1} n_{g_k}^* (M - k) \leq m \leq \sum_{k=1}^{\theta} n_{g_k}^* (M - k) \]  
(5.9)

Inequality (5.9) implies that when a type-\( m \) data call leaves the BS, all the type-1 to type-\((\theta - 1)\) calls and some of the type-\( \theta \) calls can be upgraded to type-\( M \) calls. If \( \sum_{k=1}^{M-1} n_{g_k}^* (M - k) \leq m \) which means that all the type-1 to type-\((M - 1)\) calls can be upgraded to type-\( M \) calls, we set \( \theta \) to \( M \). If \( \beta = 0 \) which means that there is no data call in the BS or all the data calls are type-\( M \), we set \( \theta \) to -1.

(4a.1) \( \theta = -1 \)

\[ q_{s \rightarrow s'} = I_{\theta=-1} n_{g_m} m \mu_g, n_v' = n_v, n_{g_m}' = n_{g_m} - 1, n_{g_k}' = n_{g_k} (k \in M - \{m\}) \]  
(5.10)
(4a.2) \( 1 \leq \theta \leq M - 1 \)

\[
q_{s \rightarrow s'} = I_{1 \leq \theta \leq M-1} n_{g_{\mu \gamma}} m_{\mu \gamma}, n_{v} = n_{v}, n_{g_{k}} = 0 (k = 1, 2, \ldots, \theta - 1), n'_{g_{\theta}} = n_{g_{\theta}} - \delta - 1, \\
n'_{g_{\theta+\delta r}} = n_{g_{\theta+\delta r}} + 1, n'_{g_{M}} = n_{g_{M}} + \sum_{k=1}^{\theta-1} n_{g_{k}} + \delta, n'_{g_{k}} = n_{g_{k}} \text{(other index } k \text{ in M)}
\]

(5.11)

where

\[
\delta = \left[ m - \sum_{k=1}^{\theta-1} n_{g_{k}} (M - k) \right] \frac{1}{M - \theta}
\]

is the number of type-\( \theta \) calls that can be upgraded to type-\( M \) calls. \( \delta_{r} = \delta - \delta(M - \theta) \)
is the number of channels that can upgrade a type-\( \theta \) call to type-\( (\theta + \delta_{r}) \).

(4a.3) \( \theta = M \)

\[
q_{s \rightarrow s'} = I_{\theta=M} n_{g_{\mu \gamma}} m_{\mu \gamma}, n_{v} = n_{v}, n'_{g_{k}} = 0 (k \in M - \{M\}), n'_{g_{M}} = \sum_{k=1}^{M} n_{g_{k}} - 1
\]

(5.12)

The transition rate \( q_{s \rightarrow s} \) can be obtained as

\[
q_{s \rightarrow s} = - \sum_{s' \in S, s' \in S} q_{s \rightarrow s'}
\]

(5.13)

Using equations (5.2) to (5.8) and (5.10) through (5.13), we can obtain the generator matrix \( Q \) for the \((M + 1)\)-dimension Markov chain. The steady-state probability
matrix $\mathbf{H}$ is obtained by solving the following linear equations \[43\]

$$
\begin{align*}
\Pi e &= 1 \\
\Pi Q &= 0
\end{align*}
$$

(5.14)

where $\mathbf{e}$ is a unit column vector. The linear equations can be solved using a numerical method introduced in \[43\].

From the steady-state probability $\pi_s$ of the Markov chain, the handoff voice call arrival rate $\lambda_{vh}$ can be expressed as

$$
\lambda_{vh} = \sum_{s \in \mathcal{S}} n_v \pi_s \mu_{cr}
$$

(5.15)

Since $\lambda_{vh}$ and $\pi_s$ are mutually related, we use an iterative algorithm introduced in chapter 4 to compute $\lambda_{vh}$ and $\pi_s$.

For schemes DASV+RAS and DASVP+RAS, the voice call is blocked if in the BS there is no channel available ($C_F = 0$) and all data calls are type-1. For DASP+RAS, the voice call is blocked when $C_F = 0$. Thus the voice call blocking probability $P_v$ can be written as

$$
P_v = \begin{cases} 
\sum_{n_v+n_g=C_s \in \mathcal{S}} \pi_s, & \text{for DASV+RAS and DASVP+RAS} \\
\sum_{n_v+C_k^{(1)} kn_{bh}=C_s \in \mathcal{S}} \pi_s, & \text{for DASP+RAS}
\end{cases}
$$

(5.16)

For schemes DASP+RAS and DASVP+RAS, the data packet is dropped if there
is no free channel \((C_F = 0)\) and all data calls all type-1. For DASV+RAS, the data packet is dropped when \(C_F = 0\). The data packet dropping probability \(P_g\) is given as

\[
P_g = \begin{cases} 
\sum_{n_v + n_g = C, s \in S} \pi_s, & \text{for DASP+RAS and DASVP+RAS} \\
\sum_{n_v + \sum_{k=1}^{M} k n_{g_k} = C, s \in S} \pi_s, & \text{for DASV+RAS}
\end{cases}
\] (5.17)

The average data packet transmission time \(T_g\) is expressed as

\[
T_g = \frac{\sum_{s \in S} \left( \sum_{k=1}^{M} n_{g_k} \right) \pi_s}{\lambda_g (1 - P_g)}
\] (5.18)

The channel utilization \(u\) is given by

\[
u = \frac{\sum_{s \in S} \left( n_v + \sum_{k=1}^{M} k n_{g_k} \right) \pi_s}{C}
\] (5.19)

To compare the performance of different schemes under the QoS requirement of the system for voice call blocking probability \(P_v\) and data packet dropping probability \(P_g\), we use a system award factor \(Q\) \cite{14} as expressed below

\[
Q = \alpha_1 (1 - P_v) + \alpha_2 (1 - P_g)
\] (5.20)

where \(\alpha_1\) and \(\alpha_2\) are two weight factors and \(\alpha_1 + \alpha_2 = 1\). The factors \(\alpha_1\) and \(\alpha_2\) weight the importance of the voice call blocking probability and data packet dropping probability for the system's QoS. They are determined by the system's overall revenue.
and service objectives [21]. Higher value of $Q$ indicates better performance.

5.4 Numerical Results

In this section, we evaluate the performance of dynamic resource allocation schemes with channel de-allocation and re-allocation based on the analytic model. The traffic parameters $\lambda_v$, $\lambda_d$, $\mu_{cr}$ and $\mu_d$ are normalized by $\mu_{ch}$ [22, 23]. We assume the total number of channels in a cell $C$ equals to 7. We investigate the performance of the dynamic resource allocation schemes with DAS/RAS under different voice call and data packet traffic loads $\rho_v$ and $\rho_d$. We provide here the performance of DRA, DASV, DASP and DASVP obtained in chapter 4 for the purpose of comparison.

Figures 5.1 to 5.4 show the performance metrics $P_v$, $P_g$, $T_g$ and $u$ under four traffic conditions. The parameters are set as $\mu_{ch} = 1/180$, $\mu_{cr} = 0.2\mu_{ch}$, $\mu_d = 100\mu_{ch}$. $M$ varies from 1 to 7. Four traffic conditions considered here are: (1) light voice and data traffics ($\rho_v = 1$ and $\rho_d = 1$); (2) heavy voice traffic and light data traffic ($\rho_v = 5$ and $\rho_d = 1$); (3) light voice traffic and heavy data traffic ($\rho_v = 1$ and $\rho_d = 5$); (4) heavy voice and data traffics ($\rho_v = 5$ and $\rho_d = 5$). From Figures 5.1 though 5.4, all dynamic resource allocation schemes have the same performance for $M = 1$. This is because there is no channel de-allocation and re-allocation can be applied so all the schemes with DAS and RAS converge to DRA.

Figures 5.1(a) through 5.1(d) plot the voice call blocking probability $P_v$ as a function of $M$ under different traffic conditions. From the results, it can be seen
that when employing RAS, DSAV+RAS and DASVP+RAS have lower $P_v$ compared to their counterparts, i.e., DASV and DASVP, respectively. However, DASP+RAS results in higher $P_v$ than DASP. That can be explained as follows. RAS is used to re-allocate the idle channels to the degraded data calls to improve the transmission time and the available channels in the BS are reduced. Upon arriving of a new voice call, DASP+RAS can not service the call if there is no free channel in the BS because channel de-allocation can not be applied for the voice call. For DASV+RAS and DASVP+RAS, channel de-allocation can be used to service the call reducing the probability of voice call blocking. For all traffic conditions, the performance of $P_v$ from the worst to the best are DASP+RAS, DASP, DRA, DASVP, DSAVP+RAS, DSAV and DASV+RAS. For DASP+RAS, $P_v$ increases along with $M$. This is because more channels are used to service data packet as $M$ increases and channel de-allocation can not be applied for voice call arrival. For DASV+RAS and DASVP+RAS, $P_v$ almost does not change when $M$ changes from 2 to 7.

Figures 5.2(a) through 5.2(d) plot the data packet dropping probability $P_g$ as a function of $M$ under different traffic conditions. In terms of $P_g$, one can observe that DASP+RAS and DASVP+RAS always outperform their counterparts DASP and DASVP, respectively. DASV and DASV+RAS exhibit same performance when both voice and data loads are light or heavy. When voice load is heavy and data load is light (Fig. 5.2(b)), DASV+RAS has slightly higher $P_g$ compared to that of DASV. On the other hand, DASV has higher $P_g$ when voice load is light and data load is heavy.
load is heavy (Fig. 5.2(c)). An interesting result shown in Fig. 5.2(c), i.e., when data load is heavy and voice load is light, DASVP+RAS has lower $P_g$ than DASP. The reason behind this is that RAS speeds up the transmission of data packets thus more channels are available for the upcoming new calls in a fixed time interval. Under heavy data load and light voice load these freed channels are mainly used for data packet transmission. As $M$ increases, $P_g$ for DASV+RAS increases because data
service needs more channels and new data packets are more likely to be dropped. For DASP+RAS and DASVP+RAS, $P_g$ decreases as $M$ increases because more channels can de-allocated for data service as $M$ becomes larger.

Fig. 5.3(a) to 5.3(d) plot the average data packet transmission time $T_g$ as a function of $M$ under different traffic conditions. For all traffic conditions, the scheme employing RAS has smaller $T_g$ compared with the corresponding scheme without

85
Figure 5.3: Average data packet transmission time $T_g$ of dynamic resource allocation schemes. (a) $\rho_v = 1$ and $\rho_g = 1$, (b) $\rho_v = 5$ and $\rho_g = 1$, (c) $\rho_v = 1$ and $\rho_g = 5$, (d) $\rho_v = 5$ and $\rho_g = 5$.

RAS (DASV+RAS vs. DASV, DASP+RAS vs. DASP, DASVP+RAS vs. DASVP) because RAS allocates the idle channels to degraded data call and reduce the data packet transmission time. For schemes with channel de-allocation/re-allocation, the $T_g$ performances from the worst to the best are DASVP+RAS, DASP+RAS and DASV+RAS. For all schemes, $T_g$ decreases as $M$ increases because more channels are used to service the data requests.
Figures 5.4(a) though 5.4(d) plot the channel utilization $u$ as a function of $M$ under different traffic conditions. Evidently RAS increases the channel utilization and the schemes with RAS outperform those without RAS. This is due to the fact that RAS attempts to fully utilize channel resources by allocating the idle channels to degraded data calls. Among all the schemes, we can find that DASVP+RAS achieves the best channel utilization among all the schemes under different traffic conditions while DRA has the worst performance in channel utilization. When voice traffic
is heavy and data traffic is light (Fig. 5.4(b)), DASV+RAS has a better channel utilization than DASP+RAS. On the other hand, when data traffic is heavy and voice traffic is light (Fig. 5.4(c)), DASP+RAS outperforms DASV+RAS.

Figures 5.5(a) though 5.5(c) plot the system award factor $Q$ of dynamic resource allocation schemes with DAS/RAS as a function of the weighting factor $\alpha_1$ and data traffic load $\rho_g$ with parameters $\mu_{ch} = 1/180$, $\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$. We only present the results of $M = 4$. For other $M$ values, we observe the similar performance. From Fig. 5.5(a) and 5.5(b), one can conclude that under same traffic loads the system award for DASV+RAS increases while for DASP+RAS it decreases as $\alpha_1$ increases. The reason of that is the migration of the system's service object from the data packet to the voice call as the weight factor $\alpha_1$ increases. Since DASV+RAS has the low voice call blocking probability and high data packet dropping probability, it has higher system award for larger $\alpha_1$. For DASP+RAS with high voice call blocking probability and low data packet dropping probability, the system award is high for small $\alpha_1$. It can be seen that the weighting factor $\alpha_1$ has no effect on DASVP+RAS under the same traffic load. The result is because voice call and data packet have the same blocking probabilities in DASVP+RAS. For all the schemes, the system award decreases when the data traffic load increases because higher traffic loads result in higher voice call and data packet blocking probabilities.

Figure 5.6 shows the system award $Q$ for different dynamic resource allocation schemes with weight factor $\alpha_1 = 0.1, 0.5$ and 0.9. Other parameters are set as $M =$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
4, $\mu_{ch} = 1/180$, $\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$, $\rho_v = 2.5$ and $\rho_g = 2.5$. It can be seen that generally the scheme with RAS has a higher system award than the corresponding
scheme without RAS. Among all the schemes, DASP+RAS has the best system award for $\alpha_1 = 0.1$ while DASV+RAS is the best for $\alpha_1 = 0.9$. This is because the system emphasizes the data packet dropping probability more for small $\alpha_1$ while the voice call dropping probability is more important with large $\alpha_1$. When $\alpha_1 = 0.5$, the system has no preference for voice call or data packet. In this case, DASVP+RAS achieves the highest system award as can be seen from Fig. 5.6. Thus the decision on using which dynamic resource allocation scheme has to be made based on the QoS requirements of system on voice call and data packet. In this way the maximum system award can be obtained.
5.5 Conclusion

In this chapter, we analyze the performance of dynamic resource allocation schemes with channel de-allocation and re-allocation. An analytic model with general data packet channel requirement is derived to evaluate the performance of DASP+RAS, DASVP+RAS and compare them with DRA, DASV, DASP, DASVP and DASV+RAS. Among all the schemes, DASV+RAS yields the lowest $P_v$ while DASP+RAS offers the lowest $P_g$. DASVP+RAS displays the highest channel utilization and also a good balance between $P_v$ and $P_g$ as DASVP. Moreover, DASVP+RAS allows for significantly lower $T_g$ compared with that of DASVP. In terms of the system award $Q$, the schemes with RAS outperform their corresponding schemes without RAS. When the voice call blocking probability is the preferred performance measure of the system (i.e. weighting factor $\alpha_1$ is large), DASV+RAS offers the best system award. On the other hand, DASP+RAS achieves the highest system award if the data packet dropping probability is preferred by the system (i.e. weighting factor $\alpha_1$ is small). When the system has no specific preferences, DASVP+RAS should be the scheme of choice for yielding the best system award.
CHAPTER 6

A NEW DYNAMIC RESOURCE ALLOCATION SCHEME FOR QOS PROVISIONING IN THE INTEGRATED VOICE/DATA WIRELESS NETWORKS

6.1 Introduction

In the integrated voice/data networks, to ensure the required QoS for voice and data, resource allocation scheme has to use the scarce radio resource optimally. Dynamic resource allocation is believed to be a judicious solution for the problem. Many strategies have been proposed in the literature for dynamic resource allocation [8, 9, 10, 22, 23, 29, 40, 41, 42, 55]. In Chapter 4, we proposed two new channel de-allocation schemes for dynamic resource allocation. Furthermore, we took into account channel re-allocation with the proposed channel de-allocation scheme in Chapter 5. The performances of these dynamic resource allocation schemes are analyzed using an analytic model and compared with other schemes including DRA [26], DASV [8] and DASV+RAS [55].

In the above-mentioned studies, the new and handoff voice calls are not differentiated. In the real systems, handoff calls always have a higher priority over new voice calls. This is because termination of a former is more noticeable, hence more annoy-
ing for users than blocking of new calls. In this chapter, we enforce the priority of handoff voice calls and propose a new dynamic resource allocation scheme employing channel reservation, channel de-allocation/re-allocation for voice call (DASV+RAS) and packet queue for QoS provisioning in the integrated voice/data wireless networks.

In the proposed QoS-based scheme, channel reservation is used to lower the forced termination probability of the handoff voice call by reserving certain channels only for handoff voice calls [6, 27, 54]. DSAV+RAS can lessen both the new and handoff voice call blocking probabilities as it is demonstrated in [8] and the previous chapter. Packet queuing is used for reducing the data packet dropping probability [10, 23]. The proposed scheme is targeting the adaptation of QoS requirements of the system on both the new/handoff voice call blocking probability and packet dropping probability which is measured using a system award factor $Q$.

To study the system performance under the proposed scheme, we derive an analytic model. In contrast with the models presented in [8, 9, 22, 23] which for the sake of analytical simplicity assume a specific maximum numbers of channels for data packet transmission, we derive a model with generalized maximum data channel requirement as in [10, 55]. That offers more versatility to the performance analysis.

The rest of this chapter is organized as follows. In the next section, we introduce the dynamic resource allocation scheme with channel reservation, channel de-allocation/re-allocation for voice call and packet queue. The analytical model for evaluating the performance of the proposed scheme is developed in Section 6.3.
Numerical results are presented and discussed in Section 6.4, and followed are the conclusions drawn in Section 6.5.

6.2 QoS-based Dynamic Resource Allocation Scheme

We consider the integrated voice/data wireless network is homogeneous such that we only need to analyze one cell case. Assume the base station (BS) of each cell has $C$ channels shared by the voice calls and data packets. The maximum number of channels for data packet transmission is $M$. A type-$m$ data call is the data packet transmitted using $m (m = 1, \ldots, M)$ channels. Assume the number of free channels in the system is $C_F$ and $C_F = C - n_{vn} - n_{vh} - \sum_{m=1}^{M} mn_{gm}$, where $n_{vn}$ is the number of new voice calls in service, $n_{vh}$ is the number of handoff voice calls, $n_{gm}$ is the number of type-$m$ ongoing data packet transmissions.

The proposed scheme is explained in Figure 6.1. Denote $g$ as the number of channels reserved for handoff voice calls. The remaining $(C - g)$ channels are shared by new/handoff voice calls and data packets. The size of the packet queue is $B$, and the number of data request buffered in the packet queue is denoted as $n_{PQ}$.

The state of the system changes according to the following six events: 1) new voice call arrival, 2) handoff voice call arrival, 3) data packet arrival, 4) new voice call completion, 5) handoff voice call completion and 6) data packet completion. For the arrived new voice call, if $C_F > g$, the call will be served. The handoff voice call will be served if $C_F > 0$. Upon arrival of data packet, $M$ channels are allocated if
$C_F \geq M + g$. If $g < C_F < M + g$, then $(C_F - g)$ channels are allocated to the data call. If $C_F \leq g$ and the number of buffered data calls is less than $B$, the data call will be buffered in the packet queue. Otherwise, the data packet is dropped. For a new arrived voice call, if $C_F \leq g$, channel de-allocation allows one channel from an ongoing type-$m$ ($m > 1$) data call (or degradable data call) to service the arrived voice call. If no degradable data call exists, the arrived new voice call will be blocked. For handoff
voice call, if $C_F = 0$, channel de-allocation will be applied to service the call if there exists a degradable data call in the system. Otherwise, the handoff voice call will be forced to terminate. Note that the channel is de-allocated from the data call with the highest number of channels, e.g. a type-$m$ data call can be degraded if there is no type-$q$ ($q = m + 1, \ldots, M - 1, M$) data call in the BS. Upon the channel release due to the voice call termination or handoff or completion of the packet transmission, if there are degraded calls in the system and no data call buffered in the packet queue, the released channels will be re-allocated to upgrade the transmission of these calls. The re-allocation is performed using the "worst degraded first upgrading" policy [55]. That is, a type-$m$ data call can be upgraded if all type-$q$ ($q = 1, 2, \ldots, m - 1$) data calls have been upgraded to type-$M$. If there are data calls buffered in the packet queue, the released channels are used to service the buffered data call instead of re-allocating for degraded on-going data calls. The algorithm of the proposed scheme is illustrated in Table 6.1.

To demonstrate the performance of the proposed scheme, we compare it with three other dynamic resource allocation schemes described below.

1. Scheme 1 (referred as DRA1) is the same as the DRA proposed in [26];
2. Scheme 2 (referred as DRA2) is the same as the DASV+RAS in [55];
3. Scheme 3 (referred as DRA3) is similar to the proposed scheme except the channel de-allocation/re-allocation for voice call is not applied.
**Case 1: New Voice Call Arrival**

If \( CF > g \)
- Accept the new voice call.

else
  - If (exist a degradable data call)
    - /* Channel de-allocation for voice call */
      - Accept the new voice call;
      - De-allocate one channel from degradable data call to the new voice call;}
  else
    - Reject the new voice call;

**Case 2: Handoff Voice Call Arrival**

If \( CF > 0 \)
- Accept the handoff voice call.

else
  - If (exist a degradable data call)
    - /* Channel de-allocation for voice call */
      - Accept the handoff voice call;
      - De-allocate one channel from degradable data call to the handoff voice call;}
  else
    - Reject the handoff voice call;

**Case 3: Data Packet Arrival**

If \( CF \geq M + g \)
- Accept the data packet;
  - Allocate \( M \) channels to the data packet;

else if \( g < CF < M + g \)
- Accept the data packet;
  - Allocate \( (CF - g) \) channels to the data packet;

else
  - If \( n_{PQ} < B \)
    - Accept the data packet and buffer in the packet queue;
  else
    - Reject the data packet;

**Case 4: New Voice Call Completion**

If \( n_{PQ} > 0 \)
- Serve a data call in the packet queue with one channel;

else
  - If (exist degraded data calls)
    - /* channel re-allocation */
      - one channel is used to upgrade a degraded data call;

**Case 5: Handoff Voice Call Completion**

If \( n_{PQ} > 0 \)
- Serve a data call in the packet queue with one channel;

else
  - If (exist degraded data calls)
    - /* channel re-allocation */
      - one channel is used to upgrade a degraded data call;

**Case 6: New Voice Call Completion**

If \( n_{PQ} > 0 \)
- Serve a data call in the packet queue with \( m \) channels;

else
  - If (exist degraded data calls)
    - /* channel re-allocation */
      - \( m \) channels are used to upgrade degraded data calls;

---

Table 6.1: Algorithm of the proposed dynamic resource allocation scheme

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
6.3 Analytic Model

For the purpose of performance analysis of the proposed dynamic channel allocation scheme, we develop an analytic model. We assume that the new voice call, the handoff voice call and the data packet all follow the Poisson process with arrival rate \( \lambda_{vn} \), \( \lambda_{vh} \) and \( \lambda_{g} \), respectively. The voice call holding time and the cell residence time are assumed to be exponentially distributed with means \( 1/\mu_{ch} \), \( 1/\mu_{cr} \), respectively. Then the channel holding time for voice call is exponentially distributed with the rate \( \mu_{v} = \mu_{ch} + \mu_{cr} \). If one channel is allocated to a data packet, the packet transmission is assumed to follow the exponential distribution with the value of mean as \( 1/\mu_{g} \). The handoff of the data packet transmission is not considered as the reason described in Chapter 4.

The dynamic resource allocation scheme is modeled as a \((M + 3)\)-dimension Markov process. A state in this process is denoted as \( s = (n_{vn}, n_{vh}, n_{g1}, n_{g2}, \ldots, n_{gM-1}, n_{gM}, n_{PQ}) \). The state space \( S \) of the Markov process is given by

\[
S = \left\{ s = (n_{vn}, n_{vh}, n_{g1}, n_{g2}, \ldots, n_{gM-1}, n_{gM}, n_{PQ}) : \right. \\
0 \leq n_{vn} + n_{vh} + \sum_{k=1}^{M} kn_{gk} \leq C - g, 0 \leq n_{vn} \leq C - g, \\
0 \leq n_{vh} \leq C - g, 0 \leq n_{gk} \leq \left\lfloor \frac{C-g}{k} \right\rfloor (k = 1, 2, \ldots, M), n_{PQ} = 0 \\
\bigcup \left\{ s = (n_{vn}, n_{vh}, n_{g1}, n_{g2}, \ldots, n_{gM-1}, n_{gM}, n_{PQ}) : \right. \\
C - g < n_{vn} + n_{vh} + \sum_{k=1}^{M} kn_{gk} \leq C - g, 0 \leq n_{vn} \leq C - g, 0 \leq n_{vh} \leq C, \\
0 \leq n_{gk} \leq \left\lfloor \frac{C}{k} \right\rfloor (k = 1, 2, \ldots, M), n_{PQ} \geq 0 \left. \right\} \right\}
\]

\[(6.1)\]
Denote the steady state probability for state \( s \) as \( \pi_s \). For all states \( s \in S \), \( \sum_{s \in S} \pi_s = 1 \). To find the steady state probability matrix \( \Pi \), we need to obtain the generator matrix \( Q = [q_{s \to s'}]_{s \in S, s' \in S} \), where \( q_{s \to s'} \) \( (s \in S, s' \in S) \) is the transition rate from state \( s = (n_{\text{vn}}, n_{\text{vh}}, n_{g_1}, \ldots, n_{g_{M-1}}, n_{g_M}, n_{pQ}) \) to state \( s' = (n'_{\text{vn}}, n'_{\text{vh}}, n'_{g_1}, \ldots, n'_{g_{M-1}}, n'_{g_M}, n'_{pQ}) \). We use the same definitions of \( I_A, n_{ga}, n_{gq} \) and \( M \) as Chapter 5.

To obtain the transition rate \( q_{s \to s'} \), we consider six cases according to the arrival and completion events of new/handoff voice call and data packet.

**Case 1: New voice call arrival**

(1a) \( C_F > g \)

\[
q_{s \to s'} = I_{C_F > g} \lambda_{\text{vn}}, n'_{\text{vn}} = n_{\text{vn}} + 1, n'_{\text{vh}} = n_{\text{vh}}, n'_{g_k} = n_{g_k}(k \in M), n'_{pQ} = 0 \tag{6.2}
\]

(1b) \( C_F \leq g \) and \( \alpha \geq 2 \)

\[
q_{s \to s'} = I_{C_F \leq g} I_{\alpha \geq 2} \lambda_{\text{vn}}, n'_{\text{vn}} = n_{\text{vn}} + 1, n'_{\text{vh}} = n_{\text{vh}}, n'_{g_{\alpha-1}} = n_{g_{\alpha-1}} + 1, \]
\[
n'_{g_{\alpha}} = n_{g_{\alpha}} - 1, n'_{g_{q}} = n_{g_{q}}(k \in M - \{\alpha - 1, \alpha\}), n'_{pQ} = 0 \tag{6.3}
\]

**Case 2: Handoff voice call arrival**

(2a) \( C_F > 0 \)

\[
q_{s \to s'} = I_{C_F > 0} \lambda_{\text{vh}}, n'_{\text{vn}} = n_{\text{vn}}, n'_{\text{vh}} = n_{\text{vh}} + 1, n'_{g_k} = n_{g_k}(k \in M), n'_{pQ} = 0 \tag{6.4}
\]
(2b) $C_F = 0$ and $\alpha \geq 2$

$$q_{s \rightarrow s'} = I_{C_F=0} I_{\alpha \geq 2} \lambda_{vh}, n'_{vm} = n_{vm}, n'_{vh} = n_{vh} + 1, n'_{g_{\alpha-1}} = n_{g_{\alpha-1}} + 1,$$

$$n'_{g_{\alpha}} = n_{g_{\alpha}} - 1, n'_{g_k} = n_{g_k} (k \in M - \{\alpha - 1, \alpha\}), n'_{PQ} = 0$$  \hspace{1cm} (6.5)

**Case 3: Data packet arrival**

(3a) $C_F \geq M + g$ which means that $M$ channels can be allocated to the data packet transmission.

$$q_{s \rightarrow s'} = I_{C_F \geq M + g} \lambda_{g}, n'_{vn} = n_{vn}, n'_{vh} = n_{vh}, n'_{g_M} = n_{g_M} + 1,$$

$$n'_{g_k} = n_k (k \in M - \{M\}), n'_{PQ} = 0$$  \hspace{1cm} (6.6)

(3b) $g < C_F < M + g$ which means that $(C_F - g)$ channels can be allocated for data transmission.

$$q_{s \rightarrow s'} = I_{g < C_F < M + g} \lambda_{g}, n'_{v} = n_{v}, n'_{vh} = n_{vh}, n'_{g_{C_F - g}} = n_{g_{C_F - g}} + 1,$$

$$n'_{g_k} = n_k (k \in M - \{C_F - g\}), n'_{PQ} = 0$$  \hspace{1cm} (6.7)

(3c) $C_F \leq g$, $B > 0$ and $n_{PQ} < B$, the packet is buffered in the packet queue.

$$q_{s \rightarrow s'} = I_{C_F \leq g} I_{B > 0} I_{n_{PQ} < B} \lambda_{g}, n'_{vn} = n_{vn}, n'_{vh} = n_{vh}, n'_{g_k} = n_{g_k} (k \in M), n'_{PQ} = n_{PQ} + 1$$  \hspace{1cm} (6.8)

**Case 4: New Voice call completion**

(4a) $n_{PQ} > 0$, which means that a data call buffered in packet queue can be served
with one channel.

\[ q_{s\rightarrow s'} = I_{n_{PQ} > 0} n_{vn} \mu_v, n'_{vn} = n_{vn} - 1, n'_{vh} = n_{vh}, n'_{g1} = n_{g1} + 1, \]

\[ n'_{gk} = n_{gk} (k \in M), n'_{PQ} = n_{PQ} - 1 \]

(6.9)

\[ (4b) \ n_{PQ} = 0 \text{ and } \beta > 0 \] which means that no data call is buffered in the packet queue and the released channel can be re-allocated to upgrade a degraded data packet transmission.

\[ q_{s\rightarrow s'} = I_{n_{PQ} = 0} I_{\beta > 0} n_{vn} \mu_v, n'_{vn} = n_{vn} - 1, n'_{vh} = n_{vh}, n'_{g3} = n_{g3} - 1, \]

\[ n'_{g3+1} = n_{g3+1} + 1, n'_{gk} = n_{gk} (k \in M - \{\beta, \beta + 1\}), n'_{PQ} = 0 \]

(6.10)

(4c) \ n_{PQ} = 0 \text{ and } \beta = -1, \] which means that there is no degraded data packet transmission and no data call buffered in packet queue.

\[ q_{s\rightarrow s'} = I_{n_{PQ} = 0} I_{\beta = -1} n_{vn} \mu_v, n'_{vn} = n_{vn} - 1, n'_{vh} = n_{vh}, n'_{gk} = n_{gk} (k \in M), n'_{PQ} = 0 \]

(6.11)

Case 5: Handoff voice call completion

\[ (5a) \ C_F \geq g \text{ and } n_{PQ} > 0 \] which means that the freed channel is not in the range of channels reserved for handoff voice call and a data call buffered in the packet queue can be served with one channel.

\[ q_{s\rightarrow s'} = I_{C_F \geq g} I_{n_{PQ} > 0} n_{vh} \mu_v, n'_{vn} = n_{vn}, n'_{vh} = n_{vh} - 1, n'_{g1} = n_{g1} + 1, \]

\[ n'_{gk} = n_{gk} (k \in M), n'_{PQ} = n_{PQ} - 1 \]

(6.12)
(5b) $C_F \geq g, n_{PQ} = 0$ and $\beta > 0$ which means that the freed channel is not in the range of channels reserved for handoff voice call and no data call is buffered in the packet queue. The released channel can be used to serve a degraded data call in the system.

\[
q_{s \rightarrow s'} = I_{C_F \geq g} I_{n_{PQ} = 0} I_{\beta > 0} n_{v_{h_\mu_v}}, n'_{v_{n}} = n_{v_{n}}, n'_{v_{h}} = n_{v_{h}} - 1, n'_{g_{\beta}} = n_{g_{\beta}} - 1,
\]
\[
n'_{g_{\beta+1}} = n_{g_{\beta+1}} + 1, n'_{g_{k}} = n_{g_{k}}(k \in M - \{\beta, \beta + 1\}), n'_{PQ} = 0
\]  \hspace{1cm} (6.13)

(5c) $C_F \geq g, n_{PQ} = 0$ and $\beta = -1$ which means that the freed channel is not in the range of channels reserved for handoff voice call but there is no degraded data call in the system and no data call buffered in the packet queue.

\[
q_{s \rightarrow s'} = I_{C_F \geq g} I_{n_{PQ} = 0} I_{\beta = -1} n_{v_{h_\mu_v}}, n'_{v_{n}} = n_{v_{n}}, n'_{v_{h}} = n_{v_{h}} - 1, n'_{g_{k}} = n_{g_{k}}(k \in M), n'_{PQ} = 0
\]  \hspace{1cm} (6.14)

(5d) $C_F < g$, which means that the released channel is in the range of channels reserved for handoff voice call that cannot be used for re-allocation or serving the queued data call.

\[
q_{s \rightarrow s'} = I_{C_F < g} n_{v_{h_\mu_v}}, n'_{v_{n}} = n_{v_{n}}, n'_{v_{h}} = n_{v_{h}} - 1, n'_{g_{k}} = n_{g_{k}}(k \in M), n'_{PQ} = n_{PQ}
\]  \hspace{1cm} (6.15)

**Case 6: Type-$m$ data packet completion**

(6a) $n_{PQ} > 0$ which means that the released $m$ channels can be used to serve a data packet.
call buffered in the packet queue.

\[ q_{s \to s'} = I_{n_{PQ} > 0} n_{gm} m \mu_g, n'_{vn} = n_{vn}, n'_{vh} = n_{vh}, n'_{gk} = n_{gk} (k \in M), n'_{PQ} = n_{PQ} - 1 \]  \hfill (6.16)

(6b) \( n_{PQ} = 0 \), which means that the released \( m \) channels can be used to upgrade degraded data calls in the system.

Upon the type-\( m \) data call completion, the new state becomes \( s_1 = (n^*_{vn}, n^*_{vh}, n^*_{g1}, \ldots, n^*_{gm}, \ldots, n^*_{gM}, n^*_{PQ}) \), where \( n^*_{vn} = n_{vn}, n^*_{vh} = n_{vh}, n^*_{gk} = n_{gk} - 1, n^*_{gk} = n_{gk}, k \in M - \{ m \} \), \( n^*_{PQ} = n_{PQ} \). We introduce an index \( \theta \) (1 \( \leq \theta \leq M - 1 \)) such that

\[ \sum_{k=1}^{\theta-1} n^*_{gk} (M - k) \leq m \leq \sum_{k=1}^{\theta} n^*_{gk} (M - k) \]  \hfill (6.17)

Inequality (6.17) implies that when a type-\( m \) data call leaves the BS, all the type-1 to type-(\( \theta - 1 \)) calls and some of the type-\( \theta \) calls can be upgraded to type-\( M \) calls.

If \( \sum_{k=1}^{M-1} n^*_{gk} (M - k) \leq m \) which means that all the type-1 to type-(\( M - 1 \)) calls can be upgraded to type-\( M \) calls, we set \( \theta \) to \( M \). If \( \beta = 0 \) which means that there is no data call in the BS or all the data calls are type-\( M \), we set \( \theta \) to -1.

(4a.1) \( \theta = -1 \)

\[ q_{s \to s'} = I_{\theta = -1} n_{gm} m \mu_g, n'_{vn} = n_{vn}, n'_{vh} = n_{vh}, n'_{gk} = n_{gk} - 1, \]
\[ n'_{gk} = n_{gk} (k \in M - \{ m \}), n'_{PQ} = 0 \]  \hfill (6.18)
\[ (4a.2) \quad 1 \leq \theta \leq M - 1 \]

\[ q_{\theta \to s'} = I_{1 \leq \theta \leq M - 1} n_{gm} m \mu_g, n_{vn} = n_{vn}, n_{vh} = n_{vh}, n_{g_k} = 0 (k = 1, 2, \ldots, \theta - 1), \]

\[ n_{g\theta - \delta - 1} = n_{g\theta + \delta} + 1, \]

\[ n_{gM} = n_{gM} + \sum_{k=1}^{\theta-1} n_{g_k} + \delta, n_{g_k} = n_g (\text{other index } k \text{ in } M), n_{PQ} = 0 \]

where

\[ \delta = \left[ \frac{m - \sum_{k=1}^{\theta-1} n_{g_k} (M - k)}{M - \theta} \right] \]

is the number of type-\( \theta \) calls that can be upgraded to type-\( M \) calls. \( \delta_r = \delta - \delta(M - \theta) \) is the number of channels that can upgrade a type-\( \theta \) call to type-\( \theta + \delta_r \).

\[ (4a.3) \quad \theta = M \]

\[ q_{s \to s'} = I_{\theta = M} n_{gm} m \mu_g, n_{vn} = n_{vn}, n_{vh} = n_{vh}, n_{g_k} = 0 (k \in M - \{M\}), \]

\[ n_{gM} = \sum_{k=1}^{M} n_{g_k} - 1, n_{PQ} = 0 \]

The transition rate \( q_{s \to s} \) can be obtained as

\[ q_{s \to s} = - \sum_{s' \neq s, s' \in S, s' \neq s} q_{s \to s'} \]

From equations (6.2) - (6.16) and (6.18) through (6.21), we can derive the generator matrix \( Q \) for the \((M + 3)\)-dimension Markov chain. To obtain the steady-state probability matrix \( \Pi \), we need to solve the linear equation \( \Pi e = 1 \) and \( \Pi Q = 0 \), where \( e \) is a unary column vector. This is done using a numerical method introduced.
Knowing the steady-state probability \( \pi_s \) of the Markov chain, we can calculate the handoff voice call arrival rate as

\[
\lambda_{vh} = \sum_{s \in S} (n_{vn} + n_{vh}) \pi_s \mu_{cr}
\]  

(6.22)

Since the steady-state probability \( \pi_s \) and the handoff call arrival rate \( \lambda_{vh} \) are mutually related, an iterative algorithm is applied to compute \( \pi_s \) and \( \lambda_{vh} \) as the one used in Chapter 4.

To measure the performance of the proposed scheme, we use the following performance metrics - new voice call blocking probability \( P_{vn} \), handoff voice call forced termination probability \( P_{vh} \), data packet dropping probability \( P_g \) and channel utilization \( u \).

The new voice call will be blocked if the number of free channels \( C_F \leq g \) and the data calls in the BS are all type-1. Then, the new voice call blocking probability \( P_{vn} \) is represented as

\[
P_{vn} = \sum_{n_{PQ}=0}^{B} \sum_{n_{vn} + n_{vh} + n_{g_1} \geq C - g, n_{g_k}=0(k \in \mathbb{M} - \{1\}), s \in S} \pi_s
\]  

(6.23)

The handoff voice call will be forced to terminate if the number of free channels \( C_F = 0 \), and the data calls in the BS are all of type-1. The handoff voice call dropping
probability $P_{vh}$ is obtained as

$$P_{vh} = \sum_{n_p=0}^{B} \sum_{n_{vn}+n_{vh}+n_{g1}=C, \ n_{gk}=0(k\in M-\{1\}), s\in S} \pi_s$$  \hspace{1cm} (6.24)

The data packet will be dropped if the number of free channels $C_F \leq g$ and the packet queue is full. The data packet dropping probability $P_g$ is then represented as

$$P_g = \sum_{n_{vn}+n_{vh}+n_{g1} \geq C-g, \ n_p=B, s\in S} \pi_s$$  \hspace{1cm} (6.25)

The channel utilization $u$ can be expressed as

$$u = \frac{\sum_{s\in S} \left( n_{vn} + n_{vh} + \sum_{k=1}^{M} k n_{gk} \right) \pi_s}{C}$$  \hspace{1cm} (6.26)

It should be noted that the above-mentioned performance metrics are influenced by the reserved channel capacity $g$ and the packet queue size $B$. To measure the QoS of the system, we use a system award $Q$ [10, 14] which is expressed as

$$Q = \alpha(1 - P_{vn}) + \beta(1 - P_{vh}) + \gamma(1 - P_g)$$  \hspace{1cm} (6.27)

where $\alpha$, $\beta$ and $\gamma$ are weighting factors which indicate the contribution of $P_{vn}$, $P_{vh}$ and $P_g$ to the system’s QoS, respectively. Notice that $\alpha + \beta + \gamma = 1$. The weighting factors are determined by the system’s overall revenue and service objectives [21].
larger $Q$ indicates higher performance of the scheme.

6.4 Numerical Results

Based on the derived analytic model, we can evaluate the performance of the proposed dynamic resource allocation scheme. We normalize the parameters $\lambda_{vn}$, $\lambda_g$, $\mu_{cr}$ and $\mu_g$ by $\mu_{ch}$ as done in [22, 23]. The number of channels $C$ in the BS is assumed to be 7.
6.4.1 Effect of Packet Queuing

We first investigate the effect of packet queue size $B$ on the performance of the proposed scheme. Figure 6.2(a) to 6.2(d) plot the data packet dropping probability $P_g$, new voice call blocking probability $P_{vn}$, handoff voice call forced termination probability $P_{vh}$ and channel utilization $u$ versus data traffic load $\rho_g$ for different packet queue size $B$. The other parameters are set as $M = 2$, $g = 1$, $\mu_{ch} = 1/180$, $\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$ and $\rho_v = 2$. From the results, we can observe that all the parameters, i.e., $P_g$, $P_{nv}$, $P_{nh}$ and $u$ increase as the data traffic load becomes heavier. For the packet queue size $B$, larger $B$ results in lower $P_g$ but higher $P_{vn}$ and $P_{vh}$. Also from Figure 6.2(d), we can conclude that the channel utilization increase as $B$ becomes larger.

6.4.2 Effect of Channel Reservation

By reserving some channels only for handoff voice calls, the channel reservation scheme attempts to reduce the probability of their forced termination. Figure 6.3(a) to 6.3(d) shows the effect of the reserved channel number $g$ on the performance under various voice traffic load $\rho_v$. The other parameters are set as $M = 2$, $B = 4$, $\mu_{ch} = 1/180$, $\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$ and $\rho_g = 2$. It can be observed that all the performance metrics increase as voice traffic becomes heavier. The results also demonstrate that as the number of reserved channels increases, data packet dropping probability $P_g$ and new voice call blocking probability $P_{vn}$ are increasing, but handoff voice call forced termination probability $P_{vh}$ is decreasing. Figure 6.3(d) indicates
6.4.3 Effect of $M$

Since the analytic model we developed is for general maximum number of requested data channels $M$, we can investigate the effect of $M$ to the performance of the proposed scheme. Figures 6.4(a) to 6.4(d) plot the effect of $M$ under different data traffic load $\rho_g$. The other parameters are set as $B = 4$, $g = 1$, $\mu_{ch} = 1/180$, the decreasing of channel utilization $u$ as $g$ becomes larger.

Figure 6.3: Effect of channel reservation on different voice traffic load $\rho_v$. (a) $P_g$, (b) $P_{vn}$, (c) $P_{vh}$, (d) $u$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 6.4: Effect of M for different data traffic load $\rho_g$. (a) $P_g$, (b) $P_{vn}$, (c) $P_{vh}$, (d) $u$.

$\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$ and $\rho_v = 2$. The results indicate that larger $M$ leads to higher data packet dropping probability but lower new and handoff voice call blocking probability. This is conditioned by the use of DSAV+RAS in the proposed scheme. Thus higher $M$ will give the voice call more chances to be admitted and data packet is more likely to be dropped. From the results, it also can be observed that $M$ has a little effect on the channel utilization.
6.4.4 Performance Comparison for Different Schemes

Figures 6.5(a) through 6.5(d) compare the performance of the proposed scheme with that of other three schemes under different data traffic load $\rho_g$. The parameters are set as $M = 2$, $B = 4$, $g = 1$, $\mu_{ch} = 1/180$, $\mu_{cr} = 0.2\mu_{ch}$, $\mu_g = 100\mu_{ch}$ and $\rho_{ho} = 2$. Fig. 6.5(a) shows that DRA1 and DRA2 without packet queue capability have a higher data packet dropping probability $P_g$ compared to that of DRA3 and the proposed scheme with packet queue. DRA3 has lower $P_g$ than that of the proposed
scheme because DRA3 does not use channel de-allocation/re-allocation for voice calls. From Fig. 6.5(b), we can see that new voice call blocking probability \( P_{vn} \) is ranging from low to high for DRA2, the proposed scheme, DRA1 and DRA3. By using channel de-allocation/re-allocation for voice call, DRA2 and the proposed scheme achieves lower \( P_{vn} \). Since the proposed scheme employs the packet queue and channel reservation which result in higher \( P_{vn} \), DRA2 outperforms the proposed scheme in terms of \( P_{vn} \). DRA3 has the worst \( P_{vn} \) because it uses the packet queue and channel reservation without channel de-allocation/re-allocation for voice call. Fig 6.5(c) shows that handoff voice call forced termination probability \( P_{vh} \) ranging from low to high is DRA2, the proposed scheme, DRA3 and DRA1. From Fig. 6.5(d), one can observe that the proposed scheme achieves the best channel utilization. Followed are DRA3, DRA2 and DRA1.

### 6.4.5 System Award

We first investigate the parameters \( B \) and \( g \) of the proposed scheme to obtain the best system award under different system requirements. Figure 6.6(a) to 6.6(d) show the system award \( Q \) as a function of \( B \) and \( g \) for different system QoS requirements. The four figures correspond to four cases with varying weight factors for the system award \( Q \): (a) \( \alpha = 0.8, \beta = 0.1 \) and \( \gamma = 0.1 \), (b) \( \alpha = 0.1, \beta = 0.8 \) and \( \gamma = 0.1 \), (c) \( \alpha = 0.1, \beta = 0.1 \) and \( \gamma = 0.8 \) and (d) \( \alpha = 0.4, \beta = 0.3 \) and \( \gamma = 0.3 \). Other parameters are set as \( M = 2, \mu_{ch} = 1/180, \lambda_{vn} = 2\mu_{ch}, \lambda_{g} = 200\mu_{ch}, \mu_{cr} = 0.4\mu_{ch}, \mu_{g} = 100\mu_{ch} \), \( 0 \leq g \leq 5, 0 \leq B \leq 6 \). For Fig. 6.6(a), we can find \( g = 0 \) and \( B = 1 \) produce the
best value of $Q$. This is due to the fact that new voice call blocking probability $P_{vn}$ is the most important factor in the system award $Q$ ($\alpha = 0.8$) and larger $g$ and $B$ will result in higher $P_{vn}$. In Fig. 6.6(b), the best $Q$ is achieved for $g = 1$ and $B = 3$. In this case the system emphasizes the handoff voice call forced termination probability $P_{vh}$ ($\beta = 0.8$) and one channel is reserved only for handoff calls to reduce it. Fig

Figure 6.6: System award $Q$ vs. $B$ and $g$ for (a) $\alpha = 0.8$, $\beta = 0.1$ and $\gamma = 0.1$, (b) $\alpha = 0.1$, $\beta = 0.8$ and $\gamma = 0.1$, (c) $\alpha = 0.1$, $\beta = 0.1$ and $\gamma = 0.8$ and (d) $\alpha = 0.4$, $\beta = 0.3$ and $\gamma = 0.3$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 6.7: System award comparison for different schemes. (a) $\alpha = 0.8$, $\beta = 0.1$ and $\gamma = 0.1$, (b) $\alpha = 0.1$, $\beta = 0.8$ and $\gamma = 0.1$, (c) $\alpha = 0.1$, $\beta = 0.1$ and $\gamma = 0.8$ and (d) $\alpha = 0.4$, $\beta = 0.3$ and $\gamma = 0.3$.

6.6(c) shows the best $Q$ produced when $g = 0$, $B = 6$. In this case the system prefers data packet to voice call ($\gamma = 0.8$) and a larger packet queue size is used to reduce the data packet dropping probability $P_g$. Finally, for the case shown in Fig. 6.6(d), the system has no preferences for new/handoff voice call or data packet, we can find that the best $Q$ is achieved by $g = 0$, $B = 4$.

We then compare the system award $Q$ of the four dynamic resource allocation schemes as shown in Fig. 6.7 for the same four cases as for Fig. 6.6 with the parameters used in deriving results of Fig. 6.6. $Q$ values of DSA3 are obtained using the optimal combination of $g$ and $B$. The results demonstrate that the proposed
scheme always outperforms other three reference schemes because it is furnished by the capability of adjusting the reserved channel number $g$ and packet queue size $B$ to the system's QoS requirements.

6.5 Conclusion

In this chapter, we have proposed a new dynamic resource allocation scheme. It takes into account the system's QoS requirements on the new/handoff voice call and data packet. The scheme combines the channel reservation, channel de-allocation/re-allocation for voice call and packet queue and tries to adapt to different QoS requirements of the system by adjusting the reserved channel number $g$ and packet queue size $B$. We have developed an analytic model with general maximum data channel requirement to evaluate the performance of the proposed scheme and compare it with other dynamic resource allocation schemes. By using the system award $Q$ to measure the QoS of the system under different schemes, we have demonstrated that the proposed scheme outperforms other schemes.
In this dissertation, we focus on the two problems critical for QoS provisioning in wireless/mobile networks: location management and resource management. The major contributions of our work can be summarized in the followings.

- The hybrid location update scheme (HLU) is designed. It takes into account both the moving distance and the moving direction. We analyzed the performance of the HLU scheme based on 2D Markov walk model. Numerical results demonstrate that the proposed HLU scheme can reduce signaling cost significantly.

- By considering the MT’s mobility characteristics in different locations of the service area, we proposed a dynamic location management scheme with personalized location areas designed for each MT. The location area is formed based on the boundary crossing probabilities and the residence time of the MT in the cells. The continuous-time Markov chain (CTMC) model is employed to analyze the location management cost. Due to the computational complexity of forming the location areas with minimum location management cost, we developed a heuristic algorithm for determining the personalized location area.
with sub-optimal cost. Simulation results show that the proposed scheme offers a lower signaling cost than that obtained by some known strategies such as always-update (AU), distance-based location area (DBLA) and static location are (SLA).

- For the integrated voice/data wireless networks, we designed two new channel de-allocation schemes: channel de-allocation for data packet (DASP) and channel de-allocation for both voice call and data packet (DASVP) for dynamic resource allocation. We have derived an analytic model with general data channel requirement to evaluate the performance of the schemes. The numerical results demonstrate that DASP achieves a low data packet dropping probability but a high voice call blocking probability while DASVP offers the best channel utilization and a good balance between voice call and data packet.

- We have studied the performance of the proposed channel de-allocation schemes by employing channel re-allocation. The system award factor $Q$ is used as a QoS requirement to voice call and data packet. An analytic model is derived to evaluate the performance of the schemes. The results indicate that dynamic resource allocation schemes should be chosen according to the QoS requirements of the system.

- By taking into account the handoff voice call forced termination probability, we have proposed a new dynamic resource allocation scheme for QoS provi-
sioning in the integrated voice/data wireless networks. The scheme combines the channel reservation, channel de-allocation/re-allocation for voice call and packet queue and pursues different QoS requirements of the system by adjusting the number of the reserved channels $g$ and the size of the packet queue $B$. We have developed an analytic model with a general data channel requirement to evaluate the performance of the proposed scheme and compare it with other dynamic resource allocation schemes. By using the system award $Q$ to measure the QoS of the system for different schemes, we have demonstrated that the proposed scheme outperforms other schemes and satisfies the QoS requirements of the system.

The future research can be seen in elaborating the following issues.

- Designing a location management scheme based on a 2D Markov walk model for obtaining location areas of optimal shapes with minimum signaling cost for a given mobility pattern and traffic parameters.

- For the dynamic location management scheme with personalized location areas, we have used a heuristic algorithm to form the location area which results in sub-optimal signaling cost. One can try some meta-heuristic methods such as simulated annealing [30], Tabu search [31] or genetic algorithm [11] to attain lower cost.

- For all our works on resource management in the integrated voice/data wireless
networks, we assume the network is homogenous where all cells have equal number of channels, same traffic patterns, and same transition probabilities. In our future work, a heterogeneous network will be considered to analyze the impact of hotspot traffic caused by handoff calls [13].
BIBLIOGRAPHY


120


121


VITA

Graduate College
University of Nevada, Las Vegas

Jun Zheng

Home Address:
4441 Escondido St Apt1205
Las Vegas, Nevada 89119

Degrees:
Bachelor of Science, Electrical Engineering, 1993
Chongqing University, Chongqing, China

Master of Science, Electrical Engineering, 1996
Chongqing University, Chongqing, China

Master of Science in Engineering, Biomedical Engineering, 2001
Wright State University, Dayton, Ohio

Dissertation Title: Location and Resource Management for Quality of Service Provisioning in Wireless/Mobile Networks

Dissertation Examination Committee:
Chairperson, Dr. Emma Regentova, Ph.D.
Committee Member, Dr. Shahram Latifi, Ph.D.
Committee Member, Dr. Henry Selvaraj, Ph. D.
Graduate Faculty Representative, Dr. Kazam Tagava, Ph.D.