A paradigm for delivering a scalable and low-latency immersive voice over IP service to resource-constrained clients in distributed virtual environment

Ying Peng Que

University of Wollongong

Recommended Citation

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
A Paradigm for Delivering a Scalable and Low-Latency Immersive Voice over IP Service to Resource-constrained Clients in Distributed Virtual Environment

A thesis submitted in fulfilment of the requirements for the award of the degree

Doctor of Philosophy

from

THE UNIVERSITY OF WOLLONGONG

by

Ying Peng Que
Bachelor of Engineering (Honours Class I)

SCHOOL OF ELECTRICAL, COMPUTER AND TELECOMMUNICATIONS ENGINEERING
2007
Abstract

This thesis addresses the problem space of developing a high quality, yet scalable multi-party voice communication service to the ever more popular Distributed Virtual Environments (DVE) exemplified by the Multiplayer Online Games (MOG). The social interactive experience of DVE users can be greatly enhanced if the users feel immersed in a realistic environment via high fidelity visual scenes, auditory scenes and haptics. While recognising the primary role played by the visual scenes and the consequent rapid progresses made in the field of computer graphics, in this thesis, we investigate the network delivery of live DVE user voices which will be an important supplement to the primary visual scenes in creating the sense of immersion for the DVE users. The Immersive Voice over IP (VoIP) service is characterised by the creation of an auditory scene for each user which is the personalised mix of all the voices within that user’s hearing range. Each constituent voice stream in an auditory scene is localised (directional placement) and distance-attenuated in accordance with the visual positions of corresponding users in the virtual world. We believe the auditory scenes created for Immersive VoIP will be much more conducive to the creation of users’ sense of immersion than the currently prevalent text-chat and mono VoIP applications.

During our literature review, we first identified three key challenges faced by the Immersive VoIP service provider, i.e., the need to reduce respectively, the voice processing cost, the voice exchange bandwidth cost and the voice transmission latency. In view of the limited resources onboard the DVE clients (especially the wireless clients), the concept of server-rendered Immersive VoIP service is proposed which employs dedicated servers to complete the computationally expensive voice rendering tasks on behalf of the clients. Nevertheless, in exchange for the minimisation of client-side resource loads, a server processing scalability problem is created where each of the Auditory Scene Creation (ASC) servers could bear a prohibitively large processing load when supporting a large and dense user population. In the course of conducting subjective listening tests, we soon realised that the further away a listener is from the voice source, the greater is the voice localisation error which can be tolerated by the listener. This important result verifies the conjecture of distance-governed variable
acceptable localisation error which allows the same localised voice to be shared between nearby listening users. This conjecture allows for significant reductions in voice processing cost through the computational reuse of voice localisation results. In addition to considering the static Auditory Scene Creations, a mechanism known as the Transitional Deviation Reduction Algorithm has also been devised to address the issue that the Immersive VoIP users can be annoyed by more than 5 degrees of angular shifts between the voices localised at successive time instants with respect to the same speaking avatar.

To ensure low voice latency for the latency-critical Immersive VoIP service, we propose a two overlay hops distributed server architecture. In this architecture, the voice transmission path between any pair of communicating avatars is always two overlay hops, i.e., from the speaking client to the assigned ASC server (1st hop) then from the ASC server to the listening client (2nd hop). As shown by our simulations results, this two overlay hops distributed server architecture is capable of significantly improving the voice transmission latency from the central server architecture and the prior distributed server solution devised in (Nguyen, 2006), especially for the challenging, yet realistic DVE scenario where there is a low level of correlations between the distribution of avatars in the DVE virtual world and the distribution of physical clients in the underlying geographically dispersed network.

The objective of our server selection/assignment Linear Programming (LP) formulations is to obtain a balance between improving the voice transmission latency and mitigating the associated increases in the client-side voice upload bandwidth cost and the server-side voice processing cost. In particular, the Bandwidth Constrained Formulation reduces the client-side voice upload cost while still achieving a satisfactory level of latency performance by establishing a few ASC server sites which are very Latency Efficient, i.e., capable of meeting the acceptable latency constraints of a large number of communicating avatar pairs. However, the LP-based server assignment formulations were proven to be NP-hard and computationally expensive if executed in their original centralised format. Consequently, an alternative scheme was devised to divide the virtual world into equal size (measured in the number of avatar pairs enclosed) partitions so as to parallelise and distribute the execution of the server assignment formulations between different partitions.

We examined the impact of avatar mobility in the virtual world on the performance of our distributed server assignment solutions, in particular the solution’s ability to balance...
between improving the voice latency and reducing the associated rises in client-side upload bandwidth cost. In our simulations, we found that the ASC servers need to be reassigned at a regular frequency. Despite the significant scalability improvements made by our virtual world partitioning scheme, the optimised server assignment formulations, especially the complex, yet practical Bandwidth Constrained Formulation is still not scalable enough to meet the execution frequencies required for coping with avatar mobility. To this end, we have derived the Bandwidth Reduced Heuristic which offers much faster execution time than the optimal server assignment mechanism, at the expense of small performance loss in terms of both latency and client-side upload bandwidth cost. By combining our virtual world partitioning scheme with the Bandwidth Reduced Heuristic, we have produced a distributed server assignment solution which will enable the Immersive VoIP service to cope with frequent virtual world mobility for a wide range of DVE scenarios in terms of varying user population sizes and/or distribution densities. The only exception is a very sparse DVE which could require the ASC servers to be reassigned once per 60-70 seconds. Such a fast frequency is unrealistic considering the combined delays of server reassignment and handover. Consequently, a really sparse DVE with low intensity communication flows is probably best served by a simple Peer-to-Peer (P2P) architecture with minimum management overhead rather than the client-server architectures.
Statement of Originality

This is to certify that the work described in this thesis is entirely my own, except where due reference is made the text.
No work in this thesis has been submitted for a degree to any other university or institution.
Signed

Ying Peng Que
25 October, 2008
Acknowledgements

First of all, I would like to sincerely thank my supervisors Professor Farzad Safaei and Dr Paul Boustead whose intelligence, wisdom, perseverance, kindness and devotion to find the truth have inspired me to complete this thesis, and now onto other greater things in life. Over the past four year, I have become a better Engineer and more importantly, a better person, thanks to these two great men whose friendships I shall cherish for the rest of my life. Guys, It is been a true pleasure and privilege to have shared this amazing journey with you and all the ups and downs that went with it.

Secondly, I am extremely grateful to the Smart Internet Cooperative Research Centre which provided generous financial support to this thesis work.

Next, I would like acknowledge the special group of fellow students, past and present from the Smart Internet CRC lab. I regard the three students preceding me as my older brothers (a Chinese Culture term), Cong Duc Nguyen, Vinh Thanh Nguyen and Jeremy Pascal Brun. Cong’s thesis bears close relations to my thesis. Consequently, Cong’s thesis has provided ample inspirations and to some extent paved the way for the completion of my thesis. Vinh, being an excellent technical writer, has produced a fine thesis after which this thesis is modelled in terms of styling, formatting and structural flows. Jeremy’s PhD work is related to the Distributed Server part of this thesis. Jeremy has given valuable insights and suggestions to my thesis work on many occasions. On the other hand, I regard the two current CRC students, Xinbo Jiang and DingLiang as my comrades. Xinbo has been a constant solace of support to me and I am forever indebted to all the in depth discussions we shared concerning PhD and life in general. DingLiang has become a soul mate to me throughout the course of my PhD and will remain so for many years to come. Life will be so different after I depart from the Smart Internet CRC lab and I shall miss dearly all these past and present students.

I would also like to give special thanks to Mr Matthew Baker and Mr Mark O’Dwyer, my brothers in research at the University. I would not have embarked on this amazing
PhD journey without these two mates standing by my side. My acknowledgement also extends to the group of undergrad friends who remained in touch with me throughout my PhD days, in particular, Mr Shaun Maris and Mr Daniel Black. Special thanks also go to my close friends from China, especially, Wei Ying, Qi Zhang, Wei Peng, SiJia Chen. Although being so geographically far away, they held me and my PhD dear in their thoughts. I would like to thank all my church friends from Figtree Anglican Church who brought plenty of spice and much welcomed diversions to the dull moments of my PhD, especially towards the end.

I am very blessed with two very loving parents whose never ending love played a massive part in helping me through this emotionally gruelling rollercoaster otherwise known as the PhD. The same gratitude is also extended to all my grandparents, my surviving grandmother, uncles and aunties (in particular my Aunty ShaoHua). Collectively, these people form a circle of the dearest people to me. Most recently, that family circle has been augmented by Zheng Chen, my other half. Darling, the thought of you and the prospect of being able to spend more time with you after PhD were often the only incentives keeping me going through the final days of thesis writing.

Last but not the least, thank you to Jesus Christ who created and gave everything I have in my life, the PhD included.
Contents

1. INTRODUCTION ....................................................................................................... 1
   1.1 BACKGROUND ........................................................................................................ 1
   1.2 THESIS OBJECTIVES .......................................................................................... 3
   1.3 THESIS OUTLINE ................................................................................................. 5
   1.4 PUBLICATIONS .................................................................................................... 7
   1.5 JOURNALS UNDER REVIEW ............................................................................... 8

2. LITERATURE REVIEW ........................................................................................... 9
   2.1 INTRODUCTION ..................................................................................................... 9
   2.2 THE RISE OF DISTRIBUTED VIRTUAL ENVIRONMENT (DVE) ......................... 10
   2.3 THE DEPLOYMENT AND OPERATION OF DVE ..................................................... 12
       2.3.1. Network Architectures Supporting DVE ....................................................... 13
       2.3.2. Graphical Representation of DVE State Information ....................................... 16
       2.3.3. Maintaining Consistencies in DVE State Information ...................................... 16
   2.4 IMPORTANCE AND COMPOSITION OF DVE AUDITORY SCENES ................... 20
   2.5 SPATIAL AUDIO RENDERING TECHNOLOGIES ............................................... 21
       2.5.1. Fundamental Terminologies ............................................................................. 21
       2.5.2. Three Categories of Audio Spatialisation Techniques ....................................... 21
       2.5.3. Three Levels of Spatial Audio Rendering Fidelity in DVE ................................. 22
       2.5.4. Head Related Transfer Function (HRTF) .......................................................... 23
       2.5.5. Amplitude Panning ......................................................................................... 25
       2.5.6. Comparison between HRTF and Amplitude Panning ...................................... 25
       2.5.7. Reality-equivalence 3-D Audio Rendering Systems ........................................... 27
   2.6 NON-VOICE NETWORKED COMMUNICATION SERVICES FOR DVE .......... 28
   2.7 MONO VOICE OVER IP SERVICES FOR DVE ................................................... 29
   2.8 PRIOR IMMERSIVE VOICE OVER IP SERVICES FOR DVE ......................... 32
       2.8.1. Immersive Voice Conferencing Services ......................................................... 33
       2.8.2. Application Layer Multicast Schemes for Immersive VoIP ............................... 34
       2.8.3. Voice Processing Cost Reductions for Immersive VoIP ...................................... 36
       2.8.4. Voice Latency Reductions for Immersive VoIP .................................................. 41
   2.9 CONCLUDING REMARKS ................................................................................... 45

3. OVERVIEW OF THE IMMERSIVE VOICE OVER IP SERVICE ......................... 48
   3.1 INTRODUCTION .................................................................................................... 48
   3.2 THE CONCEPT OF IMMERSIVE VOIP SERVICE ................................................. 49
       3.2.1. Auditory Scene Creation for Immersive VoIP .................................................... 51
       3.2.2. Architecture Integration of Immersive VoIP with the underlying DVE ................ 52
   3.3 MOTIVATION FOR THIS THESIS WORK ........................................................... 54
       3.3.1. The Business Case for the Immersive VoIP Service ........................................... 54
       3.3.2. The Resource Constraints Imposed on Immersive VoIP by the DVE Clients .... 55
       3.3.3. The Properties of Different Scenarios of DVE ................................................. 56
   3.4 PROVISIONING FRAMEWORK FOR IMMERSIVE VOIP ................................. 58
   3.5 EXEMPLIFICATION OF DVE USING MULTIPLAYER ONLINE GAMES ................. 59
3.6 FRAMEWORK FOR PERFORMANCE EVALUATION

3.6.1. The Virtual World Model Applied

3.6.2. The Physical World Model Applied

3.6.3. Metric Definitions for the Three Key Constraints

4. COST REDUCTION AND RESOURCE DIMENSIONING IN VOICE PROCESSING

4.1 INTRODUCTION

4.2 SERVER-RENDERED IMMERSIVE VOIP SERVICE

4.2.1. The Brute-force Approach to Voice Localisations

4.2.2. The Voice Processing Sharing Scheme

4.3 SUBJECTIVE LISTENING TEST RESULTS

4.3.1. Prior Psychoacoustic Results

4.3.2. Assessment Method

4.3.3. Test Subjects

4.3.4. Test Setup

4.3.5. Test 1 and Test 2

4.3.6. Test 3

4.4 VOICE PROCESSING COST MINIMISATION FORMULATIONS

4.4.1. Single Filter Processing Minimisation Formulation

4.4.2. Cross-quadrant Reuse Processing Minimisation Formulation

4.4.3. Two Filters Processing Minimisation Formulation

4.5 SIMULATION RESULTS ON VOICE PROCESSING REDUCTION

4.5.1. Effect of Varying the Boundary Values of Acceptable Localisation Errors

4.5.2. Effect of Variations in Avatar Population

4.5.3. Additional Processing Reduction by Cross-quadrant Reuse Formulation

4.5.4. Additional Processing Reduction by Two Filters Formulation

4.6 PROCESSING RESOURCE DIMENSIONING FOR IMMERSIVE VOIP SERVICE

4.6.1. Unacceptable Localisation Error Minimisation Formulation

4.6.2. Two Heuristics for the Problem of Unacceptable Localisation Error Minimisation

4.6.3. Comparisons to Other Alternative Unacceptable Localisation Error Metrics and Objective Functions

4.7 SIMULATION RESULTS ON VOICE PROCESSING RESOURCE DIMENSIONING

4.7.1. Effect of Increasing Communication Density with Fixed Processing Capacities

4.7.2. Effect of Increasing Number of Avatars with Fixed Processing Capacities

4.7.3. Effect of Decreasing Processing Capacities at Given Communication Densities

4.7.4. Effect of Decreasing Processing Capacities for Given Numbers of Avatars

4.7.5. Performance and Scalability Comparisons between the Heuristics and the Optimum

4.8 CONCLUSIONS

5. IMPROVING VOICE TRANSMISSION LATENCY WHILE REDUCING THE ASSOCIATED RESOURCE COSTS IN THE IMMERSIVE VOICE OVER IP SERVICE

5.1 INTRODUCTION

5.2 THE TWO HOPS DISTRIBUTED SERVER ARCHITECTURE

5.3 THE CONJECTURE OF DISTANCE-GOVERNED VARIABLE VOICE LATENCY CONSTRAINTS

5.3.1. Subjective Listening Test Results

5.3.2. Choosing the Range of Acceptable Latency Constraints

5.4 TWO HOPS DISTRIBUTED SERVER ASSIGNMENT FORMULATIONS

5.4.1. Latency Constrained Server Assignment Formulation

5.4.2. Bandwidth Constrained Server Assignment Formulation

5.4.3. The Implications of the Two Server Assignment Formulations

5.4.4. Server Number Constrained Formulation
5.5 Simulation Results on the Performance of the Two Hops Distributed Server

5.5.1. Effect of Variations in Avatar Population.................................................................139
5.5.2. Performance Comparison between the two Distributed Server Assignment Formulations

5.5.3. Provisioning Cost Comparison between different DVE Scenarios..........................142
5.5.4. Server Assignment Objective Comparison...............................................................144
5.5.5. Investigating the Quality of Server Assignment Solutions at Different Resolutions ....146
5.5.6. Comparison between Variable Latency Constraints and Single Fixed Latency Constraints

5.5.7. The Effect of Varying the Boundary Values of the Acceptable Latency Constraints ....152
5.5.8. Effect of Varying the Number of Potential Server Sites...........................................154
5.5.9. Effect of Varying the Number of Servers Established..............................................157
5.5.10. Effect of Varying the Geographical Spread of Physical Network.............................158
5.5.11. Voice Processing Reductions over the Two Hops Distributed Server Architecture....160

6. Impact of Virtual World Mobility on the Network Delivery of Immersive VoIP Service

6.1 Introduction...................................................................................................................188
6.2 Mobility Model and General Simulation Settings.........................................................189
6.3 Continuous Maintenance of Auditory Scenes in Support of Virtual World Mobility

6.3.1. The Need for Transitional Deviation Reduction.........................................................190
6.3.2. Transitional Deviation Reduction Algorithm............................................................193
6.3.3. Scalability of the Transitional Deviation Reduction Algorithm...............................198
6.3.4. The Impact of Virtual World Mobility on Transitional Deviation Reduction ..........199

6.4 Impact of Virtual World Mobility on Distributed Server Reassignments.................203
6.4.1. Impact of Avatar Mobility on the Distribution of Avatars........................................204
6.4.2. The Impact of Avatar Mobility on the Optimised Server Assignment Solutions.........205
6.4.3. Deriving Near-Optimal but Efficient Server Assignment Heuristics.........................213
6.4.4. Performance Study of the Bandwidth Reduced Server Assignment Heuristic...........220
6.4.5. Scalability Improvement by the Bandwidth Reduced Server Assignment Heuristic....224

6.5 Related Work...............................................................................................................225

6.6 Conclusions..................................................................................................................227

7. Delivering the Immersive VoIP Service Using a Peer-to-Peer Clustering Architecture

7.1 Introduction...................................................................................................................229
7.2 Overview of the Peer Clustering Mechanism...............................................................230
7.2.1. The Two-stage Peer Clustering Process.................................................................231
7.2.2. Inter-cluster and Intra-cluster Exchanges of Voice Streams and Voice Localisations .232

7.3 Inter-cluster Routing Algorithms .................................................................................235
7.3.1. Minimum Delay Routing Algorithm........................................................................236
7.3.2. Fair Nodal Stress Routing Algorithm....................................................................237

7.4 Mathematical Formulations for the Two-stage Peer Clustering Mechanism.............239
7.4.1. Initial Cluster Formation Formulation....................................................................239
7.4.2. Final Cluster Formation Formulation ................................................................. 242
7.5 GREEDY HEURISTIC FOR THE TWO-STAGED PEER CLUSTERING PROBLEM ............ 245
  7.5.1. Motivation for the greedy clustering heuristic ..................................................... 245
  7.5.2. Heuristic Objective ............................................................................................ 246
  7.5.3. Detailed Algorithm Description ......................................................................... 247
  7.5.4. Heuristic Performance Evaluations .................................................................... 248
7.6 SIMULATION RESULTS ON CLUSTER PERFORMANCE .............................................. 250
  7.6.1. Average Bandwidth Cost Reductions .................................................................. 251
  7.6.2. The Distribution of Peer Bandwidth Costs .......................................................... 253
  7.6.3. Voice Transmission Latency .............................................................................. 255
  7.6.4. Voice Localisation (directional Placement) Error ................................................. 257
  7.6.5. Distance Attenuation Error ................................................................................ 257
7.7 THE IMPACT OF VIRTUAL WORLD MOBILITY ON CLUSTERING PERFORMANCE ............. 258
7.8 P2P VERSUS CLIENT-SERVER FOR THE NETWORK DELIVERY OF IMMERSIVE VOIP SERVICE ............................................................. 260
7.9 RELATED WORK ....................................................................................................... 261
7.10 CONCLUSIONS ...................................................................................................... 262

8. CONCLUSIONS ......................................................................................................... 264
  8.1 SUMMARY OF CONTRIBUTIONS AND RECOMMENDATIONS ..................................... 264
  8.1.1. The Concept of Server-rendered Immersive VoIP service and the Reduction of Server-Side Voice Processing Cost ........................................................................................................ 264
  8.1.2. Improving the Voice Transmission Latency with the Two_Overlay_Hops Distributed Server Architecture ........................................................................................................... 265
  8.1.3. The Impact of Virtual World Mobility on the Network Provisioning of the Immersive VoIP service ............................................................................................................................ 267
  8.1.4. Exploring the Peer-to-Peer architecture for supporting the Immersive VoIP Service ........................................................................................................................................ 268
  8.2 FUTURE WORK ...................................................................................................... 269

BIBLIOGRAPHY ................................................................................................................. 271

9. A CASE STUDY ON VOICE LOCALISATIONS USING HEAD RELATED TRANSFER FUNCTIONS ..................................................................................................................... 293
  A.1 DESCRIPTION OF THE LOCALISATION PROCESS ..................................................... 293
  A.2 RESULTS ON REAL-TIME SCALABILITY .................................................................. 294
  A.3 THE LUMPED VOICE DECODER PROCESSING DELAY ............................................... 296

10. PSEUDO-CODE OF THE GREEDY HEURISTIC FOR THE UNACCEPTABLE LOCALISATION ERROR MINIMISATION PROBLEM ......................................................... 298

11. DECOMPOSED UNACCEPTABLE LOCALISATION ERROR MINIMISATION FORMULATION ................................................................................................................. 301

12. PSEUDO-CODE FOR THE PROXIMITY-GOVERNED VIRTUAL WORLD PARTITION ALGORITHM .......................................................................................................................... 302

13. PSEUDO-CODE FOR THE GREEDY HEURISTIC TO THE PROXIMITY-GOVERNED PEER CLUSTERING PROBLEM ........................................................................................................ 306
List of Figures

Figure 2.1 The Logical Component Diagram of DVE.................................................... 10
Figure 2.2 A Screenshot of World of Warcraft............................................................... 11
Figure 2.3 The Functional Component Diagram of DVE............................................... 12
Figure 2.4 Network Architectures Applied to Support DVE.......................................... 14
Figure 2.5 Data Flow of a HRTF-based 3-D audio localisation system, Adapted from (Begault, 1994). ...................................................................................................... 23
Figure 2.6 The Perceptual Culling and Clustering Approach to Cost Amortised 3-D Audio Rendering Pipeline, Adapted from (Tsingos et al., 2004). ....................... 37
Figure 2.7 The DICE Distance-weighted Clustering Approach to Cost Amortised 3-D audio rendering pipeline, adapted from (Boustead et al., 2005) .................... 39
Figure 3.1 Communication Zones of different avatars .................................................. 49
Figure 3.2 Auditory Scene Creation for Immersive VoIP service. ................................. 51
Figure 3.3 Architecture Integration of the Immersive VoIP service with the underlying DVE. ....................................................................................................................... 52
Figure 3.4 Provisioning Framework for the Immersive VoIP service............................ 58
Figure 3.5 Cluster Avatar Distribution Model ............................................................ 60
Figure 3.6 Uniform Avatar Distribution Model............................................................ 60
Figure 4.1 The brute force approach to voice localisations ............................................ 68
Figure 4.2 Amortising the cost of voice localisations ..................................................... 69
Figure 4.3 The computational sharing of a localised voice between two listening avatars. ....................................................................................................................... 70
Figure 4.4 The linear model of relaxing the acceptable_localisation_errors with distance. .......................................................... 72
Figure 4.5 Subjective assessment of the impact of localisation errors introduced to the sole speaker ...................................................................................................... 78
Figure 4.6 Subjective assessment of the impact of localisation errors introduced to the secondary speaker. ................................................................. 78
Figure 4.7 The 2-D polar coordinates used for HRTF localisations............................... 79
Figure 4.8 The effect of relaxing the acceptable_localisation_errors on voice processing reductions ......................................................................................... 92
Figure 4.9 The effect of varying communication_densities on voice processing reductions. ................................................................................................. 94
Figure 4.10 Comparison in voice processing reductions between the two avatar distribution models................................................................. 95
Figure 4.11 Comparison in voice processing reductions between the Single Filter and the Cross-quadrant Reuse Formulations ..................................................................... 96
Figure 4.12 The effect of varying the interactive zone radii on the performance of the Two Filters Formulation .......................................................................................... 97
Figure 4.13 The performance comparison between the two localisation error metrics at the average communication_density of 30 avatars per hearing range. ..... 108
Figure 4.14 The effect of varying communication_densities on unacceptable_localisation_errors with the same voice processing capacity constraint................................................................................. 110
Figure 4.15 The effect of varying the number of avatars on unacceptable_localisation_errors with the same processing capacity constraint ................................................................. 112
Figure 4.16 The effect of varying processing capacity constraints on the unacceptable_localisation_errors incurred at given communication_densities ................................................................. 113
Figure 4.17 The effect of varying processing capacity constraints on unacceptable_localisation_errors at given numbers of avatars ................................................................. 115
Figure 5.1 The two_overlay_hops distributed server architecture .................................................. 121
Figure 5.2 The linear model of relaxing acceptable_latency_constraints ................................................................. 124
Figure 5.3 Subjective assessment of the impact of varying acceptable_latency_constraints on the sole speaker .................................................................................................................. 126
Figure 5.4 The various sources of delays within the end-to-end voice flow in the immersive_VoIP service ................................................................................................................................. 127
Figure 5.5 The latency improvements and the associated rise in client-side voice upload costs when increasing the number of proxies .................................................................................................................. 128
Figure 5.6 The percentage of avatar pairs with violated latency constraints over different communication_densities ................................................................................................................................. 141
Figure 5.7 The mean above zero latency_constraint_deviations over different communication_densities ................................................................................................................................. 141
Figure 5.8 The voice transmission latency and resource consumption cost comparisons between the four DVE Scenarios ................................................................................................................................. 143
Figure 5.9 Latency_Constrained_Server_Assignment_Solution to 10 locales in the virtual world ................................................................................................................................. 149
Figure 5.10 Latency_Constrained_Server_Assignment_Solution to individual avatars in the virtual world ................................................................................................................................. 150
Figure 5.11 Bandwidth_Constrained_Server_Assignment_Solution to individual avatars in the virtual world, at 50% reduction in total client-side upload capacity from the upper bound set by the Latency_Constrained_Formulation ................................................................................................................................. 150
Figure 5.12 Variations in Latency Penalty Metric 1 over different boundary values of acceptable_latency_constraints ................................................................................................................................. 154
Figure 5.13 The percentage of avatar pairs with latency constraints violated over different numbers of potential ASC server sites ................................................................................................................................. 156
Figure 5.14 Percentage of avatar pairs with their latency constraints violated over different levels of geographical spread of POPs ................................................................................................................................. 160
Figure 5.15 Comparison in percentage reductions in voice processing cost between different server assignment solutions ................................................................................................................................. 161
Figure 5.16 The separation of listening avatars in an audible range between different ASC servers ................................................................................................................................. 163
Figure 5.17 The difference between centralised ASC server assignment to the whole DVE and distributed ASC server assignments to different partitions of the DVE ................................................................................................................................. 165
Figure 5.18 The distribution of avatar pairs among different partitions when dividing the DVE virtual world into square grids, comparison between the uniform and the clustered avatar distributions ................................................................................................................................. 172
Figure 5.19 The distribution of avatar pairs among different partitions, comparison between limiting the partition size by the number of avatars and limiting the partition size by the number of avatar pairs ................................................................................................................................. 173
Figure 5.20 The distribution of communicating avatar pairs across different ASC servers chosen from 15 potential sites ........................................................................................................179
Figure 5.21 The distribution of communicating avatar pairs across different ASC servers chosen from 45 potential sites .................................................................................. 180
Figure 5.22 The distribution of communicating avatar pairs across different ASC servers chosen from 45 potential sites .................................................................................. 185
Figure 6.1 The Transitional Deviations experienced by two listening avatars. ................. 191
Figure 6.2 Percentage of avatar pairs having localisation error thresholds changed due to virtual world mobility. ........................................................................................................... 192
Figure 6.3 Percentage of avatar pairs separated out communication due to virtual world mobility. .......................................................................................................................... 193
Figure 6.4 Execution time comparisons between the Single Filter Voice Processing Minimisation Formulation and Transitional Deviation Minimisation Formulation, ......................................................................................................................................................... 199
Figure 6.5 Percentage of avatar pairs affected by Transitional Deviation at the low communication_density of 5 avatars per audible range ..................................................... 200
Figure 6.6 Mean Transitional Deviation at the low communication_density of 5 avatars per audible range .............................................................................................................. 201
Figure 6.7 Percentage of avatar pairs affected by Transitional Deviation at the high average communication_density of 25 avatars per audible range .................. 202
Figure 6.8 Mean Transitional Deviation at the high average communication_density of 25 avatars per audible range ....................................................................................... 203
Figure 6.9 Percentage of avatar pairs moving out of each other’s communication range from 0 to 1000 seconds. .............................................................................................. 205
Figure 6.10 Additional Metric 1 Latency Penalties caused by Avatar Mobility when applying the Latency Constrained Server Assignment Formulation ......................... 207
Figure 6.11 Additional Metric 2 Latency Penalties caused by Avatar Mobility when applying the Latency Constrained Server Assignment Formulation ......................... 208
Figure 6.12 Additional Metric 1 Latency Penalties caused by Avatar Mobility when applying the Bandwidth Constrained Server Assignment Formulation to achieve a 50% reduction in the total client-side voice upload cost from the Latency Constrained Formulation .............................................................................................................. 209
Figure 6.13 Additional Metric 2 Latency Penalties caused by Avatar Mobility when applying the Bandwidth Constrained Server Assignment Formulation to achieve a 50% reduction in the total client-side voice upload cost from the Latency Constrained Formulation .............................................................................................................. 209
Figure 6.14 Additional Metric 1 Latency Penalties caused by Avatar Mobility when applying the Bandwidth Constrained Server Assignment Formulation to achieve a 60% reduction in the total client-side voice upload cost from the Latency Constrained Formulation .............................................................................................................. 210
Figure 6.15 Additional Metric 2 Latency Penalties caused by Avatar Mobility when applying the Bandwidth Constrained Server Assignment Formulation to achieve a 60% reduction in the total client-side voice upload cost from the Latency Constrained Formulation .............................................................................................................. 210
Figure 6.16 Additional Metric 1 Latency Penalties caused by Avatar Mobility when applying the Bandwidth Constrained Server Assignment Formulation to achieve a 70% reduction in the total client-side voice upload cost from the Latency Constrained Formulation .............................................................................................................. 211
Figure 6.17 Additional Metric 2 Latency Penalties caused by Avatar Mobility when applying the Bandwidth Constrained Server Assignment Formulation to achieve a...
70% reduction in the total client-side voice upload cost from the Latency
Constrained Formulation.................................................................................211
Figure 6.18 Iterative process of applying the Bandwidth Reduced Server Assignment
Heuristic to find a sweet spot between improving voice transmission latency and
reducing the associated client-side voice upload cost........................................221
Figure 6.19 The mean additional Metric 1 Latency Penalties incurred by the Bandwidth
Reduced Heuristic over the optimised Bandwidth Constrained Formulation......222
Figure 6.20 The mean additional Metric 2 Latency Penalties incurred by the Bandwidth
Reduced Heuristic over the optimised Bandwidth Constrained Formulation......223
Figure 6.21 The mean additional voice uploads per client incurred by the Bandwidth
Reduced Heuristic over the optimised Bandwidth Constrained Formulation......223
Figure 7.1 Voice exchanges within a brute-force full-mesh P2P network. ............230
Figure 7.2 Inter-cluster and intra-cluster voice exchanges when performing Minimum
Delay Routing of voice packets. ........................................................................233
Figure 7.3 Inter-cluster and intra-cluster voice exchanges when performing Fair Nodal
Stress Routing of voice packets. ........................................................................239
Figure 7.4 Contrasting the different degrees of overlapping of Communication Zones.
...............................................................................................................................246
Figure 7.5 Average number of duplex voice exchanges per peer over different average
communication_densities....................................................................................252
Figure 7.6 The CDF of Voice Uploads at the average communication_density of
30 peers per hearing range when performing Minimum Delay Routing of voice
packets....................................................................................................................254
Figure 7.7 The CDF of Voice Downloads at the average communication_density of
30 peers per hearing range when performing Minimum Delay Routing of voice
packets....................................................................................................................255
Figure 7.8 The CDF of Voice Uploads at the average communication_density of
30 peers per hearing range when performing Fair Nodal Stress Routing of voice
packets....................................................................................................................255
Figure 7.8 The Impact of virtual world mobility on the peer bandwidth cost incurred by
the peers, the pair-wise distance constraint being set to 15 meters.................259
List of Tables

Table 3.1 The Properties of likely DVE Scenarios in the Virtual and Physical Worlds. ................................................................. 56
Table 4.1 Mean Opinion Score Chart for the Subjective Listening Test ................................................................. 74
Table 4.2 Parameters and Variables for the Single Filter Formulation .................................................................................. 81
Table 4.3 Parameters and Variables for the Cross-quadrant Reuse Formulation ............................................................... 87
Table 4.4 Parameters and Variables for the Two Filters Formulation .................................................................................. 89
Table 4.5 Parameters and Variables for the Unacceptable Localisation Error Minimisation Formulation ................................................................. 100
Table 4.6 Performance Comparisons between the two Objective Functions ........................................................................ 107
Table 4.7 Performance comparisons between the Optimum and Heuristics ........................................................................ 117
Table 5.1 Parameters and Variables for the Latency Constrained Formulation ............................................................... 130
Table 5.2 Comparison in Latency Penalty incurred between the two Server Assignment Objectives ................................................................................................................................. 145
Table 5.3 Comparison in terms of Latency Penalty Metric 1 between Servers Assigned to individual Avatars and Servers assigned to Locales ............................................................................. 146
Table 5.4 Comparison in terms of Latency Penalty Metric 2 between Servers Assigned to individual Avatars and Servers assigned to Locales ............................................................................. 147
Table 5.5 Comparison in the average number of voice streams uploaded per avatar between Servers Assigned to individual Avatars and Servers Assigned to Locales ................................................................................................................................. 148
Table 5.6 Comparison in Latency Penalty incurred between variable and fixed Latency Constraints ................................................................................................................................. 152
Table 5.7 The effect of increasing the number of ASC servers established .................................................................................. 157
Table 5.8 The percentages of cross-partition avatar pairs in different scenarios of virtual world partitioning ................................................................................................................................. 174
Table 5.9 The Latency Penalty and Average Client-side Voice Upload Cost incurred in different scenarios of Virtual World Partitioning ................................................................................................................................. 175
Table 6.1 Parameters and Variables for the Transitional Deviation Minimisation Formulation ................................................................................................................................. 195
Table 6.2 A summary of the recommended minimum frequencies of server reassignments for coping with avatar mobility in the virtual world ................................................................................................................................. 213
Table 6.3 Execution Times of the two Server Assignment Formulations .................................................................................. 215
Table 6.4 Comparisons in Execution Time between the optimised Bandwidth Constrained Server Assignment Formulation and the near-optimal Bandwidth Reduced Heuristic ................................................................................................................................. 225
Table 7.1 The network distance between different peers, measured in milliseconds (ms). ................................................................................................................................. 233
Table 7.2 An example comparing the two routing algorithms in their respective distributions of inter-cluster voice uploads and downloads ................................................................................................................................. 236
Table 7.3 Parameters and Variables for the Initial Cluster Formation Formulation ................................................................................................................................. 240
Table 7.4 Parameters and Variables for the Final Cluster Formation Formulation ................................................................................................................................. 243
Table 7.5 The performance comparisons between the greedy heuristic and the two-staged clustering solution using LP formulations ................................................................................................................................. 249
Table 7.6 The execution time comparisons between the greedy heuristic and the two-staged clustering solution using LP formulations ................................................................................................................................. 250
Table 7.7 The average percentage increase in delay caused by applying Minimum Delay Routing over the brute-force Full-mesh P2P case.................................256
Table 7.8 Additional Percentage Increase in the latency caused by applying Fair Nodal Stress Routing on top of the Minimum Delay Routing results .........................256
Table 7.9 The Impact of Virtual World mobility on the pair-wise distance between Peers, at the average communication_density of 30 peers per hearing range. .................258
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/D</td>
<td>Analogue to Digital</td>
</tr>
<tr>
<td>APU</td>
<td>Audio Processing Unit</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>AS</td>
<td>Autonomous System</td>
</tr>
<tr>
<td>ASC</td>
<td>Auditory Scene Creation</td>
</tr>
<tr>
<td>BCC</td>
<td>Binaural Cue Coding</td>
</tr>
<tr>
<td>BGP</td>
<td>Border Gateway Protocol</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CoA</td>
<td>Care of Address</td>
</tr>
<tr>
<td>CS</td>
<td>Control Server</td>
</tr>
<tr>
<td>CVE</td>
<td>Collaborative Virtual Environment</td>
</tr>
<tr>
<td>DICE</td>
<td>Dense Immersive Communication Environment</td>
</tr>
<tr>
<td>ADPCM</td>
<td>Adaptive Differential Pulse Code Modulation</td>
</tr>
<tr>
<td>DPM</td>
<td>Distributed Partial Mixing</td>
</tr>
<tr>
<td>DSP</td>
<td>Digital Signal Processing (Processor)</td>
</tr>
<tr>
<td>DTX</td>
<td>Discontinuous Transmission</td>
</tr>
<tr>
<td>D/A</td>
<td>Digital to Analogue</td>
</tr>
<tr>
<td>DVE</td>
<td>Distributed Virtual Environment</td>
</tr>
<tr>
<td>EGP</td>
<td>Exterior Gateway Protocol</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FPS</td>
<td>First Person Shooter</td>
</tr>
<tr>
<td>GT-ITM</td>
<td>Georgia Tech Internet Topology Model</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphics Processing Unit</td>
</tr>
<tr>
<td>ICC</td>
<td>Inter-channel Correlation Cue</td>
</tr>
<tr>
<td>ICE</td>
<td>Immersive Communication Environment</td>
</tr>
<tr>
<td>IGP</td>
<td>Interior Gateway Protocol</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Provider</td>
</tr>
<tr>
<td>HRTF</td>
<td>Head Related Transfer Function</td>
</tr>
<tr>
<td>ILD</td>
<td>Interaural Time Difference</td>
</tr>
</tbody>
</table>