

Performance Enhancements for Single Hop and Multi-Hop Mesched High Data Rate Wireless Personal Area Networks

A thesis submitted for the degree of Doctor of Philosophy

by

Sahibzada Ali Mahmud

School of Engineering and Design, Brunel University

November 2009

Abstract

The High Data Rate (HDR) Wireless Personal Area Networks (WPANs) typically have a limited operating range and are intended to support demanding multi-media applications at high data rates. In order to extend the communication range, HDR WPANs can operate in a wireless mesh configuration (i.e. enable multiple WPAN clusters) to communicate in a multi-hop fashion. HDR WPANs face several research challenges and some of the open key issues are limited capacity, optimum resource allocation to requesting devices and maintaining Quality of Service (QoS) for real time multimedia flows. Although, there have been some scheduling algorithms proposed for HDR WPANs, the main objective is to maintain the QoS in most cases whereas efficient and fair utilization of network capacity is still largely open for research.

This thesis mainly intends to resolve the issues related to capacity of HDR WPANs such as admission control, fair allocation of Channel Time Allocations (CTAs), improvement in capacity through transmission power control, and efficient utilization of time by each flow. A technique which re-orders the time slots to reduce queuing delay for meshed WPANs is also proposed and evaluated.

The first contribution aims to improve peer-to-peer connectivity in case of two or more independent piconet devices by proposing an inter-PAN communication framework that is augmented by an admission control strategy to handle the cases when the superframe capacity is congested. The queued devices are prioritized by proposing a parameter called the Rejection Ratio. The second contribution consists of a resource allocation framework for meshed WPANs. The main objectives are to reduce the control traffic due to high volume of channel time reservation requests and introduce an element of fairness in the channel time allocated to requesting devices. The objectives are achieved by using traffic prediction techniques and an estimated backoff procedure to reduce control traffic, and define different policies based on offered traffic for fair allocation of channel time. The centralized scheme uses traffic prediction techniques to use the proposed concept of bulk reservations. Based on the bulk reservations and resource allocation policies, the overall overhead is reduced while an element of fairness is shown to be maintained for certain scenarios. In the third contribution, the concepts of Time Efficiency and CTA switching are introduced to improve communication efficiency and utilization of superframe capacity in meshed WPANs. Two metrics known as Switched Time Slot (STS) and Switched Time Slot with Re-ordering (STS-R) are proposed which aim to achieve the purpose. The final contribution proposes and evaluates a technique called CTA overlappnig to improve capacity in single hop and meshed WPANs using tramission power control. Extensive simulation studies are performed to analyze and to evaluate the proposed techniques. Simulation results demonstrate significant improvements in meshed WPANs performance in terms of capacity utilization, improvement in fairness index for CTA allocation by upto 62% in some cases, reduction in control traffic overhead by upto 70% and reduction in delay for real time flows by more than 10% in some cases.

Acknowledgements

First and foremost, I would like to thank the Creator without the blessings and mercy of whom I would not have made it.

My parents have always been a source of inspiration and courage for me and I can never thank them or repay them enough for the virtues they have given me and their countless efforts and sacrifices that led me to this stage.

I would equally like to commend the efforts of my supervisor Prof. Hamed Al- Raweshidy who has not only dealt with me professionally during my research but also has been like my friend and elder brother. I shall never forget the level of maturity and professionalism, which he showed during the period when he lost his elder brother. Apart from learning some of the most important things about research, I also learned a lot about life from him. Throughout the period of my stay at Brunel, he showed his confidence in me by making me a part of different research activities that has been a source of encouragement for me and helped me get exposure about other dimensions of research.

I would like to acknowledge the patience and support shown by my wife and daughter who had to stay away from me for a long time so that I can concentrate on my work.

I am thankful to all my colleagues at the WNCC, and whom I worked with in the IST project MAGNET BEYOND. All of them contributed a great deal in making my stay enjoyable both academically and regarding other extra-curricular activities.

I would like to dedicate this thesis to everyone I mentioned above and all those who prayed for me throughout my research and have been supporting me especially my sponsors the Higher Education Commission of Pakistan and NWFP University of Engineering and Technology Peshawar who financed my degree and stay.

Table of Contents

1	INTRODUCTION	1
1.1	Motivation	1
1.2	Aims of Research	3
1.3	Overview of Wireless Personal Area Networks	4
1.4	High Data Rate WPANs	8
1.4.1	Architecture of HDR WPANs	9
1.4.2	Communication in HDR WPANs	11
1.4.3	Resource Allocation for Real Time Flows in High Data Rate WPANs	13
1.5	Meshed High Data Rate WPANs	15
1.5.1	Wireless Mesh Networks.....	16
1.5.2	Wireless High Data Rate Meshed WPAN Architecture.....	19
1.5.3	Communication in Meshed WPAN	20
1.6	Research Objectives and Contribution to Knowledge	22
1.6.1	Inter-PAN Communication and Admission Control for HDR WPANs	23
1.6.2	Resource Allocation Framework for HDR Meshed WPANs	25
1.6.3	Capacity and Delay Efficient Communication in HDR Meshed WPANs .	28
1.6.4	Capacity Improvements in WPANs with Power Control.....	30
1.7	Research Methodology	32
1.8	Organization of Theses	33
1.9	References	35
2	Inter-PAN Communication and Admission Control for HDR WPANs.....	37
2.1	Introduction.....	37
2.2	Related Work	39

2.3	Communication Issues in High Data Rate (HDR) WPANS.....	42
2.4	Parent-Child Communication Model and its Limitations.....	43
2.4.1	Scheduling Problems in the Parent-Child Communication Model.....	45
2.5	Inter-PAN Communication Model.....	48
2.5.1	Inter-PAN Communication.....	48
2.5.2	Inter-PAN Selection Criteria	55
2.5.3	Parent-Child Scheduling Solution.....	55
2.6	Capacity Analysis of HDR WPANS	57
2.6.1	Capacity for Voice Applications	60
2.6.2	Capacity for Video Applications.....	63
2.7	CTA Allocation and Admission Control	67
2.7.1	Description of Admission Control Procedure.....	68
2.7.2	Fair Allocation Via Rejection Ratio.....	70
2.8	Simulation Results	78
2.9	Conclusion	88
2.10	References.....	90
3	A Resource Allocation Framework for High Data Rate Meshed WPANs.....	92
3.1	Introduction	92
3.2	Related Work	94
3.3	Channel Time Reservation in Meshed WPANs.....	96
3.4	Centralized Time Allocation	98
3.5	Traffic Estimation	105
3.5.1	Matrix Balancing.....	105
3.5.2	Adjusted Assignment Problem	110
3.6	Calculation of d_{max}	118

3.7	Calculation of T_i	122
3.8	Fairness Metrics	130
3.9	Simulation Results	131
3.10	Conclusion	154
3.11	References.....	156
4	Capacity and Delay Efficient Communication in HDR Meshed WPANs	159
4.1	Introduction	159
4.2	Capacity and End-to-End Delay Issues in Meshed WPANs	161
4.2.1	Capacity Utilization in a multi-hop Path	161
4.2.2	End-to-End Delay in multi-hop Meshed WPANs	164
4.3	Communication Efficiency in Meshed WPANs.....	166
4.3.1	Deriving Time Efficiency to determine communication overhead	167
4.3.2	Analysis of Factors Effecting Time Efficiency	168
4.3.3	Algorithms for Selection of Parameters based on Time Efficiency	173
4.4	Proposed Capacity Efficient and Delay Efficient Metrics	181
4.4.1	Initial Analysis.....	183
4.4.2	STS and STS-R Metrics	188
4.4.3	Performance Evaluation of STS, STS-R and MSMR	191
4.5	Important observations regarding achievable improvements with STS and STS-R 195	
4.5.1	General Observations	195
4.5.2	Factors that affect further improvements	197
4.6	Conclusion	198
4.7	References.....	199
5	Capacity Improvements in WPANs with Power Control.....	201

5.1	Introduction	201
5.2	Time Slot Overlapping via Power Control	203
5.2.1	Handshake Procedure	204
5.2.2	Mutually Exclusive Time Slots	206
5.2.3	Adjustment via Overlap Ratio.....	207
5.3	Simulation Results	211
5.4	Conclusion	214
5.5	References.....	215
6	Conclusions and Future Work	217
6.1	Inter-PAN Communication and Admission Control for HDR WPANS	217
6.2	Resource Allocation Framework for Meshed HDR WPANS	220
6.3	Capacity and Delay Efficient Communication in Meshed HDR WPANS	222
6.4	Capacity Improvements in WPANs with Power Control.....	224
6.5	References	226

List of Figures

Figure1.1 IEEE 802.15.3 Superframe Structure.....	9
Figure 1.2 Dependant Piconets.....	11
Figure 1.3 Channel Time Request Command and Associated Fields Format	13
Figure 1.4 Channel Allocation Time Information Element (CTA IE)	15
Figure 1.5 A general overview of a Hybrid Wireless Mesh Architecture	17
Figure 1.6 Example of a possible IEEE 802.11s topology	18
Figure 1.7 An Example Meshed WPAN Architecture	19
Figure 1.8 Superframe Sharing in HDR Meshed WPANs	21
Figure 1.9 Communication Rules for a Parent-Child Model.....	24
Figure 1.10 Channel Time Reservation Flow chart for Centralized and Distributed Allocation Schemes.....	28
Figure 1.11 CTA structure for Dly-ACK Scheme.....	29
Figure1.12 Hidden Node Problem	31
Figure 2.1Superframe Format	38
Figure 2.2 Child Superframe Time Allocation	44
Figure 2.3 Time Allocation for Hierarchical Child Piconets	45
Figure 2.4 Scan procedure initiated by a device.....	49
Figure 2.5 Inter-PAN association process	50
Figure 2.6 The Inter-PAN Association Process.....	52
Figure 2.7 Inter-PAN Piconet Splitting Process.....	53
Figure 2.8 Force Inter-PAN Disassociation Process	54
Figure 2.9 Disassociation Process.....	54
Figure 2.10 Superframe Sharing in Inter-PAN Communication	56
Figure 2.11 Overheads added at MAC and Network Layers	58
Figure 2.12 CTA structure for different ACK schemes	59
Figure 2.13 Superframe overhead with different transmpter data rates.....	62
Figure 2.14 Superframe Capacity for different Frames/CTA	62
Figure 2.15 Percentage CTA overhead (MPDU size = 256 Octets)	62
Figure 2.16 Superframe Capacity against data rate	64
Figure 2.17 Comparison of Actual and Effective throughputs	64
Figure 2.18 Superframe Capacity (MPDU size = 2048 Octets).....	65
Figure 2.19 Comparison of Actual and Effective throughputs	65
Figure 2.20 CTA Overhead (MPDU size = 2048 Octets)	65
Figure 2.21 Time period of expected request.....	69
Figure 2.22 Three State strategy for CTA management and Admission Control.....	70
Figure 2.23 Algorithm based on the value of Rejection Ratio (RR)	74

Figure 2.24 Algorithm based on the value of Improved Rejection Ratio (IRR)	77
Figure 2.25 Superframe Utilization with Variable Number of Devices.....	79
Figure 2.26 Superframe Utilization based on variable Mean Arrival Rate (MAR).....	79
Figure 2.27 Number of Devices simultaneously supported in the superframe	80
Figure 2.28 Number of supported devices with variabe mean arrival rate	80
Figure 2.29 Queue size when the number of devices in the piconet change	81
Figure 2.30 Queue size when the mean arrival rate is varied.....	81
Figure 2.31 Comparison of Throughput for deterministic and non-Deterministic flows	82
Figure 2.32 Comparison of Actual and Effective Throughput	82
Figure 2.33 Comparison of Actual and Effective Throughput for only non-deterministic flows	83
Figure 2.34 Number of supported flows with high throughput requirements	83
Figure 2.35 Number of supported flows when different number of Child Piconets are allocated Channel time	84
Figure 2.36 Number of queued requests when different number of Child Piconets are supported	84
Figure 2.37 Number of queued requests when mean arrival rate is increased to 300 seconds	85
Figure 2.38 Number of supported flows with high throughput requirement from the devices	85
Figure 2.39 Queue size with high throughput requirement from the devices	86
Figure 2.40 Comparison of RR and IRR when both RT and NRT flows are considered.	87
Figure 2.41 Comparison of RR and IRR when ony RT flows are considered	88
Figure3.1 Estimated Backoff Algorithm	103
Figure3.2 General Superframe Structure and Allocated time	104
Figure3.3 The RAS Algorithm.....	109
Figure3.4 The SA Algorithm based on RAS Algorithm	110
Figure3.5 Selection Procedure	112
Figure3.6 (a)	113
Figure 3.6 (b)	113
Figure 3.6(c)	113
Figure3.7 (a)	113
Figure 3.7 (b)	113
Figure3.8 Adjustment for Assignment Algorithm	118
Figure3.9 Selection of dmax	122
Figure3.10 Algorithm Based on Satisfaction Index (SI) Value	130
Figure3.11 Superframe Utilization with Number of Intra-PAN Devices =10, Number of Neighbour MPNCs = 10	132

Figure3.12 Superframe Utilization with Number of Intra-PAN Devices =15, Number of Neighbour MPNCs = 10.....	134
Figure3.13 Superframe Utilization with Number of Intra-PAN Devices =20, Number of Neighbour MPNCs = 10.....	134
Figure3.14 A Continuous trend of Superframe Utilization based on a variety of mean arrival rates	134
Figure3.15 Comparison of per flow CTA Size between Deterministic and Non Deterministic Flows	135
Figure3.16 Traffic Estimation based on Mean Allocation Size (Scenario 1).....	136
Figure3.17 Traffic Estimation based on Mean Allocation Size (Scenario 2).....	137
Figure3.18 Traffic Estimation via RAS Algorithm (10 Iterations) (20, 15)	138
Figure3.19 Traffic Estimation via RAS Algorithm (10 Iterations) (15, 10)	138
Figure3.20 Traffic Estimation via RAS Algorithm (5 Iterations) (15, 10)	138
Figure3.21 Traffic Estimation via Adjusted Assignment Algorithm (15, 10) (Based on most Frequent Allocations).....	139
Figure3.22 Traffic Estimation via Adjusted Assignment Algorithm (15, 10) (Based on total Frequency).....	140
Figure3.23 Number of Retries with and without the Estimated Backoff Algorithm (Case 1)	141
Figure3.24 Number of Retries with and without the Estimated Backoff Algorithm (Case 2)	141
Figure3.25 CSI values based on Policy of Case 1	144
Figure3.26 CSI values based on Policy of Case 3 ($S_{min} = 0.6$) (Scenario 1).....	145
Figure3.27 CSI values based on Policy of Case 3 ($S_{min} = 0.6$) (Scenario 2).....	147
Figure3.28 CSI values based on Policy of Case 3 ($S_{min} = 0.8$) (Scenario 2).....	147
Figure3.29 CSI values based on Policy of Case 2 (Fixed number of MPNCs)	147
Figure3.30 CSI Values based on Policy of Case 2 (Variable Number of MPNCs)	148
Figure3.31 Individual SI values based on Policy of Case 2 (Fixed Number of MPNCs)	148
Figure3.32 CSI Values when AA Policy is not used (Fixed Number of MPNCs)	149
Figure3.33 CSI values for SA policy for η_e values of Table 3.2.....	149
Figure3.34 Comparison of proposed policies based on Jain's Index (high η_e values) ..	149
Figure3.35 Comparison of proposed policies based on Jain's Index (low η_e values) ..	150
Figure3.36 Results based on Theil's Index when AA Policy is used	150
Figure3.37 Results based on Theil's Index when AA Policy is not used	150
Figure3.38 Comparison of Distributed Resource Reservation with Centralized Resource Reservation.....	152
Figure 4.1 Capacity reservation along a source-destination path.....	162

Figure 4.2 An example of queuing delay at each hop	164
Figure 4.3 A tree topology for meshed WPANs.....	166
Figure 4.4 Time Efficiency values with varying Fragment Size and number of frames/CTA	169
Figure 4.5 CTA Size with varying Fragment Size and number of frames/CTA.....	170
Figure 4.6 CTA Size with varying Fragment Size and number of frames/CTA.....	171
Figure 4.7 CTA Size with varying transmission data rate and number of frames/CTA	171
Figure 4.8 Pseudo-Code for the Full Table Search (FTS) Algorithm	177
Figure 4.9 Pseudo-Code for the Partial Table Search (PTS) Algorithm.....	180
Figure 4.10 Example of queuing delay along a source-destination path.....	183
Figure 4.11 General Queuing considerations for Inter-PAN Flows	184
Figure 4.12 Queuing delay variation based on CTA offset	185
Figure 4.13 Queuing Delay and EED Based on CTA offsets	186
Figure 4.14 An ideal case assumption to check the effect of δ_T on capacity efficiency and EED	187
Figure 4.15 End to End Delay for Different number of Hops and Transmitter Data Rates.....	188
Figure 4.16 The fixed last hop offset problem	196
Figure 5.1 Flow Chart of the Handshake Procedure	206
Figure 5.2 CTA overlapping in a single hop WPAN	208
Figure 5.3 CTA Overlapping in a Meshed WPAN	209
Figure 5.4 An example re-arrangement of CTA overlapping based on CTA overlap ratio	210
Figure 5.5 An example bipartite arrangement for the solution of (5.6)	211
Figure 5.6 Simulation scenario for a single hop WPAN.....	212
Figure 5.7 Superframe Utilization based on power control with Overlap Ratio based arrangement	213
Figure 5.8 Simulation scenario for a two hop meshed HDR WPAN.....	213
Figure 5.9 Superframe Utilization based on power control with Overlap Ratio based arrangement for Meshed WPAN	214

List of Tables:

Table 1.1: The IEEE 802.15 Standard series including finalized standards as well as ongoing and active drafts	7
Table 2.1 Important Parameters of HDR WPANs	47
Table 2.2 Parameters Considered for Voice Traffic.....	60
Table 2.3 Some Attributes used in Simulation Scenarios	78
Table 3.1: Superframe Utilization based on a range of Mean arrival Rates and Number of Intra-PAN and Inter-PAN devices	133
Table 3.2: Values of Parameters for the scenario with AA Policy (Result given in Figure 3.29)	146
Table 3.3: Comparison of Values based on CSI with and without AA Policy	152
Table 3.4: Comparison of Values based on Theil's Index with and without AA Policy	153
Table 3.5: Comparison of Values based on Jain's Index with and without AA Policy	153
Table 3.6: Comparison of Values based on Atkinson's Index with and without AA Policy.....	154
Table 4.1: Values of Parameters Used for Analysis of Time Efficiency..	168
Table 4.2: Summary of values based on smaller fragment size's and constant transmission data rate of 22 Mbps	172
Table 4.3: Summary of values based on smaller fragment size's and transmission data rates of 33, 44 and 55 Mbps.....	172
Table 4.4: Summary of values based on larger fragment size's and constant transmission data rate of 22 Mbps	173
Table 4.5: Summary of values based on larger fragment size's and transmission data rates of 33, 44 and 55 Mbps.....	173
Table 4.6: Selected Parameters based on FTS algorithm	181
Table 4.7: Summary of simulation results for comparison between MSMR and STS metrics	192
Table 4.8: Summary of results with Capacity Utilization and Time efficiency	193
Table 4.9: Summary of simulation results for comparison between MSMR, STS and STS-R metrics	195

List of Acronyms:

AA	Accommodate All
AoDV	Ad-Hoc On-Demand Distance Vector
BAN	Body Area Network
BP	Beacon Period
ACK	Acknowledgement
CAP	Contention Access Period
CSI	Cumulative Satisfaction Index
CTA	Channel Time Allocation
CTAP	Channel Time Allocation Period
DEVID	Device ID
EED	End-to-End Delay
FTS	Full Table Search
HDR	High Data Rate
IE	Information Element
IFS	Inter Frame Space
IRR	Improved Rejection Ratio
LR-WPAN	Low-Rate Wireless Personal Area Network
MAC	Medium Access Control
MANET	Mobile Ad-hoc Network
MAS	Medium Access Slot
MDEV	Mesh Device
MIFS	Minimum Inter Frame Space
MLME	MAC Layer Management Entity
MPDU	MAC Protocol Data Unit
MPNC	Mesh Piconet Coordinator
MSDU	MAC Service Data Unit

MSMR	Minimized Slot Mis-ordering Protocol
OLSR	Optimized Link State Routing
PNC	Piconet Coordinator
PNID	Piconet Identifier
PTS	Partial Table Search
RR	Rejection Ratio
SA	Satisfy All
SI	Satisfaction Index
SIFS	Short Inter Frame Space
STS	Switched Time Slot
STS-R	Switched Time Slot with Re-ordering
TU	Time Unit
WLAN	Wireless Local Area Network
WMN	Wireless Mesh Network
WPAN	Wireless Personal Area Network

List of Publications

Published Articles:

1. S. Mahmud, S. Khan, H. Al-Raweshidy, and K. Sivarajah, "Mesched high data rate personal area networks," *Communications Surveys & Tutorials, IEEE*, vol. 10, no. 1, pp. 58-69, 2008.
2. Shah, I.A.; Jan, S.; Mahmud, S.A.; Al-Raweshidy, H.S., "Optimal Path Discovery with Mobility Management in Heterogeneous Mesh Networks," *Future Computer and Communication, 2009 International Conference on*, vol., no., pp.57-61, 3-5 April 2009.
3. Mahmud, S.A.; Khan, S.; Qiang Ni; Al-Raweshidy, H.S., "Capacity Issues in Mesched High Data Rate WPANs," *Advanced Information Networking and Applications - Workshops, 2008. AINAW 2008. 22nd International Conference on*, vol., no., pp.1285-1290, 25-28 March 2008.
4. S. Khan, S.A. Mahmud, K.K. Loo, H.S. Al-Raweshidy, A cross layer rate adaptation solution for IEEE 802.11 networks, *Computer Communications*, Volume 31, Issue 8, Special Issue: Modeling, Testbeds, and Applications in Wireless Mesh Networks, 25 May 2008, Pages 1638-1652, ISSN 0140-3664.
5. Khan, S.; Mahmud, S.A.; Al-Raweshidy, H.S., "A Rate-Adaptive MAC for IEEE 802.11 Networks," *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual*, vol., no., pp.463-469, 5-8 May 2008.
6. Mahmud, S.A.; Khan, S.; Al-Raweshidy, H.S., "Capacity Analysis of High Data Rate Wireless Personal Area Networks," *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual*, vol., no., pp.125-131, 5-8 May 2008.
7. Khan, S.; Mahmud, S.A.; Al-Raweshidy, H.S., "A Cross-Layer Solution for Dynamic Selection of Data-Rate in WLAN using Traffic Differentiation," *Mobile and Wireless Communications Summit, 2007. 16th IST*, vol., no., pp.1-5, 1-5 July 2007.
8. Mahmud, S.A, Khan, S., Al-Raweshidy, H.S, "Analyzing Capacity of High Data Rate WPANs Using different Physical Layer Specifications," *ICT-MobileSummit 2008 Conference Proceedings,2008*. ISBN: 978-1-905824-08-3.
9. Khan, S., Mahmud, S.A, Al-Raweshidy, H.S, "Rate Adaptation for Multi-Rate Wireless Stations: A Cognitive Approach," *ICT-MobileSummit 2008 Conference Proceedings,2008*. ISBN: 978-1-905824-08-3.
10. Khan, Shabbaz; Khan, Shoaib; Ali Mahmud, Sahibzada; Al-Raweshidy, Hamed, "Supplementary Interworking Architecture for Hybrid Data Networks (UMTS-WiMAX)," *Computing in the Global Information Technology, 2006. ICCGI '06. International Multi-Conference on*, vol., no., pp.57-57, Aug. 2006.
11. Javaid, U.; Meddour, D.-E.; Mahmud, S.A.; Ahmed, T., "Hybrid Wireless Networks - Towards an Efficient Gateway Discovery Scheme," *Global Information Infrastructure Symposium, 2007. GIIS 2007. First International*, vol., no., pp.46-51, 2-6 July 2007.

12. S. A. Mahmud, S. Khan, S. Khan, and H. Al-Raweshidy, "A comparison of manets and wmnns: commercial feasibility of community wireless networks and manets," in *Proceedings of the 1st International Conference on Access Networks (AcessNets '06)*. New York, NY, USA: ACM, 2006.

13. Shoaib Khan, Sahibzada Ali Mahmud, Shahbaz Khan, Wenbing Yeo, Franjo Cecelja; Generalized Architecture for Converged Heterogeneous Networks, *Wireless World Research Forum - proceedings, The 16th WWRF Meeting*, 26 – 28 April 2006, Shanghai, China.

Submitted Articles (In Press):

1. Mahmud, S.A, Khan, S., Al-Raweshidy, H.S, "A Resource Allocation Strategy for Meshed High Data Rate WPANs," *Submitted to IEEE Communication Letters*, July 2009.

2. Mahmud, S.A, Khan, S., Al-Raweshidy, H.S, "Centralized Resource Allocation Policies for Meshed High Data Rate WPANs," *Accepted for Publication in IET Communications*, Aug 2009.

CHAPTER 1

1 INTRODUCTION

1.1 Motivation

The convenience offered by wireless networks in the last mile connectivity over the traditional wired networks, and the improvements in the processing power and applications for handheld devices have been some of the significant reasons to witness the explosive growth in the wireless industry. The ultimate aim of ubiquitous connectivity although has been realized to a large extent, it is still in the process of evolution and there is room for further improvement. Since devices in a wireless network share a common medium, there are a number of issues which need to be considered e.g. Channel capacity or Bandwidth, simultaneous transmissions which lead to collisions, data integrity protection, Support for Quality of Service (QoS) for multimedia flows, inter-channel and intra-channel interference etc. Apart from the mentioned issues, the constant improvements in the capabilities of multimedia devices like handheld gaming devices, digital cameras, High Definition (HD) Audio and Video players and recorders etc, require very high throughputs which can become difficult to support in case of a large number of devices trying to communicate and competing for channel capacity. The method by which a device gets access to the wireless medium is crucial in determining the techniques to be used to efficiently utilize the medium as well as the flow between the source and destination devices.

Traditionally, two types of approaches have been considered for a device to get access to the wireless medium i.e. Contention Based and Contention Free. The contention based approach uses Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA). The IEEE 802.11 standard [1] which defines specifications for Wireless Local Area Networks (WLANS) is one example in which the 802.11 compliant wireless stations compete for channel occupancy. The contention free approach commonly uses Time Division Multiple Access (TDMA) in which one or more fixed or variable size time slots are assigned to a device to

communicate. Since each device has its own time slot, it does not have to compete with other devices for channel occupancy. Some examples of TDMA based networks consist of IEEE 802.16e [2] which is for Metropolitan Area Networks (MANs) (known as WiMAX) and IEEE 802.15 standard which defines specifications for High Data Rate (HDR) and Low Data Rate (LDR) Wireless Personal Area Networks (WPANs). Although, both 802.16x based networks and 802.15x based networks predominantly use TDMA as the medium access technique, there is a big difference in coverage area of both networks. 802.16x based networks span an area in km while 802.15x based networks span a much smaller area equal to a personal operating space with user centricity as one of its goals. The drivers for the two technologies also differ. 802.16 based networks are aimed at supporting users moving at high speeds even at vehicular speeds to get access to data with high data rates with QoS support. LDR WPANs are targeted towards enabling low cost, low power devices to communicate with each other while HDR WPANs are targeted to support high end multimedia applications with high throughput requirements.

The applications for HDR WPANs are various including HD audio and video streaming from one device to other devices, multiplayer gaming, patient monitoring systems which can involve transmitting high resolution images to the nearest Access Point (AP), Home automation, Office Automation, dynamic backups, multimedia home network, interconnection among handheld devices, indoor location based services etc. Since HDR WPANs have a limited operating range with single hop communication between devices, an approach to extend the range and enable multi-hop communication is via mesh networking. IEEE 802.15.5 standard defines specifications for including such multi-hop functionality in the legacy 802.15.3 and 802.15.4 networks. Although, the number of existing applications and the possibility of coming up with new innovative applications and uses of HDR WPANs are many, some of the issues which need to be addressed in order to take advantage of the full potential of single hop and multi-hop meshed HDR WPAN networks i.e. IEEE 802.15.3 [3] are:

- Lack of *effective* peer-to-peer communication between independent WPAN clusters

- Absence of *Fair scheduling* in order to allocate a fair share of time slots to requesting devices for intra-cluster communication as well as inter-cluster communication
- Minimizing *End-to-End delay* in case of multi-hop meshed WPANs which becomes more pronounced when the time slots are scheduled out of order from the source to the destination
- *Limited capacity* of superframe in which time slots are assigned to requesting devices

This thesis is the end result of research efforts put into addressing the above issues by designing, analyzing and evaluating techniques to overcome these issues either completely or to some extent.

1.2 Aims of Research

Since the research presented in this thesis tackles multiple issues, the aims are multi-fold:

1. Develop Inter-PAN communication framework to enhance and enable the effective peer-to-peer connectivity and determine a mechanism for two or more independent WPAN clusters to communicate. In order to do so, determine the feasibility of Inter-PAN communication framework and analyze and evaluate the relevant parameters such as capacity limitations. In case of capacity issues, determine and propose an admission control technique to induce some fairness in allocation of channel time to requesting devices.
2. In case of meshed WPANs, multiple WPAN clusters compete for channel time in a shared superframe. Therefore, an approach to determine the channel time requirements of each cluster with a certain number of devices becomes mandatory. After a suitable fraction of superframe time is determined for each requesting cluster, design, analyze and evaluate techniques for fair allocation of channel time based on superframe capacity and total requested time.
3. Determine the essential attributes and their effect on the communication efficiency of single hop and multi-hop WPAN flows. Based on such attributes, develop a routing metric to find the best route which is capacity efficient as well as delay efficient.

4. In order to improve the capacity of single hop as well as multi-hop WPANs, investigate power control schemes in order to determine the feasibility of spatial reuse and time slot overlapping.

1.3 Overview of Wireless Personal Area Networks

Wireless Personal Area Networks or WPANs operate in a limited area or a personal space around a user. The devices in a WPAN are generally low power and low cost devices with usually one of the devices with enough resources being the coordinator. Based on the Air Interface (AI), data rate, MAC layer access mechanisms and intended application, different standards of WPANs exist. The family of WPAN standards consists of IEEE 802.15.1 (Bluetooth), IEEE 802.15.2 (Co-existence), IEEE 802.15.3 (High Data Rate WPANs), IEEE 802.15.4 (Low rate WPANs), IEEE 802.15.5 (Extensions for multi-hop mesh functionality) and 802.15.6 (Body Area Networks or BANs). In a Bluetooth WPAN, the devices communicate in a master slave relationship in an ad-hoc communication system called a piconet. The device that provides the synchronization reference is usually termed as the Master and the rest are termed as the Slaves. In the presence of different piconets in a common operating space, each piconet usually operates on a different radio frequency channel with a different Master device. A device can participate in one or more piconets at the same time. If a device participates in more than one piconet at a time, it is said to be operating in a scatter-net. A device can act as a slave in more than one piconet but it cannot act as a master in more than one piconet. The access mechanism used is Time Division Multiplexing (TDM). Although there are constant efforts being made to enhance the supported data rates by Bluetooth, the maximum data rate is < 5 Mbps, which is insufficient for most of Real Time Applications.

In order to support Multi-Media applications, IEEE 802.15.3 defines specifications for High Data Rate (HDR) WPANs. The devices in an HDR WPAN communicate in an ad-hoc system also called a piconet and one of the devices with enough resources becomes a Piconet Coordinator (PNC). The PNC transmits a periodic Beacon frame that contains the timing information for the piconet. The PNC is also responsible for the channel time allocation for the devices to communicate. The access mechanism is Time Division Multiple Access (TDMA). The 802.15.3 HDR WPANs uses the 2.4 GHz frequency band and the

maximum transmitter data rate is 55 Mbps. Some revisions, minor optimizations and additional recommendations to clarify ambiguities for the MAC and PHYSical (PHY) layer of 802.15.3 are given in IEEE 802.15.3b. To support even higher data rates, the specifications for an alternate PHY layer based on Ultra Wide Band (UWB) are given in IEEE 802.15.3a. The maximum data rate supported by IEEE 802.15.3a is 480 Mbps. In order to support data rates in Gbps, the Task Group TG3c is working to finalize specifications to standardize IEEE 802.15.3c for millimetre wave based WPANs in the 60 GHz frequency band.

The IEEE 802.15.4 or Zigbee defines specifications for very low power devices with comparatively much lower data rates and slightly longer transmission range. It is mainly targeted for devices with very limited battery consumption requirements such as sensors. According to [4], the maximum supported data rate is up to 250 kbps. The access mechanism is mainly Carrier Sense Multiple Access with Collision Avoidance (CSMA-CA) with an optional allocation of Guaranteed Time Slots (GTS). There are two types of devices in a Low Rate WPAN (LR-WPAN), namely a Full Function Device (FFD) and a Reduced Function Device (RFD). An FFD can communicate with any other device while an RFD can only communicate with an FFD. The LR-WPAN can operate in two topologies i.e. star topology and peer-to-peer topology. In the star topology, one of the FFD's assumes the role of a PAN Coordinator to manage the piconet. In the peer-to-peer topology, the PAN Coordinator is also present but any device can communicate with any other device if they are in the transmission range of each other. Such a topology is also preferred for sensor networks and other complex topologies like Wireless Mesh Networks (WMNs). Some examples of the use of LR-WPANs for sensor networks are given in [5], [6] and [7].

The recommendations for operation in a mesh configuration for WPANs are given in IEEE 802.15.5. It has two parts in which one part defines necessary specifications for LR-WPANs and the other for HDR WPANs. The impetus for a WPAN to operate in a mesh topology is to increase the network coverage without increasing the transmit power, increase the route reliability via route redundancy, self configuration, and efficient use of device battery life. The various applications for LR-WPANs to operate in a mesh configuration according to [8] include home and industrial automation and control, security and environmental monitoring, situational awareness and asset tracking, automatic meter reading, personal health monitoring etc. For the HDR WPANs, the operation in a mesh

topology can result in network range extension, reliable communication and efficient bandwidth reuse for multi-media applications such as video conferencing and HDTV in home environment.

The BANs consist of a network of sensors either external or implanted in the human body, or both to monitor critical health parameters. Typically, the coverage area for BANs spans a little more than that of the human body. BANs are envisaged mainly for significant improvements in the health monitoring of patients with chronic diseases like Diabetes, Cardiac Diseases etc and to reduce the risk of complications which result because of delay in attending emergencies. The patients and their vital signs can be proactively monitored through the BAN that can send critical data through either a Wireless Local Area Network (WLAN) Router or another local base station in its transmission range. Since the BAN has very limited coverage, therefore, in order to realize the main purpose of proactive patient monitoring, the availability of a nearby device through which it can relay the data is very important. In some cases, patients with critical vital signs in a hospital may undergo imaging tests or scans and therefore, the BAN can connect to a nearby device to relay its data using a suitable interface to the appropriate server which has the relevant details about the patient and to notify the doctor about any recommended procedures or emergencies. However, in order to effectively communicate the details of a patient, the available interface to the BAN must be compatible with it. Furthermore, to transfer the results of some imaging tests like MRI etc, a relatively high data rate is required by the relay device. The authentication, integrity and security of transmitted data in such cases is crucial for such systems to work efficiently. The task group for making recommendations to standardize BANs i.e. IEEE 802.15.6 is Task Group 6 (TG 6). A summary of 802.15x standards is given in Table 1.1.1.

The work done in the thesis focuses on different aspects of HDR WPANs i.e. capacity, admission control, resource allocation and optimizing real time flows. The main reason is that the support for real time applications having higher throughput and strict QoS requirements give rise to significant research issues which need to be resolved for more efficient use of HDR WPANs. The improvement and advancements in the newer handheld devices especially mobile phones with a myriad of interesting and innovative applications can pave way for various implementations of HDR WPANs in the near future e.g. multimedia

streaming, multiplayer gaming, peer to peer sharing, personalized and user centric profiles for various tasks etc. A more detailed description of HDR WPANs and Meshed HDR WPANs is therefore, as follows.

Table 1.1: The IEEE 802.15 Standard series including finalized standards as well as ongoing and active drafts

Extension	Task Group	Description
802.15.1	TG 1	Specifications for an adaptation of industry defined Bluetooth for the MAC (L2CAP, LMP, and Baseband) and PHY (Radio)
802.15.2	TG 2	Specifications to quantify the mutual interference of a WLAN and WPAN. A set of coexistence mechanisms to facilitate coexistence of WLANs and WPANs
802.15.3	TG3	MAC and PHY layer specifications of HDR WPANs (11-55 Mbps). Low cost, low power solutions to address the needs of portable consumer digital imaging and multimedia applications are also provided
802.15.3b	TG 3b	An amendment to improve and clear the ambiguities in IEEE 802.15.3. Minor optimizations while preserving backward compatibility are also provided
802.15.3c	TG 3c	Developing a millimetre-wave-based alternate PHY layer for the existing 802.15.3 MAC to support data rates in excess of 1Gbps. The frequency band used is 57-64 GHz (Unlicensed)
802.15.4	TG 4	Standard to support low power, low data rate solutions for low complexity devices e.g. sensors. Some salient features of this standard include CSMA/CA access with optional support of Guaranteed Time Slots (GTS), power management to ensure low power consumption and data rates of 250, 40 and 20 Kbps
802.15.4a	TG 4a	Serves as an amendment to 802.15.4 to provide high precision ranging/location capability, high throughput and low power. It also has specifications to add scalability to data rates and support longer range
802.15.4b	TG 4b	Provides enhancements and clarifications to the 802.15.4 standard to resolve ambiguities and consider newly available frequency allocations
802.15.4c	TG 4c	A PHY amendment to 802.15.4 and 802.15.4a to address the Chinese regulatory changes which opened the 314-316 MHz, 430-434 MHz and 779-787 MHz bands for WPAN use in China
802.15.4d	TG 4d	An amendment to 802.15.4 to define a new PHY and appropriate changes to the MAC necessary to support a new frequency band (950-956 MHz) in Japan
802.15.4e	TG 4e	An amendment to 802.15.4 to add functionality and support additional applications such as Factory Automation, Process Automation, Asset Tracking, General Sensor Control (Industrial/Commercial, including Building Automation), Home Medical Health/Monitor, Telecom Application, Neighbourhood

		Area Networks etc
802.15.4f	TG 4f	To define specifications for new PHY layer and enhancements to the 802.15.4 MAC to support Active Radio Frequency Identification (RFID) mainly for location determination applications
802.15.4g	TG 4g	To create a PHY amendment to 802.15.4 to enable Smart Utility Networks (SUNs) capable of supporting large geographically diverse networks with minimal infrastructure which can consist potentially millions of end points
802.15.5	TG 5	Specifications and necessary mechanisms to enable the PHY and MAC layer of WPANs to support mesh networking. The resultant mesh network is able to provide extension of network coverage, enhanced reliability via route redundancy, easier network configuration and improved battery life
802.15.6	TG 6	Developing a standard to support low power devices which can communicate in and around the human body to facilitate medical applications such as personal health monitoring systems as well as personal entertainment systems
802.15.7	TG 7	To develop MAC and PHY layer specifications for Visible Light Communication (VLC) which uses an LED to establish communication by using the rapid switching attribute of an LED

MAC Layer – Medium Access Control Layer
PHY Layer – Physical Layer

1.4 High Data Rate WPANs

In HDR WPANs, the devices communicate in an ad-hoc communication system called a piconet. The topology is similar to a star topology with one of the devices having enough resources acting as the Piconet Coordinator (PNC). The PNC transmits a periodic beacon frame that consists of the timing information as well as other control information. The beacon frame is the first frame sent in a superframe that has a maximum duration of 65535 μ s. The superframe structure can be seen in Figure1.1 IEEE 802.15.3 Superframe Structure. The superframe consists of a beacon frame, a Contention Access Period (CAP) and a contention free period called the Channel Time Allocation Period (CTAP). The access mechanism used during the CAP is CSMA-CA while it is TDMA during the CTAP. Therefore, guaranteed time slots called Channel Time Allocations (CTAs) can be allocated to requesting devices during the CTAP. The beacon frame transmitted by the PNC has information about all the allocated CTAs that are reserved by different devices for data transfer. Efficient allocation of CTAs is mandatory in maintaining a certain level of Quality of Service (QoS) for

the multimedia flows in the piconet. Apart from the management of QoS, the PNC is also responsible for power save modes and access control to the piconet.

In order to start a piconet, a device capable enough to become a PNC scans available radio frequency channels. If a vacant channel is found, it starts sending a beacon frame. Upon reception of a beacon frame by a device, it can send an association request command to the PNC to become a member of the piconet. It is imperative for a device to be associated with the PNC in order to access any available resources in the piconet or to communicate with another member device in the piconet. Upon successful association, the PNC sends an association response command to the requesting device and assigns an 8-bit device id (DEVID) to it that uniquely identifies the device in the piconet. The PNC then broadcasts the information about the newly associated device in its beacon frame so that the already associated devices can have information about the new member. The PNC also broadcasts the information about the rest of the devices in the piconet so that the newly associated device can be informed about the available resources in the piconet and the member devices along with their capabilities.

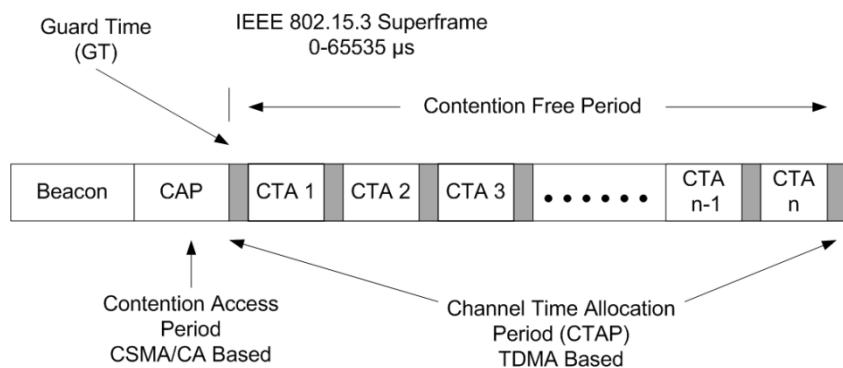


Figure 1.1 IEEE 802.15.3 Superframe Structure

1.4.1 Architecture of HDR WPANs

In HDR WPANs, the communication between the PNC and the devices, and between the devices takes place in a single hop. The coverage of WPANs is usually that of a Personal Operating Space (POS) and spans around 10m – 30m depending on the transmit power, data rate and the modulation scheme. However, the PNC has an option to extend its range by forming child piconets. The child piconets are also useful for the PNC to share its computational and memory requirements with another PNC capable device that acts as a

child PNC. The child PNC is termed as a dependant piconet while the established piconet is termed as the parent piconet. The PNC of the parent piconet is termed as the parent PNC. Each parent piconet can have multiple child piconets. In order to form a child piconet, a PNC capable device, which is a member of the parent piconet sends a request to the parent PNC. Upon acceptance of the request, the PNC capable member device forms the child piconet and becomes the child PNC. Although the child piconet is autonomous in its operations, it is called a dependant piconet because it depends on a CTA that is allocated by the parent PNC. The communication of the child piconet takes place in a CTA that is called a Private CTA. An example of a parent piconet supporting dependant piconets can be seen in Figure 1.2.

Each piconet has its own unique Piconet Identifier (PNID) by which it is recognized. The child piconet also has a unique PNID. The child PNC after the formation of the child piconet is still a member of the parent piconet. Therefore, it can communicate with any other member device in the parent piconet as well as the parent PNC. The child PNC can also communicate with any of the member devices in the child piconet. Apart from the child piconet, there is another type of dependant piconet called the Neighbour piconet. The neighbour piconet functionality is useful when the frequency spectrum is fully utilized and there is no vacant radio frequency channel available for a device to start its own piconet. Just like the child piconet, the parent piconet can have multiple neighbour piconets or it can have the child piconets as well as the neighbour piconets. The neighbour piconet is also dependant upon a private CTA allocation from the parent PNC. Unlike the child piconet, the neighbour PNC is not a member of the parent piconet and hence cannot communicate with any device in the parent piconet. In Figure 1.2, the child piconet is shown to have other child piconets and the neighbour piconet is shown to have a neighbour and a child piconet. Although there is no limit defined in [3] for a parent or a dependant piconet to have further dependant piconets, but since all of them share a common superframe, the capacity of the superframe thus becomes an important factor to consider in such scenarios.

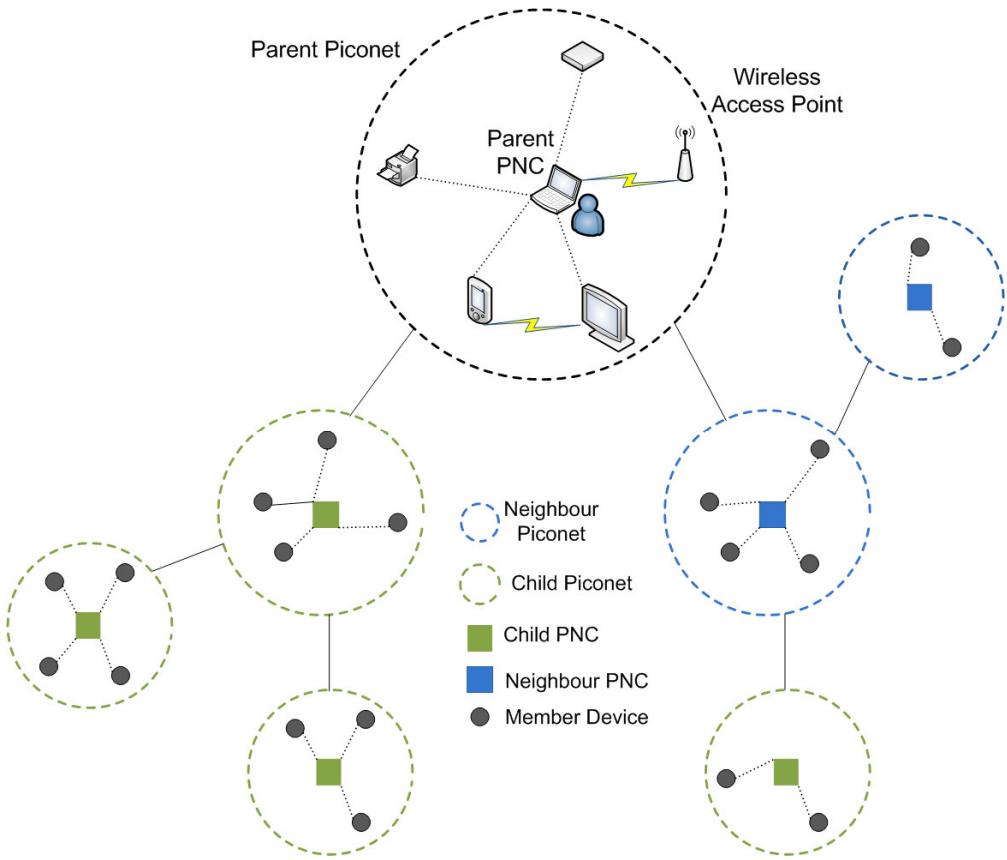


Figure 1.2 Dependant Piconets

1.4.2 Communication in HDR WPANS

In order to communicate with any device in the piconet, a device first needs to associate itself with the PNC. Before the association process, any information exchange between the device and the PNC takes place in the CAP. When a device discovers the presence of a PNC by receiving its beacon frame, it sends an association request command to the PNC using the CAP of the superframe. Upon reception of the association request command, the PNC ensures that there are enough resources available in the piconet for the requesting device. If the PNC determines that the device should be associated, it sends an association response command with appropriate fields set to indicate to the requesting device that the association process has been successful and allocates it a device ID (DEVID). The device then sends another association request command with its allocated DEVID to the PNC and the PNC acknowledges it. After the association process, the PNC sends an Information Element (IE) in its beacon frame with the relevant information about the newly associated device e.g. DEVID, its MAC address and its capabilities to all the members of the piconet to

update their lists about the member devices. The capabilities of the associated device which are broadcast by the PNC in its beacon contains the relevant information about its communication abilities e.g. supported data rates, preferred fragment size, multicast support etc.

When an associated device needs to communicate with another member device in the piconet, it can do so while using the CAP or request a CTA from the PNC. Usually CAP is used by the devices and the PNC to send commands and/or send small amounts of asynchronous data. For isochronous streams, the devices can request a guaranteed time slot in the form of a CTA from the PNC. In order to do so, the device sends a channel time request command to the PNC to reserve a CTA in the superframe. The device specifies the type of CTA it requires in the channel time request command. There are two types of CTAs, namely dynamic CTAs and pseudo-static CTAs. The dynamic CTAs can be used by the device to establish asynchronous streams as well as isochronous streams. The dynamic CTAs can be moved by the PNC to a different location in the superframe on a superframe-to-superframe basis, if required. The pseudo-static CTA is only used to establish isochronous streams in the piconet by the device.

When a PNC receives a channel time request command from a device, it checks for available time in the superframe. If there is time available, then the PNC sends a channel time response command to indicate to the requesting device that it has been allocated channel time in the superframe. The PNC announces the newly allocated CTA by sending a CTA Status IE in its beacon frame to notify the other member devices. The PNC transmits the information about all the allocated CTAs in its beacon using CTA IEs. The updated information about the new allocation is therefore, sent in the CTA IE by the PNC in the subsequent beacon frames. In case the device needs to form a child piconet, it sends a channel time request command to the PNC and requests a Private CTA from the PNC. The private CTA is a type of a pseudo-static CTA with the difference that the source ID and the destination ID in the channel time request command are set to the same value as the requesting device's DEVID to indicate to the PNC about the formation of a child piconet.

If upon the reception of a channel time request command, the PNC determines that there is not enough time available in the superframe to serve the request, the PNC sends a

channel time response command to the requesting device and indicates to it about the failure to allocate a CTA due to insufficient time.

1.4.3 Resource Allocation for Real Time Flows in High Data Rate WPANs

The HDR WPAN specifications have been made to address the QoS requirements for multi-media traffic. The PNC is responsible for maintaining the QoS for isochronous streams that are active in the piconet at a given time. The efficient allocation of CTAs in HDR WPANs is therefore, mandatory to achieve the desired QoS for a given flow. Each device informs the PNC about certain requirements of its intended flow in the channel time request command. The channel time request command is shown in Figure 1.3.

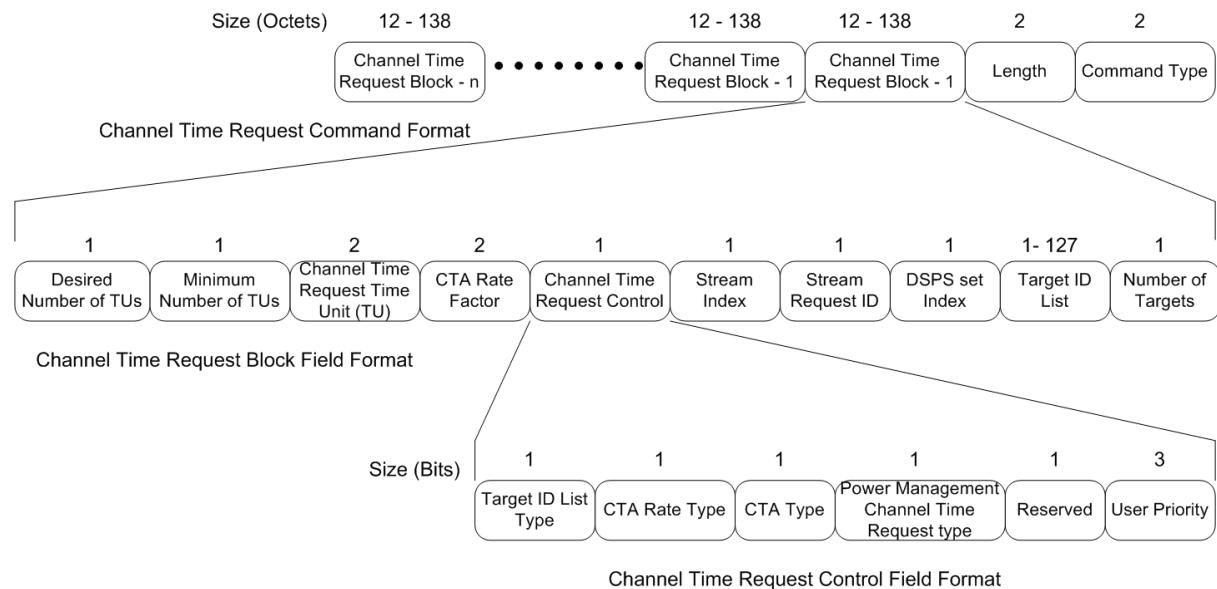


Figure 1.3 Channel Time Request Command and Associated Fields Format [IEEE 802.15.3]

Different Multimedia flows have different requirements. The most important parameters for which certain minimum values are necessary in order to maintain and achieve a satisfactory level of QoS are throughput and delay. The applications that are most sensitive to these two factors are mostly voice and video applications. The generic specifications for acceptable delay are mentioned for voice and video applications in [3]. For isochronous streams supporting voice applications, the acceptable delay and jitter should be less than 10 ms and for video applications it should be less than 100 ms. In order to

ascertain that a particular flow adheres to its delay requirements, the locations of its allocated CTAs in the superframe and the number of CTAs per flow per superframe play an important role. Each device based on its QoS requirements can request either sub-rate CTAs or super-rate CTAs from the PNC. For isochronous streams, the super-rate CTAs are used. The super-rate CTA allocation means that there are multiple CTAs allocated per superframe while a sub-rate CTA allocation means that a CTA is not allocated for each superframe but is rather allocated for alternate superframes.

In Figure 1.3, each *channel time request block* field corresponds to a channel time request. The *cta rate type* field in the channel time request control field determines whether the requested flow requires a sub-rate CTA allocation or a super-rate CTA allocation. The value of the *CTA type* field determines whether the allocation is for a pseudo-static CTA or a dynamic CTA. The value for this field is set to one if the channel time request is for a pseudo-static CTA. The *CTA rate factor* field along with the CTA rate type field determines the frequency of CTA within a superframe or from superframe to superframe. For a sub-rate CTA allocation, a CTA rate factor of 2 indicates that the device is requesting a CTA allocation every second superframe. For a super-rate CTA allocation, a cta rate factor of 8 suggests that the device is requesting a CTA allocation 8 times for the same superframe. Since the HDR WPANs have a single hop topology, the distance or time offset between successive CTAs have a great impact on the throughput and delay of the flow. The *channel time request Time Unit (TU)* field indicates the unit of time for a CTA e.g. 1 μ s, 2 μ s etc to make the allocation more efficient. The *minimum number of TUs* field and the *desired number of TUs* field are mainly used for isochronous streams and they usually denote the minimum and desired throughput requirements for the flow. Each flow is allocated a unique stream index that identifies the flow among other active flows in the piconet.

The PNC, based on the feasibility of received channel time request command, sends an appropriate channel time response command to the requesting device. In case of successful time allocation to the requesting device, the PNC places a CTA status IE into its beacon frame. The CTA status IE consists of a field called the *start beacon number* that is set to the beacon number of the first beacon where the allocated CTA of the new or modified stream is expected to appear. Apart from the CTA status IE, the PNC also places the information about the newly allocated CTA in the CTA IE. The format of the CTA IE is shown in Figure 1.4.

Each *CTA block* field in the CTA IE consists of source and destination ID fields, the stream index for the flow, the location of the CTA in the CTAP and the total CTA duration.

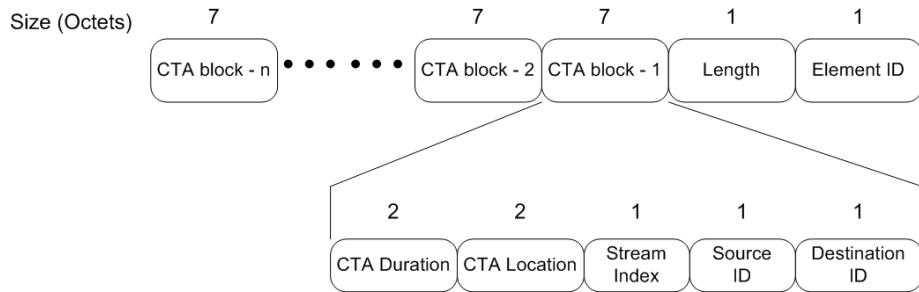


Figure 1.4 Channel Allocation Time Information Element (CTA IE)

The channel time request command is not only used by a device to request for channel time from the PNC but it is also used to modify and terminate flows. Once a flow is established and is allocated a stream index by the PNC, using that stream index, a device can request for flow modification or termination by sending a channel time request command to the PNC.

The HDR WPAN standard supports four acknowledgement (ACK) schemes, namely, No acknowledgement (No-ACK), Immediate acknowledgement (Imm-ACK), Delayed acknowledgement (Dly-ACK) and Implied acknowledgement (Imp-ACK). For isochronous streams, the Dly-ACK scheme is usually preferable due to its similarity to the sliding window technique used in TCP/IP. Therefore, the Dly-ACK scheme has been used for the analysis and simulation work done in this thesis.

1.5 Meshed High Data Rate WPANs

In order to extend the single hop topology of HDR WPANs to multiple hops and define necessary specifications for their operation in a meshed configuration, the IEEE 802.15.5 standard has been developed. Before going into the details of meshed WPAN specifications, an insight into Wireless Mesh Networks (WMNs) is presented for a better overall understanding of its main features and differences when compared to traditional Mobile Ad-Hoc Networks (MANETs).

1.5.1 Wireless Mesh Networks

Wireless Mesh Networks (WMNs) have grown in popularity in recent years due to their advantages over traditional MANETs. WMNs are self-configurable and self-healing networks that consist of mesh clients and mesh routers. In a hybrid WMN, the mobile devices are usually termed as the mesh clients while the static devices that relay the data to/from the mesh clients to a wired or wireless backhaul network are termed as wireless routers. Apart from other factors, WMNs are a viable option to reduce the cost of implementation since only a few backbone devices or wireless routers are required to have a dedicated wired high capacity connection to overlay the data of mesh clients to another mesh network or a public network such as the Internet. The mesh clients can communicate with each other directly, in a multihop fashion or through the mesh routers if the destination mesh client belongs to another mesh network.

WMNs provide more flexibility in terms of the underlying topology when compared to MANETs because WMNs are not strictly infrastructure-less. WMNs consist of a mobile part that consists of the mesh clients and the supporting part in the form of mesh routers that are usually fixed and thus have some form of infrastructure as shown in Figure 1.5. The efficiency of a WMN to a large extent depends on the routing protocol in use. In order to increase capacity and reduce throughput, the feasibility of using low cost multiple radios in a single device have been studied extensively. Some of the well known approaches are given in [9], [10], [11] and [12]. A detailed survey about WMNs in general is given in [11]. Unlike the routing protocols for MANETs which operate on a single radio interface per device, the routing protocols designed for WMNs can recognize multiple radio interfaces per device and thus with the use of dynamic channel assignment schemes [14][15], the capacity is shown to be improved and the interference is reduced in WMNs when compared to traditional MANETs.

WMNs, therefore, offer a great potential and promise in terms of extended coverage, robustness, self-configuration, self-healing, easy maintenance and low cost. IEEE has therefore been actively involved in standardizing mesh networks. Their efforts have targeted WPANS, Wireless Local Area Networks (WLANS) and WiMAX. IEEE 802.15.5 has been developed which provides mesh extensions for both high data rate (802.15.3) and LR-WPANS. In meshed WPANS, Mesh Devices (MDEVs) and Mesh Piconet Coordinators

(MPNCs) can form a tree topology or other distributed topologies. Each MPNC is connected to another MPNC in its range and is responsible for its respective piconet. The mobile MDEVs can move between different piconets and associate with different MPNCs. The MPNCs coordinate with each other in such a way that they send their beacons in Medium Access Slots (MAS) to avoid beacon collision. The communication between MDEVs and MPNCs of different piconets is called Inter-PAN communication while communication between devices in the same piconet is called Intra-PAN communication. Since a WPAN piconet covers a relatively smaller area than a WLAN, multiple overlapping piconets can cause significant interference. It is therefore likely in future that the MDEVs are equipped with multiple radios to increase capacity and combat interference.

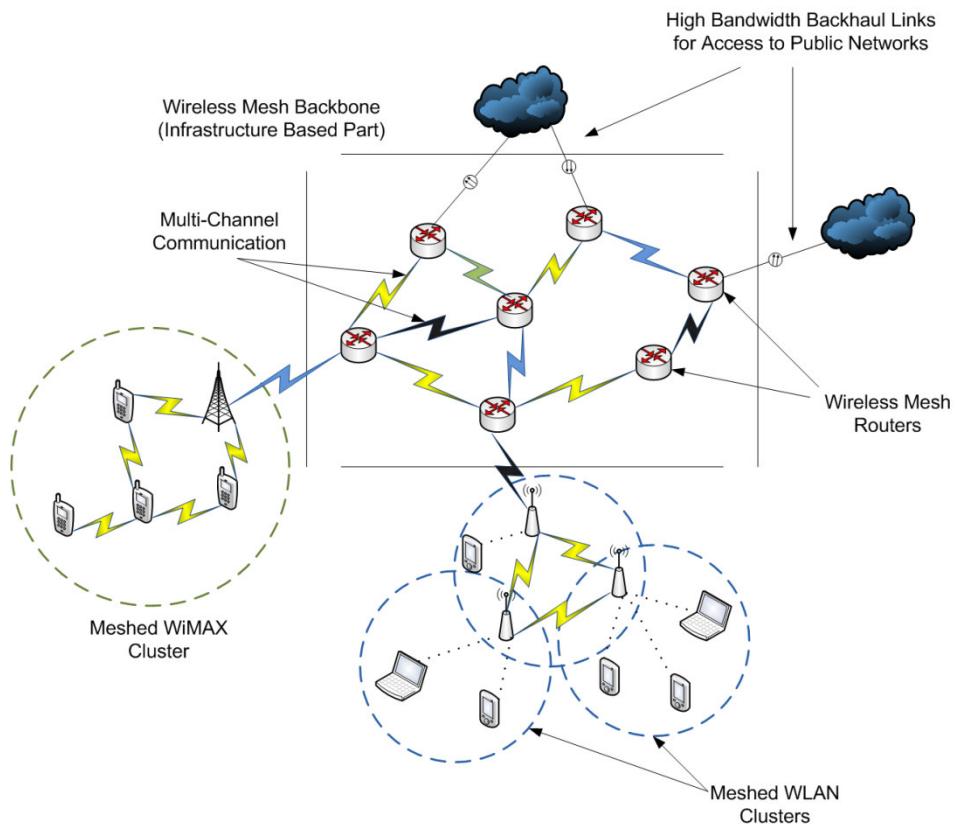


Figure 1.5 A general overview of a Hybrid Wireless Mesh Architecture

IEEE 802.11s is an amendment being developed, which integrates mesh networking services and protocols at the MAC Layer of the IEEE 802.11 Wireless LAN (WLAN) standard. It aims to create a Wireless Distribution System (WDS) with automatic topology learning and wireless path configuration. There is dynamic, radio-aware path selection in the meshed

WLAN, enabling data delivery on single-hop and multi-hop paths. In the traditional WLAN, the stations directly associate with the Access Points (APs) through a single hop connection. The APs have a wired connection to the public network (Internet) that can have an impact on cost. In meshed WLAN, Mesh Points (MPs) are added in the network (see Figure 1.6) and are connected to other MPs and Mesh Access Points (MAPs) through a wireless connection. One (or more) of the MPs performs the role of a Mesh Portal (MPP) through which the frames exit and enter a WLAN mesh. Since only the MPP has a wired connection, there is significant reduction in cost as well as increased coverage through MPs. The stations that associate with the MAP can reach the MPP in a multi-hop fashion through the MPs. To increase the capacity of the meshed WLAN, multi-channel communication is made possible and the MPs can be equipped with multiple radios. The routing protocol is also made radio aware and is called Hybrid Wireless Mesh Protocol (HWMP). HWMP is the default routing protocol for meshed WLAN and consists of an on-demand routing protocol and a pro-active tree based routing protocol. On demand routing is based on Radio Metric Ad-Hoc On-demand Distance Vector (RM-AODV). There is also an optional protocol specified for meshed WLAN called Radio Aware OLSR (RA-OLSR). It is a combination of Optimized Link State Routing (OLSR) [16] protocol and Fisheye State Routing (FSR) protocol [17].

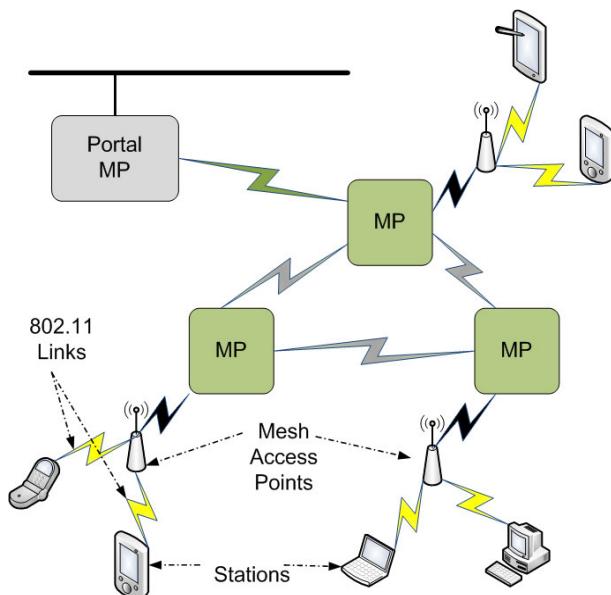


Figure 1.6 Example of a possible IEEE 802.11s topology

1.5.2 Wireless High Data Rate Meshed WPAN Architecture

In a meshed WPAN, there are mesh devices (MDEVs) and mesh PNCs (MPNCs) in addition to the legacy 802.15.3 devices and PNCs. Communication between MDEVs that belong to different independent piconets is possible in a mesh configuration. The MPNCs can communicate with each other and they can also manage their own piconets that can contain both the MDEVs and the legacy 802.15.3 devices. Thus both inter-PAN and intra-PAN communication can take place in meshed WPANs as shown in Figure 1.7. Unlike the intra-PAN communication that takes place in the legacy single hop 802.15.3 piconet in which a single PNC coordinates the channel time in the superframe, several MPNCs share the channel time in the superframe among them. For this reason, the superframe is divided into Medium Access Slots (MASs) that are of equal length. The length of the superframe is kept the same throughout the mesh network and all the MPNCs synchronize their beacons to a common reference MPNC.

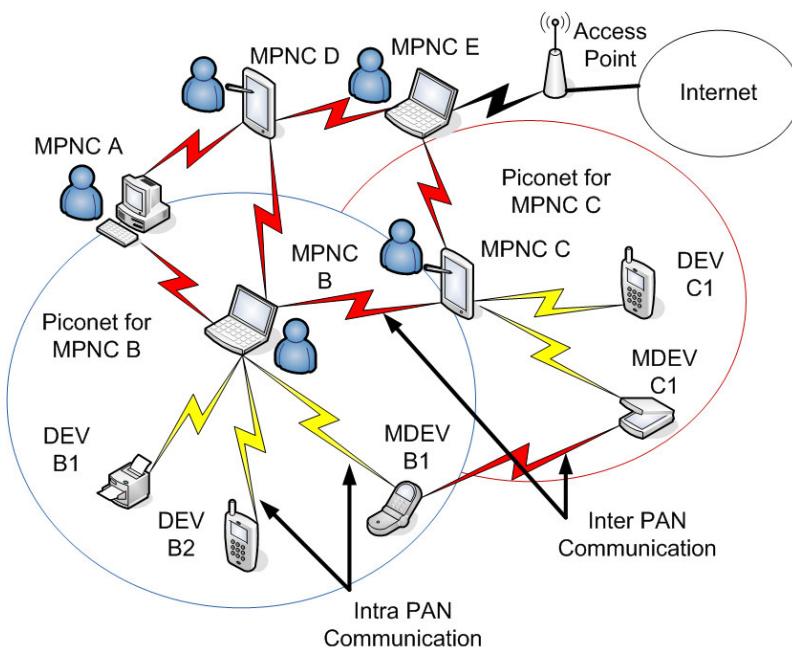


Figure 1.7 An Example Meshed WPAN Architecture

The superframe can be used for transmission of beacons from MPNCs, reception of beacons from neighbouring MPNCs, intra-PAN communication and inter-PAN communication. The beacon transmitted by an MPNC announces the presence of the source MPNC, negotiates and announces reservations, provides information about neighbour reservations and provides timing information that aids in synchronization. Inside a piconet,

the beacon provides the same information as in legacy 802.15.3 WPANs. A comprehensive overview of meshed HDR WPANs can be found in [18].

The particular topology formed by the meshed WPAN can be either a tree topology or any other implementation specific topology, since [8] provides some flexibility in this context. In case of a tree topology, the root of the tree is a device that initiates a WPAN mesh network and is called a Mesh Coordinator (MC). The MC indicates in its beacon about the tree formation and its willingness to accept other MPNCs as its children. Those MPNCs willing to join the tree send a tree association request command to the MC, which might include the potential number of descendants that they might support. All the MPNCs that become the tree members indicate their ability to relay data on behalf of their peers and/or children. After the tree is formed, the MC collects the tree information by sending a tree topology discovery command frame, to assign tree IDs accordingly. The advantage of using a tree topology is that routing is much easier and is done proactively since the topology is already known. If a source MDEV needs to send data to a destination MDEV outside its piconet, it forwards the data to its MPNC. The MPNC checks the destination tree ID if it belongs to its own tree ID block or if it belongs to one of its children MPNCs tree ID blocks. Based on this information, the MPNC either forwards the data to one of its child MPNCs or parent. The addressing scheme in a tree topology is simpler and the level of a MPNC can be determined from its tree ID block. The superframe can be shared among MPNCs of the same level that belong to the same branch of the tree.

1.5.3 Communication in Meshed WPAN

In a meshed WPAN, communication between MPNCs and within their piconets takes place in a shared superframe. Figure 1.8 shows two MPNCs sharing the superframe with each other. The duration of their superframes is the same but the beacons are shifted by an offset e.g. MAS. Each MPNC has to listen to the beacons transmitted by its neighbour MPNCs that contains important information about their intra-PAN and inter-PAN time reservations. If the beacons from different MPNCs are transmitted subsequently, the time period from the first to the last beacon in the group is defined as a Beacon Period (BP). Each MPNC can reserve time after its own beacon transmission for other MPNCs that may want to join the network. Each beacon transmission contains information about the transmission

of beacons from its neighbours. This helps in avoiding beacon reservation conflicts. For Inter-PAN communication between MPNCs, the source MPNC includes a reservation request Information Element (IE) into its beacon that indicates the TUs and possibly the rate (sub-rate or super-rate) at which it wants that amount of time. Upon the reception, the destination MPNC checks the available capacity of the superframe, and if feasible, replies with a reservation reply IE that also announces the reservation to the neighbouring MPNCs.

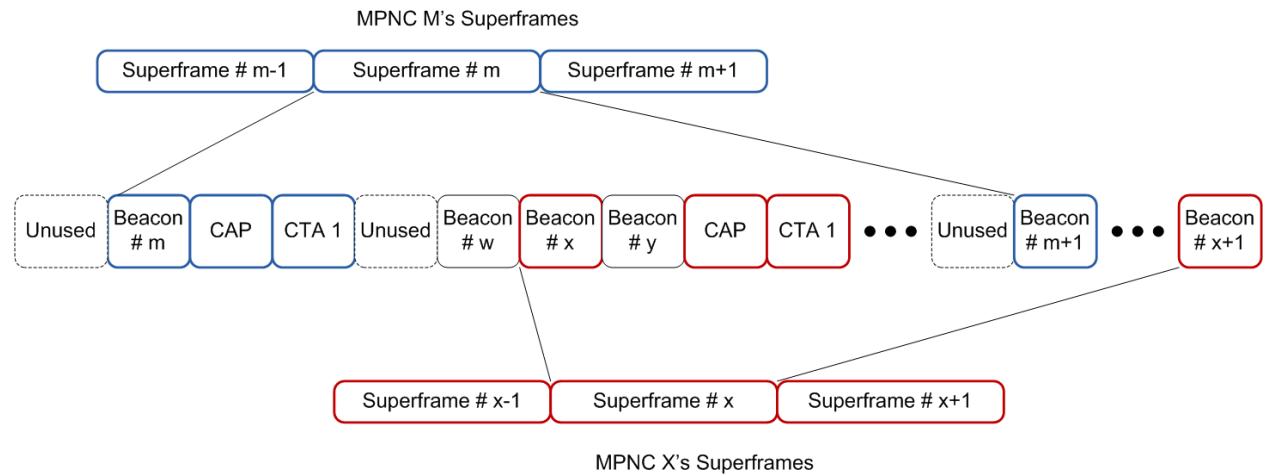


Figure 1.8 Superframe Sharing in HDR Meshed WPANS

Upon the reception of the reservation reply IE, the source MPNC announces the successful negotiation to its neighbours by including a reservation IE into its beacon which stays there until cancelled by one of the MPNCs. During the negotiation process, the neighbouring MPNCs may object to the reservation by transmitting a reservation objection IE into their beacon. When a piconet administered by a legacy 802.15.3 PNC comes in the vicinity of the WPAN mesh network, there is an issue of time sharing since the legacy PNC does not have any algorithm for beacon sharing. Therefore, appropriate time has to be allocated to the legacy PNC to send its superframe. If the legacy PNC utilizes the full duration of the superframe, then it might have to change its channel to allow the meshed WPAN to co-exist. According to [8] if a legacy PNC is found during an initial scanning procedure, the MPNC has an option to associate with the PNC and request for a pseudo static CTA.

Initially when a new MPNC intends to join a meshed WPAN, it synchronizes its slot boundary with a common MPNC within the same operating space called a reference MPNC.

If the meshed WPAN allows the formation of a BP, then the reference MPNC shifts its CAP to allow sufficient transmission time for the new MPNC to transmit its beacon. Each MPNC has an option of either using a common CAP shared with the reference MPNC or to reserve a Management CTA (MCTA) for communication within their piconets. The communication between devices and MDEVs in a piconet is carried out in the same way as described for single hops WPANs. In the case of Inter-PAN communication, the communication takes place between MPNCs in a single hop or multi hop fashion, depending on the distance between the source and destination MPNCs and the network topology. If a MDEV intends to communicate with a MDEV in another piconet, then the source MDEV has to forward the data to its MPNC. Depending on the type of routing used the source MPNC then tries to establish a route to the destination MPNC. The destination MPNC forwards the data to the destination MDEV in its piconet. Each MPNC in the multi-hop path from the source to destination reserves time in the superframe to forward data to the next-hop relay MPNC and to receive it from the previous-hop MPNC. Each MPNC uses the reservation procedure as mentioned before, to reserve time in the superframe for inter-PAN communication.

1.6 Research Objectives and Contribution to Knowledge

The work done in the thesis focuses on providing solutions for some issues present in the single hop HDR WPANs like limited superframe capacity, improving effective peer to peer connectivity and providing admission control in case of congestion. For meshed WPANs, the issue of resource allocation in case of a shared superframe is a non-trivial matter. There is a general solution provided for channel time allocation in a distributed manner in [8]. Another approach is proposed by [19] for fair reservation of channel time by MPNCs in a distributed fashion and an admission control algorithm is also presented. However, due to some disadvantages of distributed channel time reservation that are discussed later in relevant sections, a centralized approach is preferred and proposed in this thesis. In order to realize the centralized approach, traffic estimation techniques and techniques for fair utilization of resources are also proposed. Suitable metrics to measure the fairness of proposed techniques are used to quantify their effectiveness.

For multi-media traffic e.g. voice and video, support in meshed WPANs, the relevant parameters which contribute in determining the *capacity efficiency* and *time efficiency* are identified and their impact on communication overhead is analyzed. By identifying such parameters, a capacity and delay efficient routing metric called *Switched Time Slot* (STS) is proposed and evaluated which helps in finding the path with the least delay and which consumes the least capacity. An upgraded version of the metric is also proposed called *Switched Time Slot with Re-adjustment* (STS-R) which re-orders the time slots once a flow is terminated and time is released in a superframe which belongs to a link in the source-destination path.

Apart from that, a power control based approach is proposed and evaluated to eliminate the hidden node problem to some extent and hence improve capacity utilization in single hop as well and multi-hop meshed WPANs. The formal description of the problem formation and a general overview of the solutions are presented in the sub-sections that follow while the complete details along with related work, performance analysis and simulation results are given in subsequent chapters.

1.6.1 Inter-PAN Communication and Admission Control for HDR WPANs

When two or more independent piconets come in the transmission range of each other, there should be a mechanism for them to share resources and for the devices in one piconet to communicate with devices in the other piconet. Apart from the communication between independent piconets, the parent-child model of the single hop WPANs does not offer a complete and effective peer-to-peer communication strategy. In a parent child communication model, the child PNC can communicate with the parent PNC and the devices in the parent piconet since the child PNC is also a member of the parent piconet. The child PNC can also communicate with the devices in the child piconet. The member devices in the child piconet however, cannot communicate with the parent PNC or the member devices of the parent piconet even if they are in transmission range of each other. The communication model for the parent-child relationship is shown in Figure 1.9.

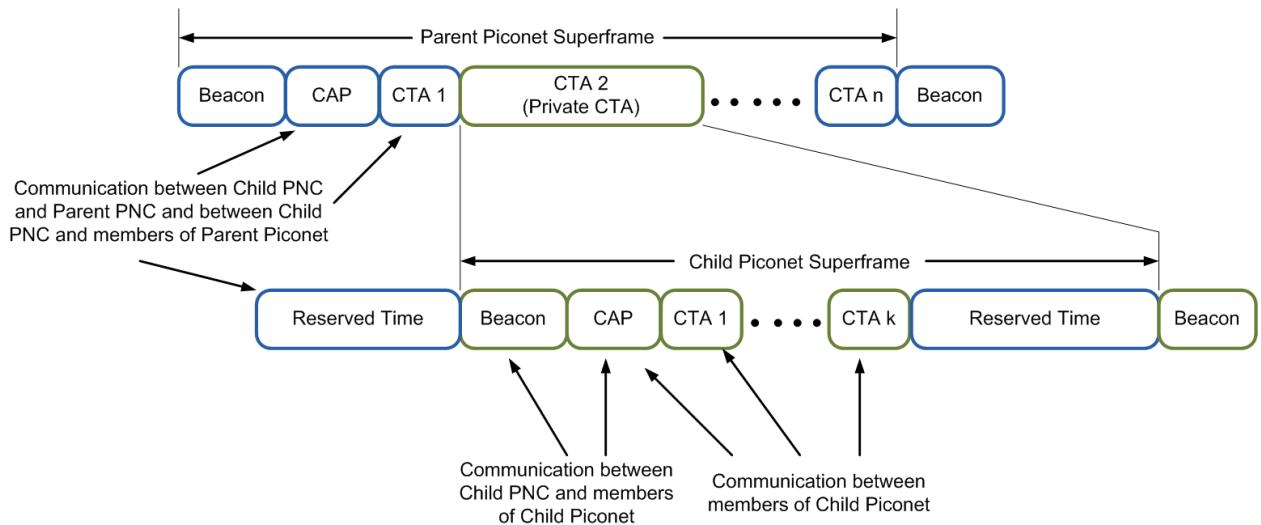


Figure 1.9 Communication Rules for a Parent-Child Model

The HDR WPAN standard [3] suggests that super-rate CTAs for isochronous streams should be spread out evenly in the superframe. Adherence to this suggestion is mandatory especially in case of voice flows because the distance or time offset between two successive CTAs for a voice flow cannot exceed 10ms in order to maintain a satisfactory level of QoS. This requirement imposes additional constraints on the parent-child model because in case of any device in the child piconet that needs to establish a voice flow requires at least 8 CTAs per superframe. Therefore, a single private CTA allocated by the parent PNC cannot suffice for such scenarios. In order to determine a practical number of child piconets that a parent can support, the relevant analysis of superframe capacity in presence of highly demanding flows is also necessary. In case of insufficient superframe capacity to cater to the requirements of all the flows, efficient admission control algorithms need to be implemented. The proposed solution to the problems discussed is presented in chapter 2 with the results of detailed analysis as well as simulations.

An Inter-PAN Communication model is proposed based on the parent-child model to establish complete peer-to-peer connectivity among devices of parent and child piconets. Necessary enhancements are made to the parent-child communication model using the flexibility provided by HDR WPAN standard and hence the proposed model is in adherence to the current specifications of the standard. Since the Inter-PAN model can also be used by independent piconets to enable their member devices to communicate with each other, a

mechanism is required to select the appropriate PNC to act as the parent PNC and the other as the child PNC. Although, a criteria is proposed in [3] for the selection of a PNC, additional parameters are used for the selection in the proposed model keeping in mind some of the constraints involved in Inter-PAN communication. An important parameter which helps determine an upper limit for the number of piconets and the number of devices per piconet that can be supported by an established parent PNC is the capacity of the superframe. In Inter-PAN communication, time is allocated from the same superframe to all the supported piconets. Therefore, detailed capacity analysis is carried out in Chapter 2. In case of many flows with high throughput requirements which can choke the superframe capacity, an admission control algorithm is also proposed in Chapter 2 along with a CTA allocation strategy. The proposed techniques are evaluated with the help of simulations.

1.6.2 Resource Allocation Framework for HDR Meshed WPANs

In HDR meshed WPANs, the superframe time is shared among MPNCs which broadcast their beacon frames in the same Beacon Period (BP). Time can be reserved in the superframe by an MPNC for either an inter-PAN flow or an intra-PAN flow. The procedure given in [8] for time reservation by an MPNC for intra-PAN flows is different from inter-PAN flows. In case the MPNC reserves channel time in the superframe for intra-PAN flows, it reserves the time in the superframe and indicates the time reserved via a CTA status IE and then a CTA IE in its beacon frame. The other MPNCs, upon checking the CTA status IE and the CTA IE in the beacon, can check if the reserved time conflicts with any of the time slots already reserved by them. In case there is a conflict, the other MPNCs can raise an objection IE in their beacon and the MPNC which reserved the time can shift the reserved time slots to another location on the superframe and then wait for an objection again. If there is no objection raised, the MPNC finalizes the reserved time and uses it for intra-PAN flows.

Since for inter-PAN flows, one MPNC might have to send or relay data to another MPNC, time is reserved in a relatively different way. In case an MPNC is a member of a Source-Destination (SD) path and it has to relay data, there are at least two CTAs it needs. One of the CTAs is to receive the data to be relayed from the previous-hop MPNC while the second CTA is to relay the data to its next-hop MPNC in the SD path list. The CTA which it requires to receive the data from its previous-hop MPNC is reserved by the previous-hop MPNC via a

channel time request command. Therefore each MPNC which need to relay the data has to reserve time to send that data to its preferred next-hop MPNC in the SD path. The MPNC does so by sending a channel time request command to the Destination MPNC (DMPNC). If the reservation is feasible to the DMPNC, the source MPNC (SMPNC) indicates the new reservation to its neighbour MPNCs through the CTA status IE in its beacon. If there is any conflict, the SMPNC is notified and the reservation has to be moved to another slot. When there is no objection raised by any of the neighbouring MPNCs about the reservation, the SMPNC can finalize it and start using the newly reserved CTA to relay data or send data to the DMPNC.

The distributed allocation although having some merits, has some drawbacks. First of all each MPNC has no knowledge about the number of devices in other piconets operated by its neighbouring MPNCs and hence no knowledge about the neighbour MPNCs channel time requirements. Therefore, if the number of devices in some piconets is high with a high frequency of intra-PAN channel time requests, the probability of conflict in the channel time reservation increases and as a result, the objection being raised by MPNCs increases. The same is true when an MPNC shares a beacon with a large number of MPNCs which can increase the frequency of inter-PAN channel time reservation requests. Another drawback is that since each MPNC does not have any idea about the channel time requirement of other MPNCs, there can be serious issues of fairness in the channel time reserved by an MPNC. Furthermore, if the number of requests for channel time per piconet is high, the control traffic in sending the channel time request commands increased. There can also be delay experienced by some devices or MPNCs when reserving channel time on request by request basis. The delay can increase when there are objections raised by other MPNCs in case of CTA conflicts.

Keeping in view the above factors about distributed channel time reservation, a centralized approach is proposed for channel time reservation in the thesis. In the centralized approach, one MPNC among each neighbourhood of MPNCs which share a common BP is selected as the reference MPNC (ref-MPNC). The ref-MPNC not only contributes in allocating superframe time to requesting MPNCs in a relatively fair manner, it also serves as a reference for other MPNCs to synchronize their time slots with. In other words, the beacon transmitted by the ref-MPNC is taken as a time offset for other MPNCs to

align their beacon transmissions with. In the centralized approach, the ref-MPNC keeps track of the time reservations in the superframe up to two hops and therefore assuring that there are no conflicts in channel time reservation. The ref-MPNC is also aware of the channel time requirements of its neighbouring MPNCs and therefore can allocate time in a fair manner. Another salient feature of the proposed approach is that the channel time is reserved in bulk by the MPNCs rather than request by request basis to reduce the amount of control overhead. To make the bulk allocation more efficient, the requirement of each MPNC is taken into account by considering the number of devices in each piconet and the traffic requirements of each device.

The requirement of each MPNC is predicted from a traffic matrix which is based on the previous history of time reservations by its member devices. The traffic matrix is used for intra-PAN flows as well as inter-PAN flows. Based on the available superframe time, appropriate algorithms are used to balance the requirement of an MPNC. An overview of the centralized and distributed channel time reservation can be seen in the flow chart for Figure 1.10.

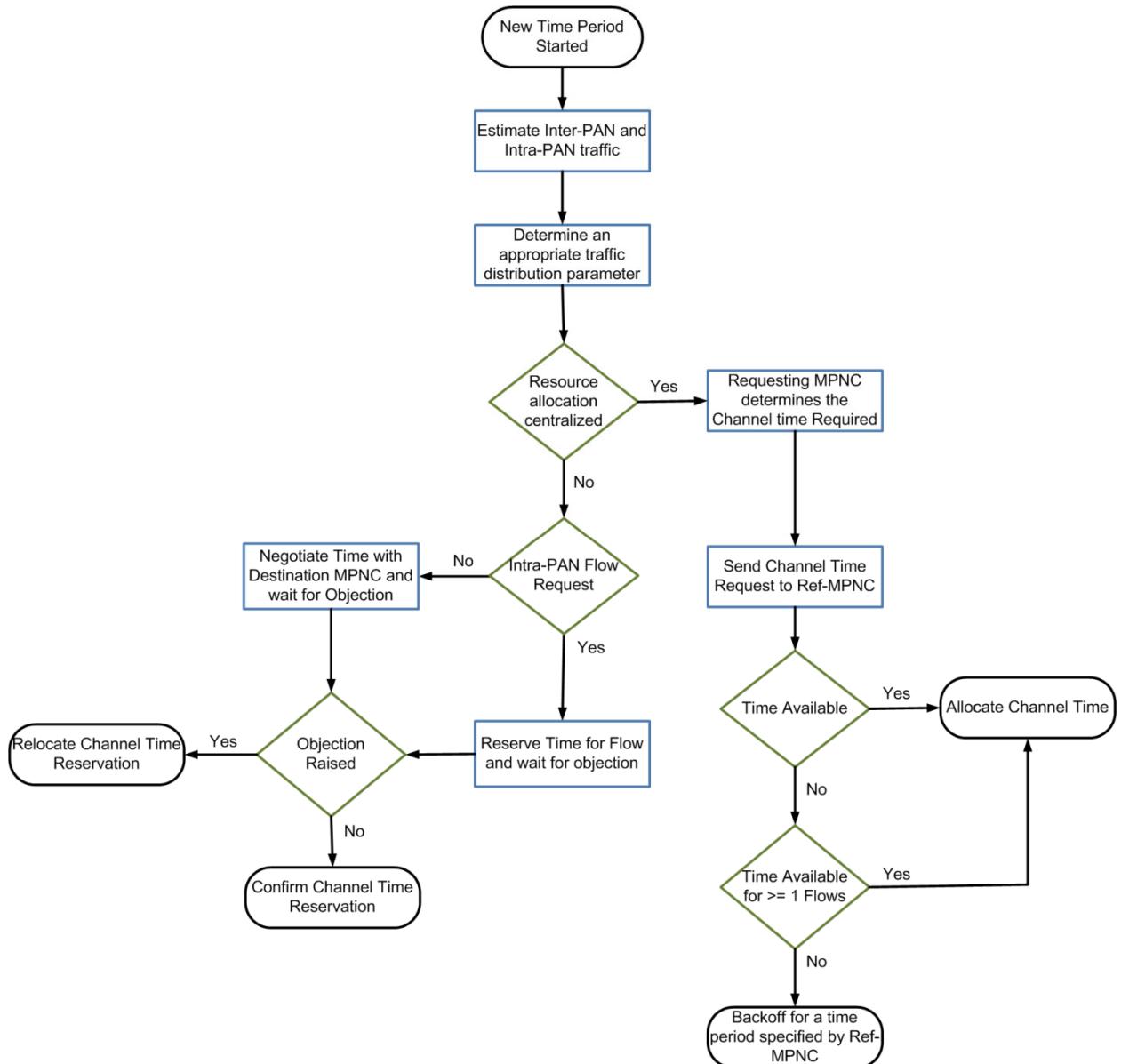


Figure 1.10 Channel Time Reservation Flow chart for Centralized and Distributed Allocation Schemes

1.6.3 Capacity and Delay Efficient Communication in HDR Meshed WPANs

In case of multi-media flows with strict delay constraints, the end-to-end delay along a multi-hop path is an important parameter which needs to be bounded to certain acceptable limits. Apart from the end-to-end delay, some video flows e.g. H.264 with required throughputs above 6 Mbps take a significant fraction of superframe time thus causing capacity issues. The time efficiency of a flow determines its utilization of superframe

capacity. A higher value of time efficiency means that the overhead for a flow is lower while a lower value of time efficiency means that the overhead for a flow is more. The overhead can be calculated for a flow on the level of the CTA or on the superframe level. For the overhead calculation of a CTA, the acknowledgement scheme in use is an important parameter to consider. Since for multi-media flows, the Dly-ACK scheme is the most preferable, the structure of a CTA based on Dly-ACK is shown in Figure 1.11.

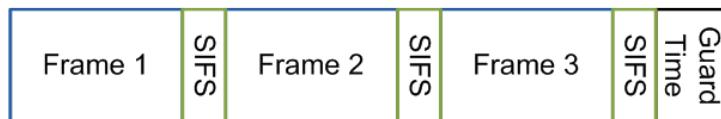


Figure 1.11 CTA structure for Dly-ACK Scheme

It can be seen in Figure 1.11 that in each CTA, there is a Short Inter-Frame Space (SIFS) used between successive frames transmitted and at the end of the CTA. There is also a Guard Time (GT) duration at the end of the CTA. Apart from the overhead of SIFS and the GT, each frame has a MAC header which also contributes to the overhead. The specifications given in [3] recommend that the MAC header should be always sent at the base data rate of 22 Mbps. The rest of the frame can be sent at any supported desired data rate. Since the MPDU along with the MAC header constitutes a MAC frame, therefore, the MPDU size, the transmitter data rate and the number of frames sent per CTA are the main parameters which decide the efficiency of time utilized by the flow on a CTA level. In Chapter 4, an approach is proposed to determine the optimal values of the parameters affecting the time efficiency for a given value of time efficiency.

In multi-hop networks, the end-to-end delay is an important performance metric used for many QoS routing protocols. Since the real time multimedia applications like voice and video are delay sensitive, the specified acceptable limits for delay and jitter in [3] are to be strictly adhered to in order to satisfy the required QoS for voice and video applications. In contention based multi-hop networks, the factors affecting the end-to-end delay are the propagation delay between each hop, processing time at each device along the SD path, queuing delay at each device and the transmission time of a frame at each hop. For TDMA based meshed HDR WPANs, the queuing delay at each hop is a predictable parameter and is not a randomly distributed variable in case of contention based multi-hop networks. The

queuing delay at each hop is the time offset of an MPNC receiving the data from a previous-hop MPNC to the relaying of the same data to the next-hop MPNC in the SD path.

The factor which has the most significant impact on the end-to-end delay is the number of hops along the SD path. The addition of each hop has an impact on all the factors i.e. queuing delay, propagation delay, processing time, transmission time etc. If a tree topology is used for meshed WPANs, then if the source MPNC (SMPNC) and the destination MPNC (MPNC) have different parent MPNCs, then they cannot form a direct path to each other and have to form a path through a common root MPNC. This can add an unnecessary number of hops to the SD path and hence have a significant impact on the end-to-end delay. A distributed topology is therefore preferred on a tree topology for routing in meshed WPANs in terms of end-to-end delay. In order to reduce the end-to-end delay, a CTA switching scheme is proposed in Chapter 4. The CTA switching scheme aims to reduce the time offset between the receiving CTA and the sending CTA for the same flow at each hop and therefore has an impact on reduction of per hop queuing delay. Based on the CTA switching, the STS metric is proposed along with its upgraded version called STS-R. These two metrics share the common goal of not only finding the most capacity efficient path by taking into consideration, the link transmission data rates, but also the total end-to-end delay after re-ordering the time slots into sequence.

1.6.4 Capacity Improvements in WPANs with Power Control

The HDR WPAN standard aims to support multi-media flows with high throughput requirements e.g. 20Mbps or more. However, in practice such high data rates are difficult to achieve due to the access method being TDMA in the CTAP. Although the transmitter data rates can go high up to 55 Mbps for [3], the actual throughput achieved at the MAC layer is much lower. The actual throughput depends on factors like number of CTAs allocated to a flow, the size of the MAC Protocol Data Unit (MPDU), and the number of MPDUs sent in each CTA. Higher values of actual throughput for a device means increased capacity utilization in the superframe or in other words more channel time required in the superframe. When the number of devices with high end throughput requirements is higher, serious capacity issues can arise. The problem is quantified even further when HDR meshed WPANs are considered because the same superframe is shared by MPNCs of different

piconets. Furthermore, in presence of multi-hop communication, the hidden node problem manifests itself. The hidden node problem can be seen in Figure 1.12 in which device B is in the transmission range of both device A and device C. However, device A and device C are not in the transmission range of each other and are unaware of each other's presence. A simultaneous transmission results in a collision and thus there is interference. In TDMA base meshed WPANs, if device A has reserved a CTA in the superframe for transmission to device B, then device B cannot transmit or receive from device C at that time. Device A cannot receive anything from any other device while transmitting to device B. Furthermore, when device B is transmitting or receiving from device C, device A cannot transmit at that time. The hidden node problem can therefore raise significant capacity issues in meshed WPANs.

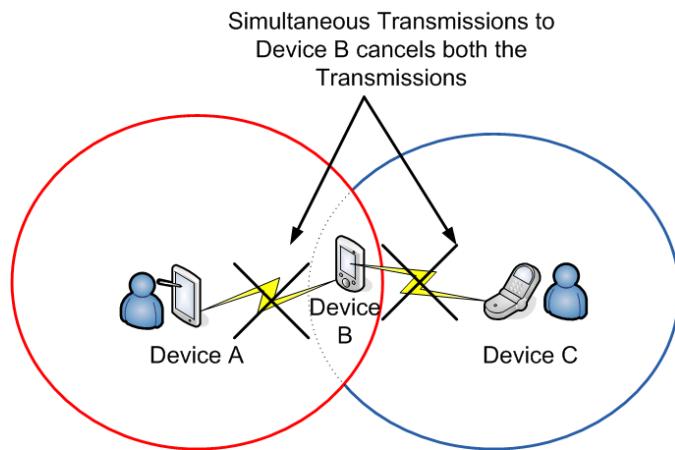


Figure 1.12 Hidden Node Problem

In order to mitigate the effect of hidden node problem on superframe capacity to some extent, a transmit power control scheme is proposed in Chapter 4 in which the possibility of using overlapping CTAs is determined. In order to investigate such a possibility, the devices which are about to initiate a flow or are already involved in a flow transmit simultaneously in pairs in certain CTAs at different transmit powers. If there is no collision on a certain power level at an acceptable SNR level, the pair of devices can use the same CTA at that power level. The same approach is also used for inter-PAN communication.

1.7 Research Methodology

A three stage strategy was followed in achieving the desired goals and objectives of the contributions made which were

- Literature Review
- Mathematical Analysis
- Use of a Simulator for verification of results

In the first stage of literature review, relevant research articles, research papers which included conference proceedings and journal papers, WPAN standards, progress and proposals of WPAN Task Groups TG3 and TG5, and different white papers on WPANs and their intended applications were studied. During the literature review phase, the most important issues of WPANs were identified and the relevant solutions were contemplated upon. The aims and objectives of the research were defined along with a timeline.

After the literature review stage, a need was felt to develop a baseline which can work as a yard stick to compare the investigations and outcomes of the proposed approaches through simulations. Therefore, the mathematical analysis of different parameters was carried out using the values defined in the HDR WPAN standards. The analysis resulted in not only providing a performance baseline but also helped in developing a different perspective into looking at the issues of capacity and resource allocation and thus alternate solutions were also thought of.

To validate the outcome of proposed ideas, the feasibility of two network simulation software's was looked into. The first simulation package was Network Simulator 2 (NS 2) and the second was the OPNET Modeller. NS 2 is an open source package and has the advantage that there are many contributed models available for MANET routing protocols, WLANs, UMTS, WiMAX etc. New models can easily be developed in NS 2 using either C++ or Tool Command Language (TCL). However, to create a simulation environment, define simulation parameters and run a simulation, a TCL script has to be written. Due to the lack of proper documentation and lack of a good debugger, OPNET Modeller proved to be a better option to develop a model of HDR WPAN and meshed HDR WPAN.

The model for HDR WPAN and a meshed HDR WPAN was developed in OPNET using C and OPNET defined APIs called kernel procedures. The specifications of the HDR WPAN standard were taken into account while development. Some changes had to be made to allow for dynamic modulation to use the appropriate modulation scheme with different transmitter data rates of 22, 33, 44 and 55 Mbps just like defined in the standard.

1.8 Organization of Theses

The contributed work along with an insight into the basic concepts and a brief overview of contributions is divided into six Chapters. Chapter 1 gives a basic introduction of WPANs, HDR WPANs and meshed HDR WPANs. It contains an overview of the issues which are tackled in the contributed work and a brief description about the proposed solutions. Some relevant papers are also referred to where appropriate. However, a more comprehensive overview relevant to each contribution and the related work is mentioned in the subsequent chapters.

Chapter 2 is about the Inter-PAN communication framework which also encompasses capacity analysis of HDR WPANs based on voice and video flows with different throughput requirements. The issues raised in supporting real time flows with the parent-child communication model are explained and a solution is proposed. A strategy for CTA allocation by the PNC and an admission control policy in case of congestion is also proposed and evaluated via simulation results.

Resource allocation in meshed WPANs is a non-trivial task due to the superframe being shared by multiple MPNCs. Apart from intra-PAN flows, time has to be reserved for inter-PAN flows in a fair manner. A centralized resource allocation scheme is proposed and evaluated in Chapter 3. Appropriate algorithms for estimating the channel time requirements of different piconets are proposed. Various scenarios based on the allocation of channel time by the ref-MPNC are discussed and analyzed keeping in view the superframe utilization.

Chapter 4 starts by investigating and analyzing the most relevant parameters i.e. transmitter data rate, fragment or MPDU size and number of MPDUs or frames transmitted per CTA which affect the communication efficiency by considering the overhead and the

capacity required for flows with certain throughput requirements. Based on the output of the analysis, a metric is proposed for meshed WPANs which not only considers the most capacity efficient path but also the path with the least end-to-end delay.

In Chapter 5, an approach is proposed to utilize the superframe capacity more efficiently by using transmit power control in order to determine the possibility of overlapping CTAs. The proposed approach takes into account an initial handshake procedure between the source and destination device to agree on a most suitable and feasible transmit power level which adheres to a certain acceptable threshold of SNR. Upon a successful handshake, the possibility of overlapping CTA's is determined by taking into account all the active flows and their possible interference.

Chapter 6 concludes the thesis by discussing the outcomes of the contributed work and possible future work in general.

1.9 References

- [1] "IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, vol., no., pp.C1-1184, June 12, 2007.
- [2] "IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1," *IEEE Std 802.16e-2005 and IEEE Std 802.16-2004/Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004)*, vol., no., pp.0_1-822, 2006.
- [3] "IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 15.3: wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs)," *IEEE Std 802.15.3-2003* , vol., no., pp. 0_1-315, 2003.
- [4] "IEEE Standard for Information technology- Telecommunications and information exchange between systems- Local and metropolitan area networks- Specific requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)," *IEEE Std 802.15.4-2006 (Revision of IEEE Std 802.15.4-2003)* , vol., no., pp. 0_1-305, 2006.
- [5] Hsin-Mu Tsai; Tonguz, O.K.; Saraydar, C.; Talty, T.; Ames, M.; Macdonald, A., "Zigbee-based intra-car wireless sensor networks: a case study," *Wireless Communications, IEEE* , vol.14, no.6, pp.67-77, December 2007.
- [6] C. Gomez, P. Salvatella, O. Alonso, J. Paradells, "Adapting AODV for IEEE 802.15.4 Mesh Sensor Networks: Theoretical Discussion and Performance Evaluation in a Real Environment," *A World of Wireless, Mobile and Multimedia Networks, International Symposium on*, pp. 159-170, 2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks(WoWMoM'06), 2006.
- [7] Gutierrez, J.A., "On the use of IEEE 802.15.4 to enable wireless sensor networks in building automation," *Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004. 15th IEEE International Symposium on* , vol.3, no., pp. 1865-1869 Vol.3, 5-8 Sept. 2004.
- [8] "IEEE Draft Recommended Practice to Standard for Information Technology-Telecommunications and information exchange between systems-Local and Metropolitan networks-Specific

requirements-Part 15.5: Mesh Enhancements for IEEE 802.15 WPANs," *IEEE Draft 15-06-0237-02-0005*, 2006.

[9] Draves, R., Padhye, J., and Zill, B. 2004., "Routing in multi-radio, multi-hop wireless mesh networks," *Proceedings of the 10th Annual international Conference on Mobile Computing and Networking. MobiCom '04*. ACM, pp.114-128. Oct. 2001.

[10] Subramanian, A.P.; Buddhikot, M.M.; Miller, S., "Interference aware routing in multi-radio wireless mesh networks," *Wireless Mesh Networks, 2006. WiMesh 2006. 2nd IEEE Workshop on* , vol., no., pp.55-63, 25-28 Sept. 2006.

[11] Guerin, M. Portmann, and A. Pirzada, "Routing metrics for multi-radio wireless mesh networks," in *Telecommunication Networks and Applications Conference, 2007. ATNAC 2007. Australasian*, 2007, pp. 343-348.

[12] Bruno, R.; Conti, M.; Gregori, E., "Mesh networks: commodity multihop ad hoc networks," *Communications Magazine, IEEE* , vol.43, no.3, pp. 123-131, March 2005.

[13] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer Networks*, vol. 47, no. 4, pp. 445-487, March 2005.

[14] H. Skalli, S. Ghosh, S. K. Das, and L. Lenzini, "Channel assignment strategies for multiradio wireless mesh networks: Issues and solutions," *Communications Magazine, IEEE*, vol. 45, no. 11, pp. 86-95, 2007.

[15] A. Raniwala, K. Gopalan, and T.-C. Chiueh, "Centralized channel assignment and routing algorithms for multi-channel wireless mesh networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 8, no. 2, pp. 50-65, April 2004.

[16] P. Jacquet, P. Mühlethaler, T. Clausen, A. Laouiti, A. Qayyum, and L. Viennot, "Optimized link state routing protocol for ad hoc networks," in *Proceedings of the 5th IEEE Multi Topic Conference (INMIC 2001)*, 2001.

[17] G. Pei, M. Gerla, and T. W. Chen, "Fisheye state routing in mobile ad hoc networks," in *ICDCS Workshop on Wireless Networks and Mobile Computing*, 2000, pp. D71-D78.

[18] S. Mahmud, S. Khan, H. Al-Raweshidy, and K. Sivarajah, "Meshed high data rate personal area networks," *Communications Surveys & Tutorials, IEEE*, vol. 10, no. 1, pp. 58-69, 2008.

[19] M.S. PARK , B. LEE , and S.H. RHEE, "Distributed Multiple Access Control for the Wireless Mesh Personal Area Networks" *IEICE Trans Inf & Syst* E91-D: 258-263.

CHAPTER 2

2 Inter-PAN Communication and Admission Control for HDR WPANS

2.1 Introduction

As discussed before in the previous chapter, the communication in HDR WPANS takes place in a superframe which consists of a beacon, a Contention Access period (CAP) and a Channel Time Allocation Period (CTAP) (Figure 2.1). The CAP is CSMA/CA based and it can be used for sending association requests and other commands if allowed by the PNC. Small amounts of asynchronous data can also be sent in the CAP by the member devices. If the devices require channel time on a periodic basis, they can request the PNC for channel time in the CTAP which is TDMA based. The Channel Time Allocation (CTA) request can be for either a sub-rate allocation (CTAs in alternate superframes) or a super rate allocation (multiple CTAs in the same superframe) depending on the type of traffic the device wants to send and its constraints e.g. frame interarrival time.

To extend the range of the piconet the IEEE 802.15.3 [1] standard allows the formation of child piconets which are dependant on channel time from the established parent piconet. The devices in the parent piconet and the parent PNC can communicate with the child PNC and vice versa. The limitation in this extention is that the devices in the child piconet cannot communicate with the devices in the parent piconet and the parent PNC. In this chapter, the emphasis is on formulation of a strategy to overcome this limitation and hence propose a suitable extention which does not have any conflict with the MAC specifications of IEEE 802.15.3 standard. The proposed approach considers establishing a link between two or more independent piconets to enable Inter-PAN communication in such a way so that one device in one piconet can communicate freely with the intended device which might be in another piconet but in its transmission range. Although by introducing such functionality,

the effective peer-to-peer communication is improved upon and the devices can have more freedom to communicate with any other device, the superframe capacity has to be shared among the piconets. Such sharing can raise capacity issues in presence of a high number of devices in one or more piconets being considered for Inter-PAN communication, especially those device which have high throughput requirements. Therefore, not only the capacity analysis considering different throughput requirements becomes mandatory to establish a baseline for further insights, but an admission control strategy also becomes desirable. The capacity analysis provides a baseline regarding the number of devices that can be supported, which use a certain traffic type to communicate. Apart from that, the usefulness of capacity analysis is further quantified by the fact that the most important parameters which affect the superframe capacity along with the overall CTA overhead per flow are also identified.

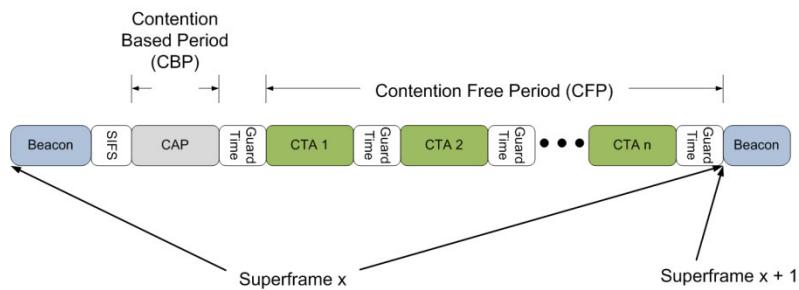


Figure 2.1 Superframe Format

The main contributions in this chapter therefore consist of

1. Identification of issues related to the proposed Inter-PAN communication framework which makes use of the Parent Child communication model
2. An Inter-PAN communication framework which encompasses the establishment of links between two independent piconets along with the disassociation process in case the communication is no longer required. A criteria to select one of the PNCs to be the parent is defined along with a solution to the Parent Child scheduling problem
3. Capacity analysis of the IEEE 802.15.3 superframe by taking into consideration commonly expected flows like voice (G.711) and video (H.264)
4. Specifications for an admission control framework in cases of congestion

5. A parameter to ensure a level of fairness in CTA allocation among devices called the *Rejection Ratio (RR)* is proposed along with its improved version called *Improved Rejection Ratio (IRR)* which allocates time to the most deserving devices which are queued during congestion as the time becomes available when some flows terminate and release their allocated channel time.

The above contributions are evaluated via simulation results considering different scenarios. The simulations also consist of capacity analysis in the presence of more diverse types of flows when compared with the analytical analysis. Also the capacity analysis via the simulation results differ from the analytical capacity analysis because in the later only a single type of traffic is considered for all the devices while in the simulations, the type of traffic for each device might differ from one another hence a relatively more realistic insight. The level of fairness due to the proposed RR approach is evaluated and compared with a case without the use of RR.

2.2 Related Work

For multimedia traffic flows, the provision of respective QoS guarantees becomes an important task to accomplish. HDR WPANs are meant for multimedia traffic flows which usually have high throughput requirements. Since the PNC is responsible for allocating channel time, scheduling flows in a way to satisfy the requirements of maximum or all flows becomes mandatory for it. The IEEE 802.15.3 standard leaves out the details of scheduling to the implementer as an open issue. As a result most of the research work related to HDR WPANs has been targeted to address the issue of scheduling, especially for MPEG-4 flows with Variable Bit Rate (VBR) traffic. The reason for it being the variable nature of the frame size in VBR based flows. Once the device requests for channel time from a PNC, it is allocated channel time equal to the time requested. If the frame size is greater than the allocated time for the device, the frame has to be queued and the device has to increase its allocation by sending a request to the PNC again. However, if the frame size is smaller than the allocated size of the device, the CTA is underutilized. The queue status of a device helps it determine the appropriate size to use in the subsequent superframe. Some scheduling schemes proposed in [2], [3], [4] and [5] are based on a feedback technique. In feedback

based approaches, a device reports its queue status or buffer status to the PNC. The PNC attempts to change the size of CTA for the device accordingly. In [2], the feedback is provided using the CAP at the beginning of the superframe while in [3] and [4], the feedback is provided at the end of the superframe using Management CTAs (MCTAs). In [5], the device sends the feedback known as a status report. The parameters conveyed to the PNC in the form of status report depend on whether the traffic is CBR or VBR. Usually the parameters consist of buffer status, packet transmission delay and PHY layer transmission data rate.

Rest of some well known scheduling approaches proposed in literature consist of [6], [7], [8], [9] and [10]. A multi-channel scheduling solution is proposed in [6] which avoids scheduling conflicts among the channels by using a MULTI-FIT algorithm. To improve the capacity of a superframe and promote efficient time utilization, those transmitter-receiver pairs which are not in the communication range of each other are allocated the same time slots in [7]. However, since the length of time slots vary based on traffic type, the feasibility of the proposed approach needs further investigation. This is true especially in a case where the requirement of one transmitter-receiver pair is substantially different from the other isolated transmitter-receiver pair. In presence of VBR traffic, where the CTA size can change dynamically, further issues such unpredictable increase or decrease in queue size of a device can lead to either in-efficient or in-sufficient utilization of its allocated space. Therefore, the criteria for allocation of the same time slots or CTAs should not just be on basis of transmitter-receiver isolation but also based on the CTA size requirements. In [8], apart from scheduling, the issues addressed are flow admission control and fairness among flows. A scheduling algorithm called Fair-Shortest Remaining Processing Time (Fair-SRPT) is proposed in [9]. The scheduling addresses the CBR and VBR traffic types and fairness is maintained on the frame level i.e. frames with higher priority are admitted first from the queue.

In most of the proposed scheduling algorithms, the performance metric used to determine the scheduler performance is usually the *Job Failure Rate* (JFR). The job failure rate is a measure of those frames which remain in the queue long enough so that their deadline to be decoded exceeds and hence they have to be dropped. Another technique

called *Frame Decodability Aware* (FDA) is proposed in [11] which uses the metric called *Decoding Failure Rate* (DFR) which is a measure of those dropped frames which cannot be decoded by the receiver. Some other proposed approaches which aim to reduce the JFR are given in [12] and [13]. The scheduling algorithms proposed so far suffer from a common drawback. In case of super rate CTAs allocated to a device, then even though it can request a larger sized or smaller sized CTA from the PNC based on the size of high priority frames which are queued, it is not guaranteed that all of its CTAs in the subsequent superframe are utilized efficiently. To elucidate this point, consider the case when a device is being allocated 8 CTAs per superframe. If the device requests the PNC to increase its CTA size because it has high priority frames to send from its queue, but the number of frames is such that they can be sent in the first 2-4 CTAs, then the remaining CTAs can be seriously miss-utilized especially if the rest of the frames with a lower priority have a much lower size.

The IEEE 802.15.3 standard supports different levels of fragmentation to ensure that the appropriate fragment size is used based on the Bit Error Rate (BER) of the channel. In [14], the frame size is changed according to the Frame Error Rate (FER) of the channel. If the FER is high, a lower frame size is preferred while the FER being lower, a higher frame size is preferred. For a time varying channel, a rate adaptation scheme is proposed in [15]. However, the issue of change in CTA size with the change in transmission data rate is not addressed. In [16] a virtual time slot scheme is proposed to improve the capacity utilization by assigning virtual slots to devices. The highest priority to transmit in the virtual slot is given to the owner of the slot. After the owner is done transmitting, the rest of the devices can transmit in their respective order of priority.

An issue which requires considerable attention is the capacity of the superframe which has an impact on the fair allocation of channel time to requesting devices. The capacity analysis of the 802.15.3 superframe is done in [17], [18], and [6] to some extent. In [17], the capacity analysis is done for voice and video flows but the in-depth effect of parameters like transmission data rate, fragment size, number of frames per CTA and number of CTAs per superframe is not discussed. Also the transmission data rate values used, are for IEEE 802.15.3a. The capacity of time-slotted ALOHA based systems is discussed in [18] but only one frame per time slot is considered with fixed slots for the same user. In [6], the same

assumption of one frame per time slot is made which can have an effect on the upper limit of throughput along with the overhead incurred per flow.

In case the number of devices which have high throughput requirements increase, the probability that the superframe capacity becomes congested, increases. In some cases, the need for an admission control algorithm becomes mandatory to decide which flows to admit. Some of the scheduling algorithms discussed, propose fairness which is on the frame level and prefer those frames which have a higher priority. The admission control algorithm given in [2] uses the product of difference between the superframe duration and the total size of expected user requests, and the fraction of time a device requires per superframe. This technique does not guarantee that the requesting device might get the minimum number of Time Units (TUs) that it requires hence no minimum throughput guarantees. The admission control scheme given in [8] is relatively straight forward and a flow is admitted if there is enough available time in the superframe, otherwise rejected. Such an admission control scheme does not guarantee fairness among requesting devices. The admission control scheme given in [19] gives priority to certain service classes or in other words those devices which request for channel time for a certain traffic type. However, the proposed approach in this chapter considers fairness from the user's perspective and is unrelated to the traffic type. The aim of the admission control scheme is to regulate the flow allocation in such a way that most of the requesting devices are satisfied. In other words, the algorithm attempts to reduce discrimination in the CTA allocation. To analyze the proposed Inter-PAN communication model to improve peer-to-peer connectivity, a thorough capacity analysis of the IEEE 802.15.3 superframe is carried out along with the effect of each parameter on capacity. The output of the capacity analysis also serves as a guideline for the design parameters as well as the suitable parameters for the simulation environment to test and evaluate the proposed concept of admission control and inter-PAN communication in this chapter.

2.3 Communication Issues in High Data Rate (HDR) WPANs

In order for a device to communicate and access resources in a piconet, it associates itself with the PNC of that piconet. Upon successful association, all the devices in the

piconet are notified about the capabilities of the newly associated device along with its supported data rates, fragment size and the assigned device ID (DEVID). A device can send either asynchronous data in small amounts to other devices or isochronous data. The PNC can allow small amounts of data to be sent by a device in the CAP if the time required by the devices is not on periodic basis. If a device requires channel time on a periodic basis for isochronous flows, then it can send a channel time request command to the PNC for guaranteed time slots in the CTAP. Since different types of traffic i.e. voice, video have different QoS requirements e.g. tolerable delay, jitter, a required data rate etc the device calculates the necessary time it requires for a CTA and the number of CTAs per superframe. The device also decides the number of frames to be sent per CTA. The device then indicates through a channel time request command to the PNC, the minimum and desired number of Time Units (TUs) it requires.

A TU is a unit of time of the CTA with a resolution of $1\mu s$ to make the allocation of time easier for the PNC. The PNC after getting the channel time request from the device checks the available capacity of the superframe and sends a channel time response command to the requesting device indicating the available number of TUs and acceptance or rejection of the request. If the PNC allocates the channel time to a device based on the available time in the superframe only, then most of the devices might not get a fair share of the superframe. The PNC while allocating the channel time should ensure an approach such that the maximum number of devices in the piconet get a fair share of superframe time based on their requirements and maintaining the QoS at the same time.

2.4 Parent-Child Communication Model and its Limitations

The high data rate WPAN standard, 802.15.3 defines the transmitter data rates (δ_T) of 11, 22, 33, 44 and 55 Mbps. The beacon and MAC headers of the frames are sent at the base rate of 22 Mbps and the rest of the payload at any of the desired values of δ_T . Since the CTAP is TDMA based, it is not possible to achieve the defined δ_T . If a device wants to send small amounts of asynchronous data in a single CTA, then it can transmit and achieve the defined data rates. For isochronous transmission, the requirement is to allocate more than one CTA per superframe (depending on the tolerable inter arrival delay) for the device. The actual throughput (δ_A) is always less than δ_T and depends on certain factors such as number

of CTAs allocated to the device, number of frames sent in each CTA and the time duration of each CTA (which depends on the required throughput). The number of devices in a piconet influences the decision of the PNC to allocate a particular number of CTAs to a device to ensure fair allocation.

Theoretically there can be 256 devices supported by the PNC in a piconet. Since some of the Device IDs (DEVIDs) are reserved for special purposes [as given in the 802.15.3 standard], the maximum number of devices that a single PNC can support as allowed by the 802.15.3 standard, is 243. The practical number of devices that a single PNC can support is, however, much lower than 243 if multimedia transfers are taking place between most of the devices. The increased number of devices also imposes additional processing overhead on the PNC. To resolve the processing burden and extend the range of piconet, the 802.15.3 standard allows the formation of child piconets which are dependant on the established parent piconet. Though the administration of the child piconets is done autonomously by a child PNC, the channel time is provided by the parent PNC from its transmitted superframe through a private CTA. It can be seen in Figure 2.2 that the time period in the superframe of a child piconet after the private CTA is reserved till the start of a subsequent private CTA in another superframe of the parent piconet. This is to keep synchronized with the time allocated by the parent PNC to the child PNC. Figure 2.2 also indicates an issue related to the child superframe about the allocation of super rate CTAs for isochronous streams with strict delay constraints. If the reserved time after the private CTA allocated to the child PNC exceeds the maximum tolerable delay for most of the real time applications, then it is not possible for the child piconet to support them. If the reserved time is decreased by increasing the time allocation of the private CTA, it can disturb the CTA allocations in the parent piconet.

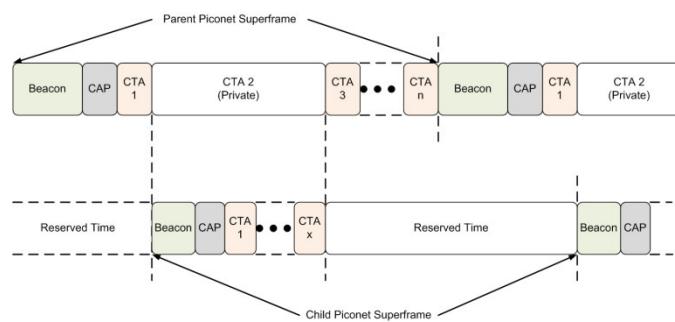


Figure 2.2 Child Superframe Time Allocation

2.4.1 Scheduling Problems in the Parent-Child Communication Model

As explained before, it is difficult in a parent child relationship to maintain QoS for certain multi-media applications especially voice. The issue is further aggravated if the formation of a child piconet within a child piconet is considered. Figure 2.3 shows us such a formation which has a two level hierarchy.

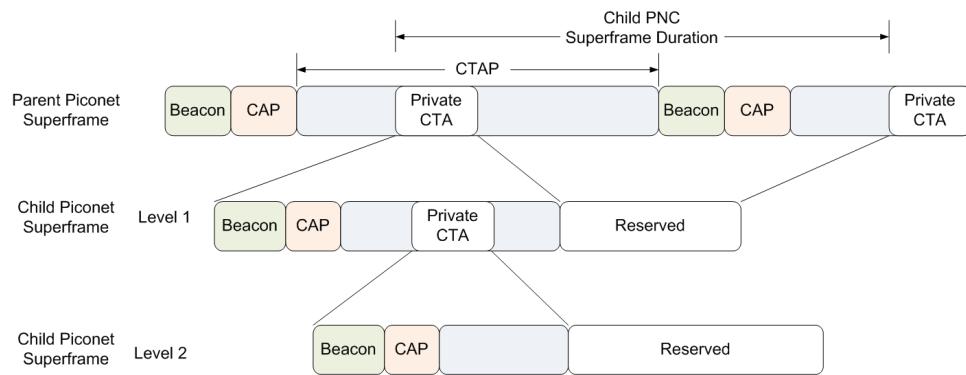


Figure 2.3 Time Allocation for Hierarchical Child Piconets

If η_{PS} is the duration of the parent piconet superframe and η_{C1S} is the duration of the superframe for the level 1 child piconet, then

$$\eta_{C1S} = \eta_{PCTA1} + \eta_{(RSVD)1} = \eta_{PS}$$

$$\eta_{PCTA1} = k_1 \eta_{PS} \quad (0 < k_1 < 1)$$

Where η_{PCTA1} is the duration of a private CTA allocated to the level 1 child PNC and η_{RSVD} is duration of the reserved time in η_{C1S} for η_{PS} till the start of next successive private CTA for the level 1 child PNC. Similarly for the level 2 child piconet which is formed within the level 1 child piconet as follows:

$$\eta_{C2S} = k_2 \eta_{PCTA1} + \eta_{(RSVD)2} = \eta_{PS}$$

$$\eta_{PCTA2} = k_2 \eta_{PCTA1} \quad (0 < k_2 < 1)$$

Where η_{C2S} is the duration of level 2 child piconet superframe and η_{PCTA2} is the duration of a private CTA allocated to the level 2 child PNC by the level 1 child PNC while $\eta_{(RSVD)2}$ is the reserved time for the level 1 child piconet and the parent piconet. The value of k_2 is less than 1 to indicate that $\eta_{PCTA2} < \eta_{PCTA1}$. The number of super rate CTAs allocated to a device which is sending a real-time traffic depends on the maximum tolerable delay and jitter for

that particular traffic type and the required throughput ∂_R . Since the super rate CTAs are evenly spread throughout the superframe, the duration of a private CTA allocated to a child PNC is a significant factor to determine if the parent and child piconet can support a particular real time traffic type with specific requirements of maximum tolerable delay and jitter. If X_{MTD} denotes the value of maximum tolerable delay and jitter for a particular real time traffic type, then the superframe can be split into logical partitions to make time allocations easier. The smallest partition size is taken to be equivalent of the strictest requirement for delay and jitter, which according to [1] is for voice (<10ms). If $X_{MTD(min)}$ denotes the minimum compulsory logical partition size for the superframe, then

$$X_{MTD} = n X_{MTD(min)}$$

Where $n \in \mathbb{Z}^+$ (set of positive integers) and the value of n is set to 1 to indicate voice applications. The number of logical partitions can be found out by

$$\text{Number of Partitions } (N_p) = \frac{\eta_{PS} - (\tau_{PNC} + \eta_{CAP})}{X_{MTD(min)}} \quad (2.1)$$

where τ_{PNC} is the beacon overhead and η_{CAP} is the CAP duration. The expression $\eta_{PS} - (\tau_{PNC} + \eta_{CAP})$ gives us the CTAP duration which is divided by $X_{MTD(min)}$ into a number of partitions. If the value of $X_{MTD(min)}$ is taken to be 8ms, then the superframe is split into 8 partitions, each of approximately 8ms. Once the superframe is partitioned, the time can be allocated much easily for real time applications keeping the boundaries of logical partitions in view. However, the time allocation for a private CTA should be done very carefully as it can have a significant effect on isochronous streams with super rate CTA allocations.

If $\eta_{RSVD} > X_{MTD(min)}$ and $\eta_{PCTA1} > X_{MTD(min)}$, then both the parent piconet and the child piconet cannot support voice applications as required. If $\eta_{RSVD} < X_{MTD(min)}$ and $\eta_{PCTA1} > X_{MTD(min)}$, then child piconet can support voice applications but the parent piconet cannot. In order for both the parent piconet and child piconet to support voice applications (since they have the strictest upper limit on tolerable delay and jitter) the following two conditions must be true:

$$\eta_{RSVD} < X_{MTD(min)} \text{ and } \eta_{PCTA1} < X_{MTD(min)}$$

It can be shown that the above two conditions cannot be true at the same time if the child PNC is allocated only a single private CTA. Since $\eta_{RSVD} < X_{MTD(min)}$ and $\eta_{PS} = \eta_{C1S}$, with $\eta_{PCTA1} + \eta_{RSVD} = \eta_{PS}$. If η_{RSVD} is assumed to be equal to $X_{MTD(min)}$, then $\eta_{PCTA1} = \eta_{PS} - X_{MTD(min)}$. This

means that $\eta_{PCTA1} = N_p - 1$ and thus takes the major portion of the parent superframe. Therefore the parent piconet cannot support voice applications. The same theory can be applied to other traffic types as well. This shows us that since the level 1 child piconet cannot support voice applications, there is no possibility for a level 2 child piconet or above to support multimedia applications for a single private CTA allocation. The increase or decrease in δ_T determines the length of the CTA required to send a particular type of traffic. With higher values of δ_T , the overhead per CTA increases but the capacity of superframe also increases due to the reduced size of CTAs required by devices. IEEE 802.15.3a defines an alternate physical layer based on UWB to achieve much higher data rates (Table 2.1) using the same MAC layer of 802.15.3. Even higher data rates are proposed in IEEE 802.15.3c in Gbps for the 60 GHz frequency band. Although by using much higher data rates, the capacity of the superframe is increased and much smaller CTA durations can be used using frame aggregation, the spacing of the super rate CTAs depending on the factor $X_{MTD(min)}$ does not change.

Table 2.1 Important Parameters of HDR WPANs

Parameters	Values			Units
	IEEE 802.15.3	IEEE 802.15.3a	IEEE 802.15.3c	
Superframe Duration	0 - 65535	0 - 65535	128000	μs
CAP Duration	0 - 65535	0 - 65535	0 - 65535	μs
SIFS	10	10	2.5	μs
MIFS	2	2	0.05	μs
Supported Data Rates	11, 22, 33, 44, 55	55, 80, 110, 160, 200, 320, 480	*HRT > 2000 – 4679 **MRT: 100 – 2000 ***LRT < 100	Mbps
Fragment (MPDU) size including FCS	64, 256, 512, 1024, 1280, 1536, 1792, 2048, 4024 (802.15.3a), 65535 (802.15.3c)			Octets

*HRT: High Rate Transmission

**MRT: Medium Rate Transmission

***LRT: Low Rate Transmission

Based on mentioned observations, it is concluded that the main limitation of Inter-PAN communication through the Parent-Child model is the capacity of the superframe. This issue can further aggravate when the Inter-PAN communication process involves more than two

piconets. Therefore, after proposing the Inter-PAN communication process, the results of extensive capacity analysis are presented in order to quantify the importance of an admission control scheme and fair CTA allocation for Inter-PAN communication.

2.5 Inter-PAN Communication Model

The main idea to enable Inter-PAN communication between independent piconets is to relax the limitation of devices in different piconets which cannot communicate with each other. The Parent-Child communication model which is defined in the IEEE 802.15.3 standard, is mainly to extend the communication range of a piconet and alleviate the processing burden of a single PNC. However, the mechanism of devices in the child piconet to communicate with the devices in the parent piconet is not defined in the standard. Therefore, in order to enable effective peer to peer connectivity among independent piconets, the following important considerations are taken into account for the proposed Inter-PAN communication model:

1. Seamless merging and splitting of two or more piconets
2. All devices in different piconets are able to communicate with each other provided they are in the transmission range of each other
3. The Inter-PAN communication process is strictly in accordance with the specifications of IEEE 802.15.3 and only the reserved fields are used to accomplish the task
4. The scheduling issues proposed in case of the Parent-Child Model are resolved so that the QoS for real time applications is not affected

When two or more piconets merge for the purpose of Inter-PAN communication, the existing Intra-PAN information such as the active associations (of member devices) and communication (allocated CTAs) is retained so that ongoing Intra-PAN flows are not affected. All devices in Inter-PAN communication can communicate with each other directly but the channel time is monitored by the respective PNCs to avoid CTA conflicts.

2.5.1 Inter-PAN Communication

The Inter-PAN communication process is started when a PNC discovers the existence of another nearby piconet in its communication range. The discovery can be as a result of beacon reception by the PNC with a different Piconet Identifier (PNID) than its own or it can

be as a result of active or passive scanning of different available radio frequency channels by either the PNC or one of the member devices. If both the piconets are operating on the same radio frequency channel, then they can discover each other with the respective beacon reception. However, if both the piconets are operating on different radio frequency channels, then they can discover each other with the help of periodic scanning either performed by the PNC or one of its devices. The scanning process for a device (abbreviated as DEV in the figure) is summarized in Figure 2.4. The notations which are used in all of the figures related to Inter-PAN communication are the same as those in the IEEE 802.15.3 for convenience. The term DME stands for *Device Management Entity* and MLME stands for *MAC Layer Management Entity* which are the logical modules as defined by [1].

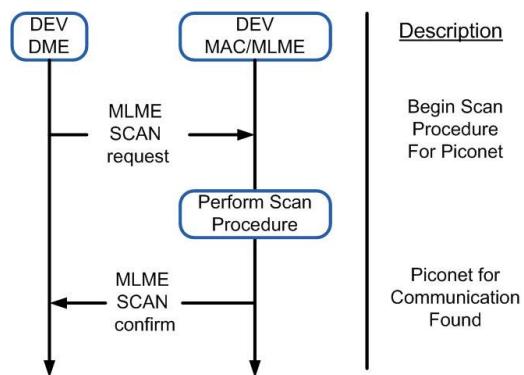


Figure 2.4 Scan procedure initiated by a device

The PNC allocates a CTA to the device in order to perform channel scanning. If the PNC initiates the channel scanning, it goes into a silent mode where it suspends the beacon transmission for one or more beacon intervals but it is not allowed to suspend beacon transmission for more than twice the parameters called *aMinChannelScan* [1]. When the desired piconet for communication is discovered, one of the PNCs needs to be associated with the other PNC as a member device. The criteria for PNC selection for Inter-PAN communication is proposed in section 2.5.2, which is in addition to the criteria defined by [1]. After the selection criteria, one of the PNC is selected and the other sends an association request to the selected PNC as a device. The association process is given in Figure 2.5. For descriptive purpose, one of the PNCs which sends an association request to the other is referred to as a device (DEV) in Figure 2.5. The device ID (DEVID) in the association request is set to the Unassociated ID (UnassocID) in order to indicate to the target PNC that the

requesting device is not currently associated with it. After successful association, the member device needs to form a child piconet which encompasses its own devices.

The member device requires its own CTA in the superframe in order to run its own child piconet. It therefore sends a channel time request command to the PNC to request a private CTA. For the private CTA request, the source ID and the destination ID are kept the same so that the PNC upon reception of the request knows that its for the formation of a child piconet. If the PNC has appropriate time available in the superframe, it allocates, the requested private CTA to the device to allow it to form its own child piconet. If the PNC rejects the formation of a child PNC for any reason such as insufficient channel time or unable to allocate a pseudo-static CTA, it sends a Channel Time Response command with the Reason Code field set to '*request denied*'. In this case, inter-PAN communication is not possible and the device dissociates from the current piconet and return to its own piconet.

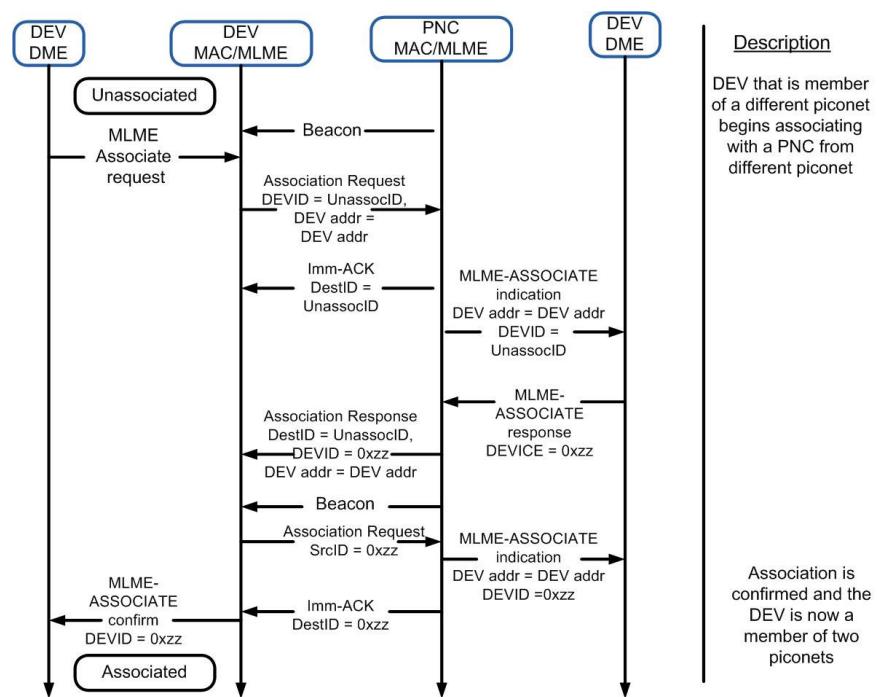


Figure 2.5 Inter-PAN association process

If the child PNC was operating its independent piconet previously on a different radio frequency channel then after associating with the other PNC and formation of its child piconet, it needs to send a beacon on the radio frequency channel on which the parent piconet is operating. Therefore, before the child PNC can begin transmitting its own beacon, it returns to its existing piconet channel and initiates moving the current channel to the

newly allocated child piconet channel. The child PNC broadcasts the *Piconet Parameter Change* Information Element (IE) with the change type set to *CHANNEL* in its current channel via its beacon for *NbrOfChangeBeacons* [1] consecutive beacons. The Piconet Parameter Change IE contains the channel index of the new channel to which the PNC intends to move the piconet, and the *Change Beacon Number* field that contains the beacon number of the first beacon with a beacon number equal to Change Beacon Number field in the previous Piconet Parameter Change IEs. The device receiving this message changes from its current channel to the new channel before the first expected beacon on the new channel. The devices do not transmit on the new channel until a beacon is correctly received on the new channel. To enable every device in the child piconet and parent piconet to communicate with one another, all members of the child and parent piconet need to associate with the respective target piconet PNCs i.e. the members of child piconet need to associate with the parent PNC while the members of the parent piconet need to associate with the child PNC.

In order to initiate Inter-PAN communication, a new command frame is used called *Inter-PAN Association Request*. The command frame is sent by either the child or parent PNC or both PNC to its members. This new command frame has a type value ‘011’ which indicates that it is a Command frame. The PNID is set to PNID of the originating piconet. The SrcID is set to the PNC’s DEVID and the DestID is either set to BcastID if the PNC requires all its members to Inter-PAN Associate or to individual DEVID if it requires only a specific device to associate. The ACK policy bit is set to ‘01’ which indicates Immediate Acknowledgement (Imm-ACK). The Inter-PAN Association Request MAC Frame payload has the form *{src PNID, dest PNID, src BSID, dest BSID, src PNC address, dest PNC address}*.

If the Inter-PAN association request is sent by the child PNC, the the source PNID (src PNID), source BSID (src BSID) and the source PNC address (src PNC address) are set to the that of the child piconet and the child PNC while the destination PNID (dest PNID), destination BSID (dest BSID) and destination PNC address (dest PNC address) are set to that of the parent piconet and the parent PNC. Based on these parameters, the devices in the child piconet start listening to the beacon of the parent PNC in order to associate with the parent PNC. The parent PNC also sends a similar Inter-PAN association request to its member devices and its member devices then start listening to the beacon of the child PNC

in order to associate with the child PNC. The Inter-PAN association process is given in Figure 2.6 in which the child PNC is referred to PNC-1 while the parent PNC is referred to PNC-2.

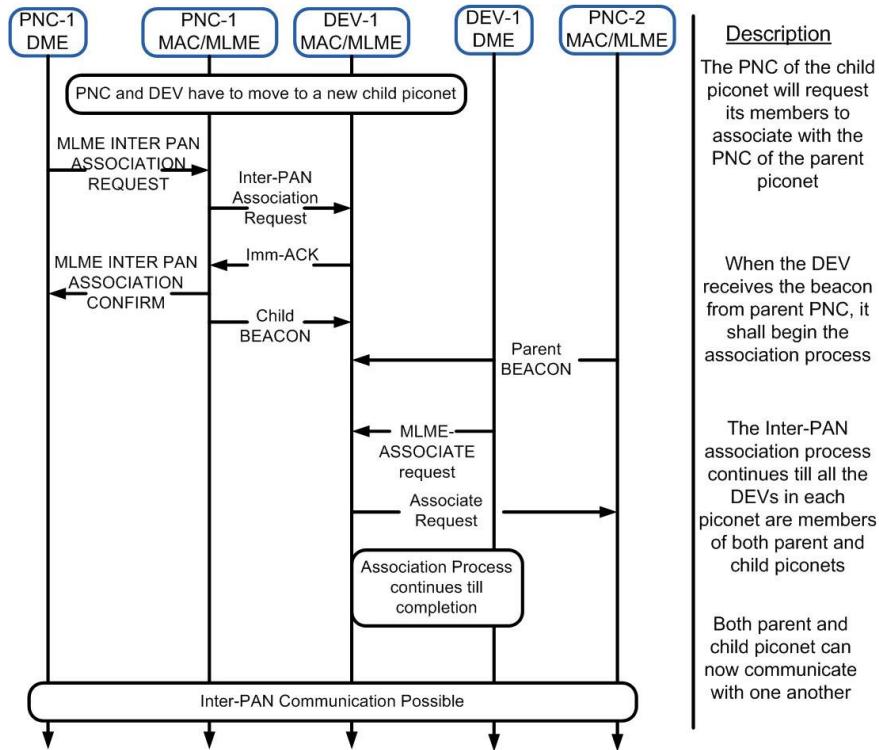


Figure 2.6 The Inter-PAN Association Process

Once the devices are associated with the target PNCs, the communication can commence normally just like defined in [1]. If for security reasons, the Inter-PAN communication needs to be limited to one way i.e. either from parent piconet to child piconet or from child piconet to parent piconet only, then in such a case, the non-conformant PNC sends an association response to the requesting device with the result code *Inter-PAN communication not possible*. This additional result code is used by taking advantage of the reserved fields in the association response frame.

The process of Inter-PAN splitting or Inter-PAN dis-association is an extension to the procedure in which a dependant piconet ends its relationship with its parent PNC. If two or more piconets are finished communicating with each other, the devices need to be informed in order to dis-associate themselves from the PNCs of the target piconets. Since the Inter-PAN communication can involve more than two piconets, the splitting might be between two or more piconets. The devices needed to be informed appropriately so that they know which piconet to dis-associate from and whether to dis-associate from one, or all

of the piconets. In order to ensure a seamless splitting, a new command frame called *Piconet Splitting Request* is used. Either the child PNC sends the Piconet Splitting Request to the parent PNC or vice versa requesting to split from one another. The Piconet Splitting Request command frame has a type value ‘011’ which indicates that it is a Command frame. This process is described in Figure 2.7. The PNID is set to PNID of the originating piconet. The SrcID is set to the PNC’s DEVID of the originating piconet and the DestID is set to PNC’s DEVID of the destination piconet. The ACK policy bit is set to ‘01’ which indicates Immediate Acknowledgement (Imm-ACK). The MAC Frame payload is kept empty. Upon receiving Piconet Splitting Request command frame, both the child and the parent PNCs start informing their devices to disassociate themselves from the inter-PAN associated piconets. The new command frame for dis-association is called *Force Inter PAN Disassociation Request* (See Figure 2.8). The PNID is set to PNID of the originating piconet. The SrcID is set to the PNC’s DEVID of the originating piconet and the DestID is set to BcastID of its member piconet. The ACK policy bit is set to ‘01’ for Immediate Acknowledgement (Imm-ACK). The MAC command frame for the Force Inter-PAN Disassociation Request the following form {src PNID, dest PNID, src BSID, dest BSID, src PNC address, dest PNC address, Mass Disassociation}. The new field which is used in the Force Inter-PAN Disassociation Request is the Mass Disassociation field which is one bit in size. If set to zero, it means that the devices only dis-associate themselves from a single piconet which is specified in the command. Otherwise, if set to one, it indicates to the member devices to dis-associate themselves from all the piconets to which they are associated with.

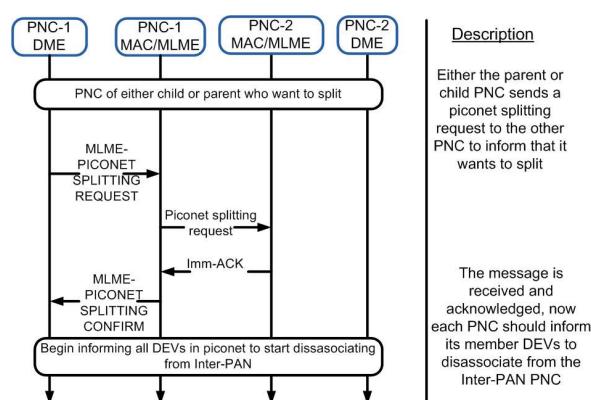


Figure 2.7 Inter-PAN Piconet Splitting Process

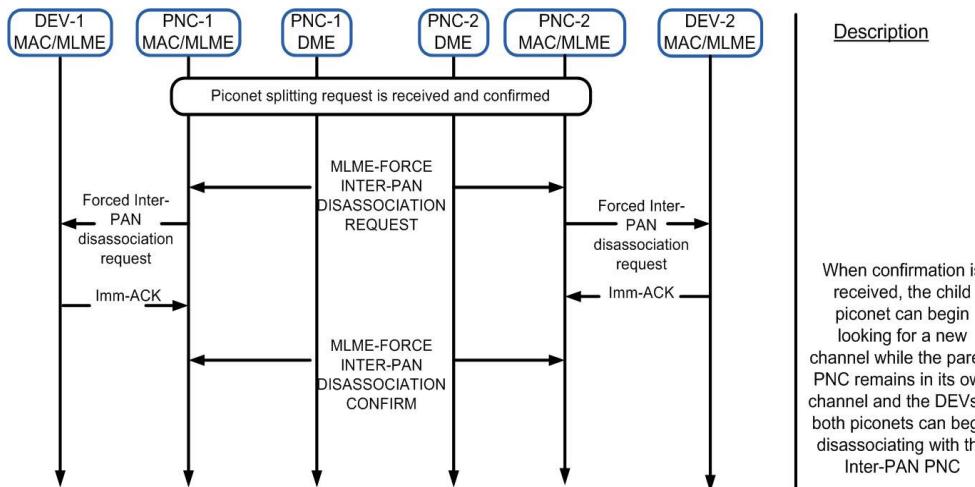


Figure 2.8 Force Inter-PAN Disassociation Process

When devices receive the Force Inter-PAN Disassociation Request message, they initiate the disassociation process with the PNCs of destination piconets. The disassociation process is similar to the one described in [1] and shown in Figure 2.9. A new Reason Code in the Disassociation Request called *Inter-PAN Split* is created for the piconet splitting procedure. The Reason Code uses one of the reserved fields that are available in the Disassociation Request fields. The parent piconet remains in its own channel once the piconet splitting request is initiated while all child piconets either shut down which joined as a dependant piconet, or move to a different channel. If the child piconet decides to maintain its piconet, it begins scanning for a new channel to move its network.

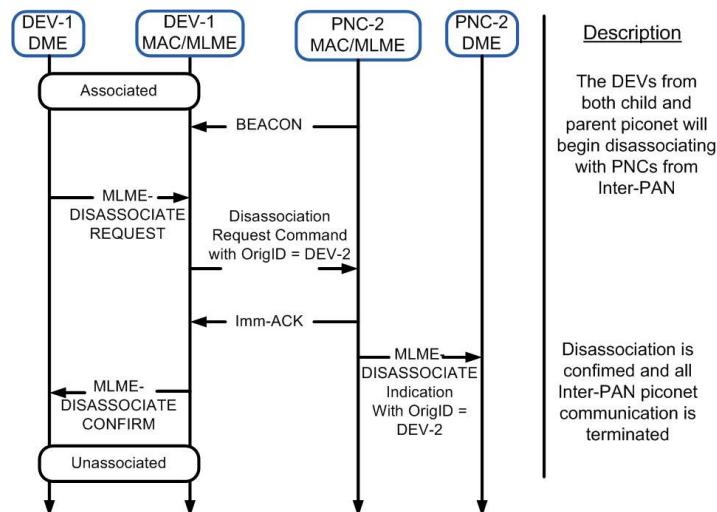


Figure 2.9 Disassociation Process

2.5.2 Inter-PAN Selection Criteria

When a PNC receives a beacon from another PNC having a different PNID, it includes the PNC capabilities IE in its subsequent beacon frame. For inter-PAN communication, since capacity is a major issue and both the devices acting as PNCs are PNC capable, four extra parameters are defined and used in the analysis and simulations apart from the ones mentioned in [1]. The four parameters are number of supported child piconets, number of active devices, type of traffic being communicated by the devices and their CTA durations, and the PNID. The PNC which already has dependent child piconets is given preference. If none of the PNCs are already supporting child piconets, then the PNC which has higher number of active devices communicating in its piconet is given preference. If the number of devices is the same, then the PNC with the highest superframe utilization is given preference. If none of the above is applicable, then the PNC with the higher PNID becomes the parent PNC. Since both of the PNCs perform this comparison, therefore the child PNC sends an association request to the parent PNC. The child PNC also informs its member devices and starts including the parent piconet IE in its beacon. The child piconet calculates the total utilized time in its superframe along with extra time (500 μ s more per superframe in the simulation model in case it receives new requests apart from already allocated channel time) that it requires and sends the request to the parent PNC. If the child piconet has devices that are communicating voice or video traffic, the parent PNC allocates 8 private CTAs to the child PNC. If there is no device with voice or video traffic in the child piconet and no device in the child piconet is capable or have any intention to send such traffic in future, then the parent PNC allocates a single private CTA to the child PNC. The efficiency of CTA utilization, CTA overhead, number of devices that can be supported are given in a section 2.6.

2.5.3 Parent-Child Scheduling Solution

When one of the PNCs associates with the parent PNC as the child PNC, it sends a channel time request command to the parent PNC for channel time. The source device ID (DEVID) and the destination DEVID are the same in the channel time request command so that the parent PNC can determine that it is a request for the private CTA from the child PNC. Upon reception of the channel time request command, if there is enough capacity in

the superframe of the parent PNC, it accepts the request of the child PNC and sends a channel time response command. If no device in either the parent piconet or the child piconet is supporting any real time traffic with a particular value of $X_{MTD(min)}$, then the parent PNC allocates a single private CTA to the child PNC. However, if either a device in the parent piconet or the child piconet or in both the piconets intends to request channel time for real time traffic with a certain value of $X_{MTD(min)}$, then if the parent PNC allocates a single private CTA to the child PNC, then the QoS for the device in the parent piconet and child piconet can be affected as explained. Since the upper limit on the tolerable delay and jitter for voice applications are the strictest, the CTAP of the superframe is partitioned into equal sized slots called Medium Access Slots (MASs). The concept of dividing the superframe into MASs is defined in [20], however, an appropriate size is not specified. We define the size of the MAS to be 8 ms so that the QoS for voice applications can be supported easily and conveniently. If the maximum size of superframe is considered i.e. $65535 \mu s$, there has to be atleast 8 CTAs per superframe for a certain device to support voice applications. Therefore, the value of N_p comes to be 8 and if there are any flows in the child piconet which intend to communicate voice flows, then a minimum of 8 private CTAs are allocated to the child piconet. In this way, the scheduling problem in the parent-child model is relaxed. However, allocating more than one private CTA to a child piconet contributes towards the capacity issue. Since the QoS requirements of video are more relaxed than voice [1], the CTA rate factor for video traffic can be in factors of 2 per superframe according to the throughput requirements and the available capacity in the superframe. The proposed structure of the superframe when inter-PAN communication is considered is shown in Figure 2.10.



Figure 2.10 Superframe Sharing in Inter-PAN Communication

It can be seen in Figure 2.10 that there is a Beacon Period (BP) in which the parent and child PNCs send their beacon. The BP can be extended in presence of multiple piconets and more than two beacons can be sent in it. A single CAP is shared between the parent and child piconets for simplicity so that the inter PAN association requests by the devices from

either the parent PNC or the child PNC can be sent in it. When the parent PNC receives a request for a private CTA from the child PNC, it checks the requested CTA duration η_{CTA} and compares it with the available time in all of the 8 MAS durations (η_{MAS}). If $\eta_{CTA} < \eta_{(MAS-A)i}$ (where $\eta_{(MAS-A)}$ is the available time in a MAS and the index i indicates the MAS number and $1 \leq i \leq 8$), then the parent PNC accepts the channel time request from the child PNC. If there are devices in the child piconet which intend to request time for voice traffic, the parent PNC allocates 8 private CTAs to the child PNC spread evenly throughout the 8 MASs in the superframe. In this way, the QoS can be supported for devices in both the parent and child piconets subject to available capacity. For video traffic, the parent or child PNC allocates CTAs to requesting devices in factors of 2 depending on the available capacity and throughput requirement specified in the request. In the next section, we analyse and explore the capacity of high data rate WPANs in order to get some intuition about the upper limit of number of piconets in Inter-PAN communication and also the number of flows which can be supported simultaneously given a certain throughput requirement. The analysis also proves to be useful in order to determine the necessity of devising an admission control algorithm to allocate channel time to queued devices in a fair manner.

2.6 Capacity Analysis of HDR WPANs

In order to gain an insight and determine an upper limit of number of flows that can simultaneously be supported in a High Data Rate (HDR) WPAN, the relevant capacity analysis is carried out. Since the HDR WPANs are mainly targeted for multi-media applications with strict QoS constraints, the capacity analysis performed considers the requirements of voice and video flows with different throughput requirements. The results of the capacity analysis are followed up by actual simulations which also takes into account, the non-real time flows to depict a more realistic picture of the superframe capacity.

Different multi-media codec's encode data at different rates with different data rate requirements. The size of the MSDU received from the higher layers to the MAC layer varies as a result. If the size of the MSDU is larger than the largest MPDU size supported by the MAC layer, the MSDU has to be fragmented into smaller MPDUs. To simplify the fragmentation process, the MSDU is divided into equal size fragments (MPDUs). Through

the DEV capabilities field in the DEV association IE, each device indicates its preferred fragment size and supported data rates to all the member devices in the piconet. If a specific application requires a throughput of x Mbps, then the MAC layer has to at least support a throughput of $(x + \text{Layer Overhead})$ Mbps. The layer overhead can be calculated by considering the preamble added at the network layer and MAC layer as shown in Figure 2.11. Figure 2.11 indicates the fact that the Actual throughput (δ_A) should be at least more than the Effective throughput (δ_E) by an amount equal to Layer Overhead in order to support the application. Since a TDMA access mechanism is used for CTAs which are the guaranteed time slots, the throughput for a particular flow is measured by the amount of bits allocated to it in each CTA and the number of CTAs in a superframe. The value of δ_A gives us the actual throughput at the MAC layer taking into account the layer overheads whereas the value of δ_E gives us the effective throughput which should be at least equal to the throughput required by the voice or video application after excluding the extra overheads at different layers.

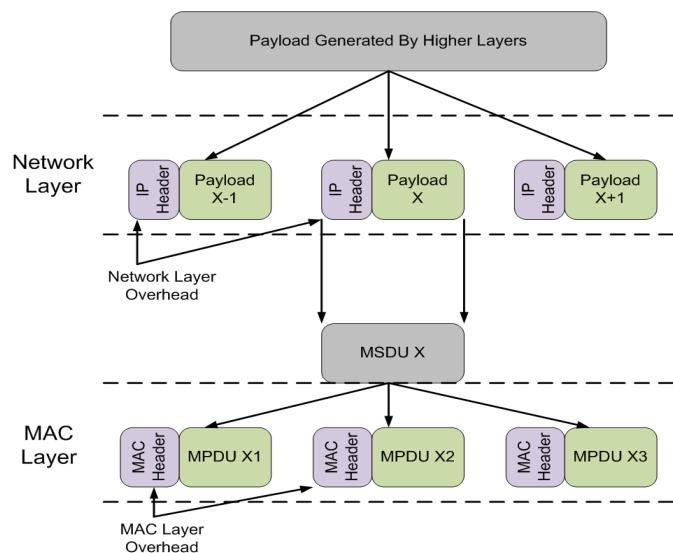


Figure 2.11 Overheads added at MAC and Network Layers

To determine the efficiency of utilization of a CTA, it is mandatory to consider the particular acknowledgement scheme which is to be used. Depending on the acknowledgement scheme a certain Inter-Frame Space (IFS) duration is used between successive frames which has an impact on the CTA overhead (Figure 2.12).

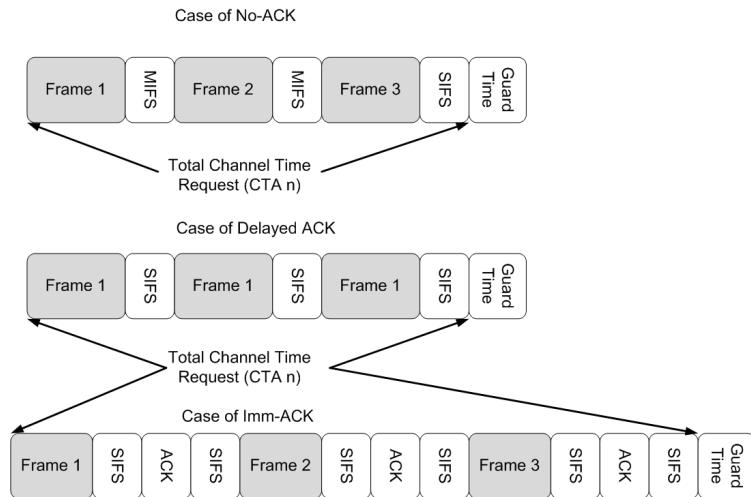


Figure 2.12 CTA structure for different ACK schemes

Since voice and video traffic is considered in the capacity analysis carried out, the Delayed Acknowledge (Dly-ACK) scheme is used. When using the Dly-ACK, either the Short Inter-Frame Space (SIFS) or the Minimum Inter-Frame Space (MIFS) can be used between successive frames. The CTA overhead when the SIFS is used between successive frames is given by

$$\tau_{DACK(SIFS)} = (\sum_{n=1}^{x_d+b} \eta_{(SIFS)n} + b \eta_{DACK} + \eta_{GT}) \quad (2.2)$$

Where x_d is the number of frames sent in the CTA, η_{SIFS} , η_{DACK} and η_{GT} are the duration of SIFS, time to send the Dly-ACK frame and the guard time respectively. The parameter b is set to one if there is a Dly-ACK frame in the CTA, otherwise it is set to 0. The total time allocated to each device (η_D) in the superframe is given by

$$\eta_D = \sum_{i=1}^x \eta_{(CTA)i} \quad (2.3)$$

Where η_{CTA} is the duration of a single CTA allocated to the device and x is the total number of CTAs allocated to the device in the superframe. The actual data rate (∂_A) of the device can be given by

$$\partial_A = \frac{\sum_{i=1}^x \eta_{(CTA)i}}{\eta_S} * \partial_T \quad (2.4)$$

The effective data rate (∂_E) at which the actual payload is delivered is given by

$$\partial_E = \frac{\sum_{i=1}^x [\eta_{CTA} - \tau_{CTA}]_i}{\eta_S} * \partial_T \quad (2.5)$$

Where τ_{CTA} is the CTA overhead for each CTA, which includes the IFS durations, ACKs and the GT. Since the 802.15.3 MAC is a TDMA MAC, the following relationship holds

$$\partial_E < \partial_A < \partial_T \quad (2.6)$$

Table 2.2 Parameters Considered for Voice Traffic

S. No.	Parameters	Value
1.	Frame Duration	20 ms
2.	Number of Frames/ Second	50
3.	Size of Each Frame	160 Octets
4.	Network (IP) Layer Overhead	40 Octets
5.	Fragment (MPDU) Size Considered	256 Octets
6.	Data Rate for G.711	64 Kbps

In order to make sure that the applications running on the devices are running smoothly during communication, it should be made sure that $\partial_E \geq \partial_R$. Where ∂_R is the minimum required throughput for a voice or video application. To calculate the capacity of an 802.15.3 superframe, voice and video traffic is considered since it consumes most of the networks resources. For voice, the G.711 codec is considered and for video, H.264 is considered. The following sections focus on the results analyzed when considering the two traffic types.

2.6.1 Capacity for Voice Applications

The values for different parameters used for calculating the capacity when using the G.711 codec are given in Table 2.2. It should be noted from Table 2.2 that the IP overhead is 40 Octets and the number of frames per second is 50. Therefore, an additional 16 Kbps is required to send the IP layer overhead apart from the 64 Kbps data rate which is for voice payload. Since the $MSDU\ size < MPDU\ size$, therefore an MPDU size of 256 Octets is chosen to send the MSDU which is 200 Octets in size. The MPDU size of 256 Octets also includes the Frame Check Sequence (FCS) which is 4 Octets in length. While sending the MSDU, an additional overhead of MAC header has to be taken into account also. Therefore, the total throughput that needs to be supported is 106.4 Kbps. Since the superframe size considered is 65535 μ s, there are 1/65535 μ s superframes in one second. Therefore, the time required per superframe for a device with a required throughput of 106.4 Kbps is

$$\partial_R \times \frac{1}{\partial_T} \times \frac{1}{\eta_s} \quad (2.7)$$

The maximum tolerable delay and jitter for voice applications should be < 10 ms according to the IEEE 802.15.3 standard. Therefore, the available channel time i.e. CTAP is divided by 8 to limit the delay and jitter to be less than 10ms. The PNC sends the beacon at a base rate of 22 Mbps; therefore, the overhead is the same for devices operating at different transmitter data rates. When a device has been allocated a CTA by the PNC, then depending on the number of frames sent in the CTA and the transmitter data rate, the superframe overhead (τ_s) increases or decreases. Although η_{MIFS} , η_{SIFS} and η_{GT} remain the same with the increase or decrease in the transmitted data rate, the time required to send the MPDUs (η_{MPDU}) within a CTA increases or decreases respectively. Since with the increase in transmitted data rates, the ratio of (η_{MPDU}) to τ_{CTA} decreases, the overhead increases. This can be seen in Figure 2.13 where the percentage superframe overhead is plotted against data rates of 22, 33, 44 and 55 Mbps with different number of frames sent in a CTA. As the number of frames per CTA increase, the ratio of (η_{MPDU}) to τ_{CTA} also increases and hence lesser overhead.

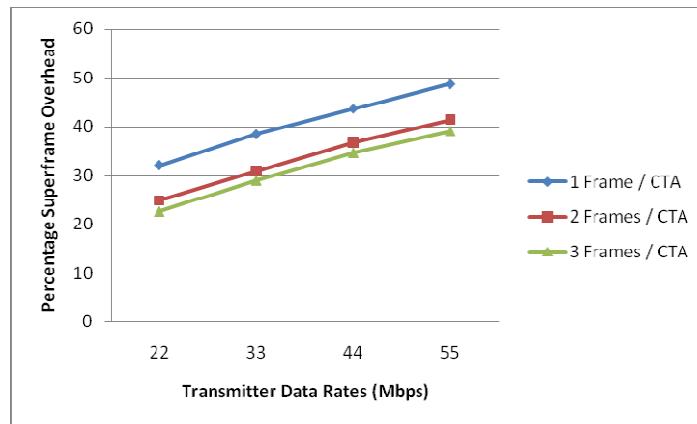
The superframe overhead apart from the transmitter data rate, also depends on the number of MPDUs sent in the CTA and is calculated using the following equation

$$\tau_s = \tau_{PNC} + (\tau_{CTA} \times N_{CPS} \times N_D) \quad (2.8)$$

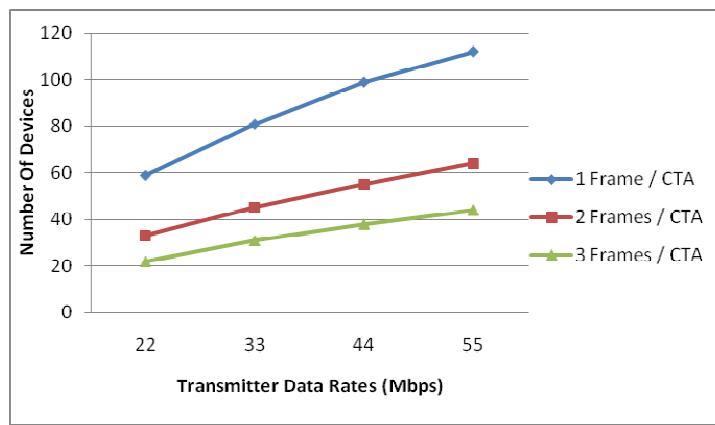
Where τ_s is the superframe overhead, N_{CPS} denotes the number of CTAs per superframe and N_D denotes the number of supported devices. The CTA overhead τ_{CTA} is calculated by

$$\tau_{CTA} = \eta_{CTA} - (\eta_{MPDU} \times x_d) \quad (2.9)$$

Where η_{MPDU} is the time required to send an MPDU and x_d denotes the number of MPDUs per CTA. Increasing δ_T also has an advantage. The CTA duration (η_{CTA}) decreases with the increase in δ_T and as a result, the superframe capacity increases. This can be shown in Figure 2.14 where the superframe capacity is plotted against δ_T with different number of frames per CTA. Figure 2.15 shows that the CTA overhead increases with the increase in the transmitter data rate. The reason is that with the increase in δ_T , the time required to send the MPDU decreases but the IFS remains constant. Furthermore the time required to send the MAC header remains the same as it is always sent at the base rate of 22 Mbps. The CTA overhead decreases when the number of frames per CTA is increased.



**Figure 2.13 Superframe overhead with different transmitter data rates
(MPDU Size = 256 Octets)**



**Figure 2.14 Superframe Capacity for different Frames/CTA
(MPDU Size = 256 Octets)**

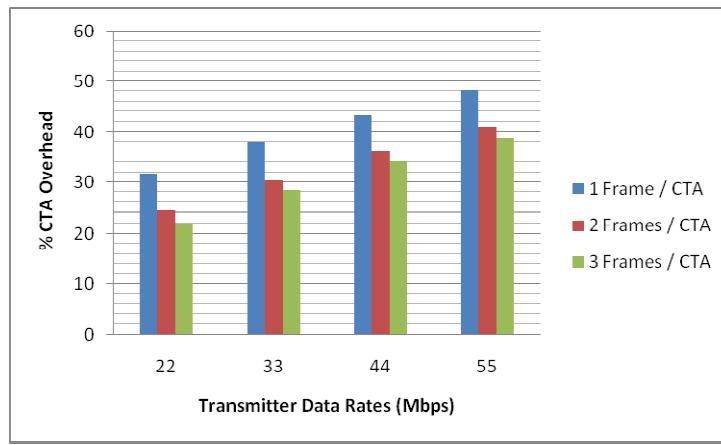
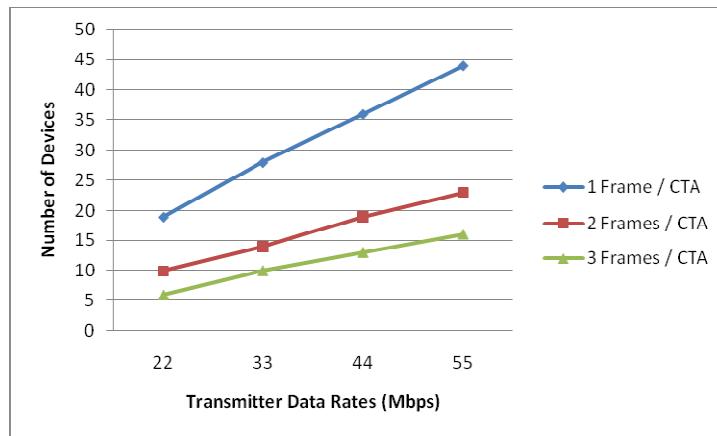


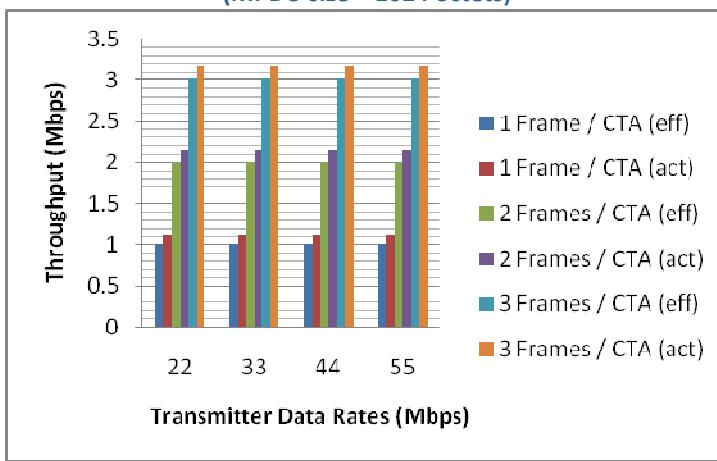
Figure 2.15 Percentage CTA overhead (MPDU size = 256 Octets)

2.6.2 Capacity for Video Applications

Since the requirements for video traffic are more resource intensive than voice, the capacity for video traffic is analyzed in order to find an upper limit of throughput. For video traffic, 4 different levels of H.264 are considered for mobile content (3G video), Internet/ Standard Definition (SD), High Definition (HD) and Full High Definition (HD). Each level has different throughput requirements and number of frames sent per second. When sending mobile content at a resolution of 176 by 144 and a frame rate of 24 fps, the throughput required is about 160 Kbps. The average size of each frame comes up to 834 Octets. If the IP overhead is considered, the frame size becomes 874 Octets. The nearest fragment size of 1024 Octets is used to efficiently carry an MSDU size of 874 Octets. If the MAC layer overhead is taken into account, δ_R becomes ≈ 200 Kbps. Since the maximum tolerable delay and jitter for video applications should be less than 100 ms, there is more flexibility is assigning super rate CTAs to video applications depending on the required throughput. Figure 2.16 shows us the superframe capacity when considering mobile content with an MPDU size of 1024 Octets. Different number of frames are sent per CTA and the capacity of superframe is analyzed. δ_A is also mentioned in Figure 2.17 where it is shown that by increasing δ_T , δ_A only increases by 2 %. However there is a two fold increase in δ_A by sending more frames in the CTA. Also δ_A can be increased or decreased by increasing or decreasing the number of CTAs in the superframe. The value of δ_A does not change with the increase in δ_T due to the TDMA MAC format. Since the number of bits sent per superframe remains the same for a device, therefore δ_A remains the same. For Internet/ Standard Definition (SD), HD and full HD, the throughput requirements are much more than those for 3G mobile content. The *MSDU size > MPDU size* and therefore an MPDU size of 2048 is chosen. Required throughputs considered are 2, 6 and 8 Mbps.



**Figure 2.16 Superframe Capacity against data rate
(MPDU size = 1024 octets)**



**Figure 2.17 Comparison of Actual and Effective throughputs
(MPDU Size = 1024 Octets)**

Figure 2.18 shows the capacity of superframe when an MPDU size of 2048 is considered for up to 4 frames / CTA. It can be seen that for lower values of δ_T e.g. 22 Mbps, only 5 devices can be supported for 2, 3 and 4 frames per CTA. Figure 2.19 shows that for the same 2, 3 and 4 frames per CTA, δ_A achieved when δ_T is 22 Mbps, δ_A is up to 8 Mbps. Therefore it can be noted that a practical limit of 8 Mbps can be set for the superframe (When the full duration of 65535 μ s is used) when the number of devices is low i.e. 5 – 10 in the piconet. To achieve fairness among higher number of devices, the upper limit should be further dropped. Figure 2.20 shows that the increase in CTA overhead is relatively less when compared with the use of smaller MPDU sizes.

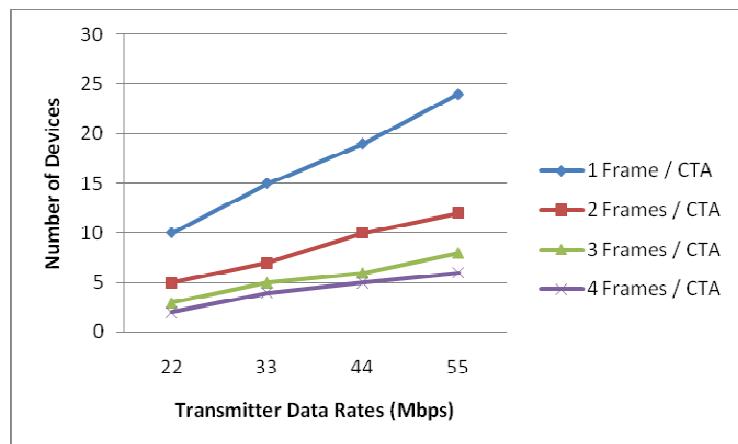


Figure 2.18 Superframe Capacity (MPDU size = 2048 Octets)

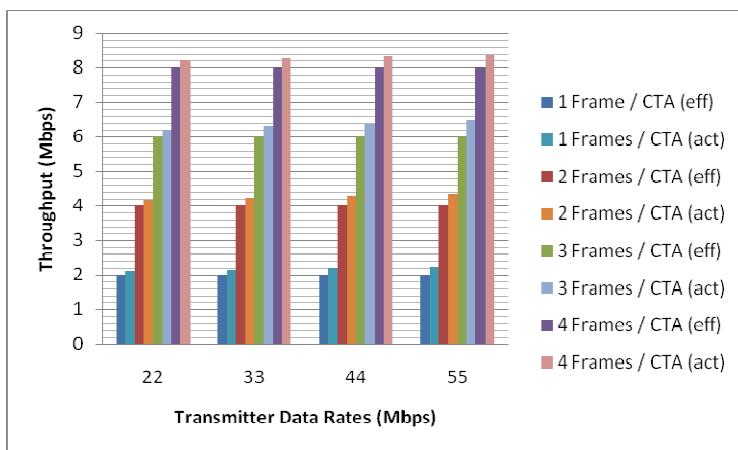


Figure 2.19 Comparison of Actual and Effective throughputs (MPDU Size = 2048 Octets)

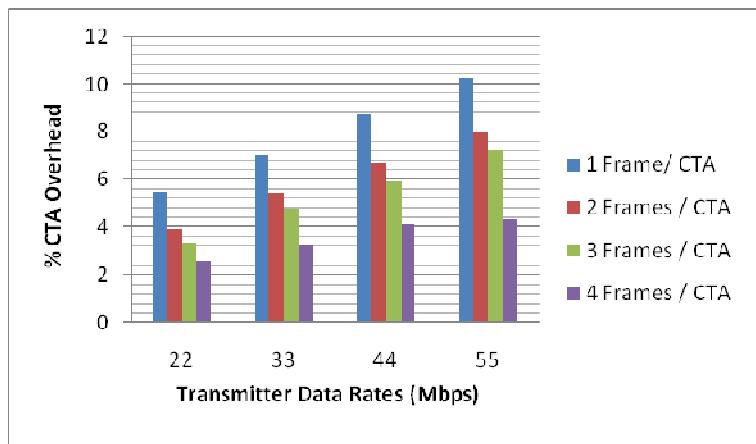


Figure 2.20 CTA Overhead (MPDU size = 2048 Octets)

For the 2.4 GHz physical layer specifications in [1], a MIIFS duration is $2\mu\text{s}$ and a SIFS duration is $10\mu\text{s}$. Keeping the size of MAC Protocol Data Unit (MPDU) constant, increasing the transmission data rate does not increase the throughput but rather increases the CTA overhead. The CTA overhead is increased because by increasing the transmission data rate, the time required to send the payload decreases and hence the overall size of the CTA is reduced. However, the MAC header is always sent at the base rate and the duration of the IFSs and the Guard Time (GT) remain the same. Therefore, the ratio of payload duration in the CTA to the total duration of the CTA decreases indicating the increase in CTA overhead. The CTA overhead can be reduced by sending more number of frames per CTA and by increasing the size of the MPDU transmitted in each frame. By increasing the transmission data rate, even though the CTA overhead increases, the capacity also improves since a lower duration is required to send a frame and thus the CTA duration reduces. Since the bit duration is equal to $1/\delta_T$, increasing the transmission data rate shall reduce the bit duration and hence reduce the total CTA duration. In a TDMA format, if the number of frames sent per CTA and the number of CTAs per superframe allocated to a device remain the same, then there is a fixed number of bits sent per superframe by a device. Increasing the transmission data rate only reduces the bit duration but the total number of bits sent remains the same for a device. Therefore, the throughput can be increased by either sending more frames per CTA or by sending more CTAs per superframe for the same device.

By increasing the transmission data rates the number of devices that can be supported might increase. This is because with the reduction in CTA size due to increase in transmission data rate, more capacity becomes available in the superframe and hence the possibility for number of devices sharing the superframe increases. The number of devices that can be supported varies with the number of frames sent per CTA. Although the capacity of the superframe improves by increasing the transmission data rate, the CTA overhead increases. The reason being that with the reduction in CTA duration with increased transmission data rate, the duration of the IFSs remains the same. Furthermore, the MAC header is always sent at the base rate of 22Mbps and the duration to send the MAC header remains the same. Thus the ratio of time duration to send the actual payload bits to the CTA duration decreases. By sending more frames per CTA, the ratio of time

duration to send the payload to the total CTA duration increases and therefore the CTA overhead decreases. Keeping everything else constant, the CTA overhead varies based on the type of acknowledgement (ACK) scheme used as well.

The capacity analysis indicates that by increasing the transmission data rate, the actual throughput and the effective throughput remain the same. The effective throughput does not consider the overhead bits added at the higher layers and the MAC and physical layer. The actual throughput considers all the overhead bits added at the higher layers along with the MAC and physical layer. Therefore it is more than the effective throughput. Also it can be noticed that by increasing the frames send per CTA, higher throughputs can be achieved. The size of the MPDU contributes in determining the throughput achieved.

2.7 CTA Allocation and Admission Control

Until now, we have analyzed the HDR WPAN in terms of its theoretical capacity and throughput. The traffic considered has been voice and video traffic. However, to be more realistic, Some devices in the piconet might be able to send/receive only a single type of traffic e.g. DVD players, mp3 players, fax machines, printers etc. while most of the devices are expected to be able to send/receive multiple traffic types e.g. Cell phones, laptops, PDAs, Desktops etc. The total number of devices in a piconet depends whether a specific WPAN belongs to a particular user i.e. Home based WPANs, Office based WPANs or the WPAN is not owned by a particular user i.e. Resources offered/shared by a public WPAN. A specific user based WPAN might have different traffic characteristics than a public WPAN because the user can personalize the communication in the WPAN. Furthermore, there might be restrictions on the resources in a user owned WPAN. A public WPAN can be demand based and the volume of traffic, the number of devices, and the time of traffic might vary based on the time of the day. For further insights and more pragmatic analysis, we evaluate the WPAN based on a mixture of flows and determine the need for an admission control algorithm.

2.7.1 Description of Admission Control Procedure

In this section we provide a theoretical basis for the admission control requirements and in the subsequent section we describe the RR and IRR Algorithms. If there are X associated devices in the piconet at a particular time instant and T is the total number of expected devices in the piconet, then T consists of those devices which are user owned and those devices which are not user owned but can communicate in the piconet. Therefore a fixed number of devices which are user owned can be assumed to be always present in the piconet. Three types of traffic are considered i.e. Voice (V_o), Video (V_i), and Data traffic (D_o) which mainly consists of peer to peer file sharing and downloads from other sources. Let P_{V_o} , P_{V_i} and P_{D_o} be the percentages of devices expected to request for voice, video and data traffic respectively such that $P_{V_o} \rightarrow T_{V_o}$, $P_{V_i} \rightarrow T_{V_i}$ and $P_{D_o} \rightarrow T_{D_o}$ where $T_{V_o} = \{t_{V_o1}, t_{V_o2}, \dots, t_{V_on}\}$, $T_{V_i} = \{t_{V_i1}, t_{V_i2}, \dots, t_{V_in}\}$ and $T_{D_o} = \{t_{D_o1}, t_{D_o2}, \dots, t_{D_on}\}$ and $T = T_{V_o} \cup T_{V_i} \cup T_{D_o}$. If some of the devices have already been allocated channel time, then depending on the type of traffic they are communicating, they either belong to A_{V_o} , A_{V_i} or A_{D_o} and $A = A_{V_o} \cup A_{V_i} \cup A_{D_o}$. Similarly those devices which have not been allocated channel time shall belong to N_{V_o} , N_{V_i} or N_{D_o} such that $N_{V_o} = T_{V_o} - A_{V_o}$, $N_{V_i} = T_{V_i} - A_{V_i}$ and $N_{D_o} = T_{D_o} - A_{D_o}$. The total number of devices, $T = N \cup A$. The request for channel time from a device can either be for a deterministic (Non Real Time flow) duration or non-deterministic (Real Time flow) duration. Let A_D and A_{Nd} represent the number of devices which have been allocated channel time for deterministic and non-deterministic time durations respectively such that $A_D \subset A$ and $A_{Nd} \subset A$. An example of traffic types with deterministic duration is downloading data files. Since the size of the file is known, the total duration of the session to download the complete file can be determined by taking into account the channel time allocated to a device. The average throughput of a device and the duration of a data transfer session depends on whether it has been allocated minimum number of TUs (min TUs) by the PNC or the desired number of TUs (des TUs). Based on the average throughput, the expected time for a session to end is calculated for deterministic flows. Similarly, an example of a non-deterministic flow can be a voice conversation or real time video streaming.

The arrival of requests from devices for channel time to a PNC is approximated by the Poisson mean $\mu = \lambda t$, where λ is the arrival rate of requests and t is the time interval of interest. For the same WPAN at a particular time, the homogeneous Poisson arrival rate is

considered. When a request for channel time is granted to a device, it moves from N to A and the sets are updated. Based on the type of traffic (i.e. voice, video, and data), the probability of that type of request in the future decreases until some devices finish their sessions and add back to N . If a request for voice traffic is received by the PNC, then the set N_{Vo} is decremented by one element and A_{Vo} is incremented by one. Thus the corresponding probability of requests for voice traffic reduces and is given by

$$\rho_{Vo} = \left(\frac{n(N_{Vo})}{n(N)} \right) \quad (2.10)$$

Let t_T be the time interval in which all the devices in the WPAN are expected to send channel time requests to the PNC. Let t_i be the start of the time interval and t_f be the end of the time interval t_T (Figure 2.21). If the superframe capacity exceeds the threshold value of $k\eta_{CTAP}$, (where $0 < k < 1$) at a particular time instant t_x , then it is checked where t_x lies. If $t_x < \frac{t_T}{2}$ and $n(N) \neq 0$, then the expected time to be requested is checked.

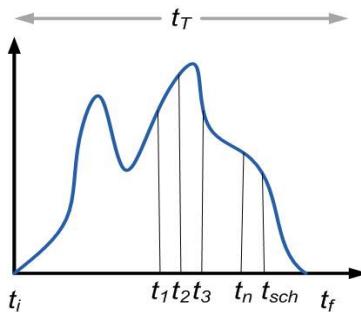


Figure 2.21 Time period of expected request

To calculate the expected time η_{etd} , the mean value of 10 previous capacity allocations to each device in $n(N)$ is taken and their sum is equated to η_{etd} . If $\eta_{etd} > \eta_{ASC}$ where η_{ASC} is the available superframe capacity, that means that the expected traffic requirements exceed the available superframe capacity and hence the need for an admission control algorithm.

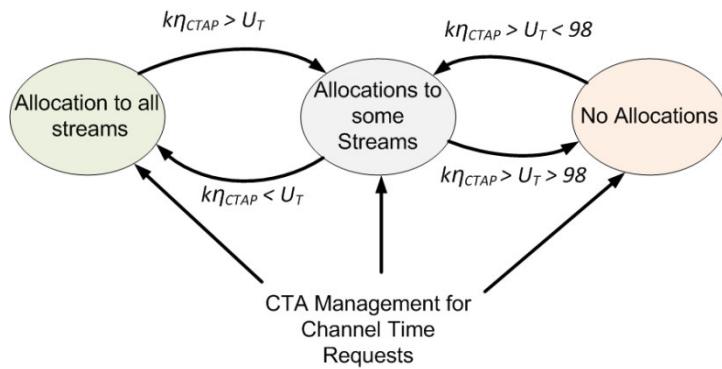


Figure 2.22 Three State strategy for CTA management and Admission Control

A three state strategy (Figure 2.22) is used for the admission control scheme. In state i , the PNC allocates channel time to all the flows until a certain superframe utilization threshold U_T is reached after which the PNC goes to state j . After the PNC transits into the second state (state j), it indicates in its beacon through a capacity Information Element (IE) to the devices about the capacity constraints. In state j , the PNC checks if there are any queued requests. If there are no queued requests and the PNC receives a request for an allocation with minimum (min) TUs and desired (des) TUs specified other than a VBR flow, then it allocates only the min TUs to that flow in state j . When the capacity of the superframe is more than 98% occupied, the PNC no longer accepts channel time requests and transits into state k which is the blocking state. All the requests received in state k are queued by PNC. The PNC remains in state k , unless some of the flows end and release their allocated channel time. Upon sufficient available time after flow terminations, the PNC transits back to state j if the condition $k\eta_{CTAP} > U_T < 98$ is true. In state j , the PNC checks for queued requests and serves the queued requests. The manner in which the queued requests are served is important for reasons of maintaining fairness in CTA allocations to requesting devices. In the subsequent section, the proposed schemes RR and IRR are introduced which are used for prioritizing and serving the queued requests.

2.7.2 Fair Allocation Via Rejection Ratio

In cases where the available superframe capacity is insufficient to meet the minimum throughput requirements of a request, the request is queued. This can happen frequently in presence of congestion when either the number of devices with very high throughput requirement, increases or the total number of devices which compete for channel time

increase. When the PNC is in state k , it does not accept any more requests but rather queues them. In case of a flow termination, time is released back to the superframe by the PNC and it can serve a request from the queue provided the released time is sufficient to meet the requirements of any of the queued devices. One option to determine which device from the queue to serve is by first come first serve basis provided the requirements of channel time for that device are met. The PNC can search for the queued devices iteratively unless it finds the suitable devices whose requirements can be met. However, this approach might not be fair to some devices because there might be a number of devices which have not been allocated channel time at all and therefore this can induce discrimination in the CTA allocation approach.

In order to serve the queued devices in a relatively fair manner by the PNC, a number of simple metrics are considered with the parameters in consideration being the total number of requests, total number of rejections and total number of allocations for each device. If we denote the total number of requests by r_t , the total number of rejections by r_r and the total number of allocations by a_t , the following metrics are considered for evaluation.

- a. $\frac{r_r}{a_t}$ or $\frac{a_t}{r_r}$
- b. $r_r - a_t$ or $a_t - r_r$
- c. $\frac{r_r}{r_t}$

The first metric considered is either the ratio of total rejections to total allocations or vice versa. In case the ratio of total allocations, a_t to total rejections r_r is considered, then it can provide us with an insight into whether a device has been allocated more allocations or more rejections. If

- $\frac{a_t}{r_r} > 1 \rightarrow$ The number of allocations exceeds the number of rejections
- $\frac{a_t}{r_r} < 1 \rightarrow$ The number of rejections exceeds the number of allocations
- $\frac{a_t}{r_r} = 0 \rightarrow$ The device has not been allocated channel time
- $\frac{a_t}{r_r} = 1 \rightarrow$ The number of allocations and number of rejections are equal

Therefore, the value of the first metric can give us an indication about the priority of a device when a queued device has to be served. Devices with a higher $\frac{a_t}{r_r}$ ratio shall have a lower priority while those with a lower value shall get a higher priority. However, the metric $\frac{a_t}{r_r}$ and $\frac{r_r}{a_t}$ can raise computational problems if the value of denominator is zero. The second

metric does not have the *divide by zero* problem and can indicate the priority of a queued device on basis of its rejections and allocations. If we consider $r_r - a_t$, a general insight related to the priority of a device be obtained by checking if

$$r_r - a_t = +ve$$

$$r_r - a_t = -ve$$

$$r_r - a_t = 0$$

A positive value indicates a higher priority and the negative value indicates a lower priority. The third case arises when the two terms are equal. A problem with this metric is that the third case might give us an ambiguous indication about the priority of a device. To elucidate upon this argument, consider a case when one device has the values of r_r and a_t equal to 10 while another device has the values for the same parameters equal to 5. In such a case it becomes difficult to distinguish between the two devices since the metric shall give a value of zero for both the cases.

The third metric which is the ratio of total rejections to the total requests gives us a bounded value from 0 to 1. Therefore it is simply the Rejection Ratio (RR) of a device. A value of 0 indicates that the device has no rejections and hence a lower priority. A value of 1 indicates that the device has a higher priority because it has no allocations. To get a better intuition from the RR of a device, it can be written as

$$\frac{r_r}{r_r + a_t} \because r_t = r_r + a_t \quad (2.11)$$

If the value of RR is 1, that means that the number of allocations for that device equal to zero and therefore $r_r = r_t$. Therefore, the algorithm based on RR prioritizes the queued devices upon considering the following cases:

Case A (1st Priority): $r_r^i = 0$ and $a_t^i = 0$

Case B (2nd Priority): $r_r^i > 0$ and $a_t^i > 0$

Case C (3rd Priority): $r_r^i = 0$ and $a_t^i > 0$

Where r_r^i and a_t^i are the number of rejections and the number of allocations for the i^{th} queued device. The value of RR is calculated to determine priority if Case B is true and no devices fall in the category of Case A. The reason is that the value of RR is only non-zero if Case B is true. The algorithm to determine the priority of a queued device and serve its request on the basis of RR value is given in Figure 2.23.

RR Algorithm Start

```

Begin
Initialize parameters
Number of devices = n
Number of queued devices =  $N_q$ 
ctr1= 0, ctr2= 0, ctr3 = 0
A[i]=0  $\forall i = 1, 2, 3 \dots n$ 
B[i]=0  $\forall i = 1, 2, 3 \dots n$ 
C[i]=0  $\forall i = 1, 2, 3 \dots n$ 
R[i] =0  $\forall i = 1, 2, 3 \dots n$ 
AT[i] = 0  $\forall i = 1, 2, 3 \dots n$ 

for (i = 1; i <=  $N_q$ ; i++)
{
    get  $q_s^i \in \vec{Q}$ 
    if ( $q_s^i \leq \eta_{rel}$ )
    {
        if ( $r_r^i == 0 \&& a_t^i == 0$ )
        {
            ctr1++;
            A[ctr1] = i;
        }

        if ( $r_r^i > 0 \&& a_t^i > 0$ )
        {
            calculate RR for  $i^{th}$  flow i.e.  $R_r^i$ 
            ctr2++;
            B[ctr2] = i;
            R[i] =  $R_r^i$ ;
        }

        if ( $r_r^i == 0 \&& a_t^i > 0$ )
        {
            ctr3++;
            C[ctr3] = i;
            AT[i] =  $a_t^i$ ;
        }
    }

    if (ctr1 > 0)
        Select A[1];

    if (ctr1 == 0 && ctr2 > 0)
    {
        If(ctr2 == 1)
        Select B[1]

        If(ctr2 > 1)
        Select maximum value in R[i]
    }
}

```

```

if (ctr1 == 0 && ctr2 == 0 && ctr3 > 0)
{
    if (ctr3 == 1)
        Select C[1]
    If (ctr3 > 0)
        Select minimum value in AT[i]
}

if (ctr1 == 0 && ctr2 == 0 && ctr3 == 0)
    Exit; (No feasible flows exist)
end

```

Figure 2.23 Algorithm based on the value of Rejection Ratio (RR)

Algorithm description for RR:

- Initialize parameters i.e. initialize vectors \vec{A} , \vec{B} , \vec{C} , \vec{R} and \vec{A}_T , $ctr1$, $ctr2$ and $ctr3$ where \vec{A}_T is the vector such that for the number of allocations of i^{th} device $a_t^i \in \vec{A}$, and $ctr1$, $ctr2$ and $ctr3$ are three counters. The vector \vec{R} consists of the RR values of queued requests while the vectors \vec{A} , \vec{B} , \vec{C} store the DEVIDs of the devices that fall into the A, B or C priority classes respectively
- Check if the queued device request can be served i.e. enough time has been released to serve the request from the queue vector \vec{Q} with the request size for the i^{th} queued device being q_s^i and the released CTA time being η_{rel}
- If the request is feasible i.e. $q_s^i \leq \eta_{rel}$, then check which priority class the device falls into based on its values of r_r and a_t
- Increment $ctr1$ if any devices fall in Class A priority list represented by vector \vec{A} , $ctr2$ for devices in \vec{B} and $ctr3$ for devices in \vec{C}
- If any device falls in \vec{B} , determine the rejection ratio value R_r^i , place it in the vector \vec{R} and increment the counter $ctr2$
- In case there are no rejections for the i^{th} device i.e. r_r^i is zero but the number of allocations for it i.e. a_t^i is non zero then increment $ctr3$
- If both the values of r_r and a_t are zero for a device, that means the device has not been allocated time before and its first request for channel time has been queued. In such a case serve the first element of vector \vec{A} if $ctr1$ is non-zero
- Serve the queued request based on the maximum value in the vector \vec{R} if $ctr2$ is non zero
- Serve the queued request based on the minimum value in the vector \vec{A}_T if $ctr1$ and $ctr2$ are zero but $ctr3$ is non zero
- If no such device is found which can be served based on the released time i.e. $ctr1$, $ctr2$ and $ctr3$ are equal to 0, then exit the algorithm

Although the RR value proves to be effective at times, there is a problem in case when the value of $a_t = 0$ and $r_r > 0$ for more than one device. In such a case $r_t = r_r$ and therefore

the value of $RR = 1$. If both the devices have a different number of rejections with the first device having a greater value of r_r than the second device, the RR value would not be able to differentiate between the two devices in terms of priority. This situation might become true for more than two devices when the number of devices increases i.e. in case of higher device density. To overcome this issue, an *Improved Rejection Ratio* (IRR) is proposed which takes into account, the number of rejections for feasible devices and gives priority to those devices which have value of $a_t = 0$ and $r_r > 0$. In case no such devices are found, it prioritizes on basis of RR value. Therefore, using the IRR technique, the issue encountered in the RR technique is resolved and those devices which have $a_t = 0$ are identified and prioritized. There are two cases in which the value of $a_t = 0$ which in turn depends on the value of r_r . The priority criterion in case of IRR is as follows:

Case A (1st Priority): $r_r^i > 0$ and $a_t^i = 0$

Case B (2nd Priority): $r_r^i = 0$ and $a_t^i = 0$

Case C (3rd Priority): $r_r^i > 0$ and $a_t^i > 0$

Case D (4th Priority): $r_r^i = 0$ and $a_t^i > 0$

The algorithm to determine the priority of a queued device and serve its request on the basis of IRR value is given in Figure 2.24.

IRR Algorithm Start

Begin

Initialize parameters

Number of devices = n

Number of queued devices = N_q

$ctr1 = 0, ctr2 = 0, ctr3 = 0, ctr4 = 0$

$A[i] = 0 \forall i = 1, 2, 3 \dots n$

$B[i] = 0 \forall i = 1, 2, 3 \dots n$

$C[i] = 0 \forall i = 1, 2, 3 \dots n$

$D[i] = 0 \forall i = 1, 2, 3 \dots n$

$R[i] = 0 \forall i = 1, 2, 3 \dots n$

$A_T[i] = 0 \forall i = 1, 2, 3 \dots n$

$Z[i] = 0 \forall i = 1, 2, 3 \dots n$

for ($i = 1; i \leq N_q; i++$)

{

 get $q_s^i \in \vec{Q}$

 if ($q_s^i \leq \eta_{rel}$)

 {

 If ($r_r^i > 0 \&& a_t^i == 0$)

```

{
ctr1++;
A[ctr1] = i;
Z[i] = rri;
}

if (rri == 0 && ati == 0)
{
ctr2++;
B[ctr2] = i;
}

if (rri > 0 && ati > 0)
{
calculate RR for ith flow i.e. Rri
ctr3++;
C[ctr3] = i;
R[i] = Rri;
}

if (rri == 0 && ati > 0)
{
ctr4++;
D[ctr4] = i;
AT[i] = ati;
}
}

if (ctr1 > 0)
{
if (ctr1 == 1)
Select A[1];

if (ctr1 > 1)
Select maximum value in Z;
}

if (ctr1 == 0 && ctr2 > 0)
Select B[1];

if (ctr1 == 0 && ctr2 == 0 && ctr3 > 0)
{
if(ctr3 == 1)
Select C[1]

if (ctr3 > 1)
Select maximum value in R[i]
}

if (ctr1 == 0 && ctr2 == 0 && ctr3 == 0 && ctr4 > 0)
{
if (ctr4 == 1)
Select D[1]
If (ctr4 > 0)
}

```

```

Select minimum value in  $A_T[i]$ 
}

if (ctr1 == 0 && ctr2 == 0 && ctr3 == 0 && ctr4 == 0)
    Exit; (No feasible flows exist)
end

```

Figure 2.24 Algorithm based on the value of Improved Rejection Ratio (IRR)

Algorithm description for IRR

- Initialize parameters i.e. initialize vectors \vec{A} , \vec{B} , \vec{C} , \vec{D} , \vec{R} , \vec{A}_T and \vec{Z} , $ctr1$, $ctr2$, $ctr3$ and $ctr4$ where \vec{Z} , is the vector consisting of the number of rejections r_r for those devices that fall into priority class A
- The vector \vec{R} consists of the RR values of queued requests while the vectors \vec{A} , \vec{B} , \vec{C} and \vec{D} stores the DEVIDs of the devices which fall into the respective A, B, C or D priority classes respectively
- Check if the queued device request can be served i.e. enough time has been released to serve the request from the queue vector \vec{Q} with the request size for the i^{th} queued device being q_s^i and the released CTA time being η_{rel}
- If the request is feasible i.e. $q_s^i \leq \eta_{rel}$, then check which priority class the device falls into based on its values of r_r and a_t
- Increment $ctr1$ if any devices fall in Class A priority list represented by vector \vec{A} , $ctr2$ for devices in \vec{B} , $ctr3$ for devices in \vec{C} and $ctr4$ for devices in \vec{D}
- If any device falls in \vec{A} , increment $ctr1$ and store the number of rejections for that device in the vector \vec{Z}
- If any device falls in \vec{B} , increment the counter $ctr2$ and store its DEVID in vector \vec{B}
- If any device falls in \vec{C} , determine the rejection ratio value R_r^i , place it in the vector \vec{R} and increment the counter $ctr3$
- In case there are no rejections for the i^{th} device i.e. r_r^i is zero but the number of allocations for it i.e. a_t^i is non zero then increment the $ctr4$
- Serve the queued request based on the maximum value in the vector \vec{Z} if $ctr1$ is non-zero
- Serve the queued request based on the vector \vec{B} if $ctr2$ is non-zero
- Serve the queued request based on the maximum value in the vector \vec{R} if $ctr3$ is non zero
- Serve the queued request based on the minimum value in the vector \vec{A}_T if $ctr4$ is non-zero, while $ctr1$, $ctr2$ and $ctr3$ are zero
- If no such device is found which can be served based on the released time i.e. $ctr1$, $ctr2$, $ctr3$ and $ctr4$ are equal to 0, then exit the algorithm

2.8 Simulation Results

The simulation setup and the specific scenarios to evaluate different dimensions of the proposed work in this chapter were implemented in the OPNET Modeller. Two different types of Node Models were developed to emulate the function of a member device in a piconet and the PNC. To enable adaptive modulation which is specified in the IEEE 802.15.3 standard, the necessary modifications were carried out in the Physical layer framework which is called the Transmission Pipeline Stage in OPNET. The MAC layer parameters and the attributes used in different simulation scenarios are given in Table 2.3.

Table 2.3 Some Attributes used in Simulation Scenarios

No.	Attribute	Value
1	Transmission Data Rates	22, 33, 44, 55 (Mbps)
2	Number of Member Devices	30 - 40
3	Mean Flow Duration for Deterministic Flows	400 seconds
4	Mean Flow Duration for Non-Deterministic Flows	900 seconds
5	Mean Request Arrival Rates	50, 80, 120, 180, 250, 300 seconds
6	Superframe Duration	0.065535 seconds
7	Number of Child Piconets	1-3
8	Simulation Time	5 Hours

The simulation scenarios created for evaluation can be placed into three categories i.e. General Performance Analysis, Inter-PAN Communication evaluation and Admission Control with RR Evaluation. The three categories along with the discussion relevant to the important results are therefore, as follows:

General Performance Analysis:

The general performance analysis was performed to get a more realistic picture of the superframe capacity in terms of simultaneous number of devices that can be supported. The effect of increase in devices on the queue size was also one of the important factors that had to be evaluated to determine the need for an admission control strategy which can serve the requests in the queue in a fair manner. The superframe capacity utilization based on variable mean arrival rate of requests for channel time and different number of devices, gives us a good idea about using the most feasible number of devices and mean arrival rate

values to observe the performance of the admission control procedure in congestion. In Figure 2.25, the superframe utilization is given when the number of member devices in the piconet change from 30-40 with a constant mean request arrival rate of 180 seconds. The affect of changing the mean arrival rate of requests on superframe utilization can be seen in Figure 2.26. The mean arrival rate of requests in the simulation setup is based on the Poisson distribution. The rapid increase in the superframe utilization curve can be seen when the mean arrival rate is 50 seconds which means that the requests for channel time arrive more frequently.

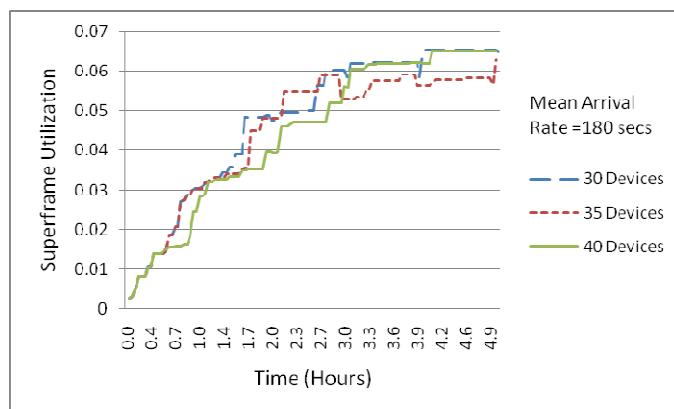


Figure 2.25 Superframe Utilization with Variable Number of Devices

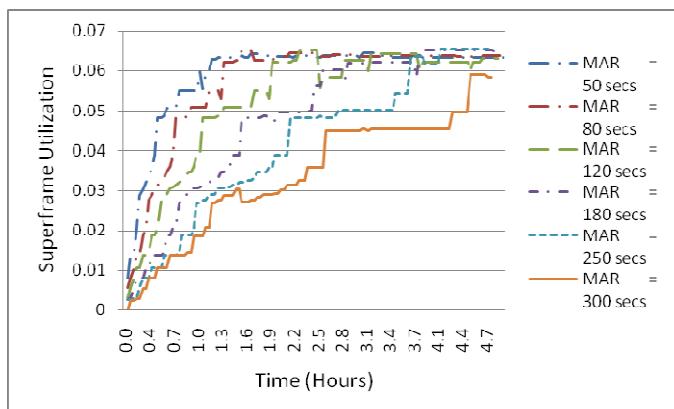


Figure 2.26 Superframe Utilization based on variable Mean Arrival Rate (MAR)

The maximum number of simultaneous devices that can be supported in the superframe depends on the CTA size of requests by the devices which relates to their required throughput and the flow duration of requests. The number of supported devices in the superframe can be seen in Figure 2.27 in which approximately 32 devices are shown to be supported simultaneously when the number of member devices in the piconet is 40. When

the number of member devices in the piconet is 30 and the mean arrival rate of requests is varied, the graph in Figure 2.28 shows that all of the flows can be accommodated by the superframe. When the number of devices in the piconet is increased above 30, it is noticed that the number of requests which get queued because of insufficient superframe capacity increases. To check the queue size when the number of devices is varied from 30-40, Figure 2.29 can be consulted in which the mean arrival rate is 120 seconds. It is obvious from Figure 2.29 that the queue size for 40 devices is significantly larger than 35 or 30 devices. When the mean arrival rate is varied, it can be observed from Figure 2.30 that the queue starts building up earlier with lower values of mean arrival data rates. Lower values of mean arrival rate suggest that the requests arrive more frequently and therefore, the superframe capacity becomes congested quicker.

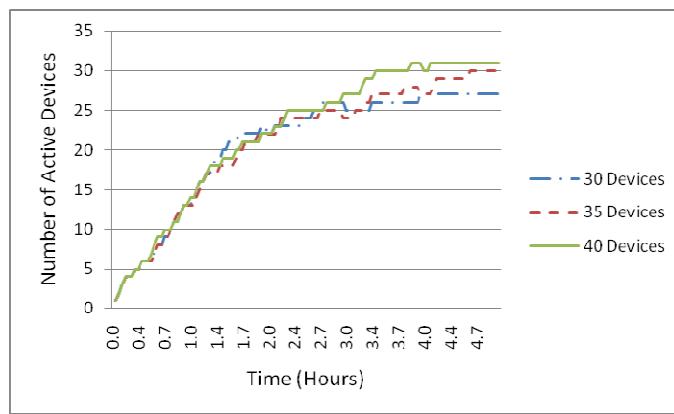


Figure 2.27 Number of Devices simultaneously supported in the superframe

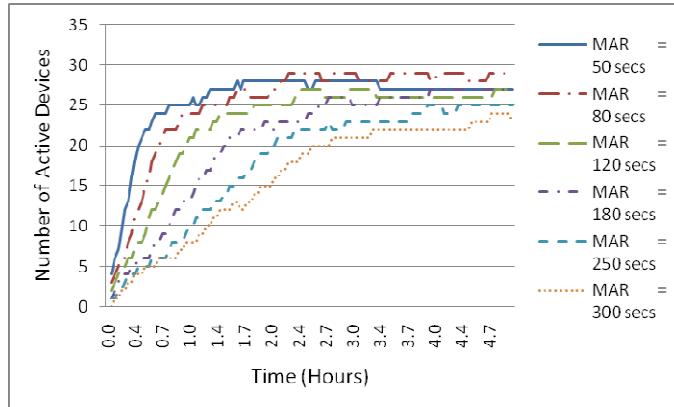


Figure 2.28 Number of supported devices with variable mean arrival rate

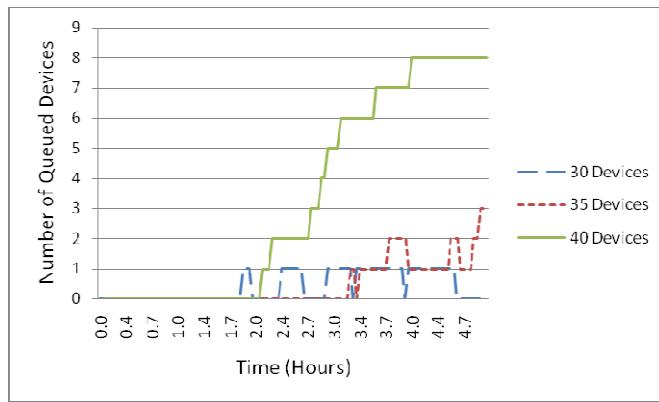


Figure 2.29 Queue size when the number of devices in the piconet change

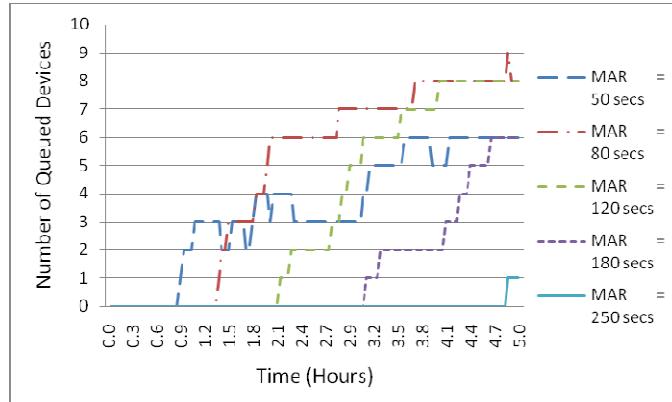


Figure 2.30 Queue size when the mean arrival rate is varied

For the simulation analysis, the majority of traffic type was considered to be Real Time (RT) or Non-deterministic flows. The reason was because HDR WPANs are intended for multi-media applications requiring high throughputs. For each flow request, the required throughput was calculated through a function in which the value of fragment size, transmission data rate and the number of frames/CTA was chosen randomly. For deterministic flows, the throughput was calculated based on the assumption that the flow requests are for file transfers. The size of the file was chosen based on a random variable, the maximum value of which was 100 MB. A comparison of throughput for deterministic and non-deterministic flows is given in Figure 2.31 from a scenario in which 30 member devices were considered with a mean arrival rate of 50 seconds. It can be noted that the maximum value of throughput for non-deterministic flows is just below 3.5 Mbps.

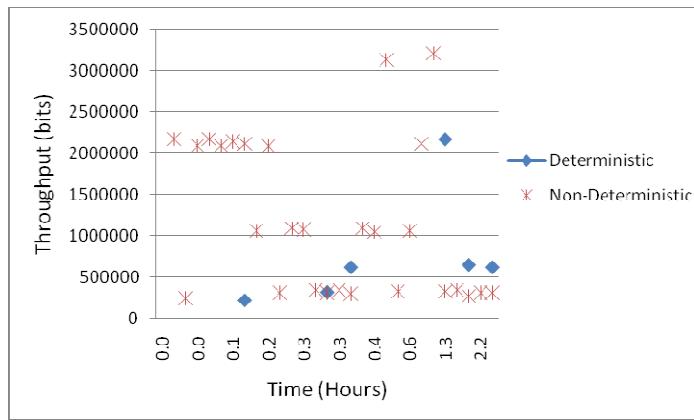


Figure 2.31 Comparison of Throughput for deterministic and non-Deterministic flows

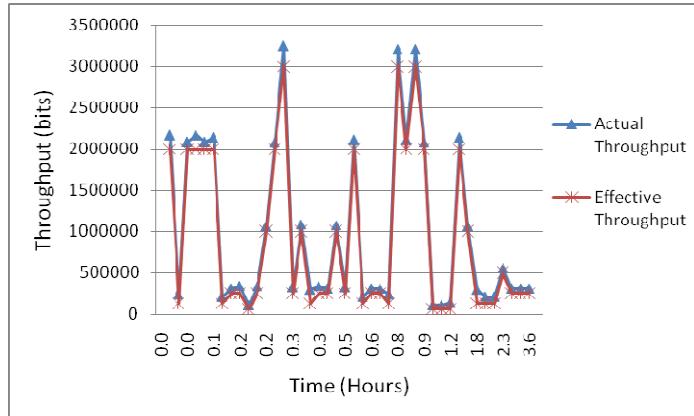
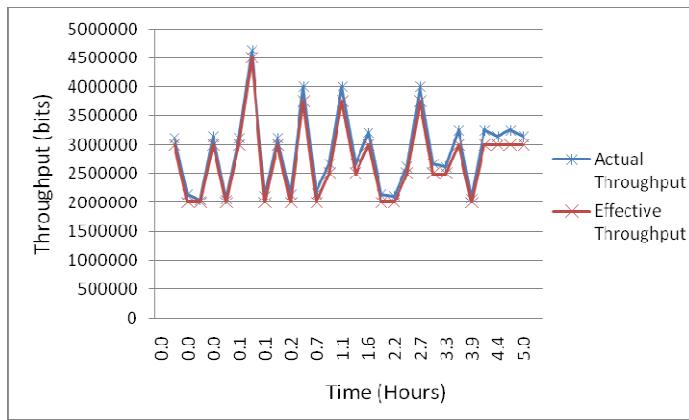
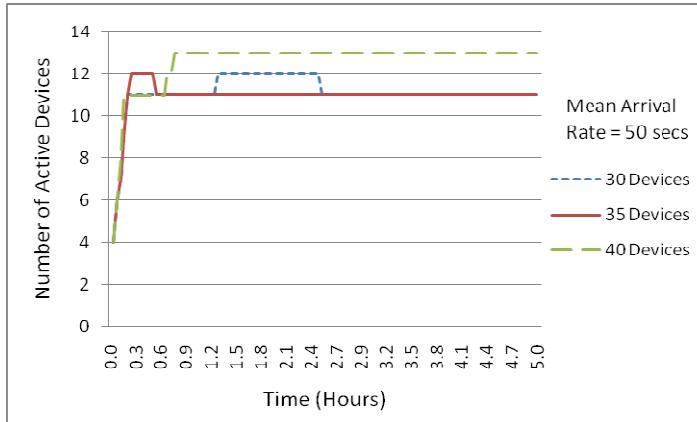


Figure 2.32 Comparison of Actual and Effective Throughput

A comparison between effective and actual throughput is given in Figure 2.32. The difference between the two curves is equal to the flow overhead at any given point on the graph. In another case, only non-deterministic flows were considered with relatively higher throughput requirements. The result can be seen in Figure 2.33. The maximum throughput in Figure 2.33 is 4.5 Mbps with some flows having throughput of 4 Mbps. The number of supported devices and the queue size when flows with high throughput requirements are considered are given in Figure 2.34 and Figure 2.35 respectively.

**Figure 2.33 Comparison of Actual and Effective Throughput for only non-deterministic flows****Figure 2.34 Number of supported flows with high throughput requirements**

It can be seen in Figure 2.34 that for a mean arrival rate of 50 seconds a significantly lower number of flows or devices is supported when compared to Figure 2.27. The number of queued requests also increases when flows with high throughput requirement are considered. For 40 member devices in the piconet, the queue size goes up to 27 requests. Therefore, it can be concluded, that the number of devices or simultaneous flows supported in the superframe is highly dependent on the throughput requirements of the requesting devices.

Inter-PAN Communication Evaluation

The important implications of the Inter-PAN communication framework proposed in this chapter are on superframe capacity and the queue size of the PNC. Therefore, the feasibility of the approach and an upper limit on the number of Child PNCs that can be supported is evaluated in this section. The number of Child piconet is kept from 1-3. Each child piconet has 5-10 member devices with variable throughput requirements. The request for channel

time from each Child PNC to the parent PNC is therefore different. If a device from the Child piconet communicates with another device in the Parent piconet, it sends a channel time request to the Parent PNC and vice versa. The Child PNC requests time from the Parent PNC to serve the intra-PAN flow requests from the member devices in the Child piconet.

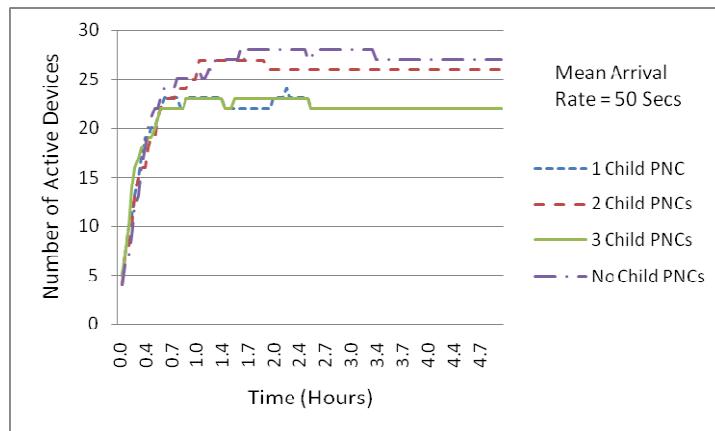


Figure 2.35 Number of supported flows when different number of Child Piconets are allocated Channel time

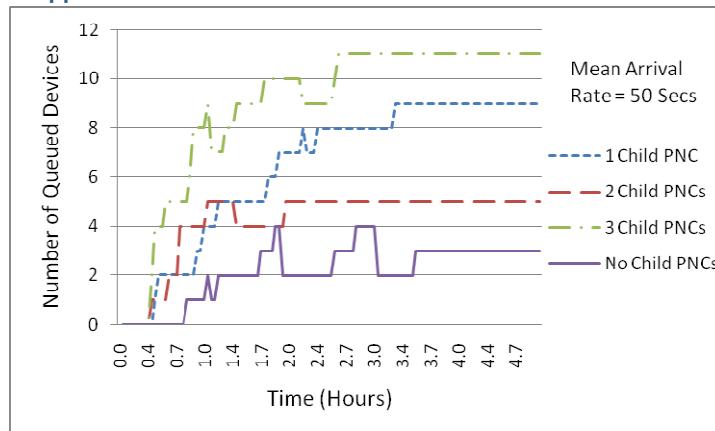


Figure 2.36 Number of queued requests when different number of Child Piconets are supported

In Figure 2.35, the number of allocated flows in the superframe is given with and without the support for Child PNCs. It can be seen that the largest number of supported flows result when only intra-PAN communication is supported. The number of supported flows seems identical when 1 Child Piconet and 3 Child piconets are considered. The reason is that the request of only one Child PNC with a channel time allocation requirement of 0.021640 seconds and 10 member devices is accepted. Similarly in Figure 2.36 the queue size is the lowest when only intra-PAN communication is considered. For both the cases of Figure 2.35 and Figure 2.36, the mean arrival rate is 50 seconds with 30 member devices in the Parent piconet. When the mean arrival rate is increased to 300 seconds, the effect on the queue size can be seen in Figure 2.37. In case of Figure 2.37, when 3 Child PNCs are

considered, only two of the Child PNCs with allocation requests of 0.021640 and 0.015620 and member devices 10 and 5 respectively, are accommodated in the superframe while the request of the third Child PNC is rejected because of insufficient superframe time.

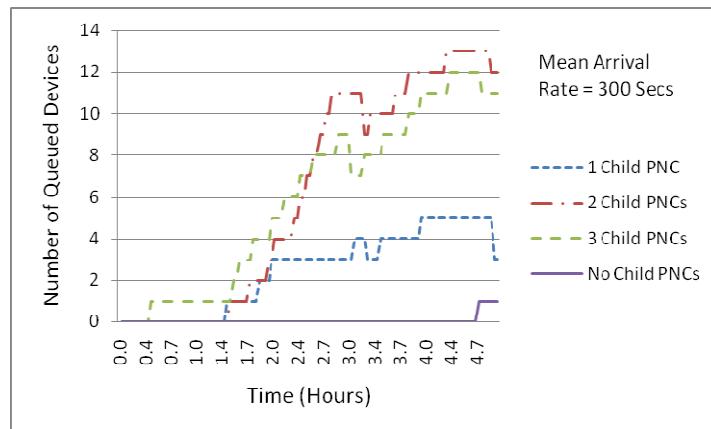


Figure 2.37 Number of queued requests when mean arrival rate is increased to 300 seconds

If the mean arrival rate is kept the same i.e. 300 seconds but the throughput requirements of devices is increased just like in the case of Figure 2.33, the number of simultaneously supported flows decreases and drops down to approximately 11 flows as shown in Figure 2.38. The queue size also increases correspondingly (Figure 2.39).

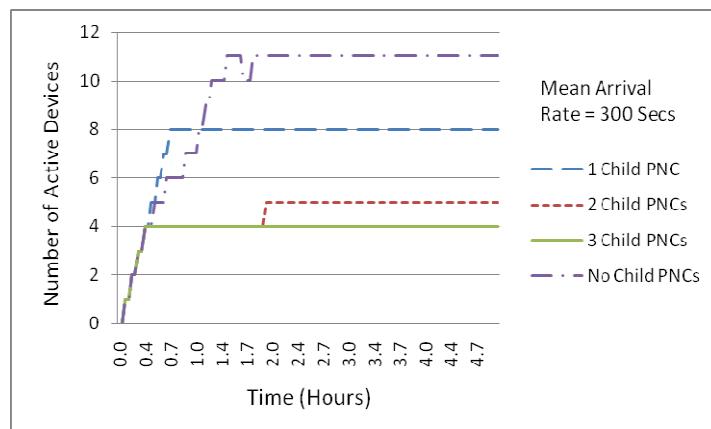


Figure 2.38 Number of supported flows with high throughput requirement from the devices

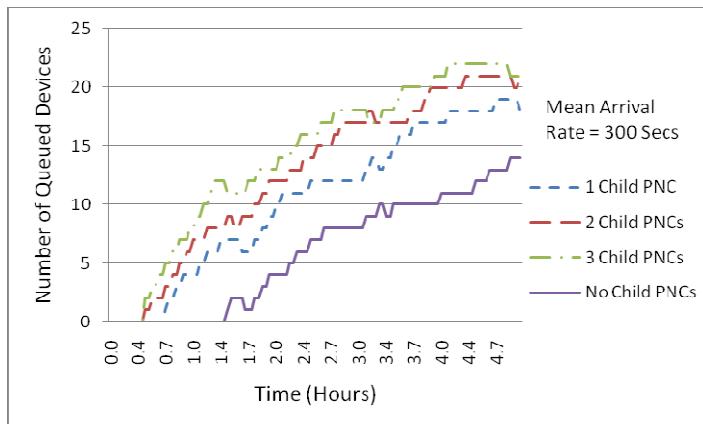


Figure 2.39 Queue size with high throughput requirement from the devices

With Inter-PAN communication, the results show that the number of intra-PAN flows that can be supported simultaneously in the superframe, decreases. The more than number of Child piconets accommodated in the superframe, then based on the channel time requirement of each Child PNC, the more the decrease in the number of intra-PAN devices that can be supported. Therefore, based on the priority of flows in a certain piconet, the PNC might allow inter-PAN communication or not. The PNC might also decide to limit the number of Child PNCs that can be supported or place an upper limit on the maximum channel time that can be allocated for inter-PAN communication in order to be fair with the member devices of its own piconet.

Admission Control based on Rejection Ratio

To evaluate the algorithms based on RR and IRR, two scenarios were considered. In the first scenario, both deterministic and non-deterministic flows were considered with the maximum throughput requirement upto 3 – 3.5 Mbps. In the second scenario, only real time i.e. non-deterministic flows were considered with relatively higher throughput requirements upto 4 – 4.5 Mbps. The mean arrival rate was 50 seconds and the number of member devices in the piconet was 40. In Figure 2.40, the performance of RR and IRR can be seen which is compared with a case without the application of these algorithms (shown as No RR in Figure 2.40). In the case when RR and IRR is not applied, requests from the queue are served on first come first served basis with the condition that only that queued request for which sufficient capacity becomes available in the superframe is served first. Figure 2.40 simply shows the distribution of number of allocations to each device. It can be easily

observed that with the application of RR, the allocations are more evenly distributed among devices when compared to the case without its application. For the case of Figure 2.40, the total number of allocations without the use RR is 42 during the 5 hours simulated time. The total number of allocations made when RR and IRR are applied is 36 and 46 respectively. The mean RR is 0.1247, 0.16667 and 0.13095 for the cases when RR is not applied, when RR is applied and when IRR is applied respectively. This indicates that with the application of IRR, the allocations are not only evenly distributed as much as possible but also the allocations from the queue are made in such a way that the total number of allocations made also increases.

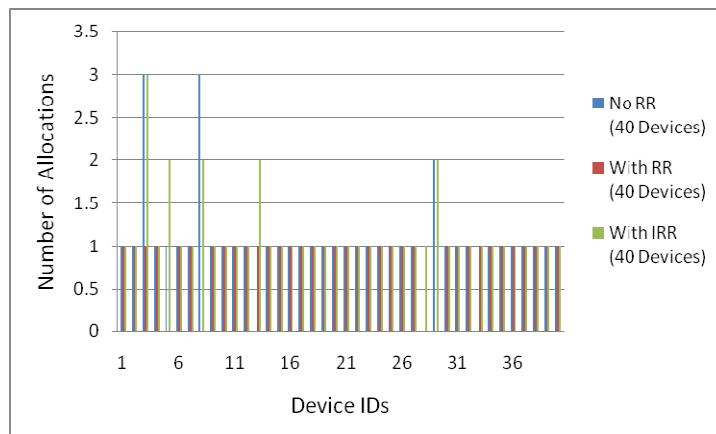


Figure 2.40 Comparison of RR and IRR when both RT and NRT flows are considered

The results of the second scenario in which only RT flows are considered with relatively higher throughput requirements are given in Figure 2.41. Once again, it can be seen that the allocations are more evenly distributed with the application of RR and IRR. However, in case of Figure 2.41, RR and IRR provide the same results. The reason for this might be the fact that since in case of taking high throughput requirements into consideration there might be deserving flows due to their higher number of rejections that need to be accommodated when a flow terminates but upon a flow termination, there is not enough time released to serve those flows. Therefore, the chance of RR and IRR serving the same request increases in such cases. Nevertheless, the results in Figure 2.40 and Figure 2.41 give a strong indication about the element of fairness the application of RR and IRR can bring when the requests for channel time need to be served in an even manner for all requesting devices by the PNC.

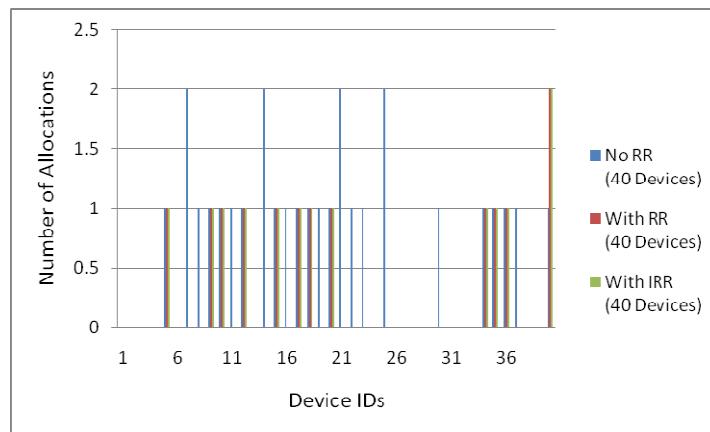


Figure 2.41 Comparison of RR and IRR when only RT flows are considered

2.9 Conclusion

The Inter-PAN Communication procedure although helps to improve peer-to-peer connectivity in HDR WPANs, there is an upper limit imposed by the limited capacity of the WPAN superframe. When the number of devices in either the Parent piconet or Child piconet increases, or if the throughput requirement of devices in either the Parent piconet or Child piconet increases, there is a substantial increase in the queue size. When the mean arrival rate is low and the requests for channel time are expected more frequently, it is shown that up to 2-3 Child piconets can be supported for up to 30 devices in the Parent piconet. The number of Child piconets that can be supported might increase if the number of devices in the Parent piconet is low with low throughput requirements. However, one of the aims in the chapter was to determine an upper limit of Child piconets that can be supported. From the results, it is evident that to increase the number of flows that can be supported along with Child piconets, capacity improvement techniques such as power control algorithms to improve spatial re-use and multi-radio devices can be considered which can operate on multiple radio frequency channels to improve capacity.

The capacity analysis and the simulation analysis to determine the capacity of the HDR WPAN superframe provides valuable insights not only to determine an upper limit on the number of flows that can be supported but also the effects of congestion and the resulting increase in the queue size. To serve the flow requests which are queued due to congestion in a fair manner, the RR and IRR based algorithms show a better performance in serving the

requests overall in a uniform manner. The RR and IRR do not give preference to any flow request on basis of its traffic type. However, in case if there are any flows that need to be dealt with on priority basis, or if there are any flows in a WPAN which are given priority by the user of that WPAN, then suitable extensions can be provided to RR and IRR in a simple way to deal with such flows in a fair manner. The extensions to RR and IRR are left as part of future work with the aim of providing more flexibility and options in the two algorithms to deal with priority based traffic.

2.10 References

- [1] "IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 15.3: wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs)," *IEEE Std 802.15.3-2003* , vol., no., pp. 0_1-315, 2003.
- [2] G. Boggia; P. Camarda; L. A. Grieco, "Scheduling channel time allocations in 802.15.3 WPANs for supporting multimedia applications," *Wireless Communications and Mobile Computing*, pp. 1530-8669, 2009.
- [3] Sun-Myeng Kim; Young-Jong Cho, "Scheduling scheme for providing QoS to real-time multimedia traffics in high-rate wireless PANs," *Consumer Electronics, IEEE Transactions on* , vol.51, no.4, pp. 1159-1168, Nov 2005.
- [4] Liu, X., Dai, Q. AND Wu, Q. 2004. Scheduling algorithms analysis for MPEG-4 traffic in UWB. In *2004 IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall*.
- [5] B.-S. Kim, S. Kim, Y. Fang, and T. Wong, "Feedback-assisted mac protocol for real time traffic in high rate wireless personal area networks," *Wireless Networks*.
- [6] Rangnekar and K. Sivalingam, "Qos aware multi-channel scheduling for IEEE 802.15.3 networks," *Mobile Networks and Applications*, vol. 11, no. 1, pp. 47-62, February 2006.
- [7] Yi-Hsien Tseng; Eric Hsiao-kuang Wu; Gen-Huey Chen, "Maximum traffic scheduling and capacity analysis for IEEE 802.15.3 high data rate MAC protocol," *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th* , vol.3, no., pp. 1678-1682 Vol.3, 6-9 Oct. 2003.
- [8] Chen, X., Xiao, Y., Cai, Y., Lu, J. AND Zhou, Z. 2006. An energy diffserv and application-aware MAC scheduling for VBR streaming video in the IEEE 802.15. 3 high-rate wireless personal area networks. *Computer Communications* 29, 3516-3526.
- [9] Mangharam, R., Demirhan, M., Rajkumar, R. AND Raychaudhuri, D. 2004. Size matters: size-based scheduling for MPEG-4 over wireless channels. In *SPIE & ACM Proceedings in Multimedia Computing and Networking*, 110–122.
- [10] Ranran Zeng; Geng-Sheng Kuo, "A novel scheduling scheme and MAC enhancements for IEEE 802.15.3 high-rate WPAN," *Wireless Communications and Networking Conference, 2005 IEEE* , vol.4, no., pp. 2478-2483 Vol. 4, 13-17 March 2005.
- [11] Moradi, S. AND Wong, V. 2007. Technique to improve MPEG-4 traffic schedulers in IEEE 802.15. 3 WPANs. In *IEEE International Conference on Communications, 2007. ICC'07*, 3782-3786.

- [12] Xi Chen; Jianhua Lu; Zucheng Zhou, "An enhanced high-rate WPAN MAC for mesh networks with dynamic bandwidth management," *Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE* , vol.6, no., pp.5 pp.-3412, 2-2 Dec. 2005.
- [13] Seung Hyong Rhee; Kwangsue Chung; Yongsuk Kim; Wonyong Yoon; Ki Soo Chang, "An application-aware MAC scheme for IEEE 802.15.3 high-rate WPAN," *Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE* , vol.2, no., pp. 1018-1023 Vol.2, 21-25 March 2004
- [14] Choi, E.C., Huh, J.D., Kim, K.S. AND Cho, M.H. 2006. Frame-size adaptive MAC protocol in high-rate wireless personal area networks. *ETRI Journal* 28, 660-663.
- [15] B. S. Kim, Y. Fang, and T. F. Wong, "Rate-adaptive mac protocol in high-rate personal area networks," in *Wireless Communications and Networking Conference, 2004. WCNC. 2004 IEEE*, vol. 3, 2004, pp. 1394-1399 Vol.3.
- [16] Doyoun Hwang; Euihyeok Kwon; Jaesung Lim, "A Virtual Slot Multiple Access for IEEE 802.15.3 High-Rate Wireless Personal Area Networks," *Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th* , vol., no., pp.1-5, 25-28 Sept. 2006.
- [17] Lin X. Cai, Xuemin (Sherman) Shen, Jon W. Mark, Lin Cai, "Capacity analysis and MAC enhancement for UWB broadband wireless access networks," *Computer Networks*, Volume 51, Issue 11, 8 August 2007, Pages 3265-3277, ISSN 1389-1286.
- [18] Medard, M.; Jianyi Huang; Goldsmith, A.J.; Meyn, S.P.; Coleman, T.P., "Capacity of time-slotted ALOHA packetized multiple-access systems over the AWGN channel," *Wireless Communications, IEEE Transactions on* , vol.3, no.2, pp. 486-499, March 2004.
- [19] Xuping Zhai; Guangguo Bi; Pingping Xu, "Call admission control with bandwidth reallocation for adaptive multimedia in high-rate WPAN," *Emerging Technologies: Frontiers of Mobile and Wireless Communication, 2004. Proceedings of the IEEE 6th Circuits and Systems Symposium on* , vol.2, no., pp. 557-560 Vol.2, 31 May-2 June 2004.
- [20] "IEEE Recommended Practice for Information technology--Telecommunications and information exchange between systems-- Local and metropolitan area networks-- Specific requirements Part 15.5: Mesh Topology Capability in Wireless Personal Area Networks (WPANs)," *IEEE Std 802.15.5-2009* , vol., no., pp.1-166, May 8 2009.

CHAPTER 3

3 A Resource Allocation Framework for High Data Rate Meshed WPANs

3.1 Introduction

Resource allocation has been one of the highly investigated topics in both wired and wireless communication networks. The term *resource* has been used to indicate different parameters of interest such as power, radio frequency channels, required bandwidth etc. Some important considerations made in most of resource allocation schemes have been their fairness, adherence to the QoS requirements, complexity, and whether it is centralized or distributed. In order to decide whether the resource allocation is centralized or distributed depends mainly on the network topology, network density and the complexity of the approach.

In this chapter, an effort has been made to work on the main aspects which are imperative in the performance of a resource allocation scheme and are sufficient to devise a framework for resource allocation. The parameters considered are traffic estimation, determining the appropriate bandwidth for each flow, decision making factors for resource allocation and fairness based on satisfaction from the requesting device's perspective. The network considered is based on the specifications of [1] for a meshed Wireless Personal Area Network in which there are a number of clusters or piconet and each cluster is controlled by a Meshed Piconet Coordinator (MPNC). Since the communication takes place in a superframe, which consists of guaranteed time slots, the MPNCs synchronize their transmission boundaries with a commonly selected MPNC called the Reference MPNC (ref-MPNC). In the proposed approach, we use the ref-MPNC for resource allocation to the requesting MPNCs. The resource in our case is therefore the *time slot* in the superframe on a basic level.

Each ref-MPNC keeps a record of its two hop neighbourhood and thus the resource allocation can be termed as centralized on the local level but distributed on the network

level since there can be many ref-MPNCs based on the size (based on number of hops) of the network. In [2], a distributed algorithm is proposed which calculates a fair share in the superframe for each requesting MPNC. It considers a cooperative approach in which the utility for each ref-MPNC is maximized in a way that the other requesting MPNC benefit as well. However, the algorithm is based on the individual requests for each flow by the MPNC. Therefore, as the number of devices in a cluster controlled by an MPNC increase, the number of requests from competing MPNCs increase and the conflict resolution (section 3.3) can take longer as a result. Furthermore, there are intra piconet (intra-PAN) flows as well as inter piconet (inter-PAN) flows with the former between devices in the same piconet or cluster and the later between MPNCs and their requirements are variable based on the traffic type.

We consider *bulk reservations* by each requesting MPNC in which it determines the time required for multiple queued and expected flows and then sends a request to the ref-MPNC. The main idea behind bulk reservation is that it reduces the frequency of channel time requests and most of the requests are dealt by an MPNC once it has been allocated appropriate channel time. In [3], a similar topology like the meshed WPANs is considered for intra-cluster packet-level resource allocation approach with effective QoS provisioning using a combination of the Karush-Kuhn-Tucker (KKT)-driven approach and a genetic algorithm (GA)-based approach. However, the appropriate traffic estimation techniques for each cluster are not taken into account.

Effective traffic estimation techniques are mandatory for the efficiency of resource allocation in the proposed approach. When an MPNC estimates its traffic, then based on its traffic requirements, it sends a request for channel time allocation (a number of guaranteed time slots in the superframe) to the ref-MPNC. If the ref-MPNC finds it feasible to allocate the requested channel time, then it allocates it, otherwise, it indicates to the requesting MPNC an estimated backoff duration after which appropriate channel time is expected to be available to cater the request. The estimated backoff duration is calculated on the basis of maximum channel holding time of each MPNC. The maximum channel holding time for each MPNC is the time for which it is expected to hold its allocated resources or channel time. The calculation of channel holding time is given in [4], [5] and [6]. Most of these approaches

are targeted towards calculating and determining the distribution of the holding times for voice traffic.

The impetus of the proposed work mainly comes from the notion that a centralized solution can lower the communication overhead as well as eliminate the channel time reservation conflicts between MPNCs that compete for channel time. Also using appropriate resource allocation policies, an element of fairness can be maintained in channel time allocation to requesting MPNCs. The contribution in this chapter consists of

1. Centralized resource allocation scheme based on the ref-MPNC
2. Estimated backoff algorithm as an alternative to random backoff by each MPNC once its request for channel time is rejected
3. Proposed traffic estimation techniques to determine an appropriate share of channel time required by each MPNC
4. Three different proposed resource allocation schemes to determine a fraction of channel time which can be allocated to a requesting MPNC
5. Evaluation of proposed schemes via simulation results and by using fairness metrics to check the fairness of proposed resource allocation schemes

3.2 Related Work

The term resource allocation has been used in a different context depending on the type and requirements of a network. In this chapter, the term resource allocation indicates the number and duration of CTAs which are allocated to a requesting MPNC. In [7], an intra-cluster resource allocation algorithm is proposed for wireless mesh networks that take into consideration power allocation, sub-carrier allocation and packet scheduling for a utility maximization problem. The traffic requirements of each device are determined by averaging their previous traffic loads periodically in a time window of 100ms. The proposed scheme is centralized and considers intra-cluster resource allocation on per flow basis. However, inter-cluster resource allocation is not considered. A resource allocation scheme for TDMA based systems is proposed in [8] that uses the maximum entropy principle to maximize the throughput of different flows. Some of the other resource allocation approaches for TDMA based systems are proposed in [9], [10] and [11]. In [9] the resource allocation scheme

considers only voice flows for low capacity wireless systems and aims to efficiently utilize the capacity. The proposed work in [10] assumes a centralized scheme for multi-hop wireless networks which are TDMA based. A static time allocation strategy is used by taking into account the channel state information. The author in [10] claims the proposed scheme to be more effective than the policy of allocating equal time to all requests. However, in [9] and [10] equal sized time slots are considered rather than variable size time slots which is the case in meshed WPANs. An end-to-end distributed resource assignment scheme for TDMA based ad-hoc networks is proposed in [11] which relies on capacity requests and grants.

The scheme proposed in [12] considers power allocation as well as user scheduling. The aim of the resource allocation policy is to maximize capacity rather than fairness. In [13] three utility based resource allocation schemes are proposed for centralized wireless networks and it is argued that optimal resource allocation depends on wireless traffic types, total available resource and channel quality. The Time Aware Resource Allocation (TARA) [14] scheme uses resource utility functions for allocating bandwidth to connections in order to maximize the accumulated utility of the system. TARA considers a centralized scheme for small networks while a distributed scheme for larger networks. In the proposed approach in this chapter, the resource allocation is done on a centralized basis for a two hop neighbourhood. Therefore, it can be termed as centralized in a local context while distributed in the global context of the meshed WPAN. A combination of genetic algorithm and Karush-Kuhn-Tucker (KKT) based centralized approach is given in [3] for intra-cluster packet-level resource allocation for mesh networks. The objective function is chosen to maximize the system throughput. The most relevant work for resource allocation in meshed WPANs is given in [2] which propose a distributed resource allocation scheme based on a utility function. The channel time is allocated to requesting devices in such a way so that the utility for the rest of the devices is maximized. The scheme in [2] considers resource allocation on per request basis. However, as the number of devices is increased in each cluster and also in presence of a dense meshed WPAN with a higher number of MPNCs, the number of requests for channel time is expected to increase which increases the probability of channel time reservation conflicts. A channel time reservation conflict can arise when two

or more devices try to reserve channel time so that their reservations overlap in the superframe. Furthermore, the overhead due to channel time reservation requests increases. In the proposed scheme, the concept of bulk reservations is used in which each MPNC predicts its channel time requirement for intra-PAN and inter-PAN flows and sends a request to reserve channel time in bulk. Since the proposed scheme is centralized, there is no possibility of channel time conflicts because a common MPNC which is chosen for time allocation keeps track of the allocations in the two hop neighbourhood.

3.3 Channel Time Reservation in Meshed WPANs

In meshed WPANs, each MPNC has to transmit its own beacon and listen to the beacons transmitted by its neighbour MPNCs. Therefore the beginning of the superframe consists of a Beacon Period (BP) in which the beacons sent by all the MPNCs sharing a superframe are accommodated and space is reserved for possibility of new MPNCs joining the same BP. The superframe time is divided into equal slots called Medium Access Slots (MASs) for ease of channel time management. The superframe size is kept the same throughout the mesh network. The member MPNCs of the same BP share the superframe time among themselves for inter-PAN and intra-PAN communication. Due to the hidden device problem [15], each MPNC has to take into account the CTA reservations in its two hop neighbourhood. Therefore, extra time has to be reserved in case of multi-hop communication due to the hidden device problem which significantly affects the superframe capacity [16]. The extra time reserved due to hidden device problem can be relieved to some extent using multiple channels through effective transmission power control schemes. The superframe time can be reserved for the following purposes in a meshed WPAN

- Transmission and reception of beacon frames
- Data frame exchange which can consist of inter-PAN flows and intra-PAN flows
- Possibly some control information like routing information packets i.e. route request, route reply etc

The time reservation scheme given in [1] takes into account channel time reservation by an MPNC through sending channel time reservation requests for each flow individually. There are two ways for reserving channel time through negotiation between a source MPNC

(SMPNC) and a destination MPNC (DMPNC) i.e. explicit and implicit. In explicit channel time reservation, the SMPNC sends a channel time request command to the DMPNC. If the negotiation is a success and there is available channel time, then the DMPNC responds with a channel time response command. After successful negotiation, the new reservation is indicated to the neighbour MPNCs through an Information Element in the beacon frame called the CTA status IE. The neighbour MPNCs can object to the new reservation if it is in conflict with the reservation of any of the neighbour MPNCs by including an appropriate IE in their beacon. If an objection is raised, the SMPNC and DMPNC can either move the CTA allocation or remove it. In order to resolve CTA conflicts, the CTA status IE can be continuously monitored by the neighbour MPNCs. If after a CTA reservation, the need for increasing or decreasing the reservation arises, the SMPNC can send a channel time request command again to the DMPNC. In [1], it is mentioned that for intra-PAN reservation, the MPNC does not negotiate with its neighbours and includes the required channel time for intra-PAN flows as already negotiated. However, it waits for an objection which can be raised by any of the neighbour MPNCs. The estimation of intra-PAN traffic requirement which determines the amount of channel time to reserve by an MPNC is not given in [1].

The idea of undergoing a reservation procedure for each individual flow can add to the probability of CTA conflicts in case of higher device density. The procedure for channel time reservation in [1] does not consider fairness in channel time allocation. Each flow allocation to an MPNC should follow a fairness criterion such that all the MPNCs can get a fair share of the superframe time. In [2] a parameter α is used to determine a fair allocation for each MPNC. However, the approach is generic and ignores details about some important factors like the distribution of inter-PAN and intra-PAN traffic and reservation conflicts which can arise. The possibility of reserving channel time in bulk by each MPNC based on its expected intra-PAN and inter-PAN traffic requirement is explored in this chapter. The idea behind bulk reservation is to reduce the possibility of CTA conflicts in case of higher device density and simplify the channel reservation process between an SMPNC and DMPNC. In bulk reservation, techniques for traffic estimation are utilized to determine the reservation requirement for an MPNC. Once the SMPNC is allocated channel time by the ref-MPNC, it

can use a portion of that time to communicate with other neighbour MPNCs without waiting for an objection by another neighbour MPNC.

3.4 Centralized Time Allocation

In the centralized approach of resourced allocation, the ref-MPNC is responsible for the channel time allocations in a fair manner to the requesting MPNCs as well as determining the inter-PAN fairness parameter α . The traffic distribution parameter ψ_d is locally defined by each MPNC based on its expected inter-PAN and intra-PAN traffic. The value of ψ_d determines the fraction of time to be allocated to inter-PAN flows such that $0 \leq \psi_d \leq 1$ and $1-\psi_d$ is the fraction of time to be allocated to intra-PAN flows. Thus an appropriate value of ψ_d can ensure a balance between channel time requirement for inter-PAN and intra-PAN flows. If the value of ψ_d is 0, it means that no inter-PAN flows are expected and the allocations are to be requested for intra-PAN traffic only. On the other hand, a value of 1 indicates that no intra-PAN traffic is expected and therefore, all the allocations can be requested for inter-PAN traffic. The expected intra-PAN traffic depends on the number of member devices in a cluster administered by an MPNC and the category of traffic intended to be transferred for each request. The expected inter-PAN traffic depends on the number of first and second hop neighbours of an MPNC and the traffic category intended for the inter-PAN flow being requested. The higher the number of neighbours of an MPNC, the higher the probability of inter-PAN requests that it can receive. Based on the expected inter-PAN and intra-PAN traffic, an MPNC calculates the total time that it needs to request from a ref-MPNC and also determine an appropriate value of ψ_d . In the centralized resource allocation approach, a requesting MPNC first requests channel time in bulk based on its requirements for a time period T_R . The block of time requested by an MPNC in bulk for a time period T_R is denoted by T_B . If at a time instant $t < T_R$, the MPNC gets some free time due to some of the flows ending and it receives either a request for inter-PAN flow or intra-PAN flow, it may request time for an individual flow in that case from the ref-MPNC depending on the resource allocation policy. After an MPNC calculates the time it requires and requests it from the ref-MPNC, the following possibilities arise based on whether there is time available in the superframe or not.

--It doesn't get any time at all

- It gets the time that it has requested
- It gets a fraction of the requested time if feasible

If we assume that some of the inter-PAN or intra-PAN flows have a higher priority, then the ref-MPNC has to ensure that at least the requesting MPNC gets enough time for the priority flows if possible. Let the capacity of the superframe be $C \mu$ sec and the amount of time allocated to other MPNCs be T_o , then the fair share of the i^{th} MPNC can be calculated as

$$T_i = \alpha(C - T_o) \quad (3.1)$$

When an MPNC initially requests the ref-MPNC for channel time, it indicates to the ref-MPNC through an IE about the desired time that it needs i.e. T_B . If the ref-MPNC is not able to allocate the requested time i.e. $T_i = T_B$, then based on the resource allocation policy, it either rejects the request or allocates it time equal to $T_i < T_B$. In case the ref-MPNC rejects the request, it estimates the availability of time in the superframe for the subsequent time period T_R such that if $T_i \leq (C - T_o)$ in any $T_R^{i+1}, T_R^{i+2} \dots T_R^{i+n}$ (if T_R^i is the present time period), then it notifies the requesting MPNC to wait for the suitable number of time periods. Each time period T_R is taken to be equal to a number of superframe durations $n\eta_S$ such that η_S is the superframe duration and $n \in \mathbb{Z}^+$ (set of positive integers). The traffic estimation techniques used for the proposed approaches are given in the next section.

If $G = (N, A)$ is used to denote a directed graph defined by a set N of n devices and a set A of m directed arcs then each of the n devices in the set N represents an MPNC. Similarly $G_p = (N_p, A_p)$ is used to represent the directed graph for the set N_p of n_p devices and a set A_p of m_p directed arcs for the cluster managed by MPNC_i. Each of the n_p devices in the set N_p represents a member device in a cluster. For each link or arc $(i, j) \in A$ if u_{ij} is the capacity which denotes the maximum flow while l_{ij} denotes the minimum amount that must flow through the link (i, j) , then the flow between two MPNCs is bounded by the total time allocated to them. More specifically, in the allocated time to an MPNC i.e. T_i , the traffic distribution parameter ψ_d defines an upper limit for the inter-PAN flows. The flow on the link $(i, j) \in A$ between two MPNCs is denoted by x_{ij} and is a sum of deterministic flows x_{ij}^d and non-deterministic flows x_{ij}^{nd} . The deterministic flows are those for which the total flow duration time can be estimated based on the value of throughput that is maintained for the

flow. The non-deterministic flows are those for which the total flow duration time cannot be determined i.e. real time flows. The minimum amount of flow that must flow through a link i.e. l_{ij} includes the time reserved for beacon transmission and reception, control traffic etc. Assuming a requesting MPNC_i have been allocated the requested time T_i then the inter-PAN flow and its relevant constraints can be defined by the following general problem description

$$\sum_{\{j:(i,j) \in A\}} x_{ij} - \sum_{\{j:(j,i) \in A\}} x_{ji} = b(i) \quad \forall i \in N \text{ and } b(i) \leq 0 \quad (3.2)$$

$$l_{ij} \leq x_{ij} \leq u_{ij} \quad \forall (i,j) \in A \quad (3.2a)$$

$$x_{ij} \leq T_i \leq u_{ij} \quad (3.2b)$$

$$x_{ij} \leq \alpha_b(C - T_o) \quad (3.2c)$$

$$x_{ij} \leq \psi_d(\alpha_b(C - T_o)) + (1 - \psi_d)(\alpha_b(C - T_o)) \quad (3.2d)$$

Or

$$x_{ij}^d + x_{ij}^{nd} \leq \psi_d(\alpha_b(C - T_o)) + (1 - \psi_d)(\alpha_b(C - T_o))$$

Where x_{ij} is the flow from MPNC_i to MPNC_j and the term $b(i)$ indicates if more traffic flows from MPNC_i to MPNC_j when $b(i) > 0$ or vice versa if $b(i) < 0$. Since an MPNC reserves time for the intra-PAN flows as well as the inter-PAN flows when it requests time from a ref-MPNC based on its traffic requirements, therefore x_{ij} is used to denote the total traffic flow which consists of the inter-PAN flows as well as the intra-PAN flows. The first term on the right side of (3.2d) i.e. $\psi_d(\alpha_b(C - T_o))$ is the upper limit on the inter-PAN traffic while the second term i.e. $(1 - \psi_d)(\alpha_b(C - T_o))$ is the upper limit on the intra-PAN traffic.

When a channel time request is sent to a ref-MPNC, it determines the fair share of the requesting MPNC by determining an appropriate value of α . If the request is received at a time instant t_o^i in the time period T_R^i , then we use t_R^i to indicate the start of the time period T_R^i and t_R^{i+1} to indicate the start of the next subsequent time period T_R^{i+1} and it also implicitly means the end of T_R^i . The status of the superframe utilization and the expected requests in the time interval $t_R^i \leq t_o^i \leq t_R^{i+1}$ are two important factors in determining an appropriate value of α_b . If there is enough available capacity in the superframe, then a fair

share of that capacity is allocated to the requesting MPNC. If there is not enough available capacity in the superframe, then the ref-MPNC negotiates a backoff duration with the requesting MPNC. The backoff duration is an integral multiple of superframe durations i.e $n\eta_S$. The ref-MPNC calculates the backoff duration by estimating the probability that enough capacity is available in the superframe to be allocated in a certain time period T_R^{i+n} . The details of traffic estimation by a requesting MPNC and the ref-MPNC are given in the subsequent section. The ref-MPNC is said to be in blocking state when the superframe capacity is exhausted. Since in the blocking state, the requesting MPNC waits for a time duration indicated by ref-MPNC and the ref-MPNC queues that request, the blocking probability can be given by the Erlang C formula

$$P_w = \frac{\frac{A^N}{N!} \frac{N}{N-A}}{\sum_{i=0}^{N-1} \frac{A^i}{i!} + \frac{A^N}{N!} \frac{N}{N-A}} \quad (3.3)$$

Where P_w is the probability that the requesting MPNC has to wait in the queue, A is the offered traffic in Erlangs and N is the superframe capacity in TUs. The value of A is determined by $A = \lambda_A h$ where λ_A is the mean arrival rate and h is the mean holding time. The value of h is taken to be the mean of the time durations allocated to the m requesting MPNCs such that $m \leq M$ and M is the number of MPNCs in the two hop neighbourhood of the ref-MPNC.

$$h = \frac{\sum_i T_i}{m} \quad \forall i = 1, 2, 3, \dots, m \quad (3.4)$$

The value of h is calculated for an integral number of time periods nT_R and the value of n depends on the duration of T_R . For the i^{th} time period T_R^i , we consider an arrival rate λ_R^i based on the poisson distribution such that the next subsequent arrival rate $\lambda_R^{i+1} \leq \lambda_R^i$. In other words, the poisson arrival rate is homogeneous within a time period T_R^i but it can vary for the subsequent time period and can be non-homogeneous within subsequent time periods. The reason for choosing different poisson arrival rates is because of the nature of expected traffic which can vary significantly throughout the day. For office meshed WPAN networks, the maximum traffic can be expected at certain times during office hours and outside the office hours, the traffic can decrease significantly. The average arrival rate λ_A used to calculate the value of offered traffic A can be calculated by

$$\lambda_A = \frac{\sum_i \lambda_R^i}{n} \quad \forall i = 1, 2, 3 \dots n \quad (3.5)$$

$$\lambda_R^i \leq \lambda_R^{i+1} \leq \lambda_R^{i+2} \dots \dots \lambda_R^{i+n-1}$$

Where the i^{th} arrival rate for a time period T_R^i is taken to be the one in which the allocation is done to the first requesting MPNC at the start of the period for which the traffic is considered to calculate the value of P_w . The calculation of P_w takes into account the current active allocations and the estimated allocations until the end of the n^{th} time period T_R^n till which it is to be calculated. Before an MPNC sends a request for channel time to the ref-MPNC, it estimates its required traffic T_B for its expected and queued flows. The requesting MPNC also indicates to the ref-MPNC about the maximum flow transfer duration d_{max} for its requested allocation T_i . The value of d_{max} is the estimated duration of the longest deterministic or non-deterministic flow whichever is greater and can either be an inter-PAN flow or an intra-PAN flow. The estimation of d_{max} is measured in integral number of superframe durations and is discussed in Section 3.6. An accurate estimation of d_{max} aids the ref-MPNC in indicating to those MPNCs a backoff duration whose requests are rejected because there is no available channel time. Therefore, the ref-MPNC keeps a record of each d_{max} value indicated by a requesting MPNC to serve those MPNCs whose requests are queued and to determine an appropriate backoff duration for each rejected MPNC. Since the backoff duration is calculated by the ref-MPNC and indicated to the queued MPNCs, it's not a random backoff duration and is called an *estimated backoff*. When a requesting MPNC calculates the value of T_B based on its expected demand, it sends the request to the ref-MPNC. If there is not enough capacity in the superframe, then based on the d_{max} values of those MPNCs for which time has been allocated, the ref-MPNC estimates a backoff duration after which enough time is expected to be available to serve that request. The requesting MPNC sets a timer based on the backoff duration and resends the request again upon expiration of the timer. The estimated backoff procedure is summarized below in Figure 3.1.

The Estimated Backoff Procedure

MPNC estimates traffic and determines the value of T_B , α , ψ_d , d_{max} and T_p

MPNC sends channel time request to Ref-MPNC and requests T_B

If ($T_B \leq (C - T_o)$)

{

```

Calculate  $T_i = \alpha(C - T_o)$ 
If ( $T_B \leq \alpha(C - T_o)$ )
Allocate channel time  $T_i$ 

else if ( $T_B > \alpha(C - T_o)$ )
{
If ( $\left| \frac{T_B - T_i}{T_B} \times 100 \right| \leq 15$ )
Allocate  $T_i$ 
}
}
else if ( $T_B \geq (C - T_o)$ )
{
If ( $\left| \frac{T_B - T_i}{T_B} \times 100 \right| \leq 15$ )
Calculate and allocate  $T_i$ 

else
{
Ref-MPNC notifies requesting MPNC about the backoff duration
MPNC back's off for the notified duration
}
}

```

Figure3.1 Estimated Backoff Algorithm

Each inter-PAN or intra-PAN flow can either be granted the minimum number of TUs or desired number of TUs. Let the duration for the minimum number of TUs for the i^{th} flow be m_i and the duration for the desired number of TUs for the i^{th} flow be M_i , then the decision to allocate m_i or M_i to a request for single flow depends on whether the requested MPNC has already been allocated channel time by the ref-MPNC or not. An example of the superframe structure is shown in Figure 3.2. If the MPNC which receives the request has already been allocated channel time in bulk by the ref-MPNC, then it makes the decision to either allocate m_i or M_i to the flow based on available time and traffic type. If the MPNC has been allocated channel time but there is no available time to allocate to the request, then it sends a channel time request to the ref-MPNC and requests time from the ref-MPNC for a single flow. If the request is for an intra-PAN flow, then the requesting MPNC decides whether to allocate m_i or M_i to the flow. If the request is for an inter-PAN flow, then the ref-MPNC decides whether to allocate m_i or M_i to the flow. The ref-MPNC also takes the decision of allocating m_i or M_i based on available superframe time and traffic type. If the inter-PAN flow request is for a multi-hop flow between a particular source-destination pair, then assuming that the path information is available, the capacity feasibility of the path is determined with

the lower constraint set to M_i . If all the links in that path can support M_i , only then it is appropriate to allocate M_i to the inter-PAN flow.

If the ref-MPNC gets channel time request for a single inter-PAN flow from MPNC_i and at the same time, it also gets a request from another MPNC_j which had backed off with $(C - T_o)$ being the released superframe time, then there is a possibility of the following situation to occur

$$T_i \leq (C - T_o) \text{ and } T_j \leq (C - T_o) \text{ such that}$$

$$m_i \leq T_i \leq M_i \text{ and } T_j = m_j$$

$$T_j \leq (C - T_o) - m_i \text{ and } T_o \not\geq T_i$$

$$T_j > (C - T_o) - M_i$$

In such a case, the ref-MPNC allocates m_i to MPNC_i to accommodate MPNC_j also. In a similar scenario we can also have a second case (which is not considered in this chapter)

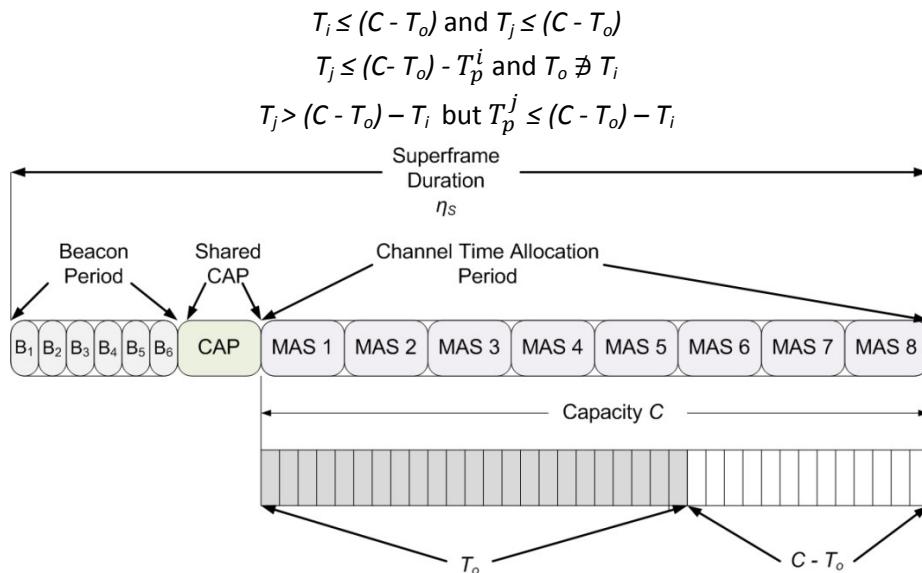


Figure 3.2 General Superframe Structure and Allocated time

Where T_p^i and T_p^j are the times required for priority flows of MPNC_i and MPNC_j, then the ref-MPNC allocates maximum channel time to that MPNC which has the lower d_{max} value. In the centralized resource allocation approach, the main advantage is that the ref-MPNC manages the resources and therefore when an MPNC is being allocated channel time, the ref-MPNC informs the other MPNCs about the allocation and it is ensured that there are no conflicts with other allocations in the one or two hop neighbourhood.

3.5 Traffic Estimation

The traffic estimation techniques in this chapter are based on the Matrix Balancing technique and the Assignment Problem based optimization.

3.5.1 Matrix Balancing

Matrix Balancing (MB) problems have been utilized in a variety of applications i.e. economics, regional planning, transportation, telecom and transport traffic estimation etc. In [17], the requirement for the estimated matrix to be symmetric is emphasized for telecommunication demand forecasting. In [18] a general reference on MB algorithms can be found. Some traffic estimation techniques and their comparison is given in [19]. Generally in MB problems, one is interested in estimating an $m \times n$ matrix D , given an $m \times n$ matrix A , that has given row sums and column sums and it approximates the matrix A in some sense e.g. comparable Eigen values, row and column sums etc. An approach which considers upper bounds on the elements a_{ij} of a matrix A is given in [20]. However, in our case, we limit the constraints to the row and column sums when estimating the traffic demand matrix D . Therefore, given an $m \times n$ nonnegative matrix $A = (a_{ij})$ and positive vectors $u \in \mathbb{R}^m$ (set of m real numbers) and $v \in \mathbb{R}^n$, a nonnegative matrix $X = (x_{ij})$ of the same dimension is estimated, such that

$$\sum_{j=1}^n x_{ij} = u_i \text{ for } i = 1, 2, \dots, m \quad (3.6)$$

$$\sum_{i=1}^m x_{ij} = v_j \text{ for } j = 1, 2, \dots, n \quad (3.7)$$

and $x_{ij} > 0$ only if $a_{ij} > 0$

$$\text{Let } A^{IPT} = \begin{bmatrix} a_{ij} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{mj} & \cdots & a_{mn} \end{bmatrix} \text{ and } A^{IT} = \begin{bmatrix} a_{ij} & \cdots & a_{ip} \\ \vdots & \ddots & \vdots \\ a_{rj} & \cdots & a_{rp} \end{bmatrix}$$

Where A^{IPT} is a square matrix of order $m \times n$ representing the time allocated to intra-PAN flows and $m = n$ while A^{IT} is a square matrix of order $r \times p$ representing the time allocated to inter-PAN flows and $r = p$. The traffic matrices are considered to be square matrices because bidirectional flows are considered. Therefore $a_{ij} \in A^{IPT}$ indicates the time reserved for a flow from device i to device j whereas $a_{ij} \in A^{IT}$ indicates the time reserved for a flow from device j to device $i \forall i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ where $a_{ij} \leq a_{ji}$. Similarly, the entries $a_{ij} \in A^{IT}$ indicate the time reserved for inter-PAN flows from MPNC i to MPNC $j \forall$

$i = 1, 2, \dots, r$ and $j = 1, 2, \dots, p$ where $a_{ij} \leq a_{ji}$. The outflow of traffic from a device contributing to the intra-PAN flows can be given as a vector r_i^{IPT} while the inflow of traffic to a device contributing to the intra-PAN flows can be given as a vector c_j^{IPT} where

$$r_i^{IPT} = \sum_j a_{ij} \quad \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n \quad (3.8)$$

$$c_j^{IPT} = \sum_i a_{ij} \quad \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n \quad (3.9)$$

$$r_i^{IPT} = [r_1 \ r_2 \ r_3 \ \dots \ r_m] \quad \text{and} \quad c_j^{IPT} = [c_1 \ c_2 \ c_3 \ \dots \ c_n]$$

$$A_T^{IPT} = r_i^{IPT} + c_j^{IPT} = [a_1 \ a_2 \ a_3 \ \dots \ a_n]$$

Where $a_i = r_i + c_i \quad \forall i = 1, 2, \dots, n$ and $a_i \in A_T^{IPT}, r_i \in r_i^{IPT}, c_i \in c_j^{IPT}$

The value of elements $a_i \in A_T^{IPT}$ indicates the superframe time required for the total inflow and outflow traffic from/to a device i . Similarly for inter-PAN flows, the elements in the vector A_T^{IT} indicate the total superframe time required for the traffic inflow and outflow from/to an MPNC i . From hereon, we shall use the term entry to refer to an element in the matrix which indicates a source-destination pair in that matrix unless otherwise specified.

In each time period T_R , the values of non-zero entries $a_{ij} \in A^{IPT}$ and $a_{ij} \in A^{IPT} \quad \forall a_{ij} \neq 0$ and $\forall a_{ij} (i \neq j)$, are taken and a sample set of observations for each entry is maintained for a certain sample set size n_s i.e. $S_{ij} = \{a_{ij}^1, a_{ij}^2, a_{ij}^3, \dots, a_{ij}^{n_s}\}_{i \neq j}$. Similarly, the values of entries in A_T^{IPT} and A_T^{IT} are also sampled at each time period T_R and a sample set of observations is maintained i.e. $S_{ij} = \{a_{ij}^1, a_{ij}^2, a_{ij}^3, \dots, a_{ij}^{n_s}\}_{i=j}$. Based on the pdf of the observations of the sample sets, two demand matrices A_D^{IPT} and A_D^{IT} are formed.

Such that

$\sum_{m \text{ or } n} a_{ij} (\forall i = j) \rightarrow \text{Estimate of total Intra-PAN traffic}$

$\sum_j a_{ij} (\forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n) \rightarrow \text{Estimate of total traffic from device } i$

$\sum_i a_{ij} (\forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n) \rightarrow \text{Estimate of total traffic to device } i$

$a_{ij} (\forall i = j) \rightarrow \text{Estimate of total traffic from/to device } i$

The non diagonal entries represent estimates of uni-directional flows while the diagonal entries represent estimates of the total traffic from/to device i . The resultant matrix has a rank which equals the total average or estimated traffic. The rest of the non-diagonal entries

represent the individual channel times occupied by flows, which are either estimated or indicate an average time for that entry.

Before the RAS algorithm is applied to A_D^{IPT} and A_D^{IT} , the diagonal entries of traffic demand matrices A_D^{IPT} and A_D^{IT} are set to zero so that only the entries representing the unidirectional flows are left which are expected to establish a link in a subsequent time period T_R . In order to determine the likelihood of a particular flow being initiated by $a_{ij} \in A_D^{IPT}$ or $a_{ij} \in A_D^{IT}$, frequency distribution matrices F^{IPT} and F^{IT} are defined such that $f_{ij} \in F^{IPT}$ and $f_{ij} \in F^{IT}$ represent the frequency of channel time requests for each corresponding entry $a_{ij} \in A_D^{IPT}$ and $a_{ij} \in A_D^{IT}$ in a time period T_R . Before an MPNC sends a channel time request to the ref-MPNC at a time instant t_o^i in the time period T_R^i , it calculates the time required in the superframe for its intra-PAN and inter-PAN flows. It takes into consideration the time required for queued requests if any and the requests expected in the period from t_o^i to t_R^{i+1} . We assume that the active devices (those already involved in data transfer) have a minimal probability to send another request for channel time in the period from t_o^i to t_R^{i+1} . If m is the total number of intra-PAN devices, r_q is the number of queued requests and r_a is the number of active devices, then the number of devices from which requests can arrive is $m - r_q - r_a = r_l$. In order to calculate the probability of requests expected in the period t_o^i to t_R^{i+1} , the individual probability for number of intra-PAN devices from $r = 1, 2, 3, \dots, r_l$ is calculated by

$$p(r, \lambda_R^{IPT}) = \frac{e^{-\lambda_R^{IPT}} (\lambda_R^{IPT})^r}{r!} \quad (3.10)$$

Where $p(r, \lambda_R^{IPT})$ is the probable outcome of r based on the poisson distribution, λ_R^{IPT} is the mean rate of intra-PAN requests and $\lambda_R = \lambda_R^{IPT} + \lambda_R^{IT}$. The value of r which has the highest probability is therefore selected as the expected number of requests in the period t_o^i to t_R^{i+1} .

After the number of expected requests for intra-PAN and inter-PAN flows is determined, the next step is to select the most probable entries in $a_{ij} \in A_D^{IPT}$ and $a_{ij} \in A_D^{IT}$ which are likely to send channel requests. Since the entries for the queued requests are already known for intra-PAN flows and inter-PAN flows, the selection of most probable entries to initiate a flow request in the time period t_o^i to t_R^{i+1} is done through the frequency matrices for intra-PAN and inter-PAN flows. For intra-PAN flows, the entries corresponding to the queued

requests are selected, those entries which correspond to the active flows are set to zero and based on the number of expected requests, the probabilities based on the frequency distribution are calculated for $f_{ij} \in F^{IPT}$. It should be noted that the probability of only those entries $f_{ij} \in F^{IPT}$ is calculated which correspond to the relevant entries $a_{ij} \in A_D^{IPT}$ leaving out the entries which indicate queued requests or active flows. After calculating the probabilities for remaining entries, the probabilities are sorted in descending order and based on the number of expected requests, r entries are then selected with the highest probabilities. We call this process the selection process. The same process is repeated for the inter-PAN flows with one exception that the entries $a_{ij} \in A_D^{IT}$ which correspond to the active flows are not set to zero. It should be noted that since the time required for each queued request is known beforehand, the time for each entry in A_D^{IPT} and A_D^{IT} corresponding to a queued request is replaced by the actual value of time required for the queued request.

When the entries from $a_{ij} \in A_D^{IPT}$ and $a_{ij} \in A_D^{IT}$ are selected and finalized, the total channel time required T_B is calculated. The total channel time T_B is calculated based on either the moving average of the channel times for selected entries or their mean. When an MPNC sends the channel time request to the ref-MPNC with channel time T_B , the ref-MPNC calculates the value of $\alpha_b(C - T_o)$ i.e. T_i . If $T_B = T_i$, then the channel time equal to T_i is allocated to the requesting MPNC. Otherwise if $T_B > T_i$, the requesting MPNC readjusts the value of T_B based on the value of T_i . The adjustment is done with the help of the RAS algorithm. The row and column constraints are defined for the RAS algorithm such that $\sum_j u_i < T_i$ and $\sum_i v_j < T_i \forall i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$. After defining the constraints, the RAS algorithm is applied to estimate new traffic matrices X_D^{IPT} and X_D^{IT} given A_D^{IPT} and A_D^{IT} . After the estimation of matrices X_D^{IPT} and X_D^{IT} , the entries $x_{ij} \in X_D^{IPT}$ and $x_{ij} \in X_D^{IT}$ are checked and adjusted. The adjustment process ensures that after balancing, the non-zero entries in the matrices $x_{ij} \in X_D^{IPT}$ and $x_{ij} \in X_D^{IT} \forall x_{ij} > 0$ indicating the expected requests, have a value which is at least greater than or equal to its corresponding entry in $a_{ij} \in A_D^{IPT}$ and $a_{ij} \in A_D^{IT} \forall a_{ij} > 0$. Therefore, after the adjustment process, each non-zero entry in the matrices X_D^{IPT} and X_D^{IT} has a value which is expected to support the QoS

requirements of the flow requested by that entry. The adjustment process is an iterative procedure and is given in Figure 3.4.

The RAS Algorithm

Step 1:

(Initialization)

$\text{Set } k = 0 \text{ and } a_{ij}^0 = a_{ij} \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$

$\text{Set } a_{ij} = 0 (\forall i = j) \text{ i.e. diagonal entries}$

Step 2:

(Row Scaling)

For $i = 1, 2, \dots, m$ define

$$\rho_i^k = \frac{u_i}{\sum_j a_{ij}^k}$$

Update A^k by ($A^k \rightarrow A_D^{IPT}$ or $A^k \rightarrow A_D^{IT}$)

$a_{ij}^k \leftarrow \rho_i^k a_{ij}^k, \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$

Step 3:

(Column Scaling)

For $j = 1, 2, \dots, n$ define

$$\sigma_j^k = \frac{v_j}{\sum_i a_{ij}^k}$$

And define A^{k+1} by

$a_{ij}^{k+1} \leftarrow a_{ij}^k \sigma_j^k, \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$

Step 4:

Replace $k \leftarrow k + 1$ and return to Step 2

Figure 3.3 The RAS Algorithm

The Selection-Adjustment (SA) Algorithm

Selection Process

(Set entries for queued requests)

Select $a_{ij} \rightarrow r_q, \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$

(For Intra-PAN Selection, entries corresponding
To active flows are set to zero)

Set $a_{ij} \rightarrow r_a = 0, \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$

(Calculate probability of expected requests)

$p(a_{ij}) = \frac{f_{ij}}{T_R}, \text{ where } T_R \text{ is in seconds and } a_{ij} > 0$

Sort probabilities in a descending order list

Select r highest entries from that list

(Calculate total time required)

$$\sum_{i,j} a_{ij} \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$$

Apply RAS Algorithm

Adjustment Process

$T_x = 0, \text{ctr} = 0$

(Adjust time for expected requests)

$\text{if } (x_{ij} \rightarrow r < a_{ij} \rightarrow r) \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$

```

{  $T_x = T_x + x_{ij}$ 
Set  $x_{ij} = 0$ 
ctr++ }
(Adjust time for queued requests)
if ( $x_{ij} \rightarrow r_q < a_{ij} \rightarrow r_q$ )
{  $x_{ij} = a_{ij}$ 
 $T_x = T_x - (a_{ij} - x_{ij})$  }
if ( $x_{ij} \rightarrow r_q > a_{ij} \rightarrow r_q$ )
{  $x_{ij} = a_{ij}$ 
 $T_x = T_x + |a_{ij} - x_{ij}|$  }
for (k = 0; k <= ctr; k++)
{ if (( $x_{ij} \rightarrow r, a_{ij} \rightarrow r$  &&  $a_{ij} < T_x$ )
{  $x_{ij} = a_{ij}$ 
 $T_x = T_x - a_{ij}$  }
}

```

Figure 3.4 The SA Algorithm based on RAS Algorithm

3.5.2 Adjusted Assignment Problem

Another approach which can be used in estimating the individual flows based on the time suggested by the ref-MPNC is the solution of a type of problem called the assignment problem. The possibilities of reaching towards a feasible solution and the techniques available to solve the assignment problems are somewhat better when compared with the MB problem. Some of the solutions for assignment problems are found in [21], [22] and [23]. The problem can be depicted as a bipartite graph $G(X \cup Y, E)$ where $x \in X, y \in Y$ and the links or arcs $x_{ij} \in E$. Each link $x_{ij} \in E$, is associated with a cost c_{ij} which usually indicates the time required for the flow in the superframe in units of microseconds apart from the selection procedure discussed further. The problem is presented as a weighted bipartite graph and can generally be stated as

$$\min(\sum_{(i,j) \in E} c_{ij}x_{ij}) \text{ or } \max(\sum_{(i,j) \in E} c_{ij}x_{ij}) \quad \forall i, j = 1, 2, 3, \dots, n \quad (3.11)$$

Subject to the constraints

$$\sum_j x_{ij} \quad \forall i, j = 1, 2, \dots, n \quad (3.11a)$$

$$\sum_i x_{ij} \quad \forall i, j = 1, 2, \dots, n \quad (3.11b)$$

$$x_{ij} \geq 0 \quad \forall i, j = 1, 2, \dots, n \text{ and } (i, j) \in E \quad (3.11c)$$

The cost $c_{ij}x_{ij}$ can either be maximized or minimized depending on the nature of the problem. The rest of the general problem structure which includes the constraints remains the same. If we assume that in each time period T_R , there is a probability that a device can

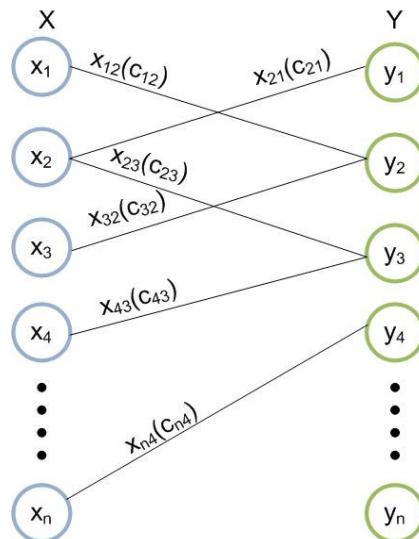
be a destination for multiple source devices i.e. more than one device tries to reserve time to send data to a common destination, and time can be reserved for only one distinct source destination pair per request, then we select the distinct source destination pair through a selection process which is different from the one described in the previous section.

3.5.2.1 Selection

Since one of the main objectives is to efficiently utilize channel time in the superframe and each MPNC is allocated a fair share of the superframe time, the parameter which is used to prefer a particular source destination pair is the transmission data rate which is used during the flow and the flow priority. Those devices which can support higher transmission data rates require relatively less channel time to send the same number of bits when compared to those devices which support lower transmission data rates. Therefore with higher data rates, lesser channel time can be utilized keeping the number of bits sent per CTA, constant. The data rates considered in our approach are those specified in [24]. An arrangement for the selection procedure is given in Figure 3.5. The cost of each link depends on the data rate and is high for lower data rates and low for higher data rates such that $c_{ij} \propto 1/\text{data rate}$. If C is the set of costs such that $c_{ij} \in C$ with c_{min} and c_{max} as the lowest and highest costs respectively, then for those $(i,j) \in E$ which represent flows with higher priority, the value of $c_{ij} = c_{min}$. For other $(i,j) \in E$ values of c_{ij} are assigned from C such that each value of c_{ij} corresponds to a particular value of data rate. Once the costs have been assigned, maximum matching of the bipartite representation of links is performed such as to minimize the overall cost i.e. $\min(\sum_{(i,j) \in E} c_{ij}x_{ij})$. If $M \subseteq E$ such that M represents the maximum matching then $\forall v \in V$ at most one link $(i,j) \in M$ is incident upon v and $V = X \cup Y$. Therefore, when we find the maximum matching subset M to minimize the cost, we also select distinct source-destination pairs with higher priority with higher data rates in a way that no destination has more than one source.

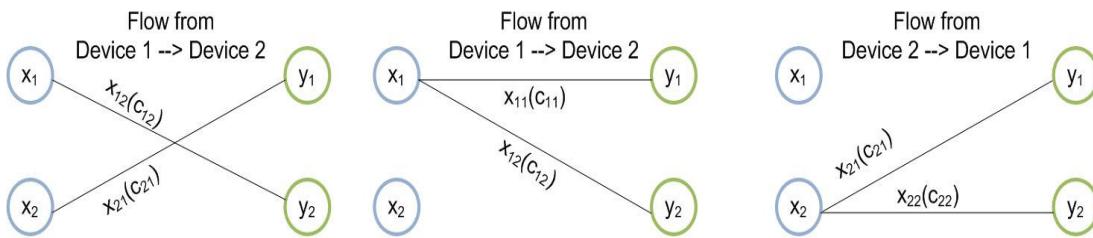
After the selection procedure, the next step is to calculate the total channel time required for the intra-PAN and inter-PAN flows. In order to determine the total channel time, both the maximum number of TUs and the minimum number of TUs are considered for the intra-PAN and inter-PAN flows. If a device is allocated minimum number of TUs for its flow, then it is enough to support its throughput requirements and it uses the minimum

superframe time possible depending on its transmission data rate. However, for some deterministic flows such as file transfers in which the total data to be transferred is known, using maximum number of TUs can reduce the total time of flow but it also requires relatively more superframe time. Furthermore, for some real time non-deterministic flows such as Constant Bit Rate (CBR) flows, using minimum number of TUs is enough to support the required throughput and using maximum number of TUs does not reduce the overall duration of flow since the flow does not depend on size of data. Therefore, the decision of using either maximum number of TUs or minimum number of TUs depends on the flow type. For some inter-PAN flows if there is a link along the source-destination multi-hop path which does not support the maximum number of TUs, then that flow is allocated the minimum number of TUs. However, if all the other links along a source-destination multi-hop path support maximum number of TUs for an inter-PAN flow, then it is assigned maximum number of TUs for its flow.



A balanced bipartite graph is used to present the selection procedure

Figure 3.5 Selection Procedure



Deterministic Intra-PAN flows or Deterministic Inter-PAN flows for which max number of TUs can be supported along the end-to-end path

Figure 3.6 (a)

Figure 3.6 (b)

Figure 3.6(c)

The problem of allocating minimum number of TUs or maximum number of TUs to a requesting device can be addressed by making different assumptions in the weighted assignment problem. In such a case, the cost c_{ij} for each link $(i, j) \in E$ is the channel time required per superframe for the flow. The value of c_{ij} can be equal to the time required for minimum TUs or maximum TUs. There are two different approaches used to allocate either minimum or maximum number of TUs to a requesting device and both are based on using bipartite graphs.

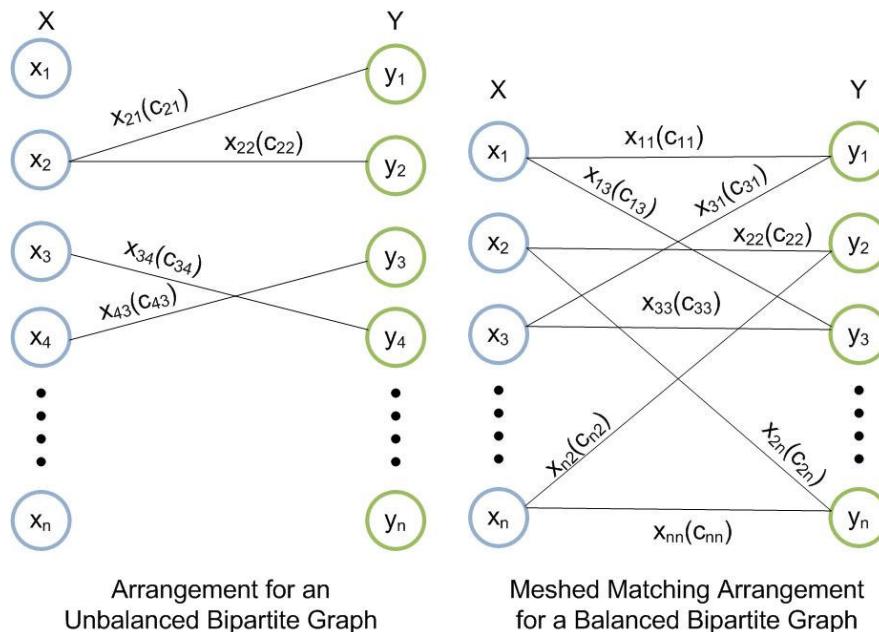


Figure 3.7 (a)

Figure 3.7 (b)

Case 1: In the first approach, the bipartite graph can either be balanced or unbalanced based on type of requests. In the case of a balanced bipartite graph $X = Y$ and $E \subseteq X \times Y$ while in an unbalanced bipartite graph $X \leq Y$ and $E < X \times Y$. Consider a case when device x_1 requests a channel time reservation for a deterministic intra-PAN flow to device y_2 and it

specifies minimum number of TUs as well as maximum number of TUs to the MPNC. In such a case since a single link in a bipartite graph cannot represent two different costs i.e. channel time for minimum number of TUs and maximum number of TUs, an arrangement like Figure3.6(a) can appeal to intuition. In Figure3.6(a), the cost c_{12} indicates the time required for maximum number of TUs while c_{21} indicates the time required for minimum number of TUs. The problem with Figure3.6(a) is that first of all it does not suggest whether the direction of flow is from device x_1 to device y_2 or vice versa and secondly in case of maximum weighted matching, both the links i.e. $(1, 2) \in E$ and $(2, 1) \in E$ might be selected. The former problem can be addressed by using the value of c_{ij} to indicate the direction of flow such that if $c_{ij} > c_{ji}$, then the direction of flow is from $(i \rightarrow j) \in E$ and if $c_{ij} < c_{ji}$, then the direction of flow is from $(j \rightarrow i) \in E$. Since we are interested to select either the maximum cost (maximum number of TUs) or a minimum cost (minimum number of TUs) for a source-destination device pair, we propose the arrangement in Figure3.6(b) and Figure3.6(c). In the proposed arrangement, if device x_i requests channel time for a deterministic intra-PAN flow to device y_j , then the link cost c_{ij} for $i = j$, represents the minimum cost while the cost c_{ij} for the link from $x_i \rightarrow y_j$ represents the maximum cost. Since in maximum matching $\forall v \in V$ at most one link $(i, j) \in M$ is incident upon v and, therefore in the proposed approach in Figure3.6(b) with x_1 having a degree 2, either the link with minimum cost or the link with maximum cost is selected. If $\exists c_{ij}$ for $i = j$, then if $\exists c_{ij} \neq 0$ but $\nexists c_{ji} \neq 0$ for $i \neq j$, it suggests a flow from $x_i \rightarrow y_j$ and if $\exists c_{ji} \neq 0$ but $\nexists c_{ij} \neq 0$ for $i \neq j$ suggests a flow from $x_j \rightarrow y_i$. However, if $\exists c_{ij} \rightarrow (i, j) \in E$ for $i \neq j$ and $\nexists c_{ij} \rightarrow (i, j) \in E$ for $i = j$, that means that the flow from $x_i \rightarrow y_j$ has either only a minimum cost or a maximum cost. This can be true especially in the case of time reservation for real time flows or some inter-PAN flows for which there is a link in the source-destination path which does not support maximum number of TUs or if all the links in the path supports maximum number of TUs. In such a case the value of the cost c_{ij} being minimum or maximum does not suggest the direction of the flow. The direction is simply given by the presence of the link between device $x_i \in X$ and device $y_j \in Y$. An example arrangement can be seen in Figure 3.7(a) in which the cost $c_{22} \neq 0$ for the link $(2, 1) \in E$ and therefore another link cost c_{21} indicates the direction of flow from device x_2 to device y_1 .

As explained before, c_{22} indicates the minimum cost and c_{21} indicates the maximum cost for the flow between x_2 and y_1 . In Figure 3.7(a), the costs c_{34} and c_{43} indicate the flows from x_3 to y_4 and from x_4 to y_3 respectively. Since for $c_{34} \neq c_{33}$ and for $c_{43} \neq c_{44}$, therefore c_{34} and c_{43} can be minimum costs or maximum costs. The time reservation requests for inter-PAN flows can also be represented in the same way. The weighted bipartite graph of Figure 3.7(a) is solved initially for the objective function $\max(\sum_{(i,j) \in E} c_{ij}x_{ij})$ for intra-PAN requests and inter-PAN requests separately and the resultant costs are added. Since after the matching procedure, the total cost represents the total required channel time required, the MPNC sends the channel request to the ref-MPNC based on the total cost calculated.

Case 2: In order to solve the problem of determining whether the maximum cost of a link or a minimum cost of a link is selected, we propose another approach. An example representation can be seen in the bipartite graph of Figure 3.7(b). A meshed arrangement of links is given between x_i and y_j . In this case, if a device x_i needs to reserve time for a flow to device y_j and both the minimum cost and maximum cost are specified for the flow request, then four links are needed to form a meshed arrangement. In Figure 3.7(b), if x_1 needs to reserve time for a flow to y_3 , links $\{(1,1), (1,3), (3,1), (3,3)\} \in E$ are established with costs c_{11}, c_{13}, c_{31} and c_{33} respectively. The value of costs $c_{ij} \forall i = j$ is the minimum number of TUs or the minimum cost while the corresponding costs c_{ij} and c_{ji} for $i \neq j$ is the maximum number of TUs or the maximum cost. Therefore from Figure 3.7(b), the costs c_{11} and c_{33} are the minimum costs while the costs c_{13} and c_{31} are the maximum costs for the flow request from x_1 to y_3 . For those time reservation requests which only requires either minimum number of TUs or maximum number of TUs, only the link $(i,j) \in E$ for $i \neq j$ between x_i and y_j is used with a cost $c_{ij}(i \neq j)$. Once the links are set up and the costs are assigned, the maximum matching procedure is used. Since in maximum matching, each vertex can at most have a degree of one, therefore, for those flows with a meshed configuration, either the link with maximum cost is selected or the link with the minimum cost. In order to elaborate further, consider Figure 3.7(b). If we consider the flow reservation request of x_1 to y_3 , then it can be noted that if the link with minimum cost c_{11} is selected, then those with maximum costs c_{13} or c_{31} cannot be selected because they are also incident on x_1 and y_1 respectively. Similarly, if the link with minimum cost c_{33} is

selected, then those with costs c_{13} and c_{31} still cannot be selected because they are also incident on y_3 and x_3 respectively. It should be however noted that if the link $(1, 1)$ with cost c_{11} is selected, then after maximum matching is performed, the link $(3, 3)$ with cost c_{33} can also be selected. This is because $(1, 1)$ and $(3, 3)$ are incident on different vertices. In the same way, if link $(1, 3)$ with cost c_{13} is selected, then there is a chance that link $(3, 1)$ with cost c_{31} is also selected after maximum matching although those links with minimum costs i.e. c_{11} and c_{33} cannot be selected. In order to resolve this redundancy, we use an adjustment procedure after the maximum matching which is explained next. In order to generalize this approach, if $\exists c_{ij} (i = j)$, then the corresponding links with costs c_{ij} and $c_{ji(i \neq j)}$ represent the maximum costs while $c_{ij(i=j)}$ represent the minimum cost for the flow request from x_i to y_j . In this approach, we assume that the direction of the flow does not matter and the time reserved indicates the total time required for the bidirectional flow between x_i and y_j while the previous proposed approach considers the flow direction as well. The same procedure is repeated to calculate the time reservation for inter-PAN flows. If the bipartite graph represents only those requests which specify both the maximum number of TUs and the minimum number of TUs, then the bipartite arrangement is fully meshed. Otherwise, if there is even one flow request which only specifies either the minimum number of TUs or maximum number of TUs, the bipartite arrangement is partially meshed. Once the weighted bipartite graph is created, maximum matching is performed initially for the objective function $\max(\sum_{(i,j) \in E} c_{ij}x_{ij})$ to calculate the feasible costs for intra-PAN flows and inter-PAN flows and the costs are added to determine the total time required which is sent to the ref-MPNC after the adjustment stage.

3.5.2.2 Adjustment

The adjustment stage is mandatory for both the cases discussed. In the first case, if there are requests which specify both the maximum number of TUs and minimum number of TUs, then for each $c_{ij} > 0 \forall i = j$, there is either a corresponding $c_{ij(i \neq j)} > 0$ or a $c_{ji(i \neq j)} > 0$. If the maximum cost is selected after maximum matching i.e. $c_{ij(i \neq j)}$ or $c_{ji(i \neq j)}$, then the relevant entry is updated in the traffic matrix i.e. A_D^{IPT} for intra-PAN traffic and A_D^{IT} for inter-PAN traffic. If the minimum cost is selected i.e. $c_{ij(i=j)}$, then the values of corresponding $c_{ij(i \neq j)}$ and $c_{ji(i \neq j)}$ are checked. If either $c_{ij} > 0$ or $c_{ji} > 0$, then its value is

replaced by link cost $c_{ij(i=j)}$ in the adjustment process. The replaced value is then used to update the matrices A_D^{IPT} or A_D^{IT} . It should be noted that, in case of the proposed approaches based on maximum matching, the expected requests are also considered and the procedure remains the same as described in the section dealing with the MB problem. The value of cost assigned to the links for expected flow requests is based on the value of either the corresponding entry of the intra-PAN traffic matrix $a_{ij} \in A_D^{IPT}$ or $a_{ij} \in A_D^{IT}$. Similar to the explanation given in the previous section on MB problem, the traffic matrices A_D^{IPT} and A_D^{IT} consists of entries, the values of which correspond to the estimates of values based on previous history. Therefore, the values of entries which correspond to the expected requests represents the mean time reserved based on a history of previous requests for that entry.

In the second proposed approach based on meshed arrangement for bidirectional flows to select maximum or minimum cost, the adjustment process is used to remove the values of redundant entries for the link costs after the maximum matching process. If after the maximum matching, a link with cost $c_{ij(i=j)} > 0$ is selected, that means that there is a probability that link with cost $c_{ji(i=j)} > 0$ is also selected. If such is the case, then the value of $c_{ji(i=j)}$ is subtracted from the total cost. Then based on the corresponding link with cost $c_{ij(i \neq j)}$, the corresponding entry $a_{ij(i \neq j)} \in A_D^{IPT}$ or $a_{ij(i \neq j)} \in A_D^{IT}$ is updated with the value of $c_{ij(i=j)}$. When the entry $a_{ij(i \neq j)}$ is updated, the entry $a_{ji(i \neq j)}$ is also set to the same value since bidirectional flows are considered. If after the maximum matching, the link with maximum cost $c_{ij(i \neq j)} > 0$ is selected, then the corresponding link with cost $c_{ij(i=j)}$ is checked. If the link exists i.e. $x_{ij}c_{ij(i=j)} > 0$, that means that the request for flow has specified both the maximum cost and minimum cost and there is a probability that the link with cost $c_{ji(i \neq j)}$ is also selected. If that is the case, the value of $c_{ji(i \neq j)}$ is subtracted from the total cost and the relevant entry $a_{ij(i \neq j)} \in A_D^{IPT}$ or $a_{ij(i \neq j)} \in A_D^{IT}$ is updated. It should be noted that all the links $(i, j) \in E$ in the bipartite graph with $i = j$ are dummy links inserted just to represent the minimum cost in both of the assignment problem approaches. If the total calculated cost is feasible, the ref-MPNC allocates it to the requesting MPNC. Otherwise, the ref-MPNC calculates a suitable value of T_i for the requesting MPNC and informs it. The requesting MPNC recalculates its required channel time based on the value

of T_i . In such a case, the objective function remains the same i.e. $\max(\sum_{(i,j) \in E} c_{ij}x_{ij})$ but a constraint is added to have an upper limit on the total cost in order to limit it to T_i . The additional constraint is given as $c_T < T_i$.

Adjustment Procedure

Case 1:

$n = \text{number of devices (inter-PAN or intra-PAN)}$

```
for (i = 0; i ≤ n; i++)
{
    for (j = 0; j ≤ n; j++)
        if (c[i][j] ≠ 0)
            // Update entry in ADIPT or ADIT
            aij = c[i][j];
    if (c[i][i] ≠ 0)
        aii = c[i][i];
    if (c[j][j] ≠ 0)
        ajj = c[j][j];
}
```

Case 2:

```
for (i = 0; i ≤ n; i++)
{
    for (j = 0; j ≤ n; j++)
        if (c[i][j] ≠ 0)
            aij = c[i][j];
    // If the reverse link is also selected, set its cost to zero
    if (c[j][i] ≠ 0)
        c[j][i] = 0;
    if (c[i][i] ≠ 0)
        aii = c[i][i];
    if (c[j][j] ≠ 0)
        c[j][j] = 0;
}
```

Figure 3.8 Adjustment for Assignment Algorithm

3.6 Calculation of d_{max}

The calculation of d_{max} is of significant importance in the proposed approach since it determines the maximum duration for which time is allocated to a requesting MPNC. Therefore, based on the value of d_{max} stored for each MPNC to which channel time T_i is allocated, the ref-MPNC can efficiently indicate suitable backoff duration to each requesting MPNC which is not allocated channel time due to unavailability of superframe time. The value of d_{max} is taken to be equal to the longest duration of either a deterministic flow or a

non-deterministic flow, whichever is greater. The duration of a flow depends on whether it is a deterministic flow or a non-deterministic flow. Usually the deterministic flows consist of data transfer, downloading and sharing files etc and are mainly the non-real time flows. Therefore the duration of deterministic flows is assumed to be less than most of the non-deterministic flows. In the proposed approach, the lognormal probability distribution for deterministic and non-deterministic flow durations is considered. The motivation for using and preferring the lognormal distribution comes from the findings in [4], [5] and [6]. In the case of lognormal distribution, a random value is used as the mean μ for the distribution and for the value of variance σ^2 . However, the range of random values used for μ and σ^2 is kept different for deterministic flows and non-deterministic flows. The duration for those flows which are expected to be initiated in the period t_0^i to t_R^{i+1} and are assumed to be deterministic can be calculated by their expected total data transfer size and their expected throughput. The throughput for deterministic flows can easily be calculated from [25].

In order to calculate d_{max} , a sample set S_{ij}^d is maintained such that the outcomes $d_{ij}^n \in S_{ij}^d$ represent the durations for up to n flows between device i and device j . Based on the values of d_{ij}^n in each set S_{ij}^d , the parameters of interest i.e. the mean μ and variance σ^2 are estimated with the help of MLE. The use of well known distributions increases the probability of getting the global maximum and avoiding the local maxima or local minima problem. If the lognormal distribution is used and $\hat{\mu}$ represents the estimated parameter for the mean of the distribution which maximizes the likelihood function, the MLE $\hat{\mu}$ for μ is

$$f(d_{ij}; \mu, \sigma) = \frac{1}{d_{ij}\sigma\sqrt{2\pi}} e^{-\frac{(\ln(d_{ij})-\mu)^2}{2\sigma^2}}, \quad d_{ij} \in \mathbb{R} \quad (3.12)$$

Where $f(d_{ij}; \mu, \sigma)$ is the function of continuous log normal pdf. Rearranging (3.12) and using the likelihood function $L(\mu, \sigma)$ based on observations $D_{ij}^1, \dots, D_{ij}^n \in S_{ij}^d$

$$f(d_{ij}; \mu, \sigma) = \frac{c}{d_{ij}\sigma} e^{-\frac{1}{2}\left(\frac{\ln(d_{ij})-\mu}{\sigma}\right)^2}, \quad C = \frac{1}{\sqrt{2\pi}}$$

Since $D_{ij}^1, \dots, D_{ij}^n$ are independent, therefore

$$f(d_{ij}^1, \dots, d_{ij}^n; \mu, \sigma) = \frac{c}{\prod_{k=1}^n d_{ij}^k \sigma} \prod_{k=1}^n e^{-\frac{1}{2}\left(\frac{\ln(d_{ij}^k)-\mu}{\sigma}\right)^2}$$

The likelihood function $L(\mu, \sigma)$ is

$$L(\mu, \sigma) = \frac{C}{\sigma^n} \frac{1}{\prod_{k=1}^n D_{ij}^k} \prod_{k=1}^n e^{-\frac{1}{2}\left(\frac{\ln(D_{ij}^k) - \mu}{\sigma}\right)^2} \quad (3.13)$$

Taking the log to get the log likelihood function $l(\hat{\mu}, \sigma)$

$$l(\hat{\mu}, \sigma) = -\sum_k \ln d_{ij}^k + \log C - n \log \sigma - \frac{\sum_{k=1}^n (\ln d_{ij}^k - \hat{\mu})^2}{2\sigma^2}$$

Taking partial derivatives w.r.t $\hat{\mu}$, and further solving, we can derive the MLE estimate for $\hat{\mu}$

$$\hat{\mu} = \frac{\sum_{k=1}^n \ln d_{ij}^k}{n} \quad (3.14)$$

Using the value of $\hat{\mu}$, we can solve and get the MLE estimate for $\hat{\sigma}^2$

$$\hat{\sigma}^2 = \frac{\sum_{k=1}^n (\ln d_{ij}^k - \hat{\mu})^2}{n} \quad (3.15)$$

Once the parameters are estimated to approximate the appropriate fit for the actual distribution, the probability of individual outcomes can be calculated. The largest mean value for the flow duration of a non-deterministic flow is compared with the largest deterministic flow duration. If μ^d is the greatest mean for the deterministic flow durations while μ^{nd} is the greatest mean for non-deterministic flow durations, and $\rho(d_{ij}^d)$ is the probability of the outcome d_{ij}^d for the deterministic flow duration while $\rho(d_{ij}^{nd})$ is the probability of the outcome d_{ij}^{nd} for the non-deterministic flow duration, then the value of d_{max} is determined by the procedure given in Figure 3.9. The notations $\min(\mu^d)$ and $\max(\mu^d)$ used in Figure 3.9 indicate the minimum average duration and the maximum average duration for the deterministic flow durations. Similar notations are used for non-deterministic flows as well.

Calculation of d_{max}

```

if( $\mu^d > \mu^{nd}$ ) // Compare Mean Allocation Size of Deterministic and Non-Deterministic Flows
{
    // Probability of Deterministic Flow Allocation size being in the range of its mean allocation size and maximum
    // of its mean allocation size is higher and similarly for the probability of Non-Deterministic Flow Allocation
    if( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) > \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 

```

```

&&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) > \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
{
if ( $\max(\mu^{nd}) > \mu^d$ )
    { compare  $\max(\mu^d)$  &&  $\max(\mu^{nd})$ 
       $d_{max}$  = Whichever is greater;
    }
else if ( $\max(\mu^{nd}) < \mu^d$ )
     $d_{max} = \mu^d;$ 
}

// Probability of Deterministic Flow Allocation size being in the range of its mean allocation size and maximum of its mean allocation size is higher but lower for Non-Deterministic Flow Allocation Size.
if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) > \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) < \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
     $d_{max} = \mu^d;$ 

// Probability of Deterministic Flow Allocation size being in the range of its mean allocation size and maximum of its mean allocation size is lower but higher for Non-Deterministic Flow Allocation Size.
if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) < \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) > \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
{
if ( $\max(\mu^{nd}) > \mu^d$ )
     $d_{max} = \mu^{nd};$ 
else
     $d_{max} = \mu^d;$ 
}

// Probability of Deterministic Flow Allocation size being in the range of its mean allocation size and maximum of its mean allocation size is lower and similarly for Non-Deterministic Flow Allocation Size.
if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) < \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) < \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
     $d_{max} = \mu^d;$ 
}

if ( $\mu^d < \mu^{nd}$ ) // Compare Mean Allocation Size of Deterministic and Non-Deterministic Flows
{
if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) > \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) > \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
{
If ( $\max(\mu^d) > \mu^{nd}$ )
    {compare  $\max(\mu^{nd})$  &&  $\mu^d$ 
       $d_{max}$  = whichever is greater;
    }
else if ( $\max(\mu^d) < \mu^{nd}$ )
     $d_{max} = \mu^{nd};$ 
}

if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) < \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) > \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
 $d_{max} = \mu^{nd};$ 

if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) > \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) < \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ 
{
if  $\max(\mu^d) > \mu^{nd}$ 

```

```

 $d_{max} = \max(\mu^d);$ 
else
 $d_{max} = \mu^{nd};$ 
}

if ( $\rho(\mu^d < D_{ij}^d < \max(\mu^d)) < \rho(\min(\mu^d) < D_{ij}^d < \mu^d)$ 
    &&  $\rho(\mu^{nd} < D_{ij}^{nd} < \max(\mu^{nd})) < \rho(\min(\mu^{nd}) < D_{ij}^{nd} < \mu^{nd})$ )
 $d_{max} = \mu^{nd};$ 
}

```

Figure 3.9 Selection of dmax

3.7 Calculation of T_i

The resource allocation policies proposed in this chapter help determine a suitable value of T_i which is the time allocated by the ref-MPNC to the i^{th} MPNC upon a bulk reservation request T_B . The value of T_i calculated for each requesting MPNC determines the relative fairness with which time is allocated by the ref-MPNC. It is therefore mandatory to calculate the value of T_i in such a way that most of the requests are allocated a fair share of superframe time. The ref-MPNC upon reception of a request for channel time T_B from a requesting MPNC checks the available superframe time before calculating and allocating time T_i to the requesting MPNC. If $T_B < (C - T_o)$, then the ref-MPNC calculates a value T_i for the requesting MPNC. The requesting MPNC recalculates its required time based on the value of T_i and notifies the ref-MPNC. The ref-MPNC then allocates the channel time to the requesting MPNC. On the other hand if $T_B > (C - T_o)$, the ref-MPNC still calculates a value T_i based on the available time. The requesting MPNC recalculates its required channel time based on the upper limit of T_i and accepts the request if its priority flows can be accommodated in T_i and rejects it otherwise and backs off. The way the ref-MPNC calculates the value of T_i is imperative in quantifying the relative fairness. Therefore, the following approaches are proposed:

- ref-MPNC calculates T_i based on utilization of superframe time
- ref-MPNC calculates T_i based on policy of accommodating all requests
- ref-MPNC calculates T_i on policy of increasing the satisfaction index of requesting MPNCs

In order to make sure that fairness is maintained in time allocation, the ref-MPNC takes into account the average time allocation of each MPNC. If R_{av} is a vector consisting of the mean time allocations for n MPNCs, then the ref-MPNC takes into account the value of $\sum_{i=1}^n r_{av}^i$, (where $r_{av}^i \in R_{av}$) in order to decide how to calculate a fair value of T_i to serve a requesting MPNC. From the requesting MPNC's perspective, the fairness estimate is called the satisfaction index s_i which is equal to

$$s_i = \frac{T_i}{T_B} \quad (3.16)$$

If there are n requests (inter-PAN and intra-PAN) which can be served in the requested time T_B and m requests which can be served in the allocated time T_i , then the value of s_i can also be written as

$$s_i = \frac{m}{n} \quad (3.17)$$

For a requesting MPNC to be 100% satisfied $T_i = T_B$ or $n = m$. The average satisfaction index for n number of MPNCs is given by

$$s_{av} = \frac{\sum_i s_i}{n} \quad (3.18)$$

The fairness of the time allocation according to [26] is

$$f(x) = \frac{(\sum_i s_i)^2}{n \sum_i s_i^2} \quad \forall i = 1, 2, \dots n \quad (3.19)$$

Where $f(x)$ is the Jains fairness index and a value of $f(x) = 1$ indicates a 100% fair system. If S is a vector consisting of the mean satisfaction index of n MPNCs as its elements, then we use \vec{S} to indicate that the elements $s_{av}^i \in S$ are arranged in descending order.

Apart from considering the mean satisfaction index for each MPNC, we use another factor in determining the fairness of time allocation which is the utilization factor γ_i for the i^{th} MPNC. The utilization factor for the i^{th} MPNC is given by $\gamma_i = T_i \times d_{max}$ where d_{max} is the total time duration for which the allocation is requested for. Therefore, the value of γ_i takes into account, the amount and the duration of the resource occupied by the i^{th} MPNC. We use the vector $\vec{\gamma}$ to indicate that it contains the mean values of $\gamma_i \forall i = 1, 2, \dots n$ in descending order. In order to derive a priority list to select MPNCs and improve their satisfaction index, increase the overall fairness and the mean satisfaction index, we use

$$\vec{P} = \vec{S} + \vec{\gamma} \quad (3.20)$$

Where \vec{P} is a vector whose elements $p_i \in \vec{P}$ are the sum of corresponding elements s_i and γ_i . If we use \overleftarrow{P} to indicate that the elements are sorted in the ascending order, then the ref-MPNC can use \overleftarrow{P} as a priority list of MPNCs whose satisfaction index and utilization factor needs to be increased in order to increase the fairness of the allocation. In order to calculate the value of T_i , the following three cases are considered.

Case 1: Calculation of T_i based on the utilization of the superframe time

In this case, when the ref-MPNC gets a request for allocation of channel time T_B from a requesting MPNC, it checks the percentage utilization of the superframe time. If the utilization is above or below a certain threshold U_T , the ref-MPNC minimizes or maximizes a cost function respectively in order to calculate the value of T_i . If x_1 and x_3 are the number of time slots used to support the intra-PAN and inter-PAN flows respectively with minimum number of TUs, x_2 and x_4 are the number of time slots used to support the intra-PAN and inter-PAN flows respectively with maximum number of TUs, the cost function and the constraints are given as

$$\min \text{ or } \max f(x) = (x_1 + x_2)N_S + (x_3 + x_4)N_S \quad (3.21)$$

$$(x_1 + x_2)N_S \leq (1 - \psi_d)T_B \quad (3.21a)$$

$$(x_3 + x_4)N_S \leq \psi_d T_B \quad (3.21b)$$

$$(x_1 + x_2)N_S + (x_3 + x_4)N_S \leq (C - T_o) \quad (3.21c)$$

$$x_1, x_2, x_3, x_4 > 0 \quad (3.21d)$$

In (3.21), the term N_S refers to the unit of time or the time resolution. In our case, the value of N_S is set to $1\mu\text{s}$ and the values considered for the superframe utilization threshold U_T are 0.5, 0.6 and 0.7. For a given value of U_T , the current utilization U_c can be simply calculated by taking the ratio of allocated superframe time to the total superframe duration. The value of U_T and U_c can then be compared to check if the threshold has been exceeded or not. If $U_T < U_c$, then the ref-MPNC maximizes the cost function i.e. $\max f(x)$, otherwise if $U_T \geq U_c$, then the ref-MPNC minimizes the cost function i.e. $\min f(x)$ to determine the value of T_i . However, in this case, fairness of time allocation to each request is not taken into account and therefore, this case is used as a baseline to compare with the

rest of the cases which take fairness into account. Once the value of T_i is calculated and it is allocated to the requesting MPNC after successful negotiation, the value of α can be calculated by

$$\alpha = \frac{T_i}{(C - T_o)} \quad (3.22)$$

Where α is the fraction of time allocated to the requesting MPNC. It should be noted that the value of α from (3.22) gives the perception of satisfaction just from the requesting MPNCs reference since it is calculated from the remaining superframe time i.e. $(C - T_o)$. In order to derive a perception of satisfaction relative to other MPNCs, a parameter α_R is used which is equal to

$$\alpha_R = \frac{T_i}{\eta_s} \quad (3.23)$$

Therefore the parameter α_R is useful in terms of indicating the relative satisfaction index and the overall fairness since it is the ratio of the allocated time T_i to the total superframe duration η_s and not the remaining superframe duration $(C - T_o)$. Another important use of the values of α and α_R can be in a situation if the ref-MPNC becomes unavailable. In such a situation the mean values of α and α_R can be used by the MPNC to request its share in a distributed environment. However, the procedure for such a situation is outside the scope of this chapter.

Case 2: Calculation of T_i based on the policy of Accommodate All (AA Policy)

In this case, we propose an approach which takes into account the average allocation size of each MPNC i.e. $r_{av}^i \in R_{av}$. The aim of this approach is to determine if all of the MPNCs can be accommodated in the superframe time or not. In order to determine the a value of T_i for each request and allocate the time fairly, the sum of average allocation of all the MPNCs is compared with the total superframe duration η_s . The feasibility of using the policy of accommodate all is determined by two possibilities which can arise after the comparison are considered which are

$$\sum_{i=1}^n r_{av}^i > \eta_s \quad (3.24a)$$

$$\sum_{i=1}^n r_{av}^i \leq \eta_s \quad (3.24b)$$

If (3.24a) is true and the sum of mean allocations of MPNCs exceeds the superframe duration by an amount η_e , then

$$\sum_{i=1}^n r_{av}^i = \eta_s + \eta_e \quad (3.25)$$

If all of the MPNCs are to be accommodated in the superframe, the amount η_e should be divided and divisions should be subtracted from the mean flows r_{av}^i such that the condition (3.24b) is satisfied. The problem is to determine suitable size of divisions to be subtracted from each value of r_{av}^i . If $r_{av}^i = r_{av}^{i+1} \forall i = 1, 2, \dots, n$ then for equal sized mean allocations, the division of η_e is straight forward and therefore (3.26) satisfies (3.24b).

$$\sum_{i=1}^n (r_{av}^i - \frac{\eta_e}{n}) = \eta_s \quad \forall i = 1, 2, \dots, n \quad (3.26)$$

However in case if $r_{av}^i \geq r_{av}^{i+1} \forall i = 1, 2, \dots, n$, then the task to determine a suitable size of division becomes non-trivial. If such a case arises, then using (3.26) to satisfy (3.24b) is not fair since the average allocation size is different for each MPNC and the percentage utilization of superframe time by each average allocation varies i.e.

$$\frac{r_{av}^i}{\eta_s} \times 100 \geq \frac{r_{av}^{i+1}}{\eta_s} \times 100 \quad \forall i = 1, 2, \dots, n \quad (3.27)$$

In order to allocation time fairly, appropriate size of division of η_e should be determined for each r_{av}^i based on the percentage superframe utilization by r_{av}^i . Before proceeding to formulate an approach for fair allocation, the following observations are mandatory to take into account.

Lemma 1: The policy of accommodate all is not feasible for some cases of $\eta_s = \eta_e$

Proof: Suppose we have the case in which $r_{av}^i = r_{av}^{i+1} \forall i = 1, 2, \dots, n$ i.e. all the mean allocation sizes are equal. Then if $\eta_s = \eta_e$, that means that $r_{av}^i = 2 \frac{\eta_e}{n}$ and therefore $(r_{av}^i - \frac{\eta_e}{n}) = \frac{\eta_e}{n}$ and $\sum_{i=1}^n (r_{av}^i - \frac{\eta_e}{n}) = \eta_e$ or η_s since $\eta_s = \eta_e$. Hence the allocated time $T_i = \frac{r_{av}^i}{2}$ and $s_i = 0.5$

Therefore, if a minimum satisfaction index of $s_i = 0.5$ is acceptable then the case of $\eta_s = \eta_e$ is feasible otherwise its not considered a viable solution to achieve fairness. If $r_{av}^i \leq r_{av}^{i+1} \forall i = 1, 2, \dots, n$ and $\eta_s = \eta_e$, then the probability of accomodating all MPNCs with a certain minimum acceptable value of s_i decreases even further. If we bound the value of T_i such that $0 \leq T_i \leq r_{av}^i$, then there are following possibilities i.e.

$$r_{av}^i \leq a_f \eta_e \quad (0 \leq a_f \leq 1) \text{ and } \sum_i a_f = 1 \quad \forall i = 1, 2, \dots, n$$

If for k requests, $r_{av}^i > a_f \eta_e$, then for the remaining $n-k$ requests $r_{av}^i \leq a_f \eta_e$ and therefore (3.24b) is not satisfied. Another point to consider is that $n-k$ devices are discriminated.

Lemma 2: When $\eta_s < \eta_e$, then there is no possibility to accomodate all the requests with $s_i \geq 0.5$

Proof: In order to prove that the case $\eta_s < \eta_e$ is not feasible for the policy of accommodate all, we first consider the case when $r_{av}^i = r_{av}^{i+1} \forall i = 1, 2, \dots, n$. The value of each mean allocation can be given as

$$r_{av}^i = \frac{\eta_s + \eta_e}{n}$$

Since $T_i = r_{av}^i - \frac{\eta_e}{n} \forall i = 1, 2, \dots, n$

$$T_i = \frac{\eta_s + \eta_e}{n} - \frac{\eta_e}{n}$$

$$T_i = \frac{\eta_s}{n}$$

Since $r_{av}^i = \frac{\eta_s}{n} + \frac{\eta_e}{n}$ and $\frac{\eta_e}{n} > \frac{\eta_s}{n}$

$$\therefore T_i < \frac{r_{av}^i}{2} \text{ and } s_i < 0.5 \forall i = 1, 2, \dots, n$$

When $r_{av}^i \leq r_{av}^{i+1} \forall i = 1, 2, \dots, n$ and $\eta_s < \eta_e$, then the possibility of accomodating all MPNCs is not viable. If we bound the value of T_i such that $0 \leq T_i \leq r_{av}^i$, then there are following possibilities i.e.

$$r_{av}^i \leq a_f \eta_e \quad (0 \leq a_i \leq 1) \text{ and } \sum_i a_f = 1 \forall i = 1, 2, \dots, n$$

If for k requests, $r_{av}^i > a_f \eta_e$, then for the remaining $n-k$ requests $r_{av}^i \leq a_f \eta_e$ and therefore (3.24b) is not satisfied. Another point to consider is that $n-k$ devices are discriminated. Keeping in view the observations about $\eta_s = \eta_e$ and $\eta_s < \eta_e$, we find it reasonable to formulate a feasible solution for the problem when $\eta_s > \eta_e$ and $r_{av}^i \leq r_{av}^{i+1} \forall i = 1, 2, \dots, n$. For a feasible solution, appropriate fractions of η_e need to be determined for each value of r_{av}^i based on its utilization of superframe i.e. $\frac{r_{av}^i}{\eta_s}$. Therefore the value of $T_i = (r_{av}^i - a_f \eta_e)$ and $(0 < a_f < 1)$. An appropriate value of a_f for each r_{av}^i , is the solution of the following optimization problem

$$\max(T_i = \sum_{i=1}^n (r_{av}^i - a_f \eta_e)) \quad (3.28)$$

$$a_f < \frac{r_{av}^i}{\eta_s} \quad (3.28a)$$

$$(r_{av}^i - a_f \eta_e) > 0 \quad (3.28b)$$

$$\sum_{i=1}^n (r_{av}^i - a_f \eta_e) \leq \eta_s \quad (3.28c)$$

$$\sum_{i=1}^n r_{av}^i = \eta_s + \eta_e \quad (3.28d)$$

Case 3: Calculation of T_i based on the policy of maximum satisfaction (SA Policy) of requesting MPNCs

An important parameter which helps to determine which policy is the best to use in a particular scenario is the threshold of satisfaction index s_i that has to be maintained for each requesting MPNC. Usually in cases where the following three conditions are true, then the policy of maximum satisfaction or requesting MPNCs is followed

$$\sum_{i=1}^n r_{av}^i = \eta_s + \eta_e$$

$$\eta_s < \eta_e \text{ or } \eta_s \ll \eta_e$$

$$s_i > s_{min} \text{ (for } \eta_s > \eta_e, \eta_s = \eta_e, \eta_s < \eta_e \text{ or } \eta_e < \eta_s \text{)}$$

In this approach, the main objectives are: (1) to allocate channel time in such a way so that for each request the satisfaction index exceeds a certain minimum threshold s_i^{min} and (2) based on the priority list vector \overleftarrow{P} , increase the overall fairness of time allocation. In order to achieve these objectives, we use the objective function (3.21) used in case 1 with additional constraints. Based on the utilization threshold of superframe time U_T , the objective function is either minimized or maximized. However, we introduce the additional constraints which are

$$\frac{(x_1+x_2)N_S + (x_3+x_4)N_S}{T_B} > s_{min} \quad (3.29)$$

The addition of constraint (3.29) ensures that the allocation to the requesting MPNC is done in such a way that its satisfaction index is above a minimum threshold s_{min} . In this approach the priority list \overleftarrow{P} is utilized from (3.20) by the ref-MPNC in order to improve the satisfaction index of those MPNCs, the mean satisfaction index of which is lower. Since the vector \overleftarrow{P} is in ascending order, the first element of \overleftarrow{P} is the one with the lowest mean satisfaction index. When the ref-MPNC calculates a value of T_i for the requesting MPNC, it

checks its mean satisfaction index. If the requesting MPNC has a mean satisfaction index which has the highest priority i.e. it is the first element of vector \overleftarrow{P} , then the ref-MPNC attempts to increase its mean satisfaction index. It does so by checking if $T_B - T_i < (C - T_o)$ for $T_B > T_i$ and if true, allocates an additional amount of time i.e. $T_B - T_i$ to the requesting MPNC to increase its satisfaction index and updates the priority vector \overleftarrow{P} accordingly. If $T_B - T_i > (C - T_o)$, then the ref-MPNC checks if $s_i > s_{av}$, where s_{av} is the last element of the vector \overleftarrow{P} and thus denotes the highest mean satisfaction index and s_i is the current satisfaction index of the requesting MPNC. If $s_i > s_{av}$, then the ref-MPNC updates the priority vector if the value of T_i allocated by the ref-MPNC is accepted by the requesting MPNC.

Therefore, in summary, the MPNC which has the highest priority gets the incentive of being allocated higher channel time and thus higher mean satisfaction index. Apart from offering this incentive, another important advantage is that by using this approach, the overall fairness of time allocation based on the mean satisfaction index of each MPNC is improved. From the requesting MPNCs perspective, it can attempt to increase its priority by the following methods when appropriate

- If the requesting MPNC is not offered its actual requested time T_B , then it can give priority to those flows which have a lower duration and thus it can update its value of d_{max} . With a lower value of d_{max} it can reduce its utilization factor and thus improve its priority.

- The MPNC which has been allocation time T_i can release some channel time if some of the flows end earlier than expected. This also brings its utilization factor down and thus increases its priority.

In the second point, if an MPNC releases some channel time, that time can be used by the ref-MPNC to allocate to other requesting MPNCs and thus increase their satisfaction index. The algorithm which is used to maintain fairness using the cumulative satisfaction index is given below.

Algorithm based on Satisfaction Index

```

if ( $s_i < s_{min}$ )
    Reject the Request (The MPNC backs off)
if (( $s_i > s_{av}$ ) || ( $s_i > s_{min}$ )){
    if ( $s_{av} > s_c^i$  ) || ( $s_{av} = s_c^i$  ){
        Allocate ( $C - T_o$ ) if  $T_B > (C - T_o)$ 
    }
}

```

```

else
    Allocate  $T_B$ 
}
if ( $s_{av} < s_C^i$ ) {
    if ( $s_{av} > 50\%$ )
        Allocate  $T_i = S_{av} * T_b$ 
    if ( $s_{av} < 50\%$ )
        Allocate  $(C - T_o)$  if  $T_B > (C - T_o)$ 
    else
        Allocate  $T_B$ 
}
}

```

Figure 3.10 Algorithm Based on Satisfaction Index (SI) Value

3.8 Fairness Metrics

In order to check whether the time allocation algorithm in use is effective in converging towards a fair system, two fairness metrics are used. The first one is a widely used metric called Jain's fairness index [26] from (3.19) which gives the overall fairness of the system based on the satisfaction index of requesting MPNCs. The second fairness metric is an economic inequality metric and is called the Theil's Index [27]. Although, the Theil's index is an economic inequality metric but since its decomposable, it is feasible to be used in our case. The asymmetric versions of Theil's Index i.e. T_T and T_L are given by

$$T_T = \ln \frac{A_{total}}{B_{total}} - \frac{\sum_{i=1}^N B_i \ln \frac{A_i}{B_i}}{B_{total}} \quad (3.30)$$

$$T_L = \ln \frac{B_{total}}{A_{total}} - \frac{\sum_{i=1}^N A_i \ln \frac{B_i}{A_i}}{A_{total}} \quad (3.31)$$

In order to use the Theil index in the perspective of proposed approach, we use A_i to be the number of devices (Intra PAN device's and expected inter-pan requests) in a cluster operated by the i^{th} MPNC while B_i is used to indicate the number of time slots allocated to the i^{th} MPNC. The notation A_{total} refers to the total number of devices in all MPNCs and B_{total} refers to the total number of time slots in a superframe. In (3.30), the equation takes the value of B as the base while in (3.31) takes the value of A as the base. A symmetric Theil's index T_s is the average of the two and is given as

$$T_s = \frac{1}{2} \sum_{i=1}^N \ln \frac{B_i}{A_i} \left(\frac{B_i}{B_{total}} - \frac{A_i}{A_{total}} \right) \quad (3.32)$$

Jain's fairness index $f(x)$ ranges between 0 and 1 with $f(x) \sim 1$ meaning a fair system while $f(x) \sim 0$ meaning a completely unfair system. With Theil index, the interpretation of its value is different. A value of 0 means perfect equality while those values which diverge from 0 indicate inequality. The Theil index gives the difference between the maximum entropy and the actual entropy of the system and a value of 0 therefore indicates maximum entropy. Higher values of Theil Index indicate lower values of entropy and thus higher inequality. A normalized version of Theil Index is given by the Atkinson's Index A [28]

$$A = 1 - e^{-T} \quad (3.33)$$

Where T is the Theil's Index and $0 < A < 1$. The Jain's fairness index indicates the overall fairness in terms of the satisfaction index of each MPNC. It doesn't take into account, the fairness of allocation based on the number of devices in each cluster. The Theil index can be used to calculate the inequality of allocations based on the number of devices in each cluster.

3.9 Simulation Results

General Observations and Selection of Basic Parameters

In order to develop simulation scenarios with the desired attributes to evaluate the traffic estimation and resource allocation strategies, the first step is to determine a practical number of intra-PAN and inter-PAN flows per cluster or piconet administered by an MPNC. In the previous chapter, the superframe capacity was analyzed keeping in view single hop intra-PAN and inter-PAN flows. However, in case of meshed WPANs, multi-hop flows are considered for inter-PAN communication. Taking this additional factor in account raises the issue of re-evaluating the superframe capacity in order to determine a suitable range for number of devices involved in intra-PAN flows as well as the mean arrival rate of requests for multi-hop inter-PAN flows. After undergoing various sets of simulation runs with different values for number of intra-PAN devices per cluster, different number of MPNCs, and mean arrival rates based on Poisson distribution for intra-PAN and inter-PAN flow requests, a summary of superframe utilization is given in Table 3.1. Table 3.1 consists of a feasible set of values for the number of intra-PAN devices per cluster which is 10, 15 and 20, number of MPNCs per two hop neighbourhood which is 10 and 15, while the poisson arrival

rate is varied from 5 mins to 15 mins for intra-PAN flows and taken to be 5, 10 and 15 mins for inter-PAN flows. The superframe utilization values given in the table represent the mean utilization observed in a 24 Hour period.

A somewhat better intuition can be sought by presenting a subset of values and the corresponding superframe utilization in Figure 3.11, Figure 3.12 and Figure 3.13. It can be clearly noticed from the 3 Figures that the mean superframe utilization increases with the increase in number of devices for the same values of Mean Poisson arrival rates. However, in some cases, a decreasing trend can be noticed as well. The reason is the highly dynamic nature of traffic requirements for a requesting device which influences the trends in superframe utilization. Although, in Figure 3.11, 3.12 and 3.13, the mean values of superframe utilization are shown in order to give a concise overall picture, the continuous trend of superframe utilization is given in Figure 3.14. The value of arrival rate is kept the same for inter-PAN and intra-PAN flows in Figure 3.14 and the continuous superframe utilization trend is observed for a total of 8 hours with the arrival rate changing after each 2 hour time period.

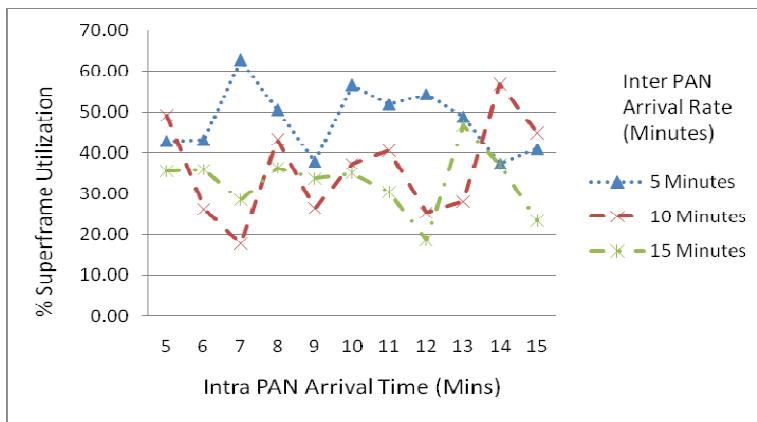


Figure 3.11 Superframe Utilization with Number of Intra-PAN Devices =10, Number of Neighbour MPNCs = 10

Since the HDR WPANs are meant for multi-media traffic, most of the requests are for non-deterministic or real time traffic type. A Pareto variable is used to serve the purpose and hence the ratio between number of non-deterministic to the deterministic flows is 80:20. In Figure 3.15, the difference between the number of non-deterministic flows as well as the difference in size of CTA's allocated to both types of traffic is clearly evident and as expected.

Table 3.1: Superframe Utilization based on a range of Mean arrival Rates and Number of Intra-PAN and Inter-PAN devices

		Average Superframe Utilization						
		Number of Intra-PAN and Inter PAN Devices (Number of Intra-PAN Devices, Number of Inter-PAN Devices)						
		(10, 10)	(15, 10)	(20, 10)	(10, 15)	(15, 15)	(20, 15)	
Poisson Arrival Rate	(Intra-PAN mean arrival Rate, Inter-PAN mean arrival Rate)	(5,5)	0.028096	0.041606	0.041439	0.018408	0.03493	0.043635
		(5, 10)	0.032221	0.030483	0.038575	0.02368	0.025475	0.034
		(5, 15)	0.023283	0.033112	0.040517	0.030457	0.029572	0.03986
		(6, 5)	0.028255	0.038094	0.044676	0.028669	0.032345	0.037242
		(6, 10)	0.017226	0.02976	0.036834	0.026124	0.028308	0.033622
		(6, 15)	0.023519	0.024785	0.039699	0.026839	0.019816	0.036535
		(7, 5)	0.041111	0.043539	0.046167	0.03742	0.031749	0.038193
		(7, 10)	0.011635	0.021615	0.040899	0.018146	0.03108	0.040765
		(7, 15)	0.018631	0.04015	0.040086	0.013707	0.030478	0.036733
		(8, 5)	0.033133	0.03292	0.03919	0.040569	0.034961	0.04612
		(8, 10)	0.028355	0.02758	0.040313	0.010829	0.029403	0.037479
		(8,15)	0.023809	0.030242	0.035487	0.026226	0.023096	0.030262
		(9, 5)	0.024809	0.043143	0.045477	0.020445	0.042575	0.046345
		(9, 10)	0.017324	0.033425	0.041497	0.017287	0.025381	0.028463
		(9, 15)	0.022086	0.028792	0.032526	0.025437	0.031421	0.041988
		(10, 5)	0.037123	0.034822	0.044403	0.0422	0.041137	0.035277
		(10,10)	0.02444	0.028858	0.038383	0.024549	0.032404	0.032706
		(10,15)	0.022929	0.017594	0.035802	0.034886	0.03121	0.037231
		(11, 5)	0.034015	0.04088	0.042367	0.020626	0.04005	0.037243
		(11,10)	0.02661	0.034935	0.040959	0.00805	0.04079	0.036217
		(11,15)	0.019907	0.031194	0.029174	0.013378	0.026455	0.035794
		(12, 5)	0.035621	0.03786	0.04913	0.036663	0.028449	0.033343
		(12,10)	0.016639	0.04017	0.040612	0.009519	0.030071	0.026446
		(12,15)	0.012196	0.029112	0.035798	0.008605	0.02982	0.039682
		(13, 5)	0.03197	0.040254	0.04167	0.02984	0.02805	0.041608
		(13,10)	0.01838	0.028021	0.032159	0.022595	0.016109	0.038419
		(13,15)	0.030586	0.029924	0.04375	0.018642	0.032703	0.034206
		(14, 5)	0.024459	0.030585	0.040541	0.042498	0.037582	0.03969
		(14,10)	0.037241	0.03803	0.033263	0.011728	0.027726	0.039774
		(14,15)	0.024342	0.027472	0.037952	0.020669	0.026386	0.025364
		(15, 5)	0.026792	0.028705	0.042841	0.030075	0.029967	0.018139
		(15,10)	0.029382	0.016674	0.034653	0.007643	0.02113	0.024688
		(15, 15)	0.015 359	0.021 467	0.031 321	0.027 816	0.032 902	0.035 526

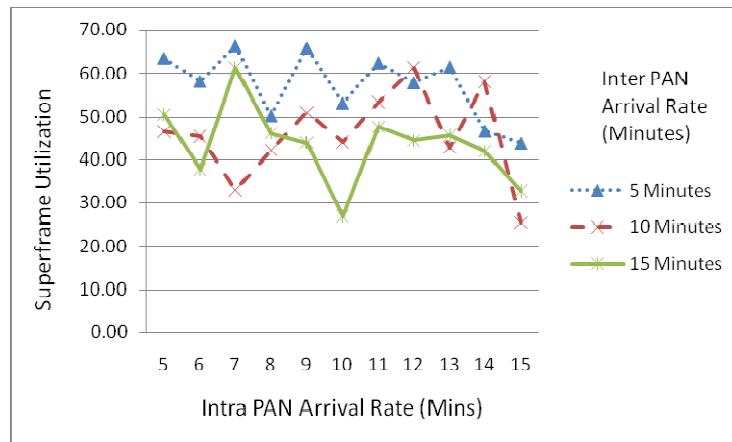


Figure 3.12 Superframe Utilization with Number of Intra-PAN Devices =15, Number of Neighbour MPNCs = 10

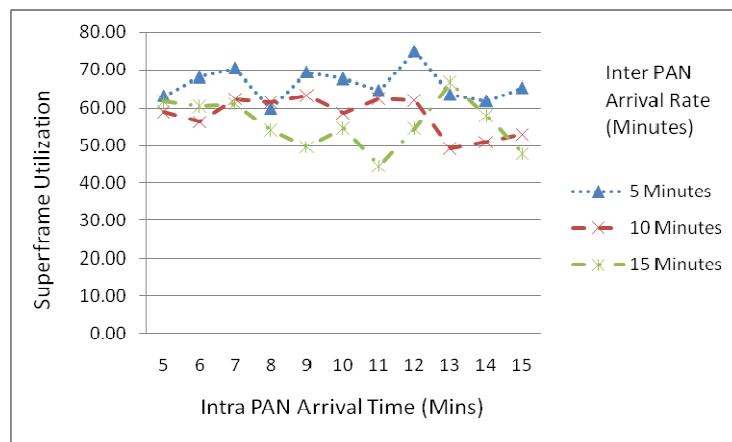


Figure 3.13 Percentage Superframe Utilization with Number of Intra-PAN Devices =20, Number of Neighbour MPNCs = 10

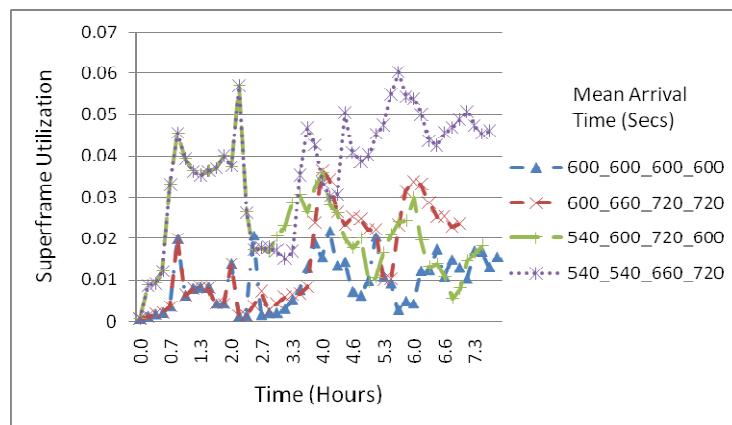


Figure 3.14 A Continuous trend of Superframe Utilization based on a variety of mean arrival rates

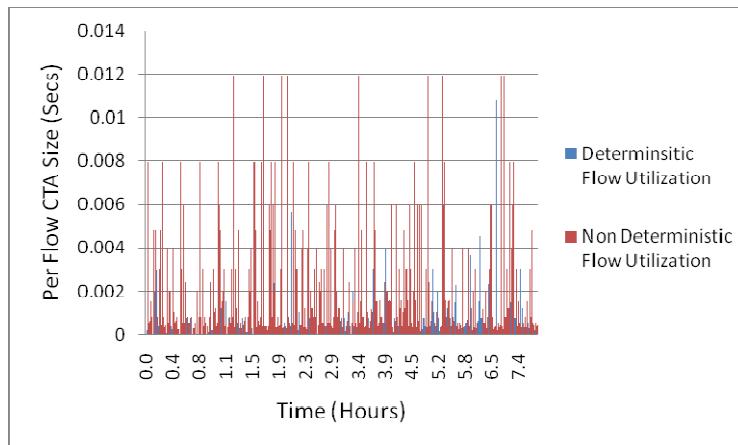


Figure 3.15 Comparison of per flow CTA Size between Deterministic and Non Deterministic Flows
Traffic Estimation

In order to determine a suitable size for a bulk reservation i.e. T_B value, each MPNC estimates its required traffic based on the information available in the demand traffic matrices \mathbf{D}_{intra} and \mathbf{D}_{inter} . The traffic estimation is done via one of the three proposed techniques in this chapter. The first technique is based on the mean allocation size or CTA size requested by a device which is expected to send a channel time request command. The second technique is based on the RAS algorithm augmented with the adjustment extension. The third technique is based on the adjusted assignment algorithm proposed in this chapter. The relevant performance evaluation of the three proposed approaches via simulation results is as follows:

Estimation based on mean allocation size:

As discussed before, once the two demand traffic matrices \mathbf{D}_{intra} and \mathbf{D}_{inter} are obtained, then based on the mean arrival rate for intra-PAN and inter-PAN flows respectively those devices which are expected to request for channel time are selected from the frequency distribution matrices \mathbf{F}_{intra} and \mathbf{F}_{inter} . The bulk reservation T_B is then simply taken as the sum of mean allocation size of selected intra-PAN and inter-PAN flows. In Figure 3.16, the estimated values are compared with the actual (simulated) values of mean superframe utilization for an MPNC. The number of intra-PAN devices is taken to be 10 and 20 in the first case while the number of inter-PAN devices or MPNCs considered is 10 and 15. It should be noted that in order to give a comprehensive view and save space, the mean superframe utilization is

shown in the figure for each value of mean arrival rate instead of the continuous superframe utilization with each mean arrival rate. The time period T_R considered for all the cases is 1 hour. From Figure 3.16, it can be seen in the first case when the number of devices is 10 for intra-PAN and inter-PAN flows, for the arrival rates of 5, 7, 13 and 15 mins, the predicted values are an over-estimation. In other words, the superframe is underutilized if the prediction is an over-estimation. In such a case, the MPNC can request for more than its requirements. However, for the arrival rate of 9 mins, the prediction is almost accurate.

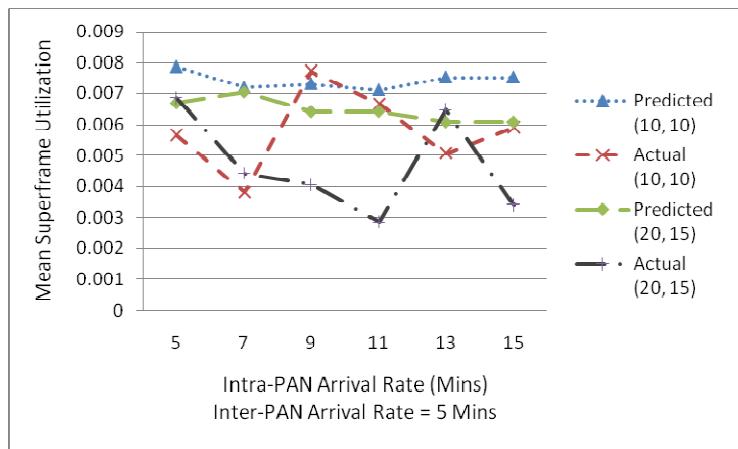


Figure 3.16 Traffic Estimation based on Mean Allocation Size (Scenario 1)

A different case of traffic estimation is also assumed based on mean allocation size. In the previous case, the allocation sizes of each intra-PAN and inter-PAN device were assumed to be random. In a practical scenario, there might be some devices in a WPAN cluster which can only be involved in a single type of flow i.e. voice flows or video flows. Based on this assumption, another scenario is evaluated in which 60% of the devices are included in the category of predicted flows which consists of voice and video flows. The result is shown in Figure 3.17 for such a scenario. When compared with Figure 3.16, it shows a relatively more consistent and smoother relationship between the predicted and simulated curves. Apart from using the predicted traffic category for 60% of the devices, the rest of the parameters like the inter-PAN and intra-PAN arrival rate and the number of devices are kept the same as in the case of Figure 3.16.

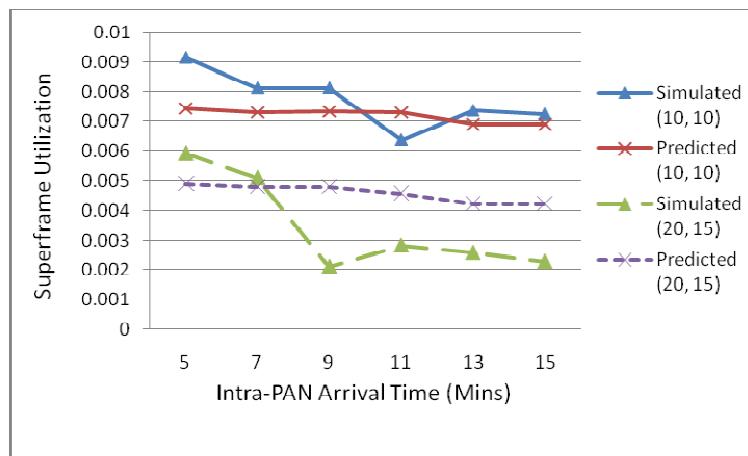


Figure 3.17 Traffic Estimation based on Mean Allocation Size (Scenario 2)

Estimation Based on the RAS Algorithm

The parameters chosen for the traffic estimation using the RAS algorithm are the same as the previous case. An important addition is the maximum number of iterations k for the algorithm. After using various values of k , two values were selected i.e. $k = 5$ and $k = 10$ which demonstrate relatively better performance in terms of reduced prediction error when compared with other values. In Figure 3.18, the number of intra-PAN and inter-PAN devices is taken to be 20 and 15 respectively, the intra-PAN arrival times are taken to be 5, 7, 9, 11, 13 and 15mins as usual while the inter-PAN arrival rates are taken to be 5mins and 10mins. The value of $k = 10$ for the case of Figure 3.18. It is clearly evident from the figure that both the cases of under-estimation and over-estimation occur in the predicted results. In Figure 3.19, the number of devices for intra-PAN and inter-PAN flows is taken to be 15 and 10 while the rest of the parameters are the same. In the second case of Figure 3.19 when the inter-PAN arrival time is 10 mins, the prediction is a substantial and consistent under-estimation. When the number of iterations $k = 5$, relatively accurate predictions can be seen in Figure 3.20.

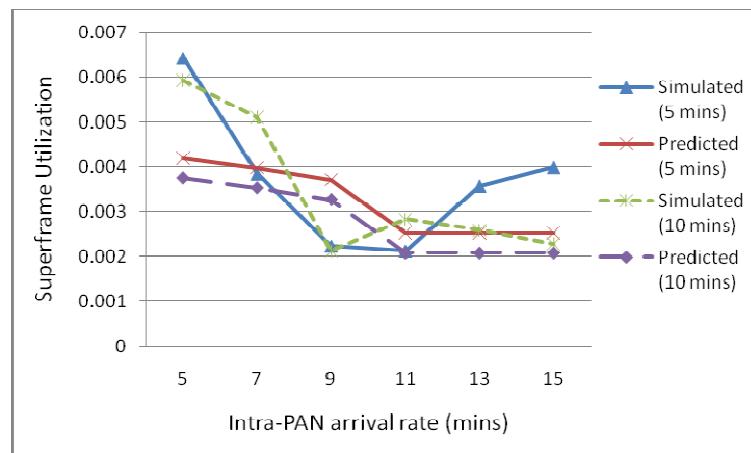


Figure 3.18 Traffic Estimation via RAS Algorithm (10 Iterations) (20, 15)

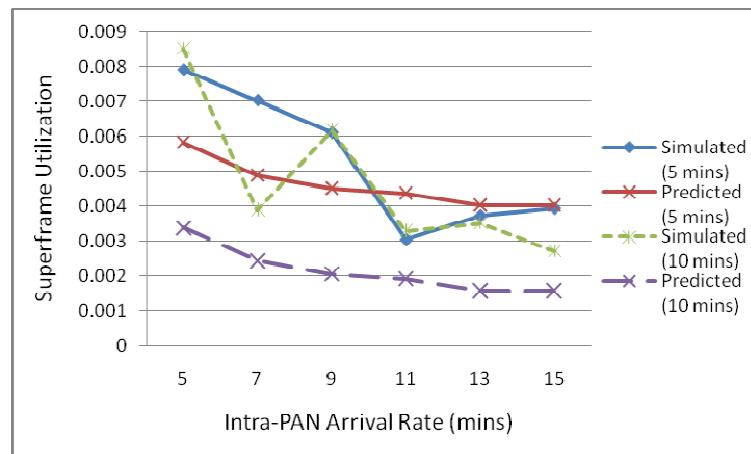


Figure 3.19 Traffic Estimation via RAS Algorithm (10 Iterations) (15, 10)

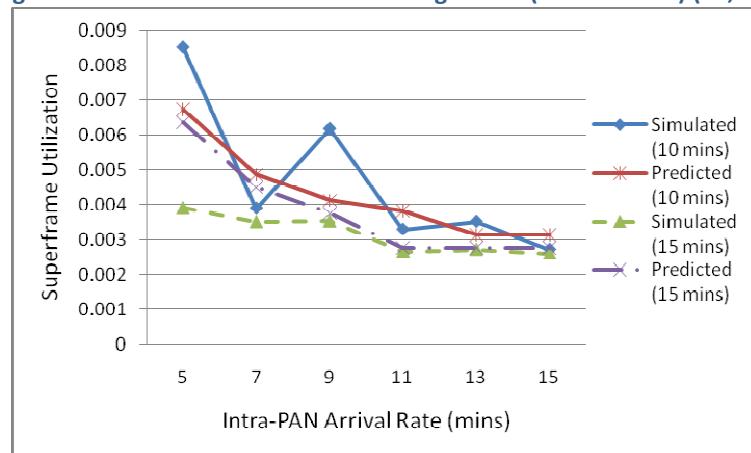


Figure 3.20 Traffic Estimation via RAS Algorithm (5 Iterations) (15, 10)

Estimation based on the Adjusted Assignment Algorithm

The adjusted assignment algorithm shows the most consistent results when compared with the previous approaches. In order to demonstrate the effectiveness of the proposed adjusted assignment problem, two cases are considered. In the first case, the expected number of inter-PAN and intra-PAN devices is selected like before based on the frequency distribution matrices. However, instead of an entry in a demand matrix representing the mean of allocated CTA size for each device, it represents the CTA size which is most frequently requested by the corresponding device to each entry. Those devices which have the highest occurrences are selected based on the mean arrival rate. The second case is the same as before and each entry in the demand matrix represents the mean allocation size of corresponding device to that entry. For both the cases, the number of intra-PAN and inter-PAN devices is taken to be 15 and 10 respectively while the inter-PAN arrival rate is taken to be 5, 10 and 15 mins. The arrival rates for intra-PAN flows are the same as before. The results for the first case can be seen in Figure 3.21 while for the second case, in Figure 3.22. It can be deduced from both figures, that the results in the first case show more consistency and accuracy when compared with the second case especially when inter-PAN arrival rate is 5 mins in the second case, there is significant divergence between the simulated and predicted values.

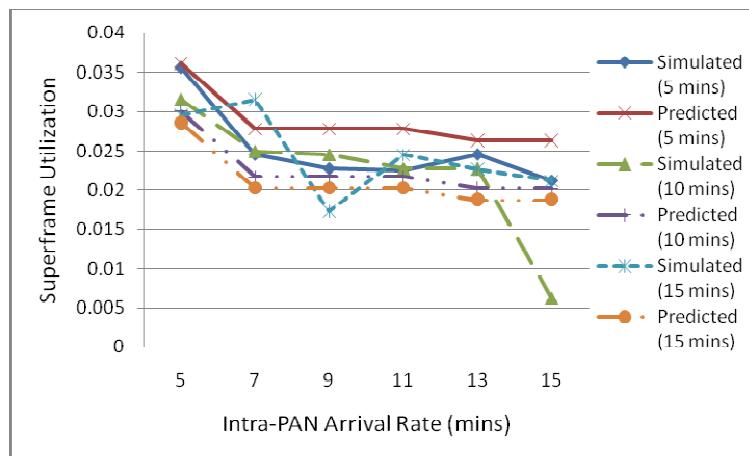


Figure 3.21 Traffic Estimation via Adjusted Assignment Algorithm (15, 10) (Based on most Frequent Allocations)

An important inference from the proposed traffic estimation approaches is that sometimes the prediction can be an over-estimation and sometimes it can be under-

estimation. In case of former, apparently it would appear from intuition that the superframe is underutilized but it might not be true in real-time cases. The reason is that the proposed approach relies on mean arrival rates. In reality, the actual arrival rate can be more than assumed and in that case over-estimation might not be that inefficient. The same is true for under-estimation when the actual arrival rate is less than the mean arrival rate. Of course if the opposite in both the cases happen, then the channel time is either underutilized, or it is over-utilized and there is congestion. The inconsistency in estimation for many of the results depicted occurs because of the multitude of parameters which dynamically vary and affect the superframe utilization. These parameters mainly consist of dynamic requirements of intra-PAN and inter-PAN flows, total duration of each flow, mean arrival rate and the number of intra-PAN and inter-PAN devices.

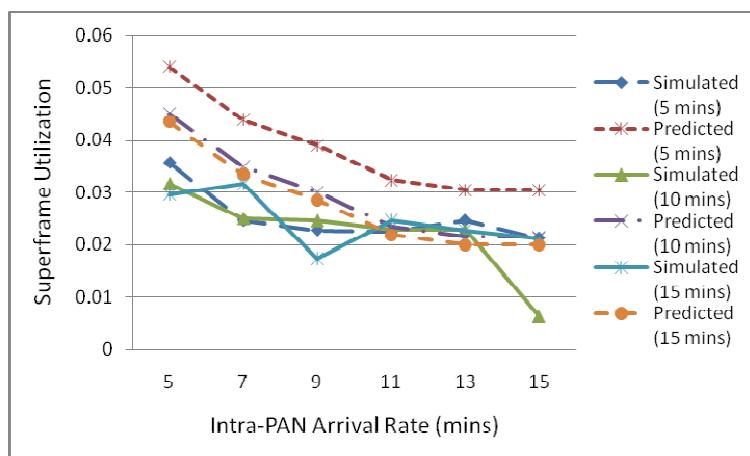


Figure 3.22 Traffic Estimation via Adjusted Assignment Algorithm (15, 10) (Based on total Frequency)

Estimated Backoff Algorithm Evaluation

In order to evaluate the effectiveness of the proposed estimated backoff algorithm, different number of MPNCs is considered. Two different simulation scenarios are defined to determine the performance. In the first case, after the d_{max} value expires for an MPNC but it still has active flows remaining in its cluster, it does not release its total allocated time but only releases the un-allocated time to the ref-MPNC. However, it does not accept further requests until its next allocation after the d_{max} value expires. Apart from that each MPNC can request for channel time from the ref-MPNC at any time and they do not have to wait until the next time period T_R starts. This situation results in a congested scenario and based on the superframe utilization; there is a higher possibility that the ref-MPNC rejects channel

time requests for requesting MPNCs. In the second case, each MPNC releases the channel time either after the end of each time period T_R or after the all the allocated flows end after expiration of d_{max} value. Therefore, at the start of each time period, the ref-MPNC usually has all of the superframe capacity available hence lesser congestion when compared with the first case.

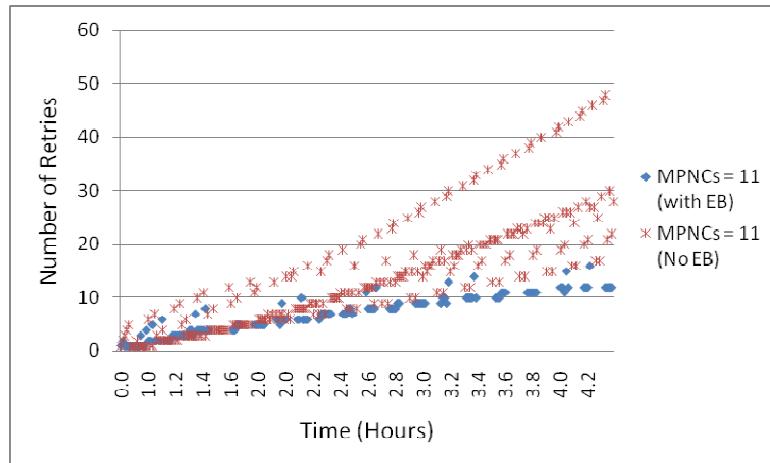


Figure 3.23 Number of Retries with and without the Estimated Backoff Algorithm (Case 1)

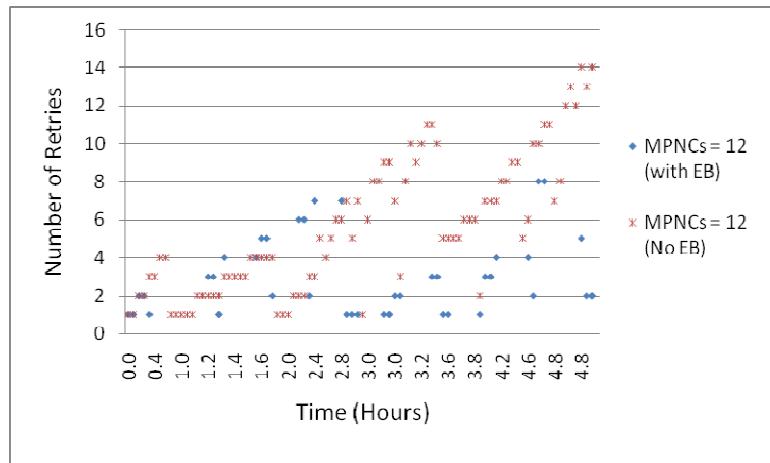


Figure 3.24 Number of Retries with and without the Estimated Backoff Algorithm (Case 2)

Figure 3.23 shows the results for the first case. The total simulation time is 5 hours and is the same for both cases. The number of MPNCs is taken to be 11 and the number of retries is compared when the Estimated Backoff (EB) algorithm is used and when the random retry approach is used. In the random retry approach, after an MPNC is rejected its channel time request, it sends another request after a random duration of time. A substantial reduction in

the number of retries can be observed in Figure 3.23. For the case of 9 MPNCs, the maximum number of retries without the EB algorithm for a given time period is 33 while with EB its 11 which is an improvement of 66%. For the case of 10 MPNCs, the improvement is 57% while for the case of 11 MPNCs, it is 66%.

In Figure 3.24, the results for the second case can be observed when 12 MPNCs are considered. The maximum number of retries in any given time period for 12 MPNCs is 14 without the EB algorithm while it is 8 with the EB algorithm hence an improvement of approximately 42%. Similarly for the case of 11 and 10 MPNCs, the improvement is 28% respectively.

The device which is expected to use a flow with duration which approximates or exceeds the value of d_{max} duration can start its flow anytime during the time period. Ideally if it starts even at the time of allocation to its MPNC, it is not necessary that the rest of the flows shall start at that time. Some of the flows might end after the d_{max} duration. Likewise if the device with d_{max} duration flow starts after the allocation time, then it can end after the d_{max} duration. In such cases, if there are on-going active flows, the MPNC does not release the total channel time. It releases whatever time is available. It however does not accept any more flows after the d_{max} value expires. The time released by the MPNC might be or might not be sufficient for a requesting MPNC after the estimated backoff. It can be more than its requirement or it can be less than its requirement. In case it's less than its requirement, then time is allocated based on the appropriate resource allocation policy in use.

Another factor is that the d_{max} value is an estimated value and it is not always accurate. The actual longest flow can either be more or less than the d_{max} value. The advantage however, remains that instead of random retries; the d_{max} value gives an educated guess to the requesting MPNCs for when the time is available as evident from the simulation results given in Figure 3.23 and Figure 3.24.

Evaluation of Proposed Resource Allocation Policies

The main parameter used for the performance evaluation of proposed resource allocation policies is the Cumulative Satisfaction Index (CSI). However, only the CSI value in itself is not sufficient enough to give an overall picture of the performance of a resource allocation scheme. Since the CSI value is based on the mean Satisfaction Index (SI) values of all the

MPNCs, it does not suggest if a certain number of MPNCs are discriminated. In order to elaborate further upon this point, consider a simple case of two MPNCs. If one of the MPNC has been allocated $T_i = T_B$ and the other MPNC is not allocated any time i.e. $T_i = 0$, then even though the SI value for the first MPNC is 1 but it is 0 for the second MPNC. The CSI value however, is 0.5. Now the CSI value of 0.5 can also be interpreted as $SI_1 = 0.5$ and $SI_2 = 0.5$ where SI_1 and SI_2 are the SI values of the first and second MPNCs respectively. Therefore, the CSI value does not give an idea about the level of discrimination for a resource allocation scheme.

In order to effectively indicate fairness and level of discrimination, the version of SI values based on the Jain's Fairness index in (19) is used. A limitation of Jain's Index is that it does not indicate the level of inequality in resource allocation to each MPNC. In meshed WPANs, each MPNC might have a different number of devices with different traffic requirements and hence a different bulk reservation size. The Theil's index and its normalized version, the Atkinson's index is helpful in determining the level of inequality in the resource allocation scheme. Therefore, the three mentioned parameters are thus sufficient to give a rather comprehensive performance evaluation of the proposed resource allocation schemes.

For each Case of proposed resource allocation policy, the total simulation time duration is kept to 5 hours i.e. 5 time periods of 1 hour duration. The number of MPNCs is also varied and a suitable number of MPNCs is selected so that the traffic requirements of all the MPNCs exceed the available superframe time. This is because it only makes sense to apply the resource allocation policies in situations where the superframe time is congested in order to get a better performance evaluation. In Figure 3.25, the resource allocation policy of case 1 is applied for 8, 10 and 12 MPNCs. The amount by which the total required time by all the MPNCs exceed the superframe capacity i.e. η_e is varied between 0.006239 secs to 0.018927 secs. In the simulation scenario, the number of intra-PAN and inter-PAN flow requests can vary per MPNC per time period. The bulk reservation size of an MPNC therefore is not necessarily the same for each time period and can change dynamically from time period to time period. It can be seen from Figure 3.25 that a good value of CSI is

maintained when the number of MPNCs is 8. When the number of MPNCs is increased to 12, the CSI drops sharply and becomes steady at a low value of 0.4.

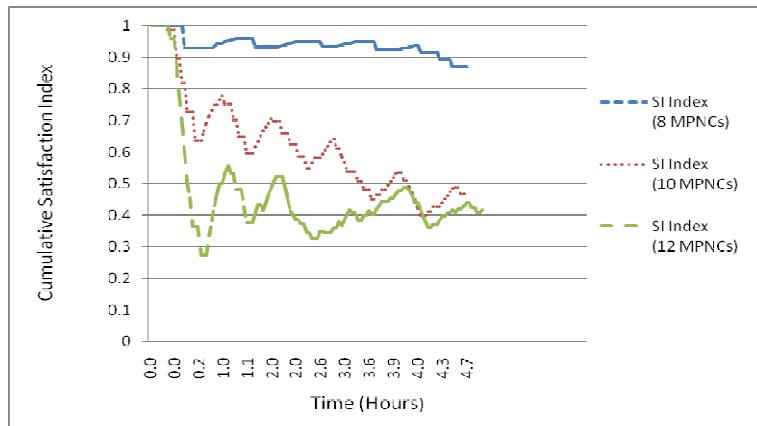


Figure 3.25 CSI values based on Policy of Case 1

When the policy of Case 3 i.e. Satisfy All (SA) Policy is applied to the same scenario which generated the results in Figure 3.26, the results can be seen in Figure 3.26. In the results of Figure 3.26, the relative improvement of the CSI is clearly evident with the application of SA policy. The minimum threshold value of SI i.e. S_{min} is taken to be 0.6. An acceptable value of CSI is maintained between 0.8 – 1.0 even when the number of MPNCs goes up to 12. Initially the value of CSI drops to approximately 0.7 for the case of 10 and 12 MPNCs but then converges to a value between 0.8-1.0. The results in Figures 3.27 and 3.28, a separate scenario is considered in which two changes are made. The first change is that the value of S_{min} is changed to 0.8 from 0.6 in Figure 3.28 and the second change is that after the expiration of d_{max} , the channel time is only released by an MPNC if there are no active flows left in its piconet. Therefore, at the start of a time period, the ref-MPNC might not have the total superframe capacity to allocate to a requesting MPNC. Another factor is that each MPNC need not wait for the start of the next time period and can request for channel time from the ref-MPNC depending on the mean inter-PAN arrival rate. Therefore, in presence of such factors, the CSI value can be seen to drop when compared to the previous scenario in which the ref-MPNC has the total superframe duration to allocate and each MPNC waits for the start of the subsequent time period to request for another allocation. However, when the value of $S_{min} = 0.6$ in Figure 3.27, a satisfactory CSI value is maintained even when the number of MPNCs is 12 even though it drops slightly below 0.6 at one point but then steadies above it for the rest of the time. In Figure 3.28, a satisfactory value above the

minimum threshold of 0.8 is only maintained when the number of MPNCs is 8. For the case of 10 MPNCs, the CSI value converges almost to 0.8 but for 12 MPNCs it remains near 0.6 for most of the time.

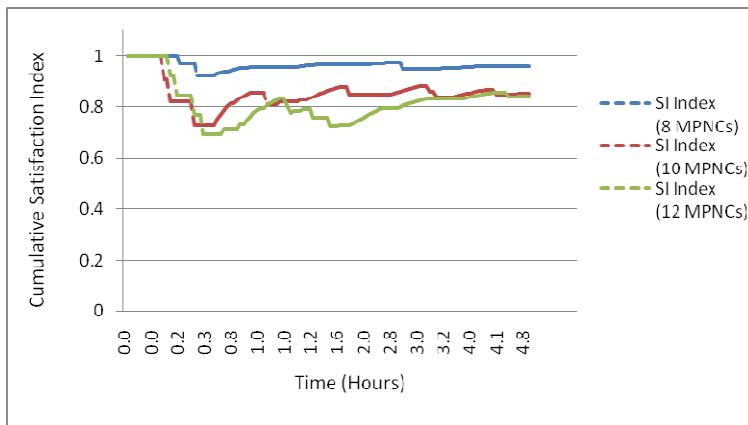


Figure 3.26 CSI values based on Policy of Case 3 ($S_{min} = 0.6$) (Scenario 1)

It can be deduced from the results of Figure 3.28 that in presence of increased traffic and congestion, the SA policy gives acceptable results for a lower threshold of 0.6 but it doesn't give acceptable results when the threshold is 0.8 and the number of MPNCs goes above 10. Therefore for cases in which the $\text{CSI} < S_{min}$ and the value of α related to Figure 3.29 , the second case of Accommodate All (AA) policy is considered. The results in Figure 3.29 show the CSI values when the number of MPNCs is changed from 3 MPNCs to 7 MPNCs. The minimum and maximum values of α and the values of α related to Figure 3.29 are given in Table 3.2. The value of CSI degrades as the number of MPNCs is increased which increases the value of α . Figure 3.30 shows the CSI values for another case when the number of MPNCs might vary in each subsequent time period. The individual SI values for each MPNC relevant to Figure 3.29 are given in Figure 3.31 for a fixed number of MPNCs per time period for the 5 Hours duration. The curves for individual SI values show a general downward trend as the number of MPNCs increase. It is important to note that for the simulation setup of AA policy, once an MPNC is allocated channel time for a time period T_R , it cannot send a request for channel time against during the same time period. The reason is that if the superframe time is sufficient for all the MPNCs, then each MPNC gets its required share. In a case when the superframe time is not sufficient i.e. $T_R < \sum_{i=1}^n T_i$, then the superframe time is divided among all the MPNCs based on a fair share. In such a case there is a probability that each MPNC gets a share lower than its required share of superframe time. Therefore, it can be

overutilized at some point and it is not feasible for it to release its allocated time before the end of the time period. In case of SA Policy, if a request by an MPNC is rejected by the ref-MPNC, the MPNC can send a request for channel time again during the same time period or the subsequent time period based on the EB algorithm. Also if an MPNC is allocated time but the allocation is under-estimation, than it can send a request for single allocations to the ref-MPNC. Therefore, the fluctuations during the time period can be seen in the graphs representing the case of SA policy being applied.

Table 3.2: Values of Parameters for the scenario with AA Policy (Result given in Figure 3.29)

Number of MPNCs	Minimum η_e	Maximum η_e	Mean Value of α for each MPNC (5 Hour Period)						
			α_1	α_2	α_3	α_4	α_5	α_6	α_7
3	0.014744	0.044091	0.40	0.34	0.26	-	-	-	-
4	0.045759	0.072594	0.27	0.22	0.25	0.27	-	-	-
5	0.064568	0.110222	0.21	0.21	0.18	0.20	0.21	-	-
6	0.079317	0.136182	0.17	0.17	0.17	0.17	0.17	0.17	-
7	0.119236	0.176955	0.22	0.18	0.20	0.18	0.10	0.11	0.10

In Figure 3.32, the CSI values depict the case when the same scenario of Figure 3.29 is simulated without the AA policy. An interesting observation is that the CSI values are almost the same as the case when AA policy is applied. When the SA policy is applied to the case of Figure 3.29, the results can be seen in Figure 3.33. The results show a downward trend in the CSI value with the application of SA policy. Based on the Jain's Fairness metric of (3.19), the difference in fairness can be seen in Figure 3.34 for the three proposed policies based on the values of η_e in Table 3.2 in which the number of MPNCs is from 3-7. Another comparison based on Jain's Fairness is shown in Figure 3.35 for lower values of η_e , for the same scenario of Figure 3.25 and Figure 3.26 in which the number of MPNCs is taken to be 8, 10 and 12. From Figure 3.34 it can be noted that the fairness of the case when no SA or AA policy is applied is better than the case when SA policy is applied but in Figure 3.35 the case with SA policy applied is better. One reason for it is that in case of scenario which generates the results in Figure 3.34 only considers relatively lower number of MPNCs i.e. 3-7 while the scenario of Figure 3.35 considers 8, 10 and 12 MPNCs. Since the Jain's fairness is sensitive to discrimination, the chance of discrimination in case of higher number of MPNCs is more.

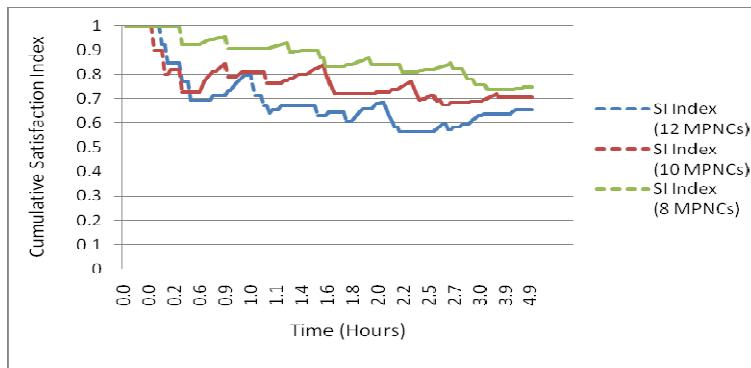
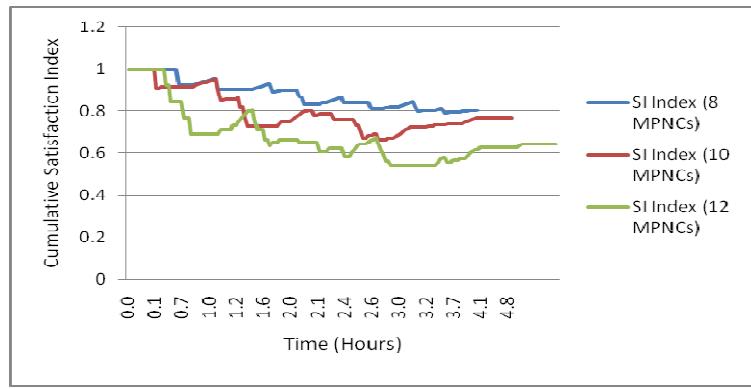
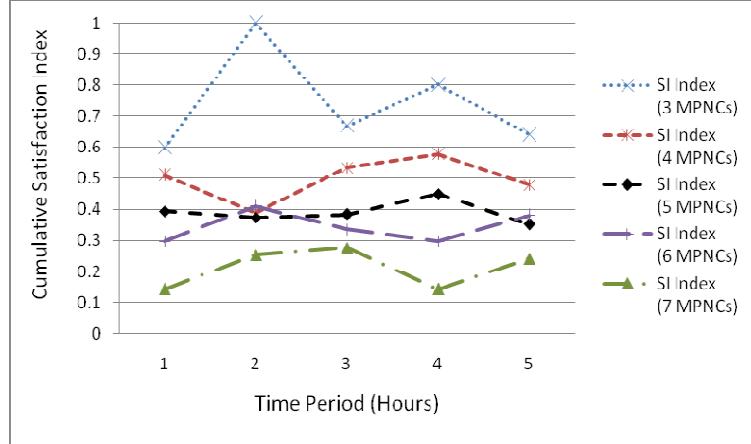
Figure 3.27 CSI values based on Policy of Case 3 ($S_{min} = 0.6$) (Scenario 2)Figure 3.28 CSI values based on Policy of Case 3 ($S_{min} = 0.8$) (Scenario 2)

Figure 3.29 CSI values based on Policy of Case 2 (Fixed number of MPNCs)

Coming back to the case of the results in Figure 3.29 and Figure 3.32 showing almost similar values of CSI with and without the application of AA policy, it should be noted that there is however a difference in fairness between the two approaches which can be seen in Figure 3.34. Another factor which separates the two approaches is the number of MPNCs which are allocated channel time. AA policy allocates time to all the requesting MPNCs, while without AA policy being applied; there is no guarantee that all the requesting MPNCs

are served hence difference in fairness. The inequality in the two approaches based on Theil's index can be observed from Figure 3.36 and Figure 3.37. A detailed comparison of the case with the application of AA policy and without the application of AA policy based on CSI values, Jain's Fairness metric, Theil's Index and Atkinson's Index is given in Table 3.3, Table 3.4, Table 3.5, and Table 3.6 respectively.

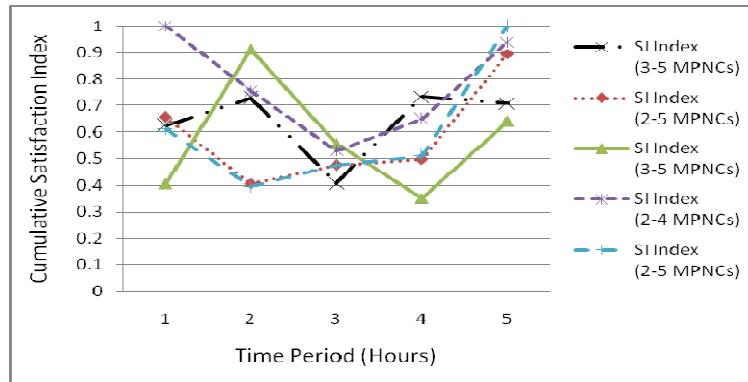


Figure 3.30 CSI Values based on Policy of Case 2 (Variable Number of MPNCs)

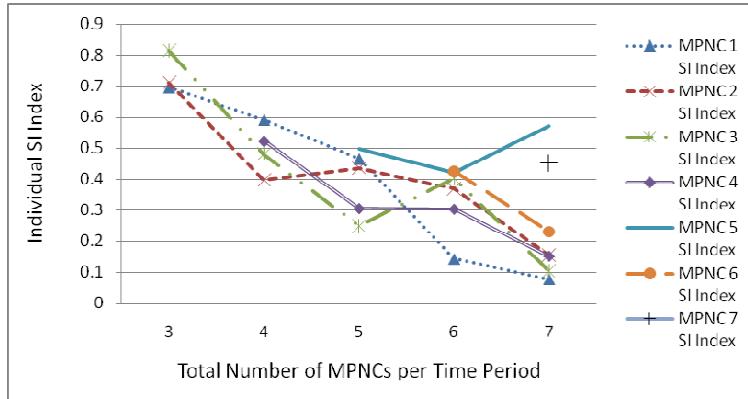


Figure 3.31 Individual SI values based on Policy of Case 2 (Fixed Number of MPNCs)

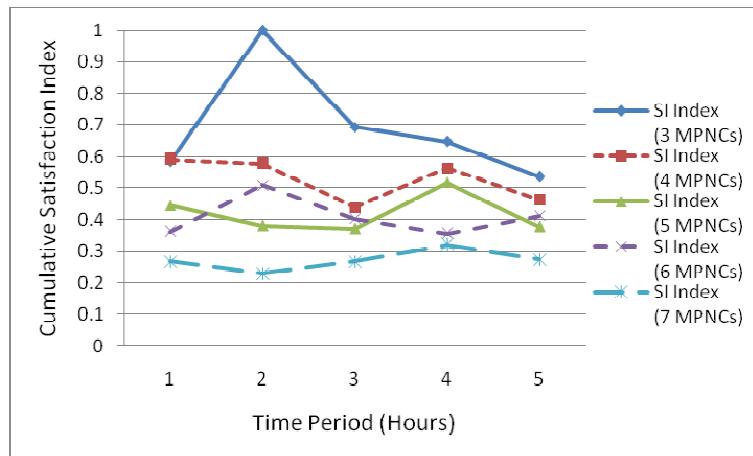


Figure 3.32 CSI Values when AA Policy is not used (Fixed Number of MPNCs)

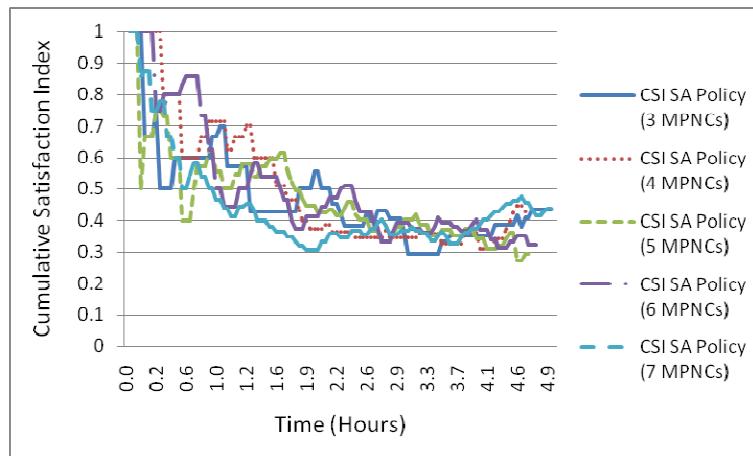


Figure 3.33 CSI values for SA policy for values of Table 3.2

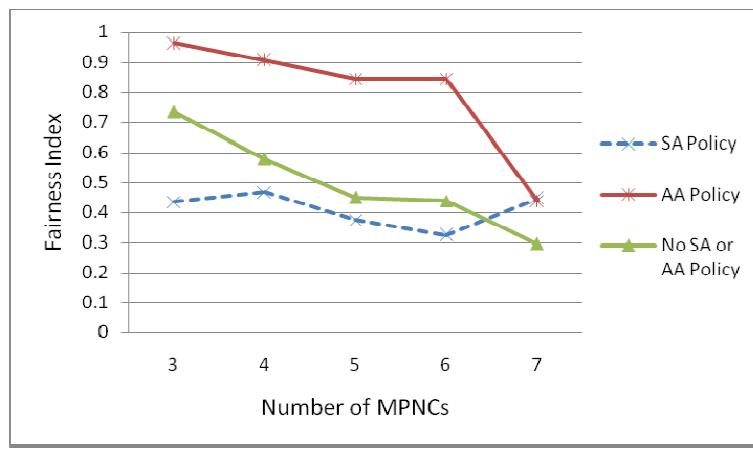


Figure 3.34 Comparison of proposed policies based on Jain's Index (high values)

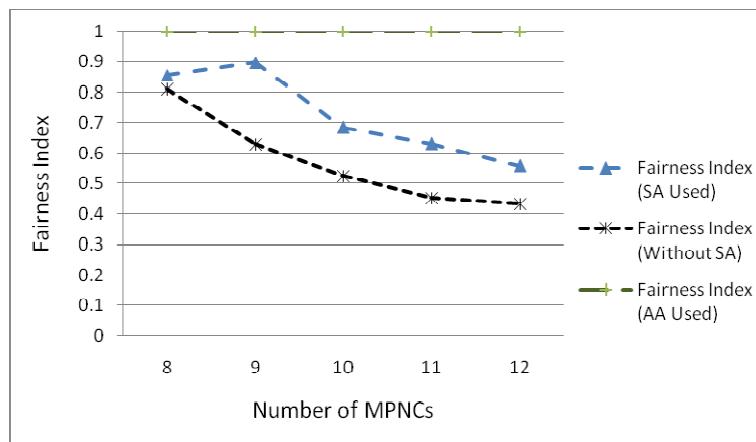


Figure 3.35 Comparison of proposed policies based on Jain's Index (low values)

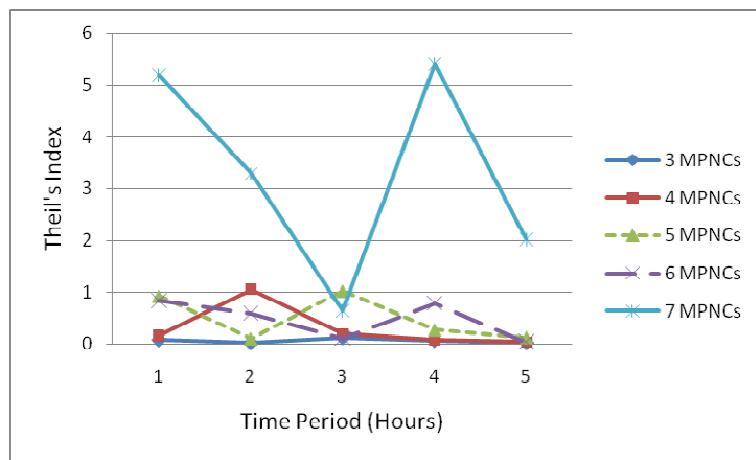


Figure 3.36 Results based on Theil's Index when AA Policy is used

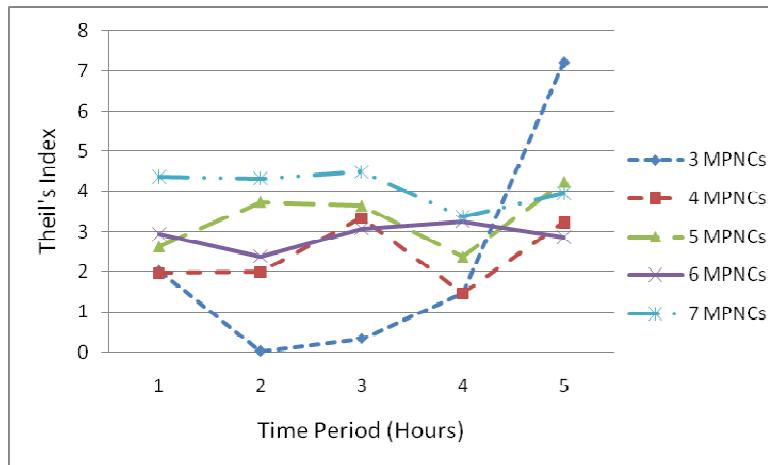


Figure 3.37 Results based on Theil's Index when AA Policy is not used

Comparison of Centralized and Distributed Allocation

In order to compare the centralized resource allocation proposed in this chapter with the distributed resource allocation scheme, two key parameters are selected for comparison. The first one is the control traffic overhead while the second one is the probability of conflicts in channel time reservations which can arise in case of distributed channel time reservation. The simulation setup for the centralized case is almost the same as before. However, for distributed resource reservation, a separate simulation model was developed. In the distributed model, each MPNC includes all its channel time allocation offsets and allocation sizes in its beacon frame. Therefore each MPNC can keep track of the respective allocations of its peer MPNC. Before reserving channel time for either an inter-PAN flow or and intra-PAN flow, a local conflict avoidance check is made by checking the offsets of already allocated reservations by the MPNC as well as its neighbours. An external conflict avoidance check is also made by checking the beacons for other peer MPNCs for their respective allocations. After the local and external conflict checks, an MPNC reserves channel time in the superframe for intra-PAN and inter-PAN flows. Since the CAP follows the Beacon Period (BP) and the allocation requests for intra-PAN and inter-PAN flows are sent in the CAP, there is a possibility that two or more MPNCs reserve channel time for a flow with overlapping reservations. In such a case, a conflict is raised and it is added to a statistic, the value of which is checked after the simulation run to determine the number of conflicts and calculate the probability of conflicts.

For the case of overhead comparison, the total control traffic is taken into account for centralized case as well as the distributed case. In the centralized case, once the MPNC gets its bulk reservation, it does not have to notify about its neighbours about reservations for intra-PAN flows. Even for inter-PAN flows, it just needs to indicate to its next hop neighbour MPNC about the flow parameters. There is no need to broadcast a reservation IE to the neighbour MPNCs to check for conflicts since the ref-MPNC ensures that conflicts are avoided. For checking the overhead as well as the conflict probability, 6 MPNCs are considered in the simulation setup for both the centralized as well as the distributed resource allocation setup. The mean arrival rate for intra-PAN and inter-PAN flows is kept the same for both cases and is taken to be 5, 6, 7, 8, 9 and 10 mins. The overhead (in bits)

comparison can be seen in Figure 3.38 in which the centralized scheme incurs much lower control overhead than the distributed scheme. For the same simulation scenario setup, the conflict probability for the arrival rates of 5, 6, 7, 8, 9 and 10mins is 0.19, 0.19, 0.14, 0.29, 0.22 and 0.41 respectively. The mean conflict probability is 0.24 which means that on average 24% allocations are in conflict for which the conflict needs to be resolved. There are no conflicts observed for the centralized case in the presence of ref-MPNC, as expected.

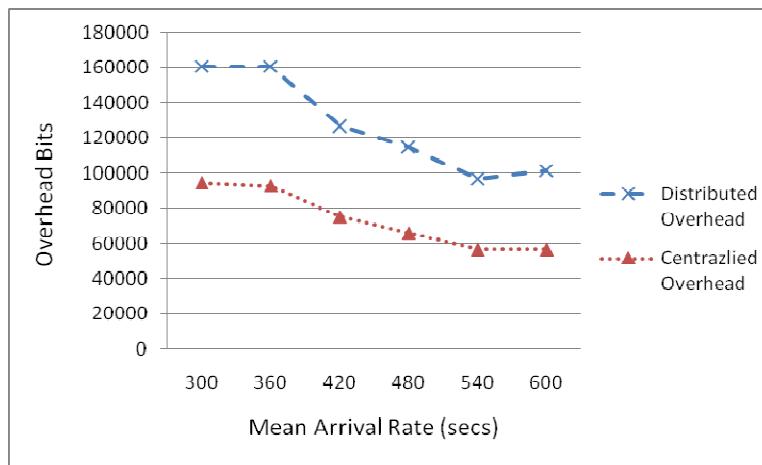


Figure 3.38 Comparison of Distributed Resource Reservation with Centralized Resource Reservation

Table 3.3: Comparison of Values based on CSI with and without AA Policy

Cumulative Satisfaction Index (CSI) When Accommodate All Policy is Used

Number Of MPNCs	3	4	5	6	7
Period 1	0.596855	0.509795	0.393114	0.296517	0.142857
Period 2	0.999188	0.389712	0.372746	0.409318	0.253289
Period 3	0.667891	0.535043	0.384028	0.334868	0.276183
Period 4	0.80106	0.577266	0.450227	0.297369	0.142857
Period 5	0.63947	0.478722	0.3515	0.380148	0.242075
Mean	0.740893	0.498108	0.390323	0.343644	0.211452

Cumulative Satisfaction Index (CSI) When Accommodate All Policy is Not Used

Number Of MPNCs	3	4	5	6	7
Period 1	0.584045	0.590714	0.445687	0.36217	0.268595
Period 2	0.999373	0.577113	0.380694	0.509272	0.229974
Period 3	0.693287	0.43939	0.371192	0.403832	0.268595
Period 4	0.647539	0.562724	0.518104	0.355948	0.320397

Period 5	0.535876	0.463991	0.376891	0.413221	0.274059
Mean	0.692024	0.526786	0.418514	0.408889	0.272324

Table 3.4: Comparison of Values based on Theil's Index with and without AA Policy

Theil's Index When Accommodate All Policy is Used					
Number Of MPNCs	3	4	5	6	7
Period 1	0.083603	0.160635	0.925414	0.830858	5.185463
Period 2	0.025543	1.054645	0.099529	0.585119	3.28496
Period 3	0.124277	0.189333	1.019174	0.103835	0.643817
Period 4	0.068918	0.080057	0.273179	0.775833	5.399812
Period 5	0.025448	0.042353	0.128168	0.029701	2.019689
Mean	0.065558	0.305405	0.489093	0.465069	3.306748

Theil Index When Accommodate All Policy is Not Used

Theil Index When Accommodate All Policy is Not Used					
Number Of MPNCs	3	4	5	6	7
Period 1	2.021561	1.964786	2.637879	2.93903	4.35424
Period 2	0.025528	1.992111	3.741688	2.391272	4.31614
Period 3	0.332552	3.335463	3.644791	3.08078	4.489622
Period 4	1.457089	1.464215	2.383619	3.249607	3.370825
Period 5	7.219139	3.227266	4.239375	2.860517	3.961124
Mean	2.211174	2.396768	3.32947	2.904241	4.09839

Table 3.5: Comparison of Values based on Jain's Index with and without AA Policy

Jain's Fairness Index When Accommodate All Policy is Used					
Number Of MPNCs	3	4	5	6	7
Period 1	0.984715	0.923359	0.792134	0.740384	0.142857
Period 2	1	0.748991	0.928917	0.828198	0.427515
Period 3	0.997344	0.937816	0.747354	0.914581	0.792348
Period 4	0.998301	0.989115	0.878649	0.781127	0.142857
Period 5	0.842223	0.938705	0.881078	0.957574	0.688662
Mean	0.964517	0.907597	0.845626	0.844373	0.438848

Jain's Fairness Index When Accommodate All Policy is Not Used					
Number Of MPNCs	3	4	5	6	7
Period 1	0.653587	0.65478	0.483965	0.387699	0.284558
Period 2	0.999999	0.635869	0.398974	0.518181	0.269861
Period 3	0.718678	0.490664	0.397605	0.449067	0.284558
Period 4	0.666085	0.613992	0.571444	0.37663	0.349005
Period 5	0.629187	0.497007	0.398502	0.459472	0.285198
Mean	0.733507	0.578462	0.450098	0.43821	0.294636

Atkinson's Index When Accommodate All Policy is Used					
Number Of MPNCs	3	4	5	6	7
Period 1	0.080204	0.148397	0.603633	0.564325	0.994403
Period 2	0.02522	0.651684	0.094736	0.442961	0.962558
Period 3	0.116864	0.172489	0.639107	0.098626	0.474716
Period 4	0.066597	0.076936	0.239044	0.53968	0.995483
Period 5	0.025127	0.041469	0.120295	0.029264	0.867303
Mean	0.062802	0.218195	0.339363	0.334971	0.858893

Atkinson's Index When Accommodate All Policy is Not Used					
Number Of MPNCs	3	4	5	6	7
Period 1	0.867551	0.859814	0.928487	0.947083	0.987148
Period 2	0.025205	0.863593	0.976286	0.908487	0.986649
Period 3	0.282909	0.964402	0.973873	0.954077	0.988775
Period 4	0.767087	0.76874	0.907784	0.961211	0.965639
Period 5	0.999268	0.960334	0.985583	0.942761	0.980958
Mean	0.588404	0.883377	0.954403	0.942724	0.981834

3.10 Conclusion

The proposed centralized resource allocation schemes perform better than the distributed resource allocation scheme in terms of reduced control overhead and channel time reservation conflict avoidance. One of the reasons for the reduction in control traffic

overhead is due to the idea of bulk reservation of channel time by an MPNC. In the case of AA policy being applied for resource allocation, an assumption made is that in case of superframe capacity being sufficient for all the MPNCs, an MPNC which is allocated its required share does not send a request for channel time allocation during the same time period in which it is allocated channel time. This assumption holds true in the case if the traffic estimation by an MPNC is either 100% accurate or it is an over-estimation. If the estimation for a bulk flow by an MPNC is an under-estimation, then even though the superframe capacity is sufficient for all the MPNCs, it might have to send a request for reservation of a single flow. Even in the rest of the cases, if the traffic estimation is an over-estimation or an under-estimation, the superframe capacity can be inefficiently utilized. Therefore, the accuracy of the traffic estimation techniques is an important factor which has substantial effects on the amount of control overhead and efficient utilization of superframe time by the MPNCs.

Another important factor to consider is that only the CSI value is not a comprehensive measure of the fairness of a resource allocation policy. As it was observed from the simulation results that by using the Jain's fairness metric and the Theil's index along with the CSI can give us a more comprehensive insight into the fairness and equality of a resource allocation scheme.

3.11 References

- [1] "IEEE Recommended Practice for Information technology--Telecommunications and information exchange between systems-- Local and metropolitan area networks-- Specific requirements Part 15.5: Mesh Topology Capability in Wireless Personal Area Networks (WPANs)," *IEEE Std 802.15.5-2009*, vol., no., pp.1-166, May 8 2009.
- [2] M.S. PARK , B. LEE , and S.H. RHEE, "Distributed Multiple Access Control for the Wireless Mesh Personal Area Networks" *IEICE Trans Inf & Syst E91-D*: 258-263.
- [3] Ho Ting Cheng; Weihua Zhuang, "Novel packet-level resource allocation with effective QoS provisioning for wireless mesh networks," *Wireless Communications, IEEE Transactions on*, vol.8, no.2, pp.694-700, Feb. 2009.
- [4] Barcelo, F.; Jordan, J., "Channel holding time distribution in public telephony systems (PAMR and PCS)," *Vehicular Technology, IEEE Transactions on* , vol.49, no.5, pp.1615-1625, Sep 2000
- [5] Jedrzycki, C.; Leung, V.C.M., "Probability distribution of channel holding time in cellular telephony systems," *Vehicular Technology Conference, 1996. 'Mobile Technology for the Human Race', IEEE 46th* , vol.1, no., pp.247-251 vol.1, 28 Apr-1 May 1996.
- [6] Bolotin, V.A., "Modeling call holding time distributions for CCS network design and performance analysis," *Selected Areas in Communications, IEEE Journal on* , vol.12, no.3, pp.433-438, Apr 1994
- [7] Ho Ting Cheng; Weihua Zhuang, "Joint Power-Frequency-Time Resource Allocation in Clustered Wireless Mesh Networks," *Network, IEEE* , vol.22, no.1, pp.45-51, Jan.-Feb. 2008.
- [8] Johansson, M.; Sternad, M., "Resource allocation under uncertainty using the maximum entropy principle," *Information Theory, IEEE Transactions on* , vol.51, no.12, pp. 4103-4117, Dec. 2005.
- [9] Murthy, K.M.S.; Catherasoo, R., "Enhanced capacity resource allocation protocol for packetized voice in a TDMA short-range wireless network. 1. Algorithm and performance," *Universal Personal Communications, 1993. Personal Communications: Gateway to the 21st Century. Conference Record., 2nd International Conference on* , vol.2, no., pp.713-717 vol.2, 12-15 Oct 1993.
- [10] Qin-yun DAI, Lu RONG, Hong-lin HU, Gang SU, "Resource allocation using time division multiple access over wireless relay networks," *The Journal of China Universities of Posts and Telecommunications*, Vol.15, Issue 3, September 2008, pp. 69-74, ISSN 1005-8885.
- [11] Vergados, D.D.; Vergados, D.J.; Douligeris, C.; Tombros, S.L., "QoS-aware TDMA for end-to-end traffic scheduling in ad hoc networks," *Wireless Communications, IEEE* , vol.13, no.5, pp.68-74, October 2006.

- [12] Gesbert, D.; Kountouris, M., "Resource allocation in multicell wireless networks: Some capacity scaling laws," *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. WiOpt 2007. 5th International Symposium on*, vol., no., pp.1-7, 16-20 April 2007.
- [13] Wen-Hsing Kuo; Wanjuin Liao, "Utility-Based Resource Allocation in Wireless Networks," *Wireless Communications, IEEE Transactions on*, vol.6, no.10, pp.3600-3606, October 2007.
- [14] Curescu, C.; Nadjm-Tehrani, S., "Time-aware utility-based resource allocation in wireless networks," *Parallel and Distributed Systems, IEEE Transactions on*, vol.16, no.7, pp. 624-636, July 2005.
- [15] C. R. Lin and J.S. Liu, "Qos routing in ad hoc wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 8, pp. 1426-1438, 1999.
- [16] Mahmud S.A, Khan S, Qiang Ni, Al-Raweshidy H.S., "Capacity Issues in Meshed High Data Rate WPANs," *Advanced Information Networking and Applications - Workshops, 2008. AINAW 2008. 22nd International Conference on*, vol., no., pp.1285-1290, 25-28 March 2008.
- [17] Clyde L Monma and T. Carpenter, "Variations of Matrix Balancing for telecommunication demand forecasts," Telcordia, 1997.
- [18] Michael H. S and Zenios S A., "A Comparative Study of Algorithms for Matrix Balancing" *Operations Research*, Vol. 38, No. 3 (May - Jun., 1990), pp. 439-455.
- [19] A. Medina, N. Taft, S. Battacharya, C. Diot, and K. Salamatian, "Traffic matrix estimation: Existing techniques compared and new directions," in *SIGCOMM, Pittsburgh*, 2002.
- [20] H. Spiess, "Biproportional matrix balancing with upper bounds," *Transportation and Network Analysis: Current Trends, Applied Optimization* vol. 63, Kluwer Academic Publishers, pp. 237–243, 2002.
- [21] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [22] K. Ahuja, T.L. Magnati, and J.B. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.
- [23] M. S. Bazaraa and J. J. Jarvis, *Linear Programming and Network Flows*, NY: Wiley, 1977.
- [24] "IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 15.3: wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs)," *IEEE Std 802.15.3-2003* , vol., no., pp. 0_1-315, 2003.

- [25] Mahmud, S.A.; Khan, S.; Al-Raweshidy, H.S., "Capacity Analysis of High Data Rate Wireless Personal Area Networks," *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual* , vol., no., pp.125-131, 5-8 May 2008.
- [26] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," DEC-TR-301, Digital Equipment Corporation, Tech. Rep., September 1984.
- [27] Theil, H. *Economics and Information Theory*, Amsterdam: North-Holland, 1967.
- [28] Amartya, S. *On economic inequality*, Oxford: Clarendon Press, 1973.

CHAPTER 4

4 Capacity and Delay Efficient Communication in HDR Meshed WPANs

4.1 Introduction

The capacity analysis results from chapter 2 indicate that devices with high throughput requirements occupy a significant fraction of superframe capacity. The increase or decrease in transmission data rate, the number of frames/CTA and the fragment size used contribute in determining the overall CTA size and the resulting CTA overhead. For a certain requirement of minimum throughput, the parameters which are used to calculate the corresponding CTA size are number of frames/CTA, Fragment (MPDU) size and the transmission data rate. Although different combinations of these parameters can result in an appropriate CTA size to be sufficient enough to support the required throughput, the resulting overhead varies. The term overhead is used in terms of the CTA overhead which consists of the InterFrame Spaces (IFS), Guard Time (GT) interval, MAC header bits per frame and any Acknowledgements (ACKs) sent/received per CTA. A suitable combination of parameters used to determine an appropriate CTA which complies with a certain minimum throughput requirement with minimum overhead, therefore becomes desirable. In order to address this issue, a parameter called *Time Efficiency* is derived and proposed to define the communication overhead of a flow based on its CTA overhead. Increasing the time efficiency for a flow means that the superframe capacity allocated to the flow is efficiently utilized and the communication overhead is minimized. However, increasing the time efficiency can result in a larger CTA size (described later in the chapter) and therefore, in case of congestion increasing the time efficiency of a flow beyond a certain limit might not be desirable. To increase the time efficiency, it is therefore mandatory to take into account, the available superframe capacity as a constraint. Based on the concept of time efficiency, two

search based algorithms are proposed which take into account the minimum throughput, maximum CTA size and minimum time efficiency values as constraints and selects the most appropriate parameters for the flow. The parameters chosen are the fragment size, transmission data rate and number of frames/CTA. It should be noted that only those values for the mentioned parameters are chosen which are supported by a device.

Another aspect investigated in this chapter is the *capacity efficiency* and *End-End-Delay (EED)* for meshed HDR WPANs. The motivation to investigate and propose a solution for the capacity efficiency is raised from the fact that in a multi-hop path in meshed WPANs, there is a possibility that one or more links along the source-destination path are congested. As a result, it might not be possible to support some flows which have a certain minimum throughput requirement on those paths having congested links. Even some paths without congested links do not imply that some links in such paths might not get congested in future. It is therefore, important to select those paths which are more capacity efficient than others. The first selection criteria for a multi-hop path based on capacity efficiency is therefore the transmission data rate which is supported by a relay device at each link throughout the source-destination path. The second selection criteria considered in the proposed approach is the EED of a path. Paths with a lower value of EED are preferred over those paths which have a higher value of EED. The EED value is of significant importance for real time applications which are delay sensitive. The factors which add substantially to the total EED along a source-destination path are the propagation delay, transmission delay and the queuing delay. For TDMA based wireless networks, the multi-path queuing delay refers to the time delay which occurs between the receiving CTA and the sending CTA for the same flow along a particular source-destination path. If the allocation of CTAs along the source-destination path exceeds the superframe capacity, the EED exceeds by one superframe duration. Therefore, the allocation pattern of CTAs for multi-hop flows in meshed WPANs not only contributes to the queuing delay but the total EED as well. In order to select the paths with a CTA pattern which results in the lowest EED, a routing metric called *Switched Time Slot (STS)* is proposed in this chapter. The STS metric not only determines the path with the lowest EED, it also re-orders and re-groups the CTAs to further minimize the total EED. Another version of STS called *Switched Time Slot with Re-adjustment (STS-R)* is also

proposed which further re-orders the CTAs if possible for a link along the source-destination path in the event of either an intra-PAN or inter-PAN flow termination. The performance of the proposed metrics is evaluated and compared with another relevant metric called Minimized Slot Misordered Routing (MSMR) [1]. Performance comparison shows the superiority of STS as well as STS-R metrics in terms of capacity efficiency as well selecting paths with lower values of EED in some cases. In the rest of the chapter the terms *slot* and *CTA* are used synonymously. This chapter contains the following contributions:

1. Derivation of Time Efficiency and analysis of factors affecting the Time Efficiency and capacity of a flow
2. Two table search based algorithms to select the most suitable parameters for a flow given the minimum throughput and CTA size constraints
3. Evaluation of two proposed capacity and delay efficiency metrics known as STS and STS-R

4.2 Capacity and End-to-End Delay Issues in Meshed WPANS

In order to quantify the feasibility of proposed contributions, an overview of the issues related to capacity and delay in meshed WPANs is given in the subsequent sections

4.2.1 Capacity Utilization in a multi-hop Path

Since the meshed HDR-WPANs extend to multiple hops, the need for routing becomes mandatory. If a source device in one piconet intends to send data to a destination device in another piconet, the MPNC of the source device's piconet has to forward data on behalf of its member device to the next MPNC along the source-destination path. Each MPNC in the multi-hop path from the source to destination has to reserve time in the superframe to forward data to the next hop relay MPNC and to receive it from the previous hop MPNC. In Figure 4.1, a multi-hop meshed WPAN is shown in which a source Mesh Device (MDEV) sends data to its MPNC (S-MPNC). The source MPNC after establishing a route to the Destination MPNC (D-MPNC) sends a reservation request to the next hop Relay MPNC (R-MPNC). In Figure 4.1, the dotted circles show the MPNCs which share the same Beacon Period (BP). Since the next hop R-MPNC is R-MPNC 1, S-MPNC forwards the data to it. R-

MPNC 1 then forwards the data to R-MPNC 2 and so on till the data is delivered to the D-MPNC. R-MPNC 1 being in transmission range of both S-MPNC and R-MPNC 2, shares the BP with both of them. Since the MPNCs which are in one hop transmission range of each other share the same superframe, time is reserved for the transmission from S-MDEV to S-MPNC, from S-MPNC to R-MPNC 1 and from R-MPNC 1 to R-MPNC 2. Although the transmission from R-MPNC 1 to R-MPNC 2 does not take place in the same shared superframe of S-MPNC, time has to be reserved for it in the superframe of S-MPNC since they are in direct transmission range of each other. Therefore, the capacity used for a single multi-hop flow along the source-destination path in the same superframe is three times that of the single hop flow to avoid the hidden node problem.

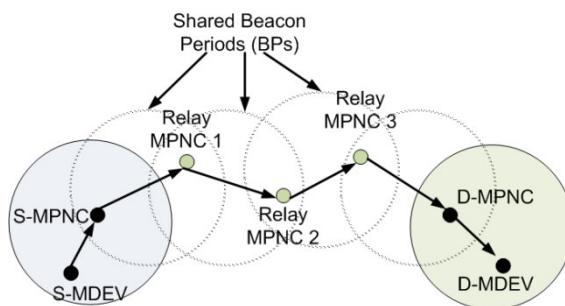


Figure 4.1 Capacity reservation along a source-destination path

The transmitter data rates defined in the IEEE 802.15.3 standard for HDR-WPANs are 11, 22, 33, 44 and 55 Mbps. The transmitter data rate used by a device is an important factor in determining the duration of a CTA required by the device. The higher the transmitter data rate, the lower is the duration of CTA required for a device to communicate. Therefore, by using higher transmitter data rates, more devices can be accommodated in the superframe and hence improved capacity. However, the CTA overhead also increases by transmitting at higher data rates. The CTA overhead increases because inside a CTA, there are InterFrame Spaces (IFSs) used between each frame transmitted¹. Each frame also consists of a MAC header which according to [2] is always sent at the base rate of 22 Mbps. At the end of each CTA there is a Guard Time (GT) interval used to make up for the propagation delay and thus

¹ As mentioned in the previous chapters, the IFS and GT duration remains constant and do not depend on transmitter data rates. However, reduced time duration is required to send the MPDU with increased transmitter data rates. Therefore, the relative CTA overhead increases.

maintain synchronization. The type of IFS used depends on the acknowledgement policy used. If the Delayed Acknowledge (Dly-ACK) scheme is considered for multi-media traffic, then either a Minimum InterFrame Space (MIFS) or a Short InterFrame Space (SIFS) is used between successive frames transmitted. The overhead in the CTA due to the Dly-ACK policy can be given by

$$\tau_{Dly-ACK(SIFS)} = (\sum_{n=1}^{x_d+b} \eta_{(SIFS)n} + b \cdot \eta_{Dly-ACK} + \eta_{GT} + x_d \cdot \eta_{MAC_HDR}) \quad (4.1)$$

Where x_d is the number of frames transmitted in the CTA, η_{SIFS} , η_{MAC_HDR} and η_{GT} are the SIFS duration, time to send the MAC header and the GT duration respectively. The CTA overhead is given by $\tau_{Dly-ACK(SIFS)}$ and the parameter b is set to one if in a particular CTA, the Dly-ACK frame is sent. If the value of b is one, then the time duration to send a Dly-ACK frame given by $\eta_{Dly-ACK}$ is also considered. The efficiency of time utilized by a device in the CTA is given by

$$Time\ efficiency\ (\eta_{eff}) = \frac{(\eta_{MPDU} \times x_d)}{(\eta_{MPDU} \times x_d) + (x_d \times \eta_{SIFS} + \eta_{GT}) + (x_d \times \eta_{MAC_HDR})} \quad (4.2)$$

Where η_{MPDU} is the time required to send an MPDU that depends on the transmitter data rate used and the size of the MPDU. Since the rest of the MPDU after the MAC header can be sent at any of the transmitter data rate, therefore the value of η_{MPDU} varies. It can be noted from (4.2) that with the decrease in the value of η_{MPDU} the time efficiency decreases. Also, with the increase in transmitter data rate, the CTA overhead increases. To decrease the CTA overhead, a higher value of x_d can be used which can increase the time efficiency and the throughput achieved by a device at the cost of consuming more superframe capacity. As a result, there is a trade-off between improving capacity and improving the time efficiency.

The size of the MPDU used and the number of frames sent per CTA are important factors in determining the actual throughput (δ_A) and effective throughput (δ_E) achieved by a device. The actual throughput considers the overall throughput achieved including the CTA overhead. The effective throughput considers the throughput without the overhead and is an important parameter in making sure that the required throughput (δ_R) by an application or a specific voice or video codec at the higher layer is maintained.

In a meshed HDR-WPAN, the capacity efficiency of a particular flow can be estimated from the sum of total time allocated to a flow on each link along a source-destination path. As explained before, the length of the CTA varies based on the value of transmitted data rate (δ_T). Higher values of δ_T results in lower values of η_{CTA} and thus improves the capacity. If the value of δ_T varies along different links along a source-destination path, then the time allocated to a flow in a superframe varies hence variation in capacity efficiency. Let C_T be the total capacity used by a particular flow along a source-destination path, then

$$C_{TH} \leq C_T = \sum_{i=1}^n c_i \leq C_{TL} \quad (4.3)$$

Where c_i is the CTA size for a flow d_i in a superframe between $R\text{-MPNC}_i$ and $R\text{-MPNC}_j$ along the source-destination path and C_{TH} and C_{TL} are the best case capacity efficiency and the worst case capacity efficiency respectively. The best case capacity efficiency, C_{TH} is achieved if all the MPNCs along a path are using the highest value of δ_T . In such a case, the CTAs allocated to d_i have a lower duration and hence time is saved along the multi-hop path. The worst case efficiency C_{TL} is achieved if the value of δ_T used by all the MPNCs along a multi-hop path is the lowest. In such a case, more superframe capacity is occupied by the flow of d_i .

4.2.2 End-to-End Delay in multi-hop Meshed WPANs

As discussed before, the factors which contribute to the EED are the queuing delay, transmission delay and propagation delay at each MPNC along the source-destination path. The upper limit on the tolerable inter-arrival delay and jitter as defined in [2] is < 10ms for voice flows and < 100ms for video flows. Therefore, it should be ensured that the propagation delay and the queuing delay do not violate the QoS requirements. An example of queuing delay in meshed HDR-WPANs is shown in Figure 4.2.

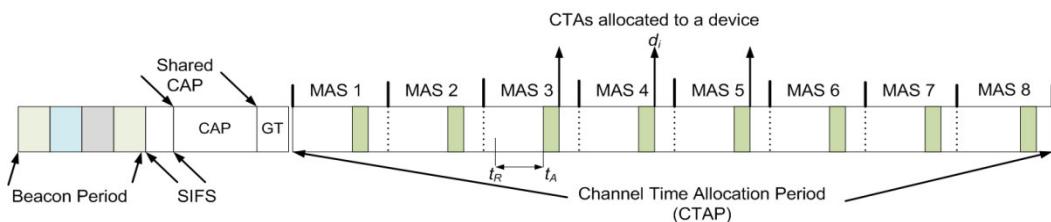


Figure 4.2 An example of queuing delay at each hop

It can be seen in Figure 4.2 that a device d_i has been allocated 8 CTAs in the superframe to transfer a real time flow. It is assumed that this allocation has been done by a Relay MPNC (R-MPNC) along a source-destination path. If the R-MPNC receives traffic which belongs to the flow d_i from its previous hop MPNC at a time instant t_R , where $t_R < t_A$ and t_A is the time instant at which the CTA_i for d_i starts, then $t_A - t_R = t_d$ and t_d is the queuing delay at a link, l_k which belongs to the source destination path for d_i . If T_d is the total queuing delay of the source-destination path such that $T_d = \sum_{i=1}^n t_{di}$ and n is the number of hops, then the following conditions are necessary to maintain the QoS for multi-media flows

- a. $T_d + PG_d + T_r < 150$ ms (for voice flows)
- b. $l_t < 10$ ms (for voice flows) and $l_t < 100$ ms (for video flows)
- c. $\partial_T > \partial_A > \partial_E \geq \partial_R$

Where PG_d and T_r are the propagation delay at each MPNC and the total transmission delay respectively while l_t is the inter-arrival time between successive frames.

The type of topology used for meshed HDR-WPANs can have an important impact on real time applications, especially in terms of the EED experienced. If a tree based topology as suggested in [2] is used, then tree based routing is the most appropriate choice for multi-hop communication. The tree topology has an advantage in terms of easier path discovery. The root of the tree is usually the first device which starts the tree and is called a Mesh Coordinator (MC). The MC then initiates the discovery of other MPNCs by including a mesh capacity IE in its beacon. Upon reception of the beacon from the MC, those MPNCs intending to join the tree, send a tree association request to the MC. Those MPNCs which cannot support another child indicate it in their beacon. Each associated MPNC is allocated a TREE ID from a TREE ID pool. The TREE ID makes it easier to determine the location of the branch where a particular destination MPNC is expected to be found. However, the tree topology can have its disadvantages in introducing additional delays in a source-destination path when the source and destination belong to difference branches of the tree. Such a situation is shown in Figure 4.3. Further details about the tree topology for meshed WPANs can be found in [3] and [4].

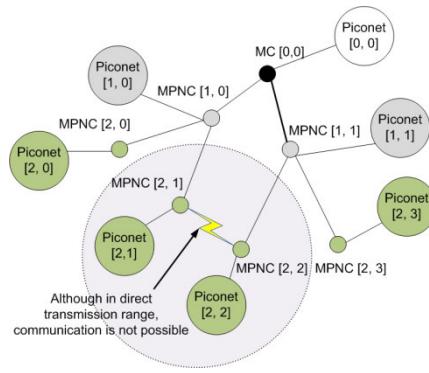


Figure 4.3 A tree topology for meshed WPANs

It can be seen from Figure 4.3 that the source and destination MPNCs have different parents and they belong to different branches of the tree. Although they are in direct transmission range of each other, they cannot communicate directly and the path must include the MC. Therefore, the number of hops in the source-destination path increases unnecessarily and the EED increases.

4.3 Communication Efficiency in Meshed WPANs

In any wireless communication system, one important aspect of increasing the communication efficiency means that the channel time or bandwidth is efficiently utilized. If the ratio of actual data sent to the actual data along the overhead is increased, that means that more actual data bits are sent relative to the overhead bits in a flow. The proposed time efficiency parameter η_{eff} takes into account the overhead bits and actual bits sent per CTA. In this section, the Time Efficiency (η_{eff}) parameter is derived and its relation to actual throughput (δ_A) and effective throughput (δ_E) is determined. The change in CTA size with the corresponding change in η_{eff} along with further analysis is also conducted. Based on the value of η_{eff} , minimum required throughput of a flow and maximum CTA size allowed for a flow, two search based algorithms are proposed to determine the essential parameters for a flow i.e. transmission data rate (δ_T), Suitable Fragment size and the number of frames/CTA

4.3.1 Deriving Time Efficiency to determine communication overhead

As discussed before, the value of ∂_A indicates the value of throughput achieved including all the overheads incurred at the MAC layer. The value of ∂_E gives us the throughput without the overheads at the MAC layer. As far as the real time applications are concerned, the value of ∂_E should at least match the required throughput ∂_R by the application layer i.e. $\partial_E \geq \partial_R$. If the Dly-ACK scheme is considered for the real time flows, the effective throughput is given by

$$\text{Effective Throughput } (\partial_E) = [\eta_{CTA} - (x_d * \eta_{SIFS} + \eta_{GT}) - (x_d * \eta_{MAC_HDR})] * \partial_T * N_{CPS} * 1/\eta_s \quad (4.4)$$

Where N_{CPS} is the number of CTAs sent per superframe and η_s is the superframe duration in seconds. The term η_{CTA} is used to indicate the total CTA duration and is given by

$$\text{CTA Duration } (\eta_{CTA}) = (\sigma_{MPDU} * x_d) / \partial_T + x_d * \eta_{SIFS} + \eta_{GT} + (x_d * \eta_{MAC_HDR}) \quad (4.5)$$

Where σ_{MPDU} is the MPDU (Fragment) size in bits. The actual throughput is given by

$$\text{Actual Throughput } (\partial_A) = N_{CPS} * \frac{1}{\eta_s} * \eta_{CTA} * \partial_T \quad (4.6)$$

Dividing (4.4) by (4.6), we get

$$\begin{aligned} \frac{\text{Effective Throughput}}{\text{Actual Throughput}} &= \frac{[\eta_{CTA} - (x_d * \eta_{SIFS} + \eta_{GT}) - (x_d * \eta_{MAC_HDR})] * \partial_T * N_{CPS} * 1/\eta_s}{N_{CPS} * \frac{1}{\eta_s} * \eta_{CTA} * \partial_T} \\ &= \frac{[\eta_{CTA} - (x_d * \eta_{SIFS} + \eta_{GT}) - (x_d * \eta_{MAC_HDR})]}{\eta_{CTA}} \end{aligned}$$

Substituting the value of η_{CTA} from (4.5)

$$= \frac{[(\sigma_{MPDU} * x_d) / \partial_T + x_d * \eta_{SIFS} + \eta_{GT} + (x_d * \eta_{MAC_HDR}) - (x_d * \eta_{SIFS} + \eta_{GT}) - (x_d * \eta_{MAC_HDR})]}{(\sigma_{MPDU} * x_d) / \partial_T + (x_d * \eta_{SIFS} + \eta_{GT}) + (x_d * \eta_{MAC_HDR})}$$

Since $\sigma_{MPDU} / \partial_T = \eta_{MPDU}$

$$= \frac{[(\eta_{MPDU} * x_d) + x_d * \eta_{SIFS} + \eta_{GT} + (x_d * \eta_{MAC_HDR}) - (x_d * \eta_{SIFS} + \eta_{GT}) - (x_d * \eta_{MAC_HDR})]}{(\eta_{MPDU} * x_d) + (x_d * \eta_{SIFS} + \eta_{GT}) + (x_d * \eta_{MAC_HDR})}$$

Further solving gives us,

$$= \frac{(\eta_{MPDU} * x_d)}{(\eta_{MPDU} * x_d) + (x_d * \eta_{SIFS} + \eta_{GT}) + (x_d * \eta_{MAC_HDR})}$$

Which is equal to the time efficiency η_{eff}

Therefore,

$$\frac{\text{Effective Throughput}}{\text{Actual Throughput}} = \eta_{\text{eff}} \quad (4.7)$$

From (4.7), it is clear that the time efficiency is simply the ratio of effective throughput to the actual throughput.

4.3.2 Analysis of Factors Effecting Time Efficiency

After determining the relationship of η_{eff} with the actual and effective throughput, it is established from (4.7) that a higher value of time efficiency η_{eff} explicitly indicates a higher communication efficiency and lower overhead. However, as discussed before, increasing the value of η_{eff} for a flow means either increasing the number of frames sent per CTA or increasing the fragment size which increases the CTA size as well. Also, when number of frames/CTA is increased, it means that the addition of each frame also adds a MAC header which contributes to the CTA overhead. This especially is evident in cases where the value of δ_T is high and the fragment size is small. This section presents detailed analysis in order to demonstrate the effect of fragment size, number of frames/CTA and transmission data rate (δ_T) on time efficiency, CTA duration and the percentage overhead. The findings in this section not only provide useful insights and motivation for proposing the table based algorithms in the subsequent section but also for the proposed metrics STS and STS-R. The values of parameters used for the analysis are given in Table 4.1.

Table 4.1: Values of Parameters Used for Analysis of Time Efficiency

	Parameter	Values used
1	Transmission Data Rate (δ_T)	22, 33, 44, 55 (Mbps)
2	Fragment Size (bits)	512, 2048, 8192, 4096, 10240, 12288, 14336, 16384
3	Number of Frames / CTA	1-5

In Figure 4.4 and Figure 4.5, the values of time efficiency and CTA size are plotted against variables values of fragment size and number of frames/ CTA. An important point to note in Figure 4.4 is that above 3 frames/CTA, the time efficiency do not show any significant increase after subsequent increase in number of frames/CTA. However, in Figure 4.5, the case is different for CTA size. A constant increase in CTA size can be seen by increasing the number of frames/ CTA. The reason is by using a larger fragment size, the actual data bits outnumber the overhead bits by a larger margin. As a result, higher values of time efficiency are obtained. Increasing either the number of frames / CTA or the

fragment size further do not increase the time efficiency by a substantial margin but increases the CTA size significantly. Therefore, in cases where the minimum throughput value is satisfied for larger values of fragment sizes, a reasonable value of time efficiency can be assumed and further increase in time efficiency value can cause capacity issues.

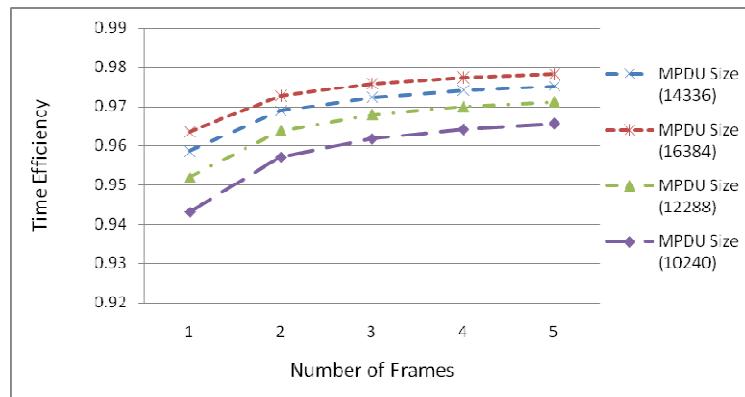


Figure 4.4 Time Efficiency values with varying Fragment Size and number of frames/CTA

When the value of δ_T is kept constant along with the number of frames / CTA, the corresponding change in the fragment size is used to calculate the percentage increase in time efficiency. Therefore, the data rate is kept at 22 Mbps, the number of frames/ CTA is kept to 1 and the fragment sizes used are 4096, 10240, 12288, 14336 and 16384. It is observed that when the fragment size is changed from 4096 to 10240, the percentage increase in the time efficiency is 7.86%. The subsequent changes in fragment sizes result in the mean percentage increase of 0.71% which is substantially smaller than the increase of fragment size from 4096 to 10240. For the same values, when the CTA size is considered, there is a percentage increase of 56.5% when the fragment size is changed from 4096 to 10240 and the subsequent changes in fragment size up to 16384 results in a mean percentage increase of 13.86%. Therefore, it can be concluded that the percentage increase in time efficiency and CTA size depends on the difference in size of the previous fragment size and the next higher fragment size.

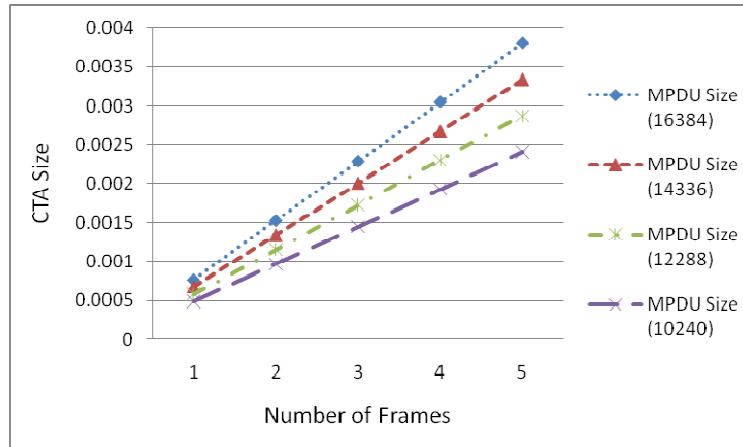


Figure 4.5 CTA Size with varying Fragment Size and number of frames/CTA

In another scenario shown in Figure 4.6, the fragment size is kept constant at 16384 and the number of frames / CTA and the value of ∂_T is varied to check the effect of transmission data rate on time efficiency. For the same scenario Figure 4.7 shows the variation in CTA size with the change in transmission data rates. It can be seen in Figure 4.6 that when the number of frames/CTA changes from 1 frame/ CTA to 2 frames/ CTA, there is a somewhat greater increase in the time efficiency when compared with higher values of frames/ CTA. Also, with the increase in the value of ∂_T , the time efficiency decreases. The reason is the relatively reduced time to send an MPDU with the IFS duration being constant, which increases the overhead. The percentage overhead per flow can easily be calculated by

$$(1 - \eta_{eff}) * 100 \text{ or } (1 - \partial_E/\partial_A) * 100 \quad (4.8)$$

For a fragment size of 4096 and 1 frame/CTA, the percentage flow overheads based on (4.8) for ∂_T values of 22, 33, 44 and 55 Mbps are 13.1, 15.9, 18.6 and 21.1% respectively. For a larger fragment size of 16384 and 1 frame/ CTA, the percentage flow overheads are 3.6, 4.5, 5.4 and 6.29% respectively. In Figure 4.7, the reduction in CTA size can be noticed with the corresponding increase in value of ∂_T . The percentage decrease in the CTA size with the increase in value of ∂_T is 32.7, 24.3 and 19.2% respectively for 5 frames/ CTA and remains the same for 1-4 frames/ CTA.

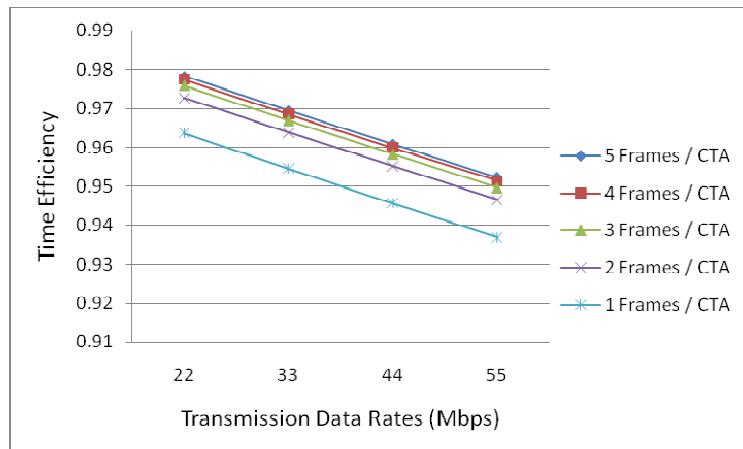


Figure 4.6 CTA Size with varying Fragment Size and number of frames/CTA

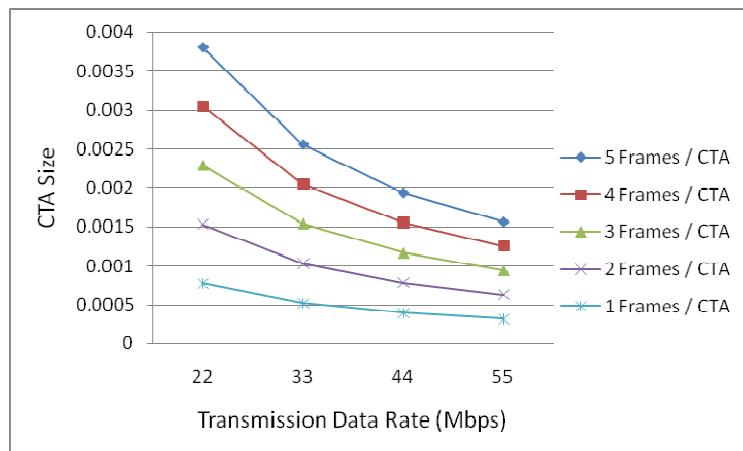


Figure 4.7 CTA Size with varying transmission data rate and number of frames/CTA

For smaller values of fragment size i.e. 512, 2048 and 8192, the percentage overhead is much greater when compared to the larger values of fragment size. For the fragment size of 512 and 1 frame/ CTA, the percentage overhead for transmission data rates of 22, 33, 44 and 55 Mbps is 47.27, 54.54, 60 and 64.3 % respectively. For the same parameters but fragment size of 8192, the percentage overhead is 7, 8.68, 10.29 and 11.84. A summary of values for time efficiency, CTA size and percentage overhead for different values of parameters in Table 4.1 is given in Table 4.2, Table 4.3, Table 4.4 and Table 4.5.

Based on the analysis, the key observations are as follows:

1. The time efficiency is much lower for smaller fragment sizes than larger fragment sizes

2. With the increase in value of ∂_T , there is less change in value of time efficiency for larger fragment sizes than smaller fragment sizes.
3. For larger fragment sizes with 2 or more frames/CTA, there is minimum change in the value of time efficiency but a substantial change in the CTA size. Therefore, when larger fragment size is used for a flow and the minimum throughput requirement is met, increasing the time efficiency can cause capacity constraints.

The percentage overhead per flow is higher when a smaller fragment size is used and lower when a higher fragment size is used.

Table 4.2: Summary of values based on smaller fragment size's and constant transmission data rate of 22 Mbps

Fragment Size (Bits)	Number of Frames/CTA	∂_T (Mbps)	η_{eff}	% Overhead	CTA Duration (η_{CTA})	∂_E (Mbps)	∂_A (Mbps)
512	2	22	0.5272	47.27	8.8E-05	0.12	0.23
	3		0.5577	44.22	0.00012	0.18	0.33
	4		0.5743	42.56	0.00016	0.25	0.43
2048	1		0.7681	23.18	0.00012	0.25	0.32
	2		0.8169	18.30	0.00022	0.50	0.61
8192	1		0.9298	7.01	0.00040	1	1.07
	2		0.9469	5.30	0.00078	2	2.11

Table 4.3: Summary of values based on smaller fragment size's and transmission data rates of 33, 44 and 55 Mbps

Fragment Size (Bits)	Number of Frames/CTA	33 (Mbps)			44 (Mbps)			55 (Mbps)		
		η_{eff}	% Overhead	CTA Duration (η_{CTA})	η_{eff}	% Overhead	CTA Duration (η_{CTA})	η_{eff}	% Overhead	CTA Duration (η_{CTA})
512	2	0.4545	54.54	6.8E-05	0.3993	60.06	5.8E-05	0.3561	64.38	5.2E-05
	3	0.4777	52.22	9.7E-05	0.4178	58.21	8.3E-05	0.3712	62.87	7.5E-05
	4	0.4903	50.96	0.0001	0.4277	57.22	0.0001	0.3793	62.06	9.8E-05
2048	1	0.7244	27.55	8.5666	0.6854	31.45	6.7E-05	0.6503	34.96	5.7E-05
	2	0.7692	23.07	0.0001	0.7267	27.31	0.0001	0.6887	31.12	0.0001
8192	1	0.9131	8.68	0.0002	0.8970	10.29	0.0002	0.8815	11.84	0.0001
	2	0.9302	6.97	0.0005	0.9140	8.59	0.0004	0.8984	10.15	0.0003

Table 4.4: Summary of values based on larger fragment size's and constant transmission data rate of 22 Mbps

Fragment Size (Bits)	Number of Frames/CTA	δ_T (Mbps)	η_{eff}	% Overhead	CTA Duration (η_{CTA})	δ_E (Mbps)	δ_A (Mbps)
10240	1	22	0.9430	5.691	0.00049	1.25	1.32
	2		0.9570	4.290	0.00097	2.50	2.61
12288	1	22	0.9521	4.788	0.00058	1.50	1.57
	2		0.9639	3.600	0.00115	3.00	3.11
14336	1	22	0.9586	4.132	0.00067	1.75	1.82
	2		0.9689	3.102	0.00134	3.50	3.61
	3		0.9724	2.754	0.00201	5.25	5.39
	4		0.9742	2.578	0.00267	7.00	7.18
	5		0.9752	2.473	0.00334	8.75	8.97
16384	1	22	0.9636	3.634	0.00077	2.00	2.07
	2		0.9727	2.725	0.00153	4.00	4.11
	3		0.9758	2.418	0.00228	6.00	6.14
	4		0.9773	2.263	0.00304	8.00	8.18
	5		0.9782	2.171	0.00380	10.00	10.22

Table 4.5: Summary of values based on larger fragment size's and transmission data rates of 33, 44 and 55 Mbps

Fragment Size (Bits)	Number of Frames/CTA	33 (Mbps)			44 (Mbps)			55 (Mbps)		
		η_{eff}	% Overhead	CTA Duration (η_{CTA})	η_{eff}	% Overhead	CTA Duration (η_{CTA})	η_{eff}	% Overhead	CTA Duration (η_{CTA})
10240	1	0.9293	7.06	0.0003	0.9159	8.40	0.0002	0.9029	9.70	0.0002
	2	0.9433	5.66	0.0006	0.9300	6.99	0.0005	0.9171	8.28	0.0004
12288	1	0.9403	5.96	0.0003	0.9289	7.10	0.0003	0.9177	8.22	0.0002
	2	0.9523	4.76	0.0007	0.9410	5.89	0.0005	0.9299	7.00	0.0004
14336	1	0.9484	5.15	0.0004	0.9384	6.15	0.0003	0.9286	7.13	0.0002
	2	0.9588	4.11	0.0009	0.9490	5.09	0.0006	0.9393	6.06	0.0005
	3	0.9624	3.75	0.0013	0.9526	4.73	0.0010	0.9429	5.70	0.0008
	4	0.9642	3.57	0.0018	0.9543	4.56	0.0013	0.9447	5.52	0.0011
	5	0.9652	3.47	0.0022	0.9554	4.45	0.0017	0.9458	5.41	0.0013
16384	1	0.9546	4.53	0.0005	0.9457	5.42	0.0003	0.9370	6.29	0.0003
	2	0.9638	3.61	0.0010	0.9551	4.48	0.0007	0.9465	5.34	0.0006
	3	0.9669	3.30	0.0015	0.9582	4.17	0.0011	0.9497	5.02	0.0009
	4	0.9685	3.14	0.0020	0.9598	4.01	0.0015	0.9513	4.86	0.0012
	5	0.9694	3.05	0.0025	0.9608	3.91	0.0019	0.9523	4.76	0.0015

4.3.3 Algorithms for Selection of Parameters based on Time Efficiency

The results from time efficiency analysis suggest that for those cases in which sufficient superframe capacity is available and a flow can benefit from higher throughput values, then a suitable strategy is to increase the time efficiency of the flow. If there is not sufficient superframe capacity available, then apart from considering the minimum throughput requirement of a flow, the maximum CTA size is an important constraint which has to be considered when deciding about the values of a suitable fragment size, number of frames/

CTA and the transmission data rate. The decision related to selection of suitable parameters for a flow is also influenced strongly by the type of traffic intended for the flow. In case of CBR traffic, increasing the throughput above a certain value does not benefit the flow but can occupy unnecessary fraction of superframe capacity. For VBR flows, the throughput is variable and the decision to select suitable parameters for the flow might vary from superframe to superframe depending on the buffer size of the source device and the priority of frames. The selection of most appropriate parameters for a flow therefore requires a minimum of two constraints i.e. minimum throughput required for the flow and the maximum CTA size which can be occupied for the flow. Based on these two constraints, a device based on its traffic requirements can decide to have one of the following objectives

- Maximize throughput ($\max(\delta_E)$)
- Maximize the time efficiency for its flow ($\max(\eta_{eff})$)
- Minimize CTA size ($\min(\eta_{CTA})$)

The first and second options can be selected when the superframe has enough capacity and there is no congestion. The third option can be considered when the superframe capacity is congested. In this section we propose two search based algorithms which take into account the throughput and capacity constraints and can be used to maximize the throughput, minimize the CTA size or maximize the time efficiency of a flow given the minimum throughput and maximum CTA size as constraints. The proposed search based algorithms can be used for real time as well as non real time flows to select suitable parameters for a flow based on its traffic requirements. The scheduling schemes proposed for HDR WPANs can also make use of the proposed algorithms in order to select appropriate values of fragment size, transmission data rate and number of frames/CTA especially in case of VBR traffic. To the best of knowledge based on the literature review done so far related to HDR WPANs, there is no such algorithm to dynamically select appropriate parameters for a flow based on its traffic requirements. For convenience, the first algorithm is called the *Full Table Search (FTS)* algorithm while the second algorithm is called the *Partial Table Search (PTS)* algorithm.

In both of the algorithms i.e. the FTS and the PTS, three sets T_x, X_d and \mathcal{F}_r are considered for parameter selection in which the values of $\delta_T \in T_x$, $x_d \in X_d$ and $\sigma_{MPDU} \in \mathcal{F}_r$. From (4.2),

(4.4) and (4.5), it is clear that the values of effective throughput (∂_E), CTA size (η_{CTA}) and time efficiency (η_{eff}) are calculated from a combination of ∂_T , x_d and σ_{MPDU} . Therefore, if the tuple $(\partial_T, x_d, \sigma_{MPDU})$ represents an ordered pair with each element corresponding to a value from the sets T_x, X_d and \mathcal{F}_r respectively, then the values from each ordered pair can be used to calculate a unique combination of ∂_E , η_{CTA} and η_{eff} . If the number of elements for T_x, X_d and \mathcal{F}_r is l, r and q respectively, then the size of the table or the total number of elements in the table is $l * r * q$. Each element corresponds to the i^{th} tuple $(\partial_T^i, x_d^i, \sigma_{MPDU}^i)$. In both of the FTS and PTS algorithms, a set of m feasible tuples is selected from the table which satisfies the given constraints. From the set of m feasible tuples, the final tuple $(\partial_T, x_d, \sigma_{MPDU})$ which corresponds to the values of ∂_E , η_{CTA} and η_{eff} that satisfies the objective function is selected. A generic version of the problem and its constraints can be given as

$$\min(\eta_{CTA}), \max(\partial_E) \text{ or } \max(\eta_{eff}) \quad (4.9)$$

Such that

$$\eta_{eff} > 0, \partial_E > 0, \eta_{CTA} > 0 \quad (4.9a)$$

$$\eta_{eff} > \eta_{eff}^{min} \text{ (optional)} \quad (4.9b)$$

$$x_d \times \sigma_{MPDU} \geq \sigma_{CTA} \text{ (Required Number of bits per CTA)} \quad (4.9c)$$

$$\text{Where } \sigma_{CTA} = \partial_R * \frac{1}{\eta_s} * 1/N_{CPS}$$

$$\partial_E \leq \partial_E + \partial_E * \varepsilon \quad (4.9d)$$

$$\partial_E \geq \partial_R \text{ (minimum required throughput)} \quad (4.9e)$$

$$\eta_{CTA} \leq \eta_{CTA}^{max} \quad (4.9f)$$

In the problem description, (4.9) is the objective function and the rest are constraints. In (4.9d), the value of ε represents, the percentage error or deviation of ∂_E from the value of ∂_R . This is useful in cases when there are many tuples which satisfy the value of ∂_R but at the same time $\partial_E \gg \partial_R$ is also true. Therefore, to avoid unnecessary values to be selected, the value of ε proves to be useful. It also helps to reduce the extra number of iterations by an amount $(m_o - m)$ where m_o is the number of feasible values when the limit of ε is not

used. Based on the objective function and identification of relevant constraints, the FTS and PTS algorithms select the most appropriate tuple in three stages. In the first stage, the table is filled based on the available number of elements of T_x , X_d and \mathcal{F}_r . In the second stage, m feasible values or tuples are selected. In the third stage, the final tuple is selected based on the objective function from the m feasible tuples. The pseudo-code for the FTS algorithm is given in Figure 4.8.

Full Table Search (FTS) Algorithm pseudo-code

```

//Initialize
 $T_x = \{\partial_T^1, \partial_T^2, \dots, \partial_T^L\}$ 
 $X_d = \{x_d^1, x_d^2, \dots, x_d^r\}$ 
 $\mathcal{F}_r = \{\sigma_{MPDU}^1, \sigma_{MPDU}^2, \dots, \sigma_{MPDU}^q\}$ 
 $g[20], h[20], k[20];$ 
 $ctr = 0;$ 
 $max\_error = \varepsilon;$ 

//Load Tables
for (g = 0; g < l; g++)
for (h = 0; h < r; h++)
for (k = 0; k < q; k++)
{
     $\eta_{CTA}[g][h][k] = \frac{(\sigma_{MPDU}^k * x_d^h)}{\partial_T^g} + (x_d^h * \eta_{SIFS}) + \eta_{GT} + (x_d^h * \eta_{MAC\_HDR});$ 
     $\partial_E[g][h][k] = (\eta_{CTA} - ((x_d^h * \eta_{SIFS}) + \eta_{GT}) - (x_d^h * \eta_{MAC\_HDR}) * \partial_T^g * N_{CPS} * 1/\eta_s);$ 
     $\eta_{eff}[g][h][k] = \frac{(x_d^h * \eta_{MPDU})}{(x_d^h * \eta_{MPDU}) + (x_d^h * \eta_{MAC\_HDR}) + (x_d^h * \eta_{SIFS}) + \eta_{GT}};$ 
}

//Get Constraints and Objective Function
get minimum Throughput ( $\partial_R$ );
get maximum CTA size ( $\eta_{CTA}^{max}$ );
get minimum time efficiency ( $\eta_{eff}^{min}$ );
get selection priority for objective function
i.e. 1 -> max ( $\partial_E$ ), 2 -> min ( $\eta_{CTA}$ ), 3 -> max ( $\eta_{eff}$ )

//Get feasible values
ctr = 0; //feasible values counter
for (g = 0; g < l; g++)
for (h = 0; h < r; h++)
for (k = 0; k < q; k++)
{
    if(( $\partial_E[g][h][k] \geq \partial_R$ ) && ( $\partial_E[g][h][k] < (\partial_R + \partial_R * max\_error)$ ))
    {
        if( $\eta_{CTA}[g][h][k] < maximum CTA size$ )
        {
            //store values of  $\partial_E[g][h][k]$ ,  $\eta_{CTA}[g][h][k]$  and  $\eta_{eff}[g][h][k]$ 
            g[ctr]=g;
            h[ctr]=h;
        }
    }
}

```

```

        k[ctr]=k;
        ctr++;
    }
}

//Select Final Values
if (max ( $\partial_E$ ))
{
for (i = 0; i < ctr; i++)
{
    Check for maximum value of  $\partial_E$  from ctr feasible values;
    Select the maximum value of  $\partial_E$ ;
}
Select the corresponding values of  $\partial_T^g$ ,  $x_d^h$  and  $\sigma_{MPDU}^k$ 
}
if (min ( $\eta_{CTA}$ ))
{
for (i = 0; i < ctr; i++)
{
    Check for minimum value of  $\eta_{CTA}$  from ctr feasible values;
    Select the minimum value of  $\eta_{CTA}$ ;
}
Select the corresponding values of  $\partial_T^g$ ,  $x_d^h$  and  $\sigma_{MPDU}^k$ 
}
if (max ( $\eta_{eff}$ ))
{
for (i = 0; i < ctr; i++)
{
    Check for maximum value of  $\eta_{eff}$  from ctr feasible values;
    Select the maximum value of  $\eta_{eff}$ ;
}
Select the corresponding values of  $\partial_T^g$ ,  $x_d^h$  and  $\sigma_{MPDU}^k$ 
}
end;

```

Figure 4.8 Pseudo-Code for the Full Table Search (FTS) Algorithm

The complexity of FTS algorithm varies based on the value of number of iterations n used. In the proposed approach for FTS algorithm, the number of elements in the transmission data rate set T_x is 4 i.e. $n(T_x) = 4$. Similarly $n(X_d) = 4$ and $n(F_r) = 8$. Therefore, in the proposed case, the total number of iterations required to load the table are $n * n * 2n = 2n^3$. The worst case complexity for the complete FTS algorithm can be given as $\mathcal{O}(4n^3 + m)$, where m is the number of feasible values which satisfies the given constraints. If $n \gg m$, then $\mathcal{O}(4n^3)$ is the worst case complexity. Based on the values given in Table 4.1, the value of n can be bounded i.e. $1 < n < 8$. In many practical scenarios, a device might not be able to support all the values of $\partial_T \in T_x$ or $\sigma_{MPDU} \in F_r$ because of its distance from the destination device and the traffic type of the flow respectively. In case the value of ∂_T is fixed for a

device, then the $n(T_x) = 1$ and the complexity of the algorithm reduces to a second degree polynomial $\mathcal{O}(4n^2 + m)$. If the values of both ∂_T and σ_{MPDU} are fixed for the device, then the complexity further reduces to $\mathcal{O}(4n + m)$. The complexity of the algorithm is therefore variable and depends on

- the distance of the device from its destination which determines the number of elements of set T_x
- the traffic type for the flow i.e. CBR, VBR which determines the number of elements of set F_r
- the number of feasible values m which is determined by the parameter value of *max_error* (ϵ) used in the FTS and PTS algorithms

The difference between the FTS and PTS algorithms lies in terms of complexity. The results returned by both algorithms are the same. The PTS algorithm in concept is the same as the FTS algorithm. The difference is in the second stage of operation in which m feasible tuples are selected. In FTS, there is a full table search in order to select the feasible values, however, in PTS (as the name suggests), there is a partial table search to reduce the number of iterations. In order to achieve reduced complexity while achieving the same results as FTS, the full table is partitioned into sub-tables. If F_T is the set consisting of $l * r * q$ ordered pairs or tuples, then in terms of sub-tables, F_T is presented as $F_T = \{S_1 \cup S_2 \cup S_3 \cup \dots \cup S_6\}$. Each sub-table S_i is a collection of tuples which consists of a number of tuples for which the ∂_E values fall in a certain range. In the proposed case, $S_1 < 1$ Mbps, 1 Mbps $< S_2 < 2$ Mbps, 2 Mbps $< S_3 < 3$ Mbps, 3 Mbps $< S_4 < 4$ Mbps, 4 Mbps $< S_5 < 5$ Mbps and 5 Mbps $< S_6 < 6$ Mbps. However, for other cases which require higher throughputs, further sub-tables can be added. In the first stage, when the table is being filled, the throughput value of each tuple is checked and based on the range it falls in, it is allocated to each sub-table. Each sub-table has a counter, associated with it which is incremented when a new entry is placed in that table.

When the algorithm enters the second stage in order to select the feasible tuples, then instead of searching the full table, only the relevant sub-table in which the required throughput range falls in, is searched. If the number of entries based on a counter for the i^{th} sub-table S_i is m_i , then the worst case complexity of the PTS algorithm can be given by $\mathcal{O}(2n^3 + m_i + m)$ where $m_i \geq m$. In case if $n \ll m_i$, then the worst case complexity is $\mathcal{O}(2n^3)$

which is clearly better than the worst case complexity of FTS algorithm which is $\mathcal{O}(4n^3)$. The pseudo-code for the PTS algorithm is given in Figure 4.9.

Partial Table Search (PTS) Algorithm pseudo-code

```

//Initialize
 $T_x = \{\partial_T^1, \partial_T^2, \dots, \partial_T^l\}$ 
 $X_d = \{x_d^1, x_d^2, \dots, x_d^r\}$ 
 $F_r = \{\sigma_{MPDU}^1, \sigma_{MPDU}^2, \dots, \sigma_{MPDU}^q\}$ 
 $ctr = 0; m_1, \dots, m_6 = 0;$ 
 $g[20], h[20], k[20];$ 
 $max\_error = \varepsilon;$ 

//Load Tables
for (g = 0; g < l; g++)
for (h = 0; h < r; h++)
for (k = 0; k < q; k++)
{
     $\eta_{CTA}[g][h][k] = \frac{(\sigma_{MPDU}^k * x_d^h)}{\partial_T^g} + (x_d^h * \eta_{SIFS}) + \eta_{GT} + (x_d^h * \eta_{MAC\_HDR});$ 
     $\partial_E[g][h][k] = (\eta_{CTA} - ((x_d^h * \eta_{SIFS}) + \eta_{GT}) - (x_d^h * \eta_{MAC_{HDR}})) * \partial_T^g * N_{CPS} * 1/\eta_s);$ 
     $\eta_{eff}[g][h][k] = \frac{(x_d^h * \eta_{MPDU})}{(x_d^h * \eta_{MPDU}) + (x_d^h * \eta_{MAC_{HDR}}) + (x_d^h * \eta_{SIFS}) + \eta_{GT}};$ 
}

if (check range 1)
{
     $S_1 \leftarrow \partial_E[g][h][k];$ 
     $g[m_1] = g;$ 
     $h[m_1] = h;$ 
     $k[m_1] = k;$ 
     $m_1++;$ 
}
.
.
.

if (check range 6)
{
     $S_6 \leftarrow \partial_E[g][h][k];$ 
     $g[m_6] = g;$ 
     $h[m_6] = h;$ 
     $k[m_6] = k;$ 
     $m_6++;$ 
}

//Get Constraints and Objective Function
get minimum Throughput ( $\partial_R$ );
get maximum CTA size ( $\eta_{CTA}^{max}$ );
get minimum time efficiency ( $\eta_{eff}^{min}$ );
get selection priority for objective function
i.e. 1 -> max ( $\partial_E$ ), 2 -> min ( $\eta_{CTA}$ ), 3 -> max ( $\eta_{eff}$ )

//Get feasible values

```

```

ctr = 0; //feasible values counter
Based on value of  $\partial_R$ , select the appropriate sub-table  $S_i$ 
for (i = 0; i < mi; i++)
{
if(( $\partial_E[g[i]][h[i]][k[i]] \geq \partial_R$ ) && ( $\partial_E[g[i]][h[i]][k[i]] < (\partial_R + \partial_R * max\_error)$ ))
{
    if( $\eta_{CTA}[g[i]][h[i]][k[i]] < maximum\ CTA\ size$ )
    {
        //store values of  $\partial_E[g][h][k]$ ,  $\eta_{CTA}[g][h][k]$  and  $\eta_{eff}[g][h][k]$ 
        g[ctr]= g[i];
        h[ctr]=h[i];
        k[ctr]=k[i];
        ctr++;
    }
}
//Select Final Values
if (max ( $\partial_E$ ))
{
    for (i = 0; i < ctr; i++)
    {
        Check for maximum value of  $\partial_E$  from ctr feasible values;
        Select the maximum value of  $\partial_E$ ;
    }
    Select the corresponding values of  $\partial_T^g$ ,  $x_d^h$  and  $\sigma_{MPDU}^k$ 
}

if (min ( $\eta_{CTA}$ ))
{
    for (i = 0; i < ctr; i++)
    {
        Check for minimum value of  $\eta_{CTA}$  from ctr feasible values;
        Select the minimum value of  $\eta_{CTA}$ ;
    }
    Select the corresponding values of  $\partial_T^g$ ,  $x_d^h$  and  $\sigma_{MPDU}^k$ 
}
if (max ( $\eta_{eff}$ ))
{
    for (i = 0; i < ctr; i++)
    {
        Check for maximum value of  $\eta_{eff}$  from ctr feasible values;
        Select the maximum value of  $\eta_{eff}$ ;
    }
    Select the corresponding values of  $\partial_T^g$ ,  $x_d^h$  and  $\sigma_{MPDU}^k$ 
}
end;

```

Figure 4.9 Pseudo-Code for the Partial Table Search (PTS) Algorithm

The results based on a few test scenarios for FTS are given in Table 4.6. Since the two algorithms differ in terms of complexity only and not on the basis of output, therefore the results in Table 4.6 represent the outcomes based on FTS only.

Table 4.6: Selected Parameters based on FTS algorithm

Scenario	Min(∂_E) Mbps	Max (η_{CTA}) seconds	Min (η_{eff})	Objective	σ_{MPDU}	x_d	∂_T Mbps	(ε)	(∂_E) Mbps	η_{eff}	η_{CTA} seconds
1	1	0.0004	-	$\min(\eta_{CTA})$	8192	1	55	0.2	1.049	0.910	0.000168
2	2	0.0004	-	$\min(\eta_{CTA})$	16384	1	55	0.2	2.049	0.953	0.000317
3	2	0.0008	-	$\max(\partial_E)$	16384	1	55	0.2	2.049	0.953	0.000317
4	2	0.0008	0.8	$\max(\partial_E)$	10240	2	55	0.4	2.549	0.929	0.000406
5	3	0.001	0.6	$\max(\eta_{eff})$	14336	2	33	0.4	3.544	0.968	0.000906
6	3	0.002	0.6	$\max(\eta_{eff})$	14336	2	22	0.2	3.541	0.978	0.001345
7	4	0.00103	0.8	$\min(\eta_{CTA})$	16384	2	55	0.2	4.049	0.954	0.000629
8	4	0.00103	0.9	$\max(\partial_E)$	12288	3	55	0.2	4.549	0.941	0.000717
9	4	0.00103	0.9	$\max(\partial_E)$	14336	3	55	0.4	5.299	0.949	0.000829
10	5	0.002	0.8	$\min(\eta_{CTA})$	10240	4	55	0.4	5.049	0.930	0.000805

4.4 Proposed Capacity Efficient and Delay Efficient Metrics

From section 4.2, it is evident that capacity efficiency and EED are two important issues to consider in meshed HDR WPANs. The proposed solution in this section consists of two metrics i.e. STS and STS-R which take into consideration the capacity efficiency of a flow in each link as well as the total EED of a flow based on mainly the queuing delay at each hop and for the full path. The metrics STS and STS-R are proposed for meshed WPANs but they can also find significant use in sensor based mesh networks which have a TDMA based MAC. Although most of the research done related to sensor networks is related to energy savings with some of the notable approaches given in [5], [6], [7], [8], [9] and [10], there are also approaches that are aimed at EED reduction [1], [11] and [12]. In [11], the proposed approach is called DMAC which is intended for data gathering trees in sensor networks.

Some of the main characteristics of DMAC consist of

- Only one packet per slot is assumed to be transmitted
- The packets are considered to have the same length
- The schedule of slots is sequenced from the child to the root device without considering any congestion
- The receiving and sending time slot duration is taken to be the same which might not be true because the transmitter data rate of the transmitter for the previous hop device and the next hop device might be different

In [12], the proposed approach has the following characteristics;

- Variable length TDMA frames are considered
- The total delay considered consists of Transmission delay and Queuing delay

- The size of the slot remains the same for each link
- Each link is scheduled independently and the effect of the schedule is not considered on the next hop

In [1], a routing metric called Minimized Slot Misordered Routing (MSMR) is proposed which selects the path which has the minimum delay due to slot misordering of devices in a multihop TDMA network. The metric prefers those paths in which the order of time slots is in alignment with the sequence of relay devices from source to destination. However, the metric does not mitigate the effect of slot misordering and neither does it re-order the slots to reduce the delay. The proposed approach with STS and STS-R has the following characteristics

- Each slot or CTA can have one or more than one frames
- The receiving and sending time slot might be or might not be the same size because of variation in transmitter data rate
- Fixed size superframe is considered
- The total delay considered consists of transmission delay, propagation delay and Queuing delay
- In STS, the effect of slot misordering can be mitigated by CTA switching
- In STS-R, there is a probability that the EED is further reduced after CTA switching if there are some intra-PAN or inter-PAN flows that terminate during the life time of the flow in consideration
- Both STS and STS-R not only consider the EED but also take into account the capacity efficiency of a path in formulating the total cost of that path

The delay which constitutes the major portion of EED is the queuing delay in meshed HDR WPANS and therefore, the focus of STS and STS-R metrics is to mitigate the effects of slot misordering as well as selecting those paths which are not only capacity efficient but have a lower EED value when compared with others. An example of queuing delay can be seen in Figure 4.10 in which t_{d1} and t_{d2} are the queuing delays for the three hop path. If for an n hop path, we have the total delay $T_d = \sum_{i=1}^n t_{di}$, then if $T_d > 0.065535$ seconds, then that means that the EED is increased by one superframe duration. Furthermore, if for any hop m , if the transmitting CTA to the next hop relay devices comes before or to the left of the

receiving CTA from a previous hop device, the EED is incremented by one superframe duration. It is therefore imperative to devise a strategy to alleviate such misordering as much as possible in case it cannot be eliminated completely hence the need for STS metric. The subsequent sections describe the various attributes relevant to the design of STS and STS-R metrics and their effect on the capacity efficiency of a path and the EED. Furthermore, the main specifications for STS and STS-R metrics are identified and their performance is evaluated and compared with MSMR.

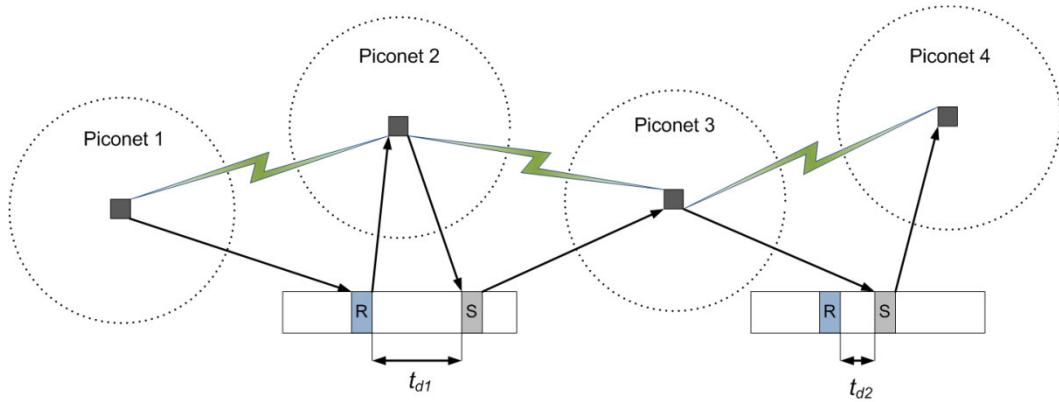


Figure 4.10 Example of queuing delay along a source-destination path

4.4.1 Initial Analysis

In this section, an effort is made to present some assumptions and relevant analysis pertinent to the proposed STS and STS-R metrics. When multi-hop communication is considered in the case of meshed HDR WPANs, then the superframe capacity used in each two hop neighbourhood consists of inter-PAN as well as intra-PAN flow reservations. The queuing delay in each hop is only relevant to inter-PAN flows because for intra-PAN flows, the communication between the source and destination device is single hop and is in the same cluster. Therefore, the position of time slots or CTAs for intra-PAN flows has no impact on the flow attributes. However, the CTA position for inter-PAN flows has important implications on the EED value in terms of queuing delay. In order to elaborate upon this, consider Figure 4.11. In Figure 4.11, the CTA numbered 1 indicates the CTA in which MPNC 1 transmits to MPNC2, CTA 2 indicates the CTA in which MPNC2 transmits to MPNC3 and so on. If the CTAs are arranged out of order like the case presented in Figure 4.11(a), then MPNC3 transmits to MPNC 4 in a CTA which is positioned after the CTA reserved for MPNC4 to transmit to MPNC5 or the destination MPNC. Therefore, unless MPNC4 receives the data

from MPNC3 in CTA 3, it cannot forward that data in CTA 4 in the same superframe. Therefore, when MPNC4 receives from MPNC3 in a superframe numbered x , it can only forward the data in that CTA to MPNC5 in superframe $x + 1$. If the CTAs are arranged in order just like shown in Figure 4.11(b), then the EED is reduced by one complete superframe duration which in the proposed case is 0.065535 seconds. In Figure 4.11(c), the intra-PAN CTA reservations are shown for the link between MPNC2 and MPNC3 as well as for the link between MPNC3 and MPNC4. If the inter-PAN CTAs are shifted to the left of the intra-PAN CTA reservations (shown in Figure 4.11(d)), then the EED is reduced by the sum of two intra-PAN CTA reservations. Such shifting of CTAs has no impact on the delay of intra-PAN flows, since they are single hop. Figure 4.11(e) shows an ideal case in which it is assumed that there are no other inter-PAN CTA reservations and the CTAs are closely spaced in time one after the other which reduces the queuing delay to a minimum.

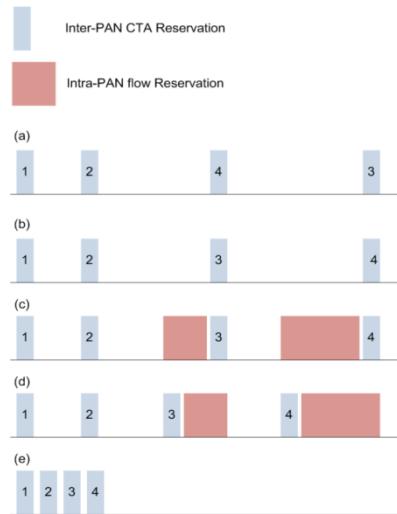


Figure 4.11 General Queuing considerations for Inter-PAN Flows

The queuing delay due to the time offset between CTAs reserved for the same flow is considered from the start of CTA reserved for transmission by MPNC_i to MPNC_{i+1} to the start of the CTA reserved for MPNC_{i+1} to transmit. An alternative approach in which the queuing delay is taken to be from the end of CTA reserved for transmission by MPNC_i to MPNC_{i+1} to the start of the CTA reserved for MPNC_{i+1} to transmit is not considered because the start of the CTA offset does not depend on the transmitter data rate for each hop while the end of the CTA can vary from hop to hop based on the transmitter data rate which might vary from hop to hop. To clarify this issue consider Figure 4.12.

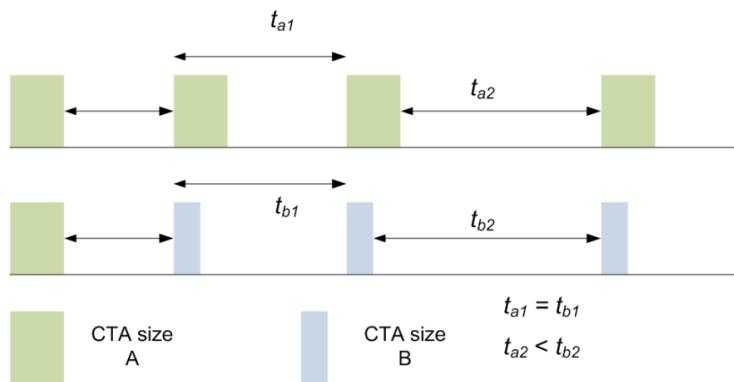


Figure 4.12 Queuing delay variation based on CTA offset

In Figure 4.12, CTA size A corresponds to a transmitter data rate of ∂_T^A while the CTA size B corresponds to transmitter data rate value of ∂_T^B such that $\partial_T^B > \partial_T^A$ and hence the smaller size of CTA B. If the queuing delay t_{a2} is considered from the end of CTA A to the start of next hop CTA A and queuing delay t_{b2} is considered from the end of CTA B to the start of next hop CTA B then $t_{b2} > t_{a2}$. However, if the queuing delay t_{a1} is considered from the start of CTA A to the start of next hop CTA A and the same for queuing delay t_{b1} for CTA B, then $t_{b1} = t_{a1}$. In other words, if the queuing delay for hop m is calculated from end of CTA_i of flow f to the start of CTA_{i+1} for the same flow along the multihop path, then in such a case those paths which have the same start offsets for CTAs reserved for flow f but with higher transmitter data rates are penalized due to higher link costs based on queuing delay even though the EED is the same or less for paths with higher transmitter data rates. The capacity efficiency of those paths which have higher transmitter data rates is also higher for a flow f . In order to clarify this issue consider Figure 4.13.

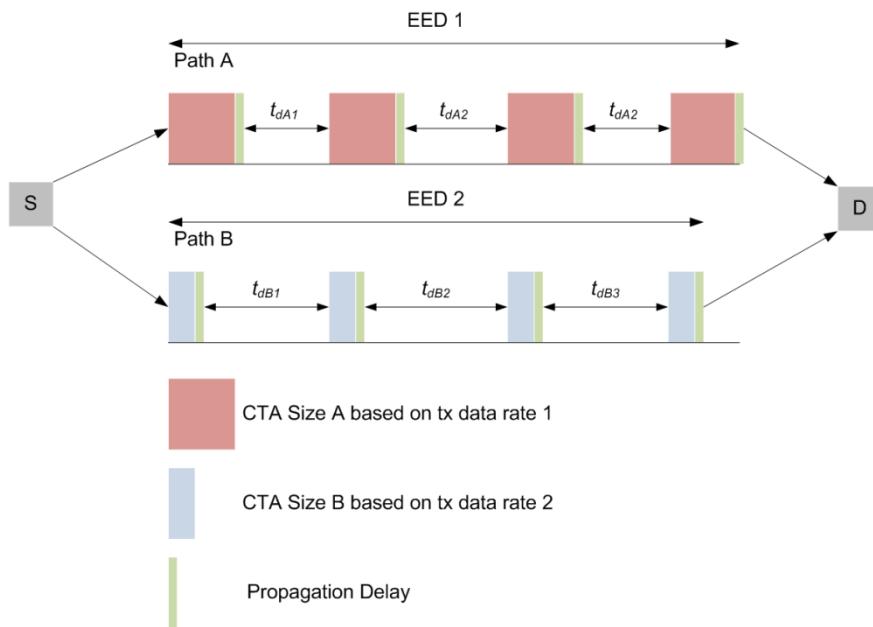


Figure 4.13 Queuing Delay and EED Based on CTA offsets

In Figure 4.13, two paths are shown on which a certain flow f can be supported from the source S to the destination D . For path A, the CTA size is larger because the transmitter data rate is lower throughout the path when compared with path B. It can be noted from Figure 4.13 that the queuing delay for each hop is lesser in path A in comparison with path B. The CTA start offsets are kept the same for both paths. Even though the total sum of queuing delays for path B is greater than the total sum of queuing delays for path A i.e. $\sum_{i=1}^n t_{dB_i} > \sum_{i=1}^n t_{dA_i}$, the total EED for path B is lower than path A. The difference in EED is only because of the last hop in which the transmission time for the last CTA is lower for path B because of smaller CTA size due to higher transmitter data rate. Since the CTA size for path B is also smaller than path A, it is more capacity efficient than path A. Therefore, in the proposed approach, the queuing delay is taken from the start of CTA at hop m for flow f to the start of CTA at hop $m + 1$ for the same flow.

If an ideal case is assumed in which only one inter-PAN flow is considered without any intra-PAN flow reservations along the source-destination path, and the CTAs for the inter-PAN flow are closely spaced one after the other, then not only the EED is lesser for the path which has higher values of δ_T per hop when compared with other paths with comparatively lower values of δ_T per hop, but the capacity efficiency is higher as well. The idea is illustrated

using three cases in Figure 4.14 in which the CTA offsets are not the same as opposed to the case in Figure 4.13 in which the CTA offsets are assumed to be the same.

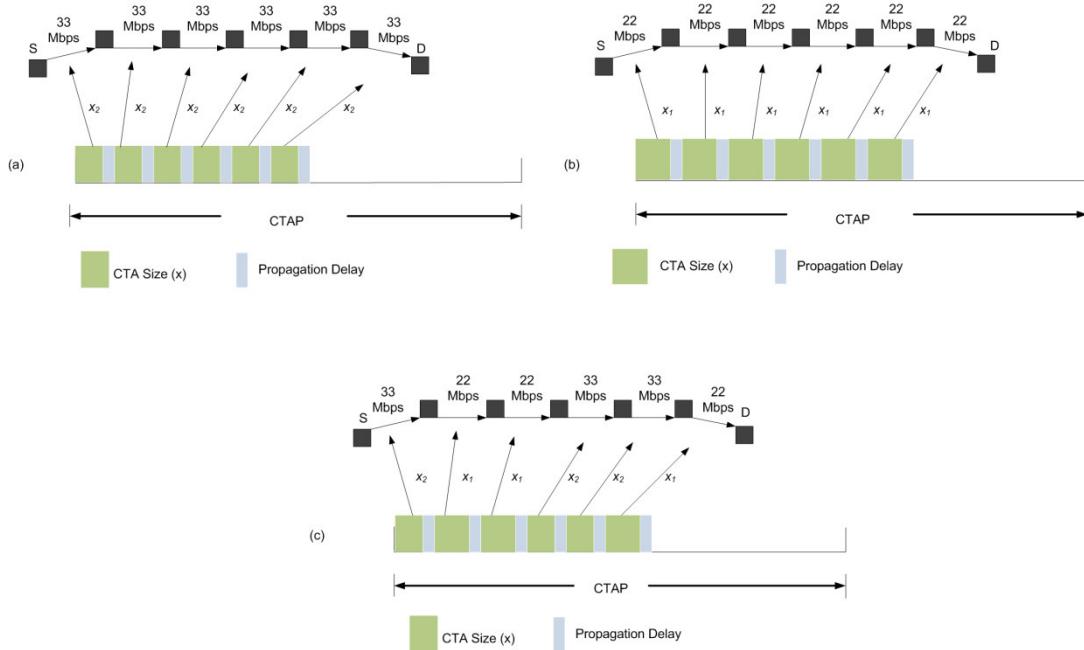


Figure 4.14 An ideal case assumption to check the effect of ∂_T on capacity efficiency and EED

The scenario depicted in Figure 4.14 assumes a flow which traverses a 6 hop path to reach the destination. In case (a) of Figure 4.14, a uniform value of $\partial_T = 33$ Mbps is assumed throughput the path, $\partial_T = 22$ Mbps in case (b) and a variable value of $\partial_T = 22\text{-}33$ Mbps in case (c) of Figure 4.14. If the EED_a , EED_b and EED_c are the end-to-end delays for case (a), (b) and (c) respectively in Figure 4.14, then $EED_a < EED_c < EED_b$. If C_T^a , C_T^b and C_T^c are used to represent the total end-to-end capacity utilized for flow f for case(a), (b) and (c) respectively in Figure 4.14 and $C_T = \sum_{i=1}^6 \eta_{(CTA)i}$, then $C_T^a < C_T^c < C_T^b$. Therefore, we can conclude that for the same flow f , if there are paths with similar CTA offsets but different data rates for one or more hops along the source-destination path, then that path which has higher values of ∂_T for greater number of hops when compared with other paths for the same CTA offsets provides a better capacity efficiency and lower EED. The EED values for different number of hops i.e. from 5 – 8 hops are given in Figure 4.15 when the transmitter data rates are varied from 22 Mbps to 55 Mbps. The results in Figure 4.15 are obtained from the simulations carried out in OPNET modeller. A fragment size of 16484 bits is considered

with 3 frames /CTA for the flow. The transmitter data rate is kept uniform throughout each hop along the source-destination path. The CTA's are arranged one after the other just like the case assumed in Figure 4.14. It can be seen in the figure that for variable CTA offsets in the ideal case, there is a substantial difference in the value of EED when the value of ∂_T is increased from 22 Mbps to 55 Mbps. As it shall be seen in the results presented in the subsequent sections, it is not actually the case in practical scenarios and the difference in EED value is not that much when the CTA offsets are the same for different paths but only the value of ∂_T is varied.

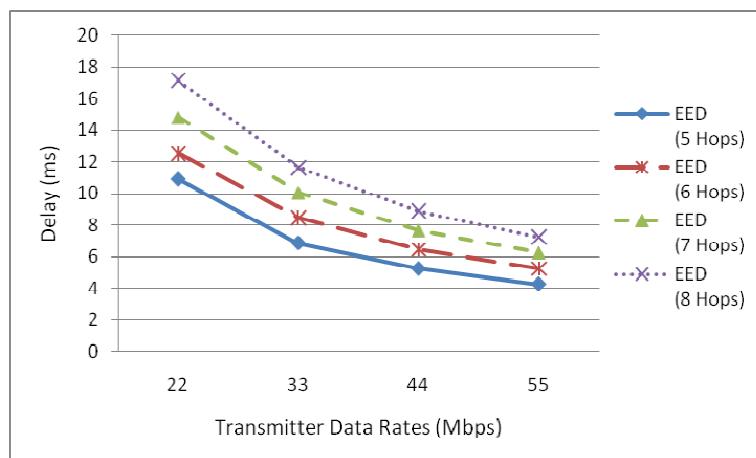


Figure 4.15 End to End Delay for Different number of Hops and Transmitter Data Rates

4.4.2 STS and STS-R Metrics

The Switched Time Slot (STS) metric takes into consideration the capacity efficiency as well as the queuing delay along each link which belongs to a particular source-destination path. The cost based on STS metric for the th link is calculated by

$$(4.10)$$

Where w_i is the weight factor and determines if the capacity efficiency is to be given preference or the queuing delay which affects the EED. The term $c_{i,j}$ denotes the cost based on STS metric for a link. The total cost for a complete path based on the STS metric is given by

$$(4.11)$$

The cost for the i^{th} link is calculated based on (4.10) before the CTA switching takes place and after the CTA switching takes place for that link. The motivation for the CTA switching originates from the case given in Figure 4.11. The CTA switching consists of two parts:

1. If time has already been reserved for the next hop MPNC, it is ensured that the CTA for transmission to the next hop MPNC is reserved after the CTA in which the data is received from the previous hop MPNC
2. If there are any intra-PAN flow reservations to the left of the CTA being reserved for the inter-PAN flow, then the intra-PAN flow reservations or CTAs are moved to the right of the inter-PAN CTA. In other words, the inter-PAN CTA is moved to the left by an amount equal to the intra-PAN CTA to reduce the queuing delay by an amount equal to the size of the intra-PAN flow reservation shifted to its right (see Figure 4.11 (c and d))

When the CTA switching takes place from source-destination assuming source is to the left and destination to the right, then the CTA along each link is shifted to the left as much as possible such that it doesn't cross the boundary of the previous hop CTA's end time whether the previous hop CTA is for the same flow or for a different inter-PAN flow. If the next hop CTA for the same inter-PAN flow is already to the left of its previous hop CTA, then it is shifted to the right of the previous hop CTA to avoid the misordering. If the value of $\phi_{STS}^p > 0.065535$ secs, then it means that more than one superframe is required to transmit a frame from the source to destination MPNC. The earliest, a CTA is reserved in a superframe, the lesser the chances for ϕ_{STS}^p to exceed the superframe duration. If $\phi_{STS}^p < 0.065535$ secs, then a transmission from source MPNC can reach the destination MPNC within one superframe duration.

The feasibility of CTA switching is an important factor to consider which determines if any improvement can be expected after the switching procedure. The feasibility is determined by considering the sum of queuing delay between the CTAs reserved for hop $m - 1$ to hop m and for hop m to hop $m + 1$ before switching and after switching. If the sum of queuing delay before switching is more, the switching is feasible, otherwise if the sum of queuing delay before switching is less or equal to the sum of queuing delay after switching, the switching is not feasible. If we consider Figure 4.10 in which t_{d1} represents the

queuing delay for hop between MPNC 1 and MPNC 2 and between MPNC 2 and MPNC 3 while t_{d2} represents the queuing delay between MPNC 2 and MPNC 3 and between MPNC 3 and MPNC 4, then the term $s = t_{d1} + t_{d2}$ is used to indicate the sum of queuing delays for hop $m - 1$ to hop m and for hop m and hop $m + 1$ before switching. The term $s' = t_{d1}' + t_{d2}'$ is used to indicate the sum of queuing delays after switching. The switching is considered feasible if and only if $s > s'$ and infeasible if $s \leq s'$. The improvement after switching for the i^{th} hop is given by $|t_{di} - t'_{di}|$. If n is the number of devices from the source to destination then $\sum_{i=1}^{n-1} |t_{di} - t'_{di}|$ gives us the total improvement in the queuing delay for the complete path.

Once the CTA switching takes place and a path is selected based on the total cost calculated via the STS metric, the relevant flow initiates and the path is maintained throughout the duration of the flow. Upon expiration of the flow, the resources i.e. the CTAs reserved for the flow are released.

A limitation of CTA switching procedure based on STS metric is that once the CTA switching is carried out and the cost calculated, the switching procedure is not repeated again. This might not be a limitation for flows with shorter duration but for flows with longer duration, there is a high possibility that one or more intra-PAN or inter-PAN flows terminate along one or more links on the source-destination path. The CTAs released by the terminated flows can create an opportunity to reduce the EED for those inter-PAN flows which have CTAs reserved to the right of the terminated CTAs. The active inter-PAN flows can therefore be shifted to the left by a time offset equal to the terminated CTA size on the relevant link thus reducing the queuing delay for the link.

In order to take advantage of CTA space released by terminated flows to reduce the EED for an ongoing active inter-PAN flow f_i , the metric STS-R is introduced. The link cost and the route or total path cost is calculated by STS-R in the same way as STS. The difference is that in STS-R, the CTA switching procedure is not only performed initially when the link cost is calculated but is repeated for that particular link on which either an intra-PAN flow is terminated or an Inter-PAN flow is terminated. If there is any inter-PAN CTA reserved for flow f_i to the right of the terminated CTA, then the CTA for flow f_i is moved to the left by a

time offset equal to the size of the terminated CTA. Therefore, when using STS-R, the probability to reduce the EED further after the initial CTA switching operation increases.

4.4.3 Performance Evaluation of STS, STS-R and MSMR

A simulation model was developed to evaluate the performance of STS and STS-R and also their respective comparison with MSMR. Static routing was considered in all the simulations and the CTA offsets were randomly chosen for all cases unless specified otherwise. All the paths from the source to destination were node disjoint [13]. The total number of hops from the source to destination was kept 7. In total 8 MPNCs were considered including the source, destination and 6 Relay MPNCs. The value of weight factor w used in (4.7) for the STS metric is taken to be 0.5 for all the simulations. Different scenarios were setup to evaluate the performance of STS, STS-R and MSMR to accumulate absolute performance indicators and to compare the three metrics. A summary of results related to total route cost for the selected paths and the EED values for MSMR and STS is given in Table 4.7. Table 4.7 shows the results of 10 different scenarios in which the transmission data rates are varied for different MPNCs along the source destination path and the path selection based on the total route cost based on the MSMR and the STS metric is taken into account. The total capacity utilized in the path selected by MSMR and in the path selected by STS along with the time efficiency of selected paths is given in Table 4.8. In all the simulations the Medium Access Slot (MAS) length is taken to be 8 ms. Therefore, in the simulation model, if any offset exceeds the limit of 8 ms, a complete superframe duration is added to the queuing delay.

From scenario 1 in Table 4.7, the total route cost based on MSMR is the same for the path selected by STS. However, when two or more paths have the same route cost, the selection is based randomly. Therefore, MSMR seems to select a different path than STS even though the route cost based on MSMR is the same. The route cost based on STS is different and lower than the one selected by MSMR. The reason is that the path selected by STS uses higher transmission data rates per hop when compared with the one selected by MSMR. Since MSMR only calculates link cost based on the queuing delay, it does not take transmission data rates into account. The EED value for the path selected by STS metric is also lower even though the CTA offset per hop is almost the same for both paths including

the last hop CTA offset. The difference in value of EED is because the last hop in the path selected by STS uses a higher value of δ_T and therefore has lower transmission delay for the last hop which reduces the overall EED value. Another advantage of STS by taking link transmission data rates is that the total capacity

Table 4.7: Summary of simulation results for comparison between MSMR and STS metrics

Scenario	Metric		Hop 1	Hop 2	Hop 3	Hop 4	Hop 5	Hop 6	Hop 7	MSMR Cost	STS Cost	EED
1	MSMR	Offset	0.001535	0.001843	0.002181	0.002489	0.002827	0.003135	0.003473	0.001938	0.00194	0.002054
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.001843	0.002181	0.002489	0.002827	0.003165	0.003473	0.001938	0.00170	0.002012
		δ_T	33	33	33	33	33	33	33			
2a	MSMR	Offset	0.001535	0.002401	0.003297	0.004164	0.005060	0.005927	0.006823	0.005288	0.00557	0.005684
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.002401	0.003297	0.004164	0.005060	0.005957	0.006823	0.005288	0.00424	0.005482
		δ_T	44	44	44	44	44	44	44			
2b	MSMR	Offset	0.001535	0.002401	0.003297	0.004164	0.005060	0.005927	0.006823	0.005288	0.00557	0.005684
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.002431	0.003327	0.004194	0.005090	0.005957	0.006853	0.005318	0.00339	0.004260
		δ_T	44	44	44	44	44	44	44			
3	MSMR	Offset	0.001535	0.003146	0.004787	0.006398	0.008039	0.009650	0.011292	0.009757	0.01041	0.07606
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.003176	0.004817	0.006428	0.008069	0.009680	0.011322	0.009787	0.00779	0.075701
		δ_T	44	44	44	44	44	44	44			
4	MSMR	Offset	0.001535	0.003146	0.004787	0.006398	0.008039	0.009650	0.011292	0.009757	0.01041	0.07606
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.003176	0.004817	0.006428	0.008069	0.009680	0.011322	0.009787	0.00814	0.075624
		δ_T	44	55	44	33	44	22	55			
5	MSMR	Offset	0.001535	0.003146	0.004787	0.006398	0.008039	0.009650	0.011292	0.009757	0.01041	0.07606
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.003146	0.004787	0.006398	0.008039	0.009680	0.011322	0.009787	0.00814	0.075624
		δ_T	22	22	22	22	22	55	55			
6	MSMR	Offset	0.001535	0.002401	0.003297	0.005034	0.005930	0.007217	0.008113	0.006578	0.00621	0.006974
		δ_T	22	22	22	22	22	22	22			
	STS	Offset	0.001535	0.002401	0.003297	0.004784	0.005930	0.007217	0.008113	0.006578	0.00621	0.006974
		δ_T	22	22	22	22	22	22	22			
7	MSMR	Offset	0.001535	0.002645	0.003785	0.005765	0.007395	0.008926	0.010973	0.009438	0.00849	0.165005
		δ_T	33	33	33	33	33	33	33			
	STS	Offset	0.001535	0.002645	0.003785	0.005516	0.006596	0.007676	0.009106	0.007571	0.00756	0.073615
		δ_T	33	33	33	33	33	33	33			
8	MSMR	Offset	0.001535	0.002645	0.005765	0.003785	0.007395	0.008926	0.010973	0.009438	0.00849	0.206552
		δ_T	33	33	33	33	33	33	33			
	STS	Offset	0.001535	0.002645	0.003785	0.005516	0.006596	0.007676	0.009106	0.007571	0.00756	0.073615
		δ_T	33	33	33	33	33	33	33			
9	MSMR	Offset	0.001535	0.003155	0.005318	0.007809	0.010859	0.012899	0.015457	0.013922	0.01252	0.079956
		δ_T	33	33	33	33	33	33	33			
	STS	Offset	0.001535	0.003155	0.005318	0.007174	0.009327	0.01108	0.012981	0.011446	0.01128	0.07748
		δ_T	33	33	33	33	33	33	33			
10	MSMR	Offset	0.001535	0.003155	0.004944	0.007434	0.010484	0.011925	0.014108	0.012573	0.01117	0.078477
		δ_T	33	33	44	33	33	55	44			
	STS	Offset	0.001535	0.003125	0.00436	0.006663	0.008625	0.009841	0.011318	0.009783	0.00929	0.075687
		δ_T	44	33	55	33	55	44	44			

utilization by the flow is decreased. This can be seen in Table 4.8 in the scenario 1 row. It can be noted that the total capacity utilized for the path selected by MSMR is 0.001945

while that of the path selected by STS is 0.00148. From hereon, the path selected in each scenario by MSMR shall be referred by path 1 while the path selected by STS shall be referred by path 2 for convenience.

Table 4.8: Summary of results with Capacity Utilization and Time efficiency

Scenario	Fragment Size (bits)	Number of Frames	MSMR C_T	MSMR Mean η_{eff}	STS C_T	STS Mean η_{eff}
1	2048	2	0.001945	81.69	0.00148	76.92
2a	8192	2	0.005855	94.69	0.00320	91.40
2b	8192	2	0.005855	94.69	0.00320	91.40
3	16384	2	0.011068	97.27	0.00580	95.51
4	16384	2	0.011068	97.27	0.00650	95.51
5	16384	2	0.011068	97.27	0.01106	97.27
6	8192	2	0.005855	94.69	0.00585	94.69
7	16384	2	0.007561	96.38	0.00756	96.38
8	16384	2	0.007561	96.38	0.00756	96.38
9	16384	3	0.011132	96.69	0.01113	96.69
10	16384	3	0.009783	95.83	0.00880	95.48

In scenario 2a, the offsets for path 1 and path 2 are different for the 6th hop and the CTA offset for the 6th hop in path 2 is 30 μ s more than the 6th hop of path 1. The route cost for path 1 and path 2 based on MSMR still comes out to be the same. The reason is that although the link cost based on MSMR for the intermediate hop is different, the last hop CTA offset is still the same for path 1 and path 2. The EED value again comes out lower for path 2 selected by STS because of the relatively higher transmission data rates of 44 Mbps used throughout path 2 when compared with path 1 hence lower transmission delay for the last hop of path 2. The capacity utilization for the flow is also lesser for path 2 for the same reason. To check the effect of last hop CTA offset on the route cost based on queuing delay, the CTA offset from the 2nd to the 7th hop in scenario 2b was shifted to the right by an amount equal to 30 μ s. The difference of cost calculated on basis of MSMR for path 1 and path 2 is 30 μ s for scenario 2b which is equal to the CTA offset difference between path 1 and path 2 for the last hop i.e. 7th hop.

A significant improvement in capacity utilization can be seen in the flow in scenario 3 (Table 4.8) which is up to 48% approximately. The time efficiency for the flow when using path 1 is only 1.81% better than path 2 for the same scenario. In scenario 4, the values of transmitter data rates are not uniform for path 2 and are different throughout the path. Path 1 however, has the same transmitter data rate of 22 Mbps for each link along the path. The route cost based on MSMR is more for path 2 than path 1 because of the final hop of

path 2 which is 30 μ s more than path 1. The value of ∂_T for the last hop of path 2 is 55 Mbps while for path 1 it is 22 Mbps. The total EED value for path 2 is still lower than path 1 because of the last hop transmission time for path 2 being lower than path 1. The results from scenario 5 are similar to scenario 4 apart from the fact that the dominant transmitter data rate throughout path 2 is 22 Mbps and therefore there is not much difference in capacity utilization of the flow when compared with scenario 4.

The route cost based on MSMR as well as for STS is the same for scenario 6. The reason is that both select the same path based on the route cost. However, after the CTA switching procedure, the queuing delay for hop 4 is decreased by moving the CTA to the left. The CTA switching procedure does not have any impact on the EED value because there is no change in the offset for the last hop. In scenario 7, again the same path is selected by MSMR and STS but after the switching procedure, the queuing delay is reduced for hop 4, 5, 6 and 7 by amounts 250, 800, 1250 and 1867 μ s respectively. The difference between the route cost based on MSMR before and after switching is 1867 μ s which is equal to the difference in the last hop after CTA switching.

In scenario 8, the same path is selected again by MSMR and STS. However, this time there is a misorder of CTAs in hops 3 and 4 i.e. the CTA offset of hop 4 is earlier than hop 3. Therefore, although the mis-ordering is removed via the CTA switching procedure, MSMR is not able to remove the misordering. As a result, there is a substantial difference in the EED value in case of MSMR i.e. 94% more than STS. Scenario 9 and 10 also show the reduced queuing delay in case of STS because of CTA switching and therefore a lower value of EED as a result.

A comparison of MSMR, STS and STS-R metric is given in Table 4.9 for 4 difference scenarios. In scenario 1, there is a difference of 750 μ s in the CTA offset for hop 4 using STS-R when compared to STS. However, there is no reduction in the EED value because the CTA offset for hop 7 remains unchanged. The difference in hop 4 for scenario 2 using STS-R is again 750 μ s while hop 7 changes by 294 μ s. A reduction in the EED value can therefore be noticed for STS-R which is the lowest value when compared with STS and MSMR. The further reduction in scenario 2 is made possible when the flow duration is increased by 200 seconds for the inter-PAN flow. In scenarios 3 and 4 and the previous scenarios, an important point to notice is that the reduction in EED value is approximately equal to the difference in the CTA offset of hop 7 for STS and hop 7 for STS-R. Although there are

differences in CTA offsets for other hops between STS and STS-R but the CTA switching which helps in reducing the EED value further when compared to STS is the CTA offset of the last hop. Therefore, it can be safely concluded that the CTA switching can only make a difference if the CTA offset of last hop can be reduced.

Table 4.9: Summary of simulation results for comparison between MSMR, STS and STS-R metrics

Scenario	Metric		Hop 1	Hop 2	Hop 3	Hop 4	Hop 5	Hop 6	Hop 7	EED 1
1	MSMR	Offset	0.00153	0.00264	0.00378	0.00576	0.00739	0.00892	0.01097	0.165005
		∂_T	33	33	33	33	33	33	33	
	STS	Offset	0.00153	0.00264	0.00378	0.00551	0.00659	0.00767	0.00910	0.073615
		∂_T	33	33	33	33	33	33	33	
2	MSMR	Offset	0.00153	0.00264	0.00378	0.00576	0.00739	0.00892	0.01097	0.206552
		∂_T	33	33	33	33	33	33	33	
	STS	Offset	0.00153	0.00264	0.00378	0.00551	0.00659	0.00767	0.00910	0.073615
		∂_T	33	33	33	33	33	33	33	
3	MSMR	Offset	0.00153	0.00315	0.00531	0.00780	0.01085	0.01289	0.01545	0.079956
		∂_T	33	33	33	33	33	33	33	
	STS	Offset	0.00153	0.00315	0.00531	0.00717	0.00932	0.01108	0.01298	0.07748
		∂_T	33	33	33	33	33	33	33	
4	MSMR	Offset	0.00153	0.00315	0.00494	0.00743	0.01048	0.01192	0.01410	0.078477
		∂_T	33	33	44	33	33	55	44	
	STS	Offset	0.00153	0.00312	0.00436	0.00666	0.00862	0.00984	0.01131	0.075687
		∂_T	44	33	55	33	55	44	44	
	STS-R	Offset	0.00153	0.00312	0.00436	0.006345	0.00862	0.009806	0.011122	0.075491
		∂_T	44	33	55	33	55	44	44	

4.5 Important observations regarding achievable improvements with STS and STS-R

Based on the Simulation Results a few important observations are given in the subsequent section followed by a list of factors which can be considered in order to create the possibility of further improvements. The information provided in the subsequent sections is important to consider to any future extensions not only to the STS and STS-R metrics but also any possible variants to reduce the capacity utilization and/or the EED for meshed WPANs or similar TDMA based sensor networks.

4.5.1 General Observations

From the simulation results, it is clear than the CTA offset of the last hop plays an important role in determining how much improvement can be achieved via CTA switching whether it

be performed once in case of STS or more than once in case of STS-R. The following points summarize the key observations which result from the simulations carried out:

1. If the last hop CTA offset or queuing delay remains unchanged, then it doesn't matter how much improvement is achieved via CTA switching in the previous hops. The total EED remains the same. Even if some of the slots are switched and the queuing delay is reduced for any intermediate hops, there is no difference in the EED unless there is any change in the last hop. Since the last hop retains the same offset, it doesn't matter if the intermediate relays switch the CTAs because the last hop still sends at its due offset to the destination hence no improvement. The situation is shown in Figure 4.16 in which the CTA switching takes place for the 5th CTA and it moves to its left by a time offset of Δt . However the total EED value of T_d remains constant because the transmission to the destination MPNC still takes place in the 6th CTA for which the time offset remains the same.

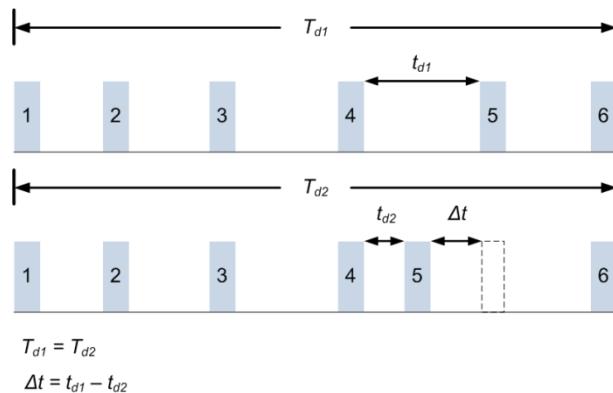


Figure 4.16 The fixed last hop offset problem

2. The amount of improvement also relates to the congestion at the final hop along the SD path. If the final hop has no congestion and there is adequate superframe capacity available. Low arrival rate and lower superframe utilization, then there are more chances that further CTA switching's can be made. The factor of mean flow durations at the final hop also plays an important role. In case of higher arrival rates but lower intra-pan or inter-PAN flow durations, there is a greater chance that room for CTA improvements can take place because of frequent flow terminations.
3. The CTA switching in $n = m - 1$ links before the m^{th} hop only has an effect equal to the change in the CTA offset of m^{th} hop. If there is no change in the m^{th} hop, the changes before the m^{th} hop in the n links has no effect on the total EED. If there is no change in the CTS

offsets of first n links but change in the final $m + k$ links by a total offset x , the total EED can be reduced by an offset q if q is the change in the final hop CTA offset of the path and if $q \leq x$ and an amount x if $q > x$. Also those paths which have intra-PAN CTA allocations to the left of the inter-PAN CTA, have more chance to improve the EED by swapping between the inter PAN and intra PAN CTAs since intra PAN flows are not affected by that since they are just single hop flows anyway.

4. The congestion in the first n links has lesser probability of affecting the total EED when compared to the congestion in the final k links. In other words, the change in the first n link count for less or nothing if there is less or no change in the final k links respectively
5. If the total flow duration of the inter-PAN flow f is such that there is a higher probability of flows ending in the final k links, then there is a higher probability of getting an improvement in the EED. In other words, if the mean flow duration of the flows in the final k links is less than the total duration or expected duration of the flow f , then the probability of flows ending before the flow f and leaving room for improvement increases.

4.5.2 Factors that affect further improvements

Some factors which affect and determine the performance improvements through CTA switching are

- Throughput of the inter-PAN flow f along its transmission data rate which determines the capacity utilization throughput the path
- Number of active inter-PAN and intra-PAN flows along the path used by flow f
- Flow duration of the flow f . With longer flow duration it is more likely that more inter-PAN or intra-PAN flows will terminate
- Offset and CTA size of the flows which are terminated
- Arrival rate of other inter-PAN and intra-PAN flows
- Flow duration of other active intra-PAN and inter-PAN flows. With relatively shorter flow durations than the inter-PAN flow f , it is likely that more flow shall terminate and creating more room for improvement via CTA switching

4.6 Conclusion

The time efficiency parameter is important in determining the communication efficiency of HDR WPAN flows. Increasing the time efficiency means that the overhead is lesser compared with the actual data bits. Increasing the time efficiency also means that either the number of frames/CTA for the flow is increased or the fragment size is increased which can also increase the capacity utilization of the flow. Also above a time efficiency value of 90%, increasing the frames/CTA and/or the fragment size have a little effect on increment in the time efficiency but the capacity utilization is increased by a greater amount. Therefore, in order to maximize the time efficiency of a flow, the capacity constraints are important to consider. The FTS algorithm takes into account the minimum throughput, maximum CTA size and minimum time efficiency as constraints and chooses the most appropriate parameters for the flow i.e. number of frames/CTA, fragment size and transmission data rate. The FTS and PTS algorithms can also find significant use in scheduling algorithms designed for handling VBR traffic in HDR WPANs and meshed WPANs. Especially when the buffer size for a device is increased and it has to request for additional time from the PNC or MPNC. With the increase or decrease in the buffer size, the requirement for CTA size changes and therefore, a device has to select the parameters accordingly. Both FTS and PTS have lower complexity which depends on the set of options that a device have in choosing the parameters, and therefore feasible for situations in which quick decisions need to be made related to varying throughput requirements of a device.

The proposed STS and STS-R metrics can provide better EED values and capacity utilization when compared with the MSMR approach which only deals with queuing delays along each hop. The STS and STS-R provide a tuneable weight factor w which can be set to the requirements of a particular flow. Although the equal weightage is given to the flows considered in the simulation for generalization, each flow supporting a particular traffic type can choose an appropriate value of w either to give more importance to delay or to the capacity utilization. The improvements in EED due to STS and STS-R depend on the CTA switching procedure, the effectiveness of which in turn depends mostly on the conditions in the final hop. The greater the reductions in the queuing delay of the final hop the higher the probability of reduction in the EED value.

4.7 References

- [1] Fan Yu, Tao Wu, Subir Biswas, "Routing with Minimized Slot Misordering for Delay Mitigation in TDMA based Sensor Networks," *Networking and Services, International conference on*, pp. 58, International Conference on Networking and Services (ICNS '07), 2007.
- [2] "IEEE standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements part 15.3: wireless medium access control (MAC) and physical layer (PHY) specifications for high rate wireless personal area networks (WPANs)," *IEEE Std 802.15.3-2003* , vol., no., pp. 0_1-315, 2003.
- [3] "IEEE Recommended Practice for Information technology--Telecommunications and information exchange between systems-- Local and metropolitan area networks-- Specific requirements Part 15.5: Mesh Topology Capability in Wireless Personal Area Networks (WPANs)," *IEEE Std 802.15.5-2009* , vol., no., pp.1-166, May 8 2009.
- [4] Mahmud, S.; Khan, S.; Al-Raweshidy, H.; Sivarajah, K., "Mesched high data rate personal area networks," *Communications Surveys & Tutorials, IEEE* , vol.10, no.1, pp.58-69, First Quarter 2008
- [5] T. van Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," In *Proceedings of the First ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Los Angeles, CA, November 2003.
- [6] Raghunathan, V.; Schurgers, C.; Sung Park; Srivastava, M.B., "Energy-aware wireless microsensor networks," *Signal Processing Magazine, IEEE* , vol.19, no.2, pp.40-50, Mar 2002.
- [7] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, "Energy-efficient, collision-free medium access control for wireless sensor networks," *Wireless Netw.*, vol. 12, no. 1, pp. 63–78, 2006.
- [8] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient mac protocol for wireless sensor networks," in *Proc. IEEE INFOCOM*, New York, NY, June 2002, pp. 1567–1576.
- [9] Z. Chen and A. Khokhar, "Self organization and energy efficient TDMA MAC protocol by wake up for wireless sensor networks," *IEEE SECON*, Oct. 2004, pp. 335-341.
- [10] Muruganathan, S.D.; Ma, D.C.F.; Bhasin, R.I.; Fapojuwo, A.O., "A centralized energy-efficient routing protocol for wireless sensor networks," *Communications Magazine, IEEE* , vol.43, no.3, pp. S8-13, March 2005.
- [11] Lu, G.; Krishnamachari, B.; Raghavendra, C.S., "An adaptive energy-efficient and low-latency MAC for data gathering in wireless sensor networks," *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International* , vol., no., pp. 224-, 26-30 April 2004.

- [12] Shuguang Cui; Madan, R.; Goldsmith, A.; Lall, S., "Energy-delay tradeoffs for data collection in TDMA-based sensor networks," *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol.5, no., pp. 3278-3284 Vol. 5, 16-20 May 2005.
- [13] Xuefei Li, Laurie Cuthbert, "On-demand Node-Disjoint Multipath Routing in Wireless Ad hoc Network," *Local Computer Networks, Annual IEEE Conference on*, pp. 419-420, 29th Annual IEEE International Conference on Local Computer Networks (LCN'04), 2004.

CHAPTER 5

5 Capacity Improvements in WPANs with Power Control

5.1 Introduction

In the previous chapters, the issue of limited capacity of superframe was addressed by trying to adopt fairness based allocations to devices (and MPNCs) or to select capacity efficient paths in meshed WPANs. These approaches help alleviate the problems due to congestion to some extent but do not help in capacity improvement. The two important and most popular techniques to improve capacity of wireless networks are multi-radio/multi-channel communication and spatial reuse through power control. If devices are equipped with multiple radios, they can communicate on different radio frequency channels at the same time. Since the frequency channel used per radio is non-overlapping with the frequency band of the channel used by another radio, significant capacity improvements can be achieved. In order to efficiently utilize multi-radio devices, the radio frequency channels allocated to one device should be ideally different than those one-hop devices in its neighbourhood with whom the device do not intend to communicate. However, to exchange control information, the existence of a common channel is mandatory among devices. The higher the diversity between allocated channels in each two hop neighbourhood, the higher are the chances to expect improvement in throughput and reduction in interference.

In [1], the performance of a multi-radio/multi-channel mesh network is evaluated and a metric called Weighted Cumulative Expected Transmission Time (WCETT) is proposed which takes into account the delay expected at each link along the source-destination path and the channel diversity along a path. The problem of channel assignment in multi-radio/multi-channel networks has been investigated extensively and the techniques lie in mainly two categories i.e. dynamic channel assignment and static channel assignment. In static channel assignment, each radio is assigned a channel permanently while in dynamic channel

assignment, the radios can switch to different available channels. Some examples of dynamic channel assignment strategies are given in [2-6] and static channel assignment strategies are given in [7-8]. A hybrid approach in which both static and dynamic schemes are considered is given in [9]. An effort is also made to estimate an upper limit on capacity in multi-radio/multi-channel based mesh networks in [9].

Power Control based algorithms have been an active topic of interest with the aim of reducing interference and increasing spatial reuse. Increasing spatial reuse means that more devices can communicate in a certain region without interfering with each other. In [10], the authors propose an integrated dynamic channel assignment and power control scheme for wireless systems. The authors claim that an integrated approach brings more improvement than using channel assignment and power control separately. Most of the power control based schemes for TDMA wireless networks have the priority to use the same time slots for those device which are not in the transmission range of each other hence better spatial reuse. The probability of overlapping time slot use is increased by decreasing the transmission power of devices to such a level that a certain level of SNR is maintained between the transmitter and receiver. Some of the schemes are proposed in [11-14] which make use of the concept of time slot overlapping for those devices which are not in the transmission range of each other. In [11] and [12], fixed size time slots are considered while in [13] and [14], the size of the time slots is variable because the approach is specific to WPANs. An important point to note is that when variable size time slots (CTAs) are assumed for the power control algorithm, the most efficient improvement in capacity results when those CTAs are overlapped the size of whom is nearly or almost the same, if not completely the same. Obviously such overlapping depends on the fact that the CTAs for the two flows must be from the devices that are not in the transmission range of each other. In this chapter, the CTA overlapping that result from power control is arranged based on the CTA overlap ratio. CTA overlap ratio helps in grouping those flows together which can be allocated the same CTA and then allocates overlapping CTAs to those flows the CTA size for which is the closest to each other. Some other relevant papers dealing with the issue of spatial reuse in general are [15-16]. The benefits and limitations of spatial reuse are

generally discussed in [15] while [16] is about determining the achievable throughput in dense networks with spatial reuse. The contribution in this chapter consists of

1. A handshake procedure in which the source and destination devices reach an agreement over the transmission data rate and transmission power based on a minimum threshold value of SNR.
2. CTA overlapping based on grouping those source-destination device pairs, the transmission range of whom is mutually exclusive. The same CTAs are allocated to those flows the difference in CTA size of which is the lowest.

5.2 Time Slot Overlapping via Power Control

Building up on the discussion related to superframe capacity in single hop and meshed HDR WPANs in the previous chapters, an approach to allocate CTAs possibly to greater number of devices in the same superframe capacity is presented. The reason for preferring the power control procedure to achieve capacity improvement over the multi-radio/multi-channel approach is because WPANs are intended for low power and low cost devices with limited battery life. The use of multiple radios per device can increase the energy consumption of the device and drain its battery life quicker than by using a single radio. To the best of knowledge, there is not much work done on the energy consumption of multi-radio devices until now. As discussed in the previous section, the difference in the proposed approach compared to the other existing approaches is that due to the use of variable sized CTAs, the difference in the CTA size of overlapping CTAs is taken into account.

The proposed scheme aims to achieve its goal in three steps. First of all a handshake procedure takes place between the source and destination device in which the transmission power and transmission data rate values are negotiated such that a minimum threshold of SNR γ_{SNR} is maintained. After the handshake procedure, the second step is to find those source-destination pairs which can transmit simultaneously in the same CTA. Finally in the third step, the CTAs are scheduled in such a way as to minimize the total overlap ratio ρ_T . Further details related to each step are given in the subsequent sections.

5.2.1 Handshake Procedure

Before the handshake procedure is initiated, the source device sends a channel time request command to the PNC to request a CTA. After the successful allocation of a CTA, the source device first of all starts a test transmission to the destination device to initiate the handshake procedure. The source device s_i sends a test frame to the destination d_j using the lowest transmission data rate i.e. 22 Mbps and starts a timer to trace the reception of an Acknowledgement (ACK) frame. The device s_i can either receive the ACK or the timer expires. Different steps are taken based on the occurrence of either the former or later which are explained below:

Case A: If ACK from d_j is received, then

1. If the acknowledgement from the receiver is received with an acceptable value of SNR which is below the threshold γ_{SNR} , the s_i sends another packet using the next higher transmitter data rate and the same process is repeated unless the value of γ_{SNR} is violated
2. The source s_i keep on increasing the transmitter data rate until the SNR value is not violated
3. Upon reaching the highest data rate with a suitable value of SNR, s_i starts decreasing the power level iteratively unless γ_{SNR} is reached
4. When the threshold SNR is reached, s_i sets the attributes and sends a confirmation to d_j
5. After d_j receives the confirmation from the s_i , it sets the transmitter and receiver data rates and the data transmission begins

Case B: If ACK from d_j is not received, then

1. When the ACK is not received upon the expiration of the timer by s_i , it resends the frame again.
2. If the ACK is still not received, then s_i checks its transmission data rate. If the transmission data rate is already the lowest i.e. 22 Mbps, it starts increasing the transmission power iteratively unless an ACK is received. If at the lowest data rate, after m number of tries or if the transmission power reaches its maximum level, and the ACK is not received, then the transmission fails

3. If after sending the second frame s_i does not get a response, it reduces its data rate if not at the lowest data rate and keeps on reducing unless it gets a response. If it reaches the lowest data rate then point 2 is repeated
4. If s_i receives a response by reducing its data rate with a suitable SNR level above a threshold, it starts reducing the transmission power iteratively unless k iterations are reached or the SNR reaches the acceptable threshold γ_{SNR} . At that point the attributes are finalized and s_i sends a confirmation frame about the attributes of data transmission to the receiver d_j

The SNR is calculated by d_j and compared with the value of γ_{SNR} . In case the threshold value of SNR is violated, d_j notifies the source s_i . The source s_i also calculates the SNR value from the ACK received from d_j and compares it with γ_{SNR} . If the value of SNR is acceptable i.e. above threshold when calculated by s_i but not acceptable i.e. lower than the threshold when calculated by d_j , then the transmission power is increased iteratively by an amount equal to p_{it} at each iteration, until the SNR is acceptable. If the SNR calculated by s_i and d_j is not acceptable, then the transmission data rate is lowered and the transmission power is increased stepwise by an amount equal to p_{it} at each iteration. If the transmission data rate is already the lowest, then point 2 of Case B applies. The handshake process is summarized in the flow chart of Figure 5.1

In single hop WPANs, the PNC has the option to specify the maximum transmission power level in the beacon frame or keep it variable. In meshed HDR WPANs, the MPNC has to transmit the beacon frame at the maximum transmit power level. However, control information and data can be sent at any negotiated power level with the peer MPNC or a device. The handshake procedure therefore applies equally to single hop WPAN devices as well as MPNCs in a meshed WPAN network. According to IEEE 802.15.3, the Immediate Acknowledgement (Imm-ACK) frame is always sent by the destination at the same transmission data rate at which it receives the frame from the source. However, the command frames and the data frames can be sent at any transmission data rate. Therefore, in our case, command frames are used for the handshake procedure.

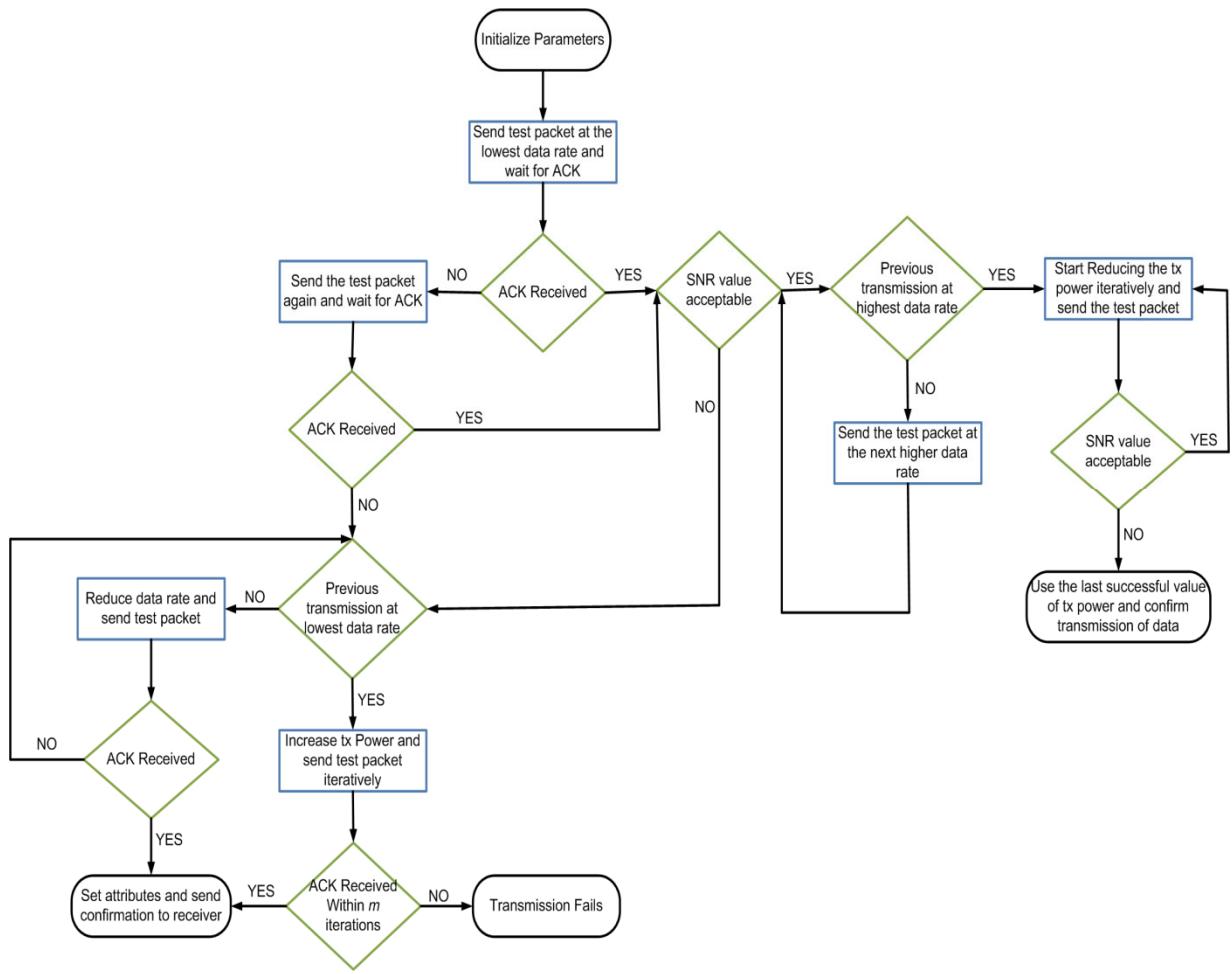


Figure 5.1 Flow Chart of the Handshake Procedure

5.2.2 Mutually Exclusive Time Slots

Once the handshake procedure is completed and the transmission data rate and transmission power levels are negotiated upon between s_i and d_j , either the PNC or the ref-MPNC determines if the particular source-destination pair can be placed in a set of active flows which can use non-overlapping CTAs. If \mathcal{F} is a vector that consists of all the active flows (all the flows in a piconet for single hop WPANs and all the flows in a two hop neighbourhood for meshed WPANs), then f_{ij} which is the newly negotiated flow between s_i and d_j is checked for possibility of CTA overlapping with $f_{r \neq i} \in \mathcal{F} \forall r = 1, 2, \dots n$. In order to check whether a subset of flows exist the transmission range of which is mutually exclusive with that of f_{ij} , mini time slots or CTAs which have a duration of 30 μ s are reserved in the superframe, the number of which equals that of the elements in the vector \mathcal{F} i.e. $n(\mathcal{F})$. In each mini CTA, the new flow from device s_i broadcasts a test frame using the

same transmission power level negotiated with d_j and waits for a reply from the flow represented by the first element of \mathcal{F} which upon reception of the broadcast frame, sends a unicast frame to s_i using the devid ID of s_i which it gets from the broadcast frame. The unicast frame is sent at the same transmission power level with which the source s_r^1 and destination d_s^1 which relate to the flow $f_{rs}^1 \in \mathcal{F}$ (first element of \mathcal{F}) transmit to each other. If s_i receives the frame from s_r^1 , it indicates that both flows cannot be accommodated in the same CTA. If s_i does not receive the frame from s_r^1 till the end of the mini CTA, it stores the device ID of s_r^1 . The list of device IDs for all the source devices in \mathcal{F} is provided to s_i by either the PNC or the ref-MPNC. Therefore if in the r^{th} mini slot, s_i does not receive any unicast frame from the source of r^{th} flow i.e. $f_{rs} \in \mathcal{F}$, it stores the device ID of s_r .

When the entire mini CTAs have been used, s_i checks the number of stored device IDs with which it can share a CTA. If the vector of stored device IDs $M_i \neq \emptyset$, then s_i sends the list to either the PNC or the ref-MPNC. The process of mini CTAs to check whether the transmissions of different flows are mutually exclusive is mandatory because it is highly likely that the flows that belong to \mathcal{F} have unequal transmission range since all the flows undergo the handshake procedure before starting the actual data transfer. Since the devices are at different locations throughout the cluster or clusters, the distance between each source-destination pair and the corresponding transmission power levels and the transmitter data rate values might differ from one flow to another. The order of mini CTAs is based on the order of flows in the vector \mathcal{F} in which the flows are arranged from the earliest flow to the most recent flow. Once the PNC or ref-MPNC receives the list of device IDs in the vector M_i from s_i , it can then determine a suitable CTA overlapping arrangement for s_i based on CTA overlap ratio.

5.2.3 Adjustment via Overlap Ratio

The handshake procedure described previously helps in using the lowest transmission power at the highest transmission data rate possible. In the CTA arrangement stage, the CTA overlap ratio is calculated for each (s_i, m_i) pair where $m_i \in M_i$. An example scenario for CTA overlapping in a single hop HDR WPAN can be seen in Figure 5.2 in which the pairs $(s_1, m_4) \rightarrow (1, 4)$ and $(s_5, m_8) \rightarrow (5, 8)$. From Figure 5.2, it can be noted that the CTA size for the flow f_{12} i.e. η_{12} is smaller than the CTA size for the flow f_{43} i.e. η_{43} . The total time

allocated for both the flows in the overlapping CTA is therefore equal to the CTA size η_{43} . The CTA overlap ratio for the overlapping CTAs of flows f_{12} and f_{43} can simply be calculated by

$$\varrho_{14} = \frac{\eta_{12}}{\eta_{43}} \quad (5.1)$$

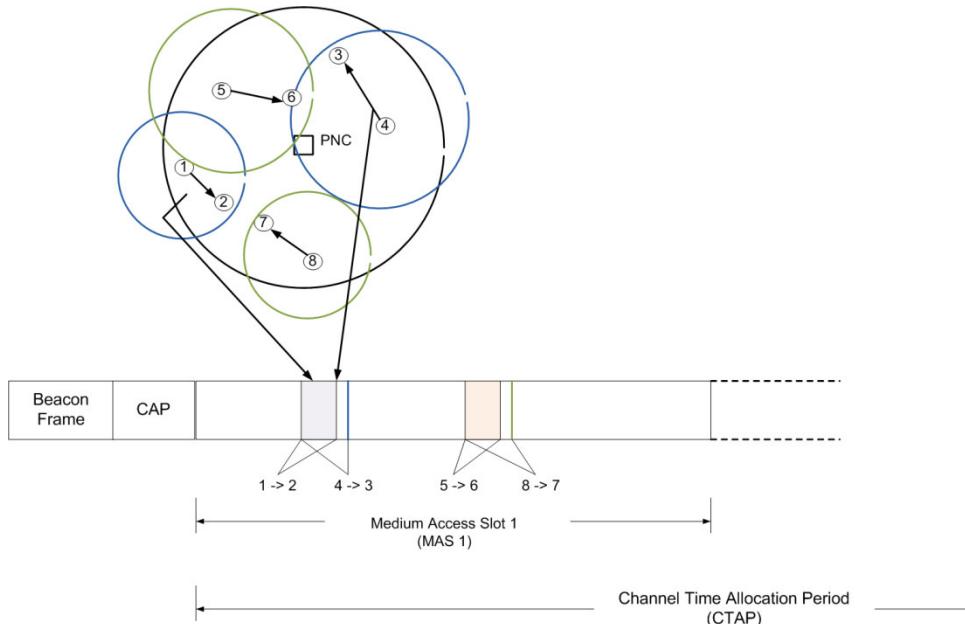


Figure 5.2 CTA overlapping in a single hop WPAN

Assuming there are only two overlapping CTAs in case of Figure 5.2, the total CTA overlap ratio can be calculated by

$$\varrho_T = \frac{\eta_{12} + \eta_{56}}{\eta_{43} + \eta_{87}} \quad (5.2)$$

If $\sum_n \eta_{ij}$ is the sum of the smaller CTA sizes in n overlappings in the superframe and $\sum_n \eta_{lm}$ is the sum of larger CTA sizes in n overlappings, then the total overhead ratio for the superframe can be calculated by

$$\varrho_T = \frac{\sum_n \eta_{ij}}{\sum_n \eta_{lm}} \quad \forall i \neq l, m \text{ and } \forall j \neq l, m \quad (5.3)$$

The total capacity improvement c_T can simply be given by

$$c_T = \sum_n \eta_{ij} \quad (5.4)$$

Another scenario which shows CTA overlapping for meshed WPANs is given in Figure 5.3. Two types for flows are shown in Figure 5.3 i.e. an intra-PAN flow between devices (S, D)

and inter-PAN flows between MPNCs (1, 2), (3, 1) and (3, 4) respectively. In case of meshed WPANs, the CTA overlapping for intra-PAN flows is determined by the MPNC of that cluster while the CTA overlapping for inter-PAN flows are determined and arranged by the ref-MPNC. From Figure 5.3, it can be seen that if MPNC 1 transmits at power p_1 to MPNC 2, then MPNC 3 cannot transmit to MPNC 4 at the same time. However, when MPNC 1 lowers the power to p_2 , then MPNC 3 can transmit to MPNC 4 at power p_2 respectively. From (5.4), the capacity improvement c_T for the scenario of Figure 5.3 is given by

$$c_T = \eta_{SD} + \eta_{12} \quad (5.5)$$

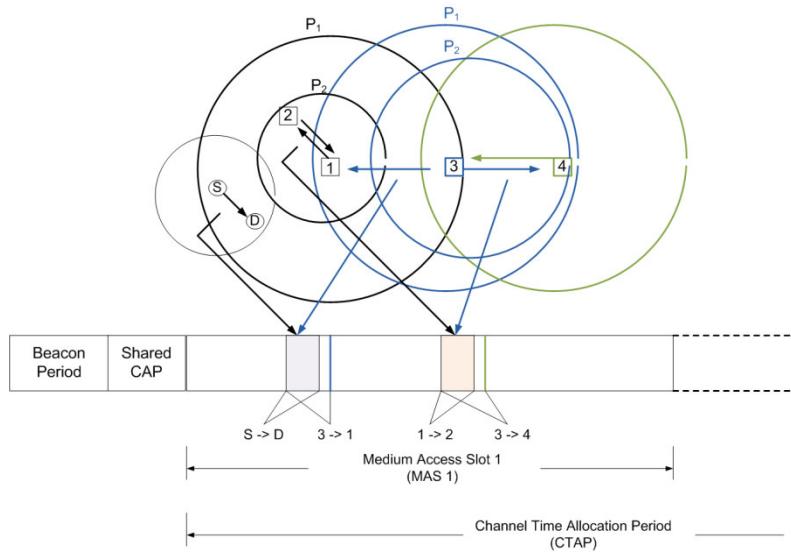


Figure 5.3 CTA Overlapping in a Meshed WPAN

The arrangement of CTA overlapping based on overlap ratio can be illustrated by a simple example given in Figure 5.4. Assuming for flow f_{12} , the corresponding vector of mutually exclusive device IDs $M_1 = \{3, 4, 5, 6\}$ while for flow f_{43} , the corresponding vector $M_4 = \{7, 8, 9, 10\}$, then the total capacity improvement for Figure 5.4(a) is

$$c_T^A = \eta_{12} + \eta_{56}$$

While the total capacity improvement for Figure 5.4(b) is

$$c_T^B = \eta_{87} + \eta_{9-10} + \eta_{12}$$

Therefore, by using CTA overlap ratio, the arrangement of CTA overlapping improves the capacity compared to c_T^A by

$$c_T^B - c_T^A = (\eta_{87} - \eta_{56}) + \eta_{9-10}$$

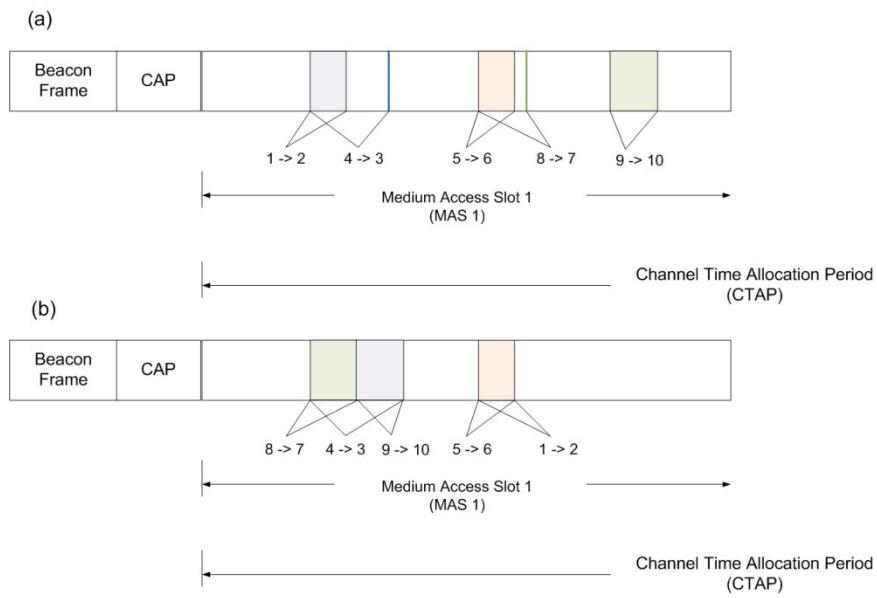


Figure 5.4 An example re-arrangement of CTA overlapping based on CTA overlap ratio

The CTA overlap procedure consists of the following steps

- When a new flow f_{ij} completes, the handshake process, then based on its vector of mutually exclusive devices M_i , the overlap ratio ϱ_{il} is calculated for each possible pair (s_i, m_l)
- Based on the highest value of ϱ_{il} for a particular pair, the CTA overlapping is selected
- If even for the pair of flows f_{ij} and f_{lm} which have been allocated the same CTA, the value of $\varrho_{il} < 1$, then in case a new flow initiates, which is mutually exclusive to both f_{ij} and f_{lm} , it can be allocated a CTA offset starting from the end of η_{ij} if $\eta_{ij} < \eta_{lm}$ and if $\eta_{ab} \leq \eta_{lm} - \eta_{ij}$ where η_{ab} is the CTA size of the new flow f_{ab} .

The procedure can be presented as the solution of the following problem

$$\max(\varrho_T) \quad (5.6)$$

The problem can be presented by a bipartite graph bipartite graph $G(X \cup Y, E)$ as shown in Figure 5.5 for an example arrangement. Each node on the left of the bipartite graph represents a flow f_{ij} ($f_{ij} \in X$) with its vector M_i ($M_i \in Y$) which consists of nodes on the right. If the flow f_{ij} is mutually exclusive to all the flows $f_{lm} \in Y$, then $M_i \cap Y = \emptyset$. In case of Figure 5.5, in which the flow $f_{12} \in X$ is mutually exclusive to only two flows $f_{56}, f_{9-10} \in Y$, then $M_i \in Y$. The cost of edge $c_{il} \in E$ corresponds to the overlap ratio ϱ_{il} . Since in maximum matching a flow f_{ij} can only connect through a distinct edge to a distinct f_{lm} ,

therefore, only a single value of ρ_{il} is selected for each pair after the matching. The total overlap ratio ρ_T is then selected based on the sum of individually selected values of ρ_{il} such that the value of ρ_T is maximized.

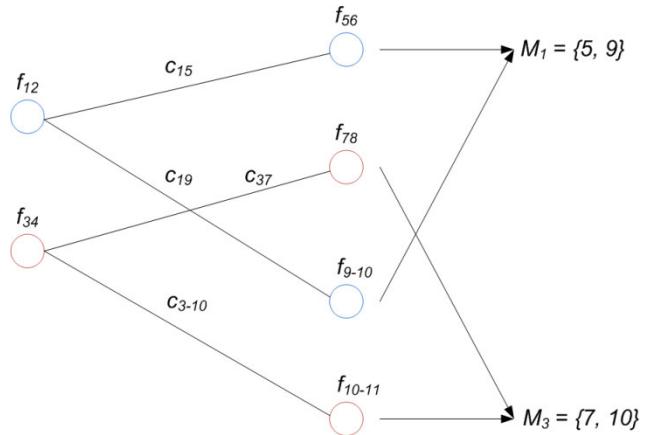


Figure 5.5 An example bipartite arrangement for the solution of (5.6)

5.3 Simulation Results

The simulation setup to check the effectiveness of the proposed approach to arrange overlapping CTAs based on CTA overlap ratio, consisted of four Node modules developed in the OPNET Modeller. The first one was a normal device module which was an extended version of a WPAN device developed for simulation purposes in previous chapters. The second was a single hop WPAN PNC, the third a MPNC and the fourth a ref-MPNC. All the modules used in the simulation scenario were upgraded versions of models used in previous simulation setups. The extensions consisted of handshake procedures in the WPAN devices, PNC and the MPNC. The PNC, MPNC and ref-MPNC were upgraded to identify and arrange the overlapping CTAs as well as re-arrange them if possible based on the CTA overlap ratio.

The simulation scenarios were chosen to check the improvement in superframe capacity due to the power control procedure as well as the proposed CTA overlap based approach. The first simulation scenario was setup to investigate such possible improvements in single hop WPANs while the second scenario was setup for a two hop meshed WPAN topology. The simulation setup for single hop WPANs can be seen in the snapshot of Figure 5.6. The single hop WPAN consists of 17 devices. The maximum throughput limit for devices is kept ≤ 4.5 Mbps. The simulation run is for 5 hours with the mean arrival rate of requests equal to

50 seconds. The superframe capacity is the maximum considered i.e. 0.065535 seconds. The Channel Time Allocation Period (CTAP) capacity for communication is taken to be 0.06400 seconds with the rest of the superframe capacity used for Beacon transmission and the control information exchange in the Contention Access Period (CAP).

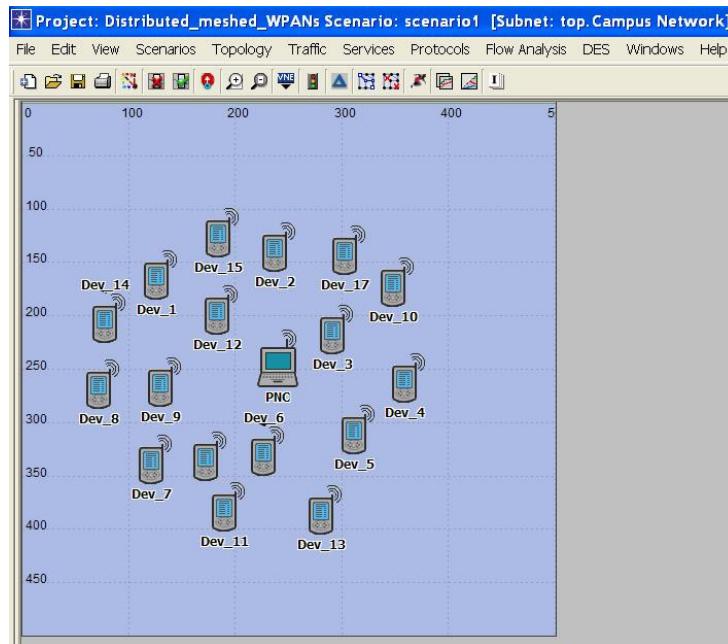


Figure 5.6 Simulation scenario for a single hop WPAN

The total superframe utilization by all the concurrent flows can be seen in Figure 5.7 in which the three cases are compared. The first case is when there is no power control algorithm; the second is with the use of power control algorithm while the third is with the power control algorithm augmented with the CTA overlap ratio based arrangement. It can be seen that in the time interval between 1.2 to 1.6 hours, the utilization increases beyond the maximum superframe capacity with the power control based and the one based on power control with overlap ratio which indicates use of overlapping CTAs to accommodate mutually exclusive flows. After 1.6 hours, the superframe utilization goes back to 0.065535 approximately which indicates that there are no mutually exclusive flows. However, in the time interval between 3.4 – 3.7 hours, the proposed approach shows further improvement over the power control based approach.

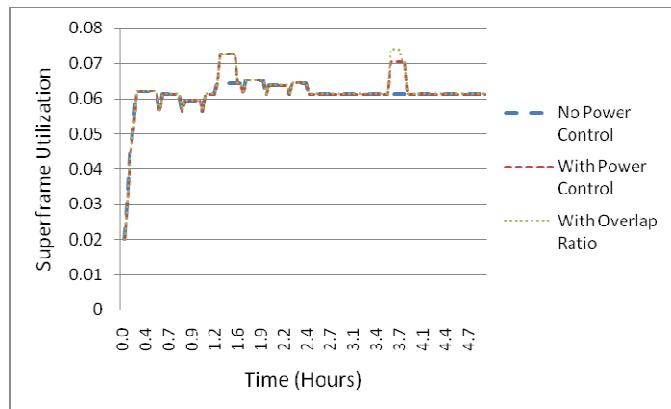


Figure 5.7 Superframe Utilization based on power control with Overlap Ratio based arrangement

The setup for a two hop meshed WPAN can be seen in Figure 5.8. The ref-MPNC monitors the allocations in the two hop neighbourhood to track any possibility of overlapping CTAs and the corresponding re-arrangement based on CTA overlap ratio.

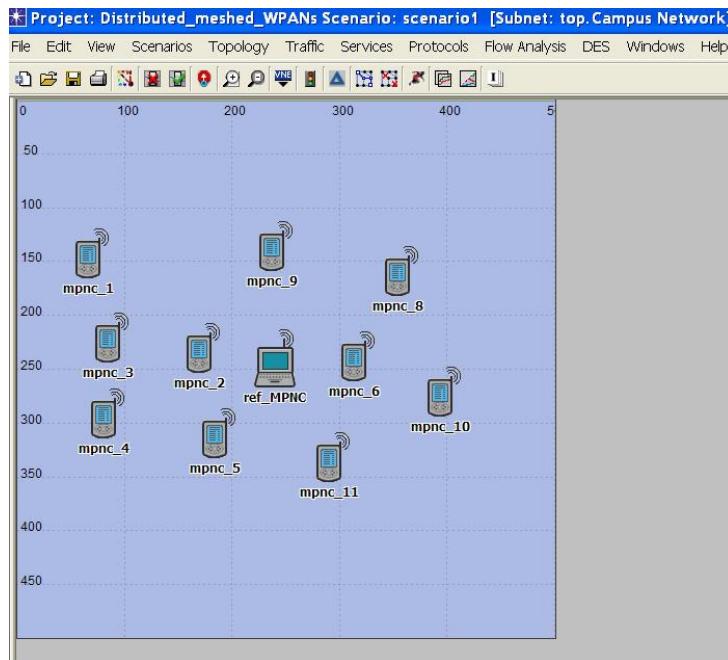


Figure 5.8 Simulation scenario for a two hop meshed HDR WPAN

The maximum throughput for intra-PAN flows is limited to 2.4 Mbps while for Inter-PAN flows it is limited to 2 Mbps. The reason is to accommodate as much MPNCs as possible within the superframe capacity. It should be noted that distributed allocation is considered in this simulation setup unlike the proposed centralized scheme in chapter 3. The mean arrival rate for intra-PAN flows is 300 seconds while the same is used for inter-PAN flows. The number of intra-PAN devices in each cluster vary from 10 - 15. The results for the

meshed HDR WPAN based setup can be seen in Figure 5.9. The three cases compared for HDR single hop WPANs are compared for the case of meshed HDR WPAN also. It can be seen that with the CTA overlap based approach, somewhat improved performance is observed when compared with the single hop approach. A reason which can justify this is that in meshed WPANs, the number of MPNCs and devices involved in inter-PAN communication is greater and therefore, the ref-MPNCs have more options to consider in re-arranging overlapping CTAs to achieve better capacity utilization. The improvement in capacity also depends on the duration of those flows which have overlapping CTAs.

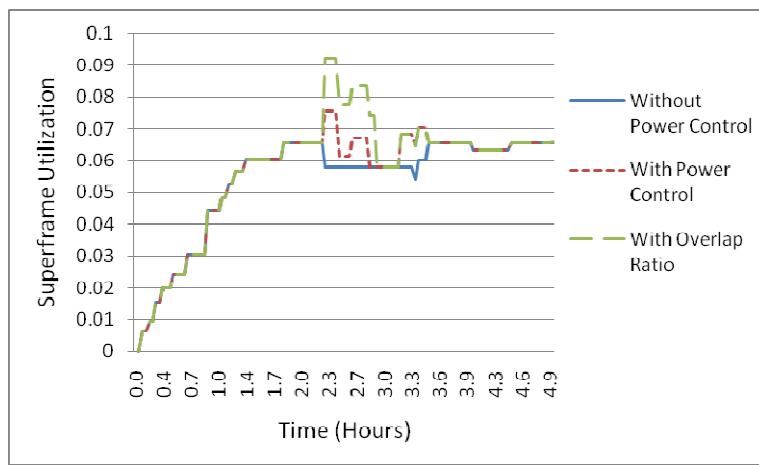


Figure 5.9 Superframe Utilization based on power control with Overlap Ratio based arrangement for Meshed WPAN

5.4 Conclusion

From the simulation analysis, it is deduced that the probability of capacity improvement in the superframe via either the CTA overlap ratio or the normal power control procedures depend on the number of overlapping CTAs. The possibility of having more overlapping CTAs depends on

- The location of devices in the piconet
- The distance (and the resulting transmission range) between the source-destination pairs

The handshake procedure proposed helps in reducing the transmission power to a minimum level possible to increase the probability of devices which can be allocated

overlapping CTAs. The re-arrangement of overlapping CTAs based on the CTA overlap ratio further ensures that not only the superframe capacity is utilized in the most efficient manner but also the spatial re-use. The preliminary investigations of the proposed approach based on the simulation results are encouraging and lead the way for further work in this area. The same procedure can be modified further to locate those overlapping CTAs for which the overlap ratio is less than 1, and hence the unutilized space can be allocated to those VBR flows which are mutually exclusive to such overlapping CTAs and therefore, instead of the devices involved in VBR flows requesting the PNC for extra CTA time when their buffer size is increased, they can send the extra traffic in the unutilized spaces in the overlapping CTAs. Such extensions are currently under investigation and further improvements are expected from them.

5.5 References

- [1] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2004, pp. 114-128.
- [2] Chih-Yung Chang; Chao-Tsun Chang; Po-Chih Huang, "Dynamic channel assignment and reassignment for exploiting channel reuse opportunities in ad hoc wireless networks," *Communication Systems, 2002. ICCS 2002. The 8th International Conference on*, vol.2, no., pp. 1053-1057 vol.2, 25-28 Nov 2002.
- [3] P. Kyasanur and N. H. Vaidya, "Routing and link-layer protocols for multi-channel multi-interface ad hoc wireless networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 10, no. 1, pp. 31-43, January 2006.
- [4] Hua Yu; Mohapatra, P.; Xin Liu, "Dynamic Channel Assignment and Link Scheduling in Multi-Radio Multi-Channel Wireless Mesh Networks," *Mobile and Ubiquitous Systems: Networking & Services, 2007. MobiQuitous 2007. Fourth Annual International Conference on*, vol., no., pp.1-8, 6-10 Aug. 2007.
- [5] A. Raniwala, K. Gopalan, and T.-C. Chiueh, "Centralized channel assignment and routing algorithms for multi-channel wireless mesh networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 8, no. 2, pp. 50-65, April 2004.
- [6] Haq, A.; Naveed, A.; Kanhere, S.S., "Securing Channel Assignment in Multi-Radio Multi-Channel Wireless Mesh Networks," *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, vol., no., pp.3111-3116, 11-15 March 2007
- [7] Das, A.K.; Vijayakumar, R.; Roy, S., "WLC30-4: Static Channel Assignment in Multi-radio Multi-Channel 802.11 Wireless Mesh Networks: Issues, Metrics and Algorithms," *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, vol., no., pp.1-6, Nov. 27 2006-Dec. 1 2006.
- [8] Tang, J.; Xue, G.; Zhang, W., "Maximum Throughput and Fair Bandwidth Allocation in Multi-Channel Wireless Mesh Networks," *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, vol., no., pp.1-10, April 2006
- [9] M. Kodialam and T. Nandagopal, "Characterizing the capacity region in multi-radio multi-channel wireless mesh networks," in *MobiCom '05: Proceedings of the 11th annual international conference on Mobile computing and networking*. New York, NY, USA: ACM Press, 2005, pp. 73-87.
- [10] Lozano, A.; Cox, D.C., "Integrated dynamic channel assignment and power control in TDMA mobile wireless communication systems," *Selected Areas in Communications, IEEE Journal on*, vol.17, no.11, pp.2031-2040, Nov 1999.
- [11] van Hoesel, L.; Havinga, P., "Collision-free Time Slot Reuse in Multi-hop Wireless Sensor Networks," *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2005. Proceedings of the 2005 International Conference on*, vol., no., pp. 101-107, 5-8 Dec 2005.
- [12] J. Gronkvist, "Novel assignment strategies for spatial reuse tdma in wireless ad hoc networks," *Wireless Networks*, vol. 12, no. 2, pp. 255-265, April 2006.
- [13] Y. Tseng, H. Wu, K. Shieh, G. Chen, "Effcient Resource Allocation for IEEE 802.15.3(a) Ad Hoc Networks," *INTELLCOMM'04*, pp. 128–142, Bangkok, Thailand, Nov. 2004.

- [14] Yi-Hsien Tseng; Eric Hsiao-kuang Wu; Gen-Huey Chen, "Maximum traffic scheduling and capacity analysis for IEEE 802.15.3 high data rate MAC protocol," *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th* , vol.3, no., pp. 1678-1682 Vol.3, 6-9 Oct 2003.
- [15] S. Max, E. Weiss, and G. R. Hertz, "Benefits and limitations of spatial reuse in wireless mesh networks," in *MSWiM '07: Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*. New York, NY, USA: ACM, 2007, pp. 244-251.
- [16] Ramaiyan, V.; Kumar, A., "On the Limits of Spatial Reuse and Cooperative Communication for Dense Wireless Networks," *Information Theory for Wireless Networks, 2007 IEEE Information Theory Workshop on* , vol., no., pp.1-5, 1-6 July 2007.

CHAPTER 6

6 Conclusions and Future Work

In this thesis, the aim of the contribution made was to resolve some of the issues of single hop and meshed HDR WPANs based on the IEEE 802.15.3 and IEEE 802.15.5 standards. The motivation for addressing the issues in HDR WPANs mainly came from relevant literature review that paved the way to identify those research bottlenecks that were not addressed thoroughly. Upon proposing different approaches in this thesis along with the performance evaluation of those approaches, it is strongly felt that there is room for further potential improvements related to the work done. Therefore, apart from the work done in this thesis, there are some other research issues and related avenues that need to be explored based on the current work. Therefore, this chapter consists of two parts; a summary of research done in each chapter and the possible future work related to the contributions in each chapter in order to make further extensions and hence create the possibilities of further performance improvements.

6.1 Inter-PAN Communication and Admission Control for HDR WPANs

The inter-PAN communication procedure proposed in Chapter 2 serves as a simple extension to the single hop HDR WPAN standard to improve the peer-to-peer connectivity between independent piconets. Without the inter-PAN extension, the communication between devices that belong to different piconets is somehow limited and cannot take place directly even though they are in the transmission range of each other. The modifications to enable the inter-PAN communication make use of the reserve fields² in the IEEE 802.15.3

² Reserve Fields are provided in order to include implementation specific commands or Information Elements to add extra functionalities in the IEEE 802.15.3 standard

standard. Therefore, the flexibility provided by the standard to make feasible extensions helps us in devising such an approach that does not violate the standard in any way. In order to check if the proposed approach is viable, appropriate capacity analysis is carried out to check the upper limit of devices that a superframe can support simultaneously. Since the analysis take only a single type of traffic into consideration at a time, a more pragmatic picture is given by the simulation results that consider dynamic traffic requirements of devices that are active and are allocated channel time by the PNC.

In case the number of devices with high throughput requirement increases, the chances of superframe capacity being congested increases. On top of that, when inter-PAN communication is considered and the same superframe capacity is used to support more than one piconet, there are serious capacity issues. Therefore, to handle the channel time requests appropriately, the proposed admission control algorithm becomes helpful. To augment the feasibility of the admission control algorithm and introduce fairness on the device level, the rejection ratio (RR) algorithm is proposed along with its extension i.e. Improved Rejection Ratio (IRR). The RR and IRR algorithms make sure that the requests are catered to in an even manner as much as possible. In other words, the objective of the RR and IRR algorithms is to make sure that the distribution of allocated requests is analogous to a uniform distribution as much as possible. The RR and IRR consider fairness on device level i.e. the fairness criterion is flat and the devices are not prioritized based on their traffic type. The proposed admission control strategy and the RR based algorithms although prove to provide satisfactory results, there are some issues which if addressed can further add to the improvement of proposed work.

1. As mentioned that the RR based algorithms do not give priority to devices based on their flow traffic type, however, in some cases it might become necessary to give priority to some devices based on the traffic type of their flow. An example in a home based WPAN can be a Television show that needs to be streamed to a recording device or a particular job that needs to be sent to the printer at a specified period of time. In office based WPANs, the minutes or recording of a meeting, or an interactive video or voice session that might need to be streamed to a user owned multimedia device has to have a higher priority than other

flows. Therefore, such devices need to be given priority when allocating channel time to requesting devices especially in cases when the superframe capacity is congested. In order to address this problem, further extensions to the RR based algorithms are in consideration.

2. The capacity analysis performed both analytically and via simulation results had a primary objective to determine an upper limit on the superframe of devices that can be supported simultaneously in the superframe. The traffic type and the flow durations of the active devices vary dynamically and therefore it is a non-trivial issue to give an exact estimate of the utilized superframe capacity when the same number of devices is considered in different piconets. Apart from that, the capacity analysis only considers a CTA size for VBR flows based on their mean throughput. This assumption serves the purpose of the proposed work in Chapter 2 but when more accurate capacity estimation is needed e.g. for scheduling, then the mean throughput of VBR flows is not a feasible parameter to consider on its own. A few approaches that deal with proposing traffic predictors and estimating VBR traffic are given in [1-5]. In [5], the VBR traffic estimation is used to allocate bandwidth in HDR WPANs. However, the approach considers a complete superframe duration as the inter-arrival time between two consecutive CTAs allocated to VBR traffic. In such cases, the flexibility with which the throughput can be allocated to devices might suffer because of considering only one frame/CTA. A somewhat better approach is therefore sought that can predict VBR traffic accurately as well as schedule the flows in such a way that different types of traffic flows can be accommodated in a flexible way with their QoS requirements intact.

3. In the related work section of chapter 2, some feedback based scheduling schemes were referred to, specifically [6-9]. The feedback based scheduling is specifically used for MPEG based VBR traffic in which the buffer size of a device can increase or decrease dynamically. In case a device has been allocated a larger CTA size than it needs, then the CTA capacity is inefficiently utilized. If a device is allocated a smaller CTA than its needs, then the buffer size for the device increases and its CTA is fully utilized. In both the cases, the device sends a channel time request command to the PNC to request for an appropriate change in its CTA size allocation. In feedback-based approaches, it is assumed that the PNC has enough superframe capacity to allocate to a device when it requires an increase in its

CTA size. Another assumption is that if a device requests for a smaller CTA size, then the capacity released by making the CTA size smaller is not allocated to another flow. Since for a VBR traffic flow, the mean rate and the peak rate is usually defined, the traffic can vary either between the mean and the peak rate, or it can drop below the mean rate. In order to restrict the frequent channel time requests by a device to the PNC to change its CTA size, the use of CAP can be investigated in cases when the VBR rate increases to a certain threshold above the capacity of the CTA size of a device, that the additional traffic which cannot be sent in the allocated CTA, can be instead sent in the CAP.

6.2 Resource Allocation Framework for Meshed HDR WPANs

In Meshed HDR WPANs, the superframe capacity is shared among MPNCs that belong to the same Beacon Period (BP). In order to ensure that the superframe capacity is not inappropriately utilized by any one or more MPNCs in the BP, a resource allocation framework is proposed in Chapter 3. The proposed scheme is centralized for a two-hop neighborhood but distributed on the network wide scale. In order to determine channel time requirement, each MPNC calculates its share by taking into account a history of previous incoming and outgoing intra-PAN and inter-PAN flows. The history is maintained in inter-PAN and intra-PAN traffic matrices. Based on the intra-PAN and inter-PAN request arrival rates, demand matrices are formulated from the traffic matrices. Using the demand matrices, either a slightly modified RAS algorithm or an adjusted assignment problem is used to estimate traffic and formulate an appropriate share for an MPNC. Based on the estimated share, an MPNC sends a channel time request to the ref-MPNC. Instead of sending channel time requests for individual flows that can significantly increase the control traffic, the estimated share takes into consideration the channel time requirements of the whole cluster or piconet controlled by the MPNC. The request is therefore called bulk reservation.

By using the bulk reservation, it is shown in chapter 3 that the control traffic overhead is reduced and the three proposed resource allocation policies are evaluated. Although the measure of fairness shown for each resource allocation policy is different, a reasonable value of fairness is shown to be maintained even in scenarios with high traffic demands

from the MPNCs. A drawback in the approach can sometimes be as a result of inefficient superframe utilization that can be because of incorrect channel time estimation for a bulk reservation by each MPNC. The inefficient superframe utilization might be a result of either over-estimation or under-estimation of channel time requirement of a cluster. The reason for occurrence of over-estimation or under-estimation is that each demand traffic matrix relies on the mean of previous allocated CTA sizes to a device. Since the actual requirement for a flow might deviate from the mean, it in turn gives rise to under-estimation or over-estimation when the complete bulk size of the reservation is calculated. Therefore, to avoid the inefficient use of superframe capacity by an MPNC, better traffic estimation techniques need to be investigated and the current proposed techniques need to be improved. The possible investigations that are envisaged to accomplish the goal are expected to encompass the following approaches.

1. Instead of the demand matrix having each element based on the mean of the previous allocated CTA sizes, an alternate approach based on the distribution of CTA size's for allocated flows can be considered. The CTA size that is the most probable based on the probability distribution is then considered as the likely candidate to be selected. In order to relate the calculation of probability to a certain distribution, the nature of distribution needs to be verified first. The commonly used comparison techniques from information theory like Kullback-Leibler divergence [10] and the Resistor-Average distance [10] (which is a symmetrical version of Kullback-Leibler divergence) can be used to check if the distribution is analogous to a parametric distribution or not. If not, then the non-parametric tests like the Kernel Density functions [11] can be used to calculate the probability of expecting a certain CTA size for the flow based on the actual distribution.
2. Cooperative approaches using utility functions can be used to allocate appropriate shares to requesting MPNCs using different traffic levels. The proposed resource allocation schemes in Chapter 3 can be compared with the utility based approaches for further insights into relative performance of proposed schemes.

6.3 Capacity and Delay Efficient Communication in Meshed HDR WPANs

Working towards achieving good communication efficiency is one of the research goals regarding communication networks. In Chapter 4, the communication efficiency is taken in the perspective of reduction in communication overhead. In case of HDR WPANs, the term time efficiency is introduced to indicate the overhead for each flow. The overhead for each flow depends on the total overhead bits in a CTA compared to the actual data bits. The time efficiency parameter is equal to the ratio of effective throughput to the actual throughput. Upon evaluation, it is revealed that those flows that use larger fragment size with multiple frames/CTA, the time efficiency is higher. Also, the time efficiency increases when the transmitter data rate is decreased for the same fragment size and number of frames/CTA. With time efficiency values of 90% and above, then using a larger fragment size, increasing the number of frames/CTA or reducing the transmitter data rate show little improvement in time efficiency but a substantial effect on the capacity utilization for the flow, which increases. Therefore, to achieve a suitable level of time efficiency, the upper limit on throughput and constraints on capacity has to be taken into account. To serve the purpose, two table search based algorithms i.e. FTS and PTS are proposed. Both of the algorithms give the same results but PTS has lower complexity (time of execution) than FTS.

To achieve efficient capacity utilization and reduce EED in multi-hop meshed WPANs, the metrics STS and STS-R are proposed which use CTA switching to reduce the queuing delay as well as take into consideration the transmitter data rates used by each MPNC to forward data to its peer MPNC along the source-destination path. The higher the data rate used by each link along a path, the higher is the preference for that path since it helps to transmit the same CTA using shorter duration hence better capacity utilization. To achieve a balance between capacity utilization and EED for a path, the tunable weight factor w is used to set the preference. The proposed STS and STS-R metrics are evaluated and their performance is compared with the MSMR metric. The results show that the STS and STS-R metrics not only prove to select capacity efficient paths but also paths with lower EED delay and can help reduce the EED further in many cases. The STS-R metric proves to provide further improvement in EED when compared to STS in some cases because it performs CTA

switching after a flow terminates in a link along the source-destination path. With STS, the CTA switching is performed just once in the beginning of route setup when the total cost of the route is being calculated. The reduction in queuing delay due to CTA switching is however dependant on not just one link but the final link or final hop. For any hop m , if there are queuing delay reductions in $m - n \quad \forall n = 1, 2, 3 \dots m - 1$ but no change in queuing delay at hop m for a flow f , then there is no change in the EED due to the reductions in $m - n$ hops. The future work related to STS and STS-R therefore is related to make further extensions in both metrics such that the probability of flows terminating in the final hops is taken into account to increase the chances of improvement by the two metrics especially STS-R. In short the following points describe the future aspirations related to the work done in chapter 4.

1. Most the scheduling schemes that deal with VBR traffic in HDR WPANs whether they be feedback based or not, send a request to the PNC to increase or decrease the CTA size based on their buffer size. The change in CTA size is requested so that those frames which have higher priority or those frames for which the deadline is about to expire are sent. For devices which have high throughput requirements and therefore have multi-rate CTAs reserved (multiple CTAs per superframe), the frequent change in CTA size might not be feasible. The reason is that for devices that change their CTA size to send the prioritized frames might run out of those frames while there are still CTAs of the same size left in the superframe for the same flow. Using the same fragment size might not be an efficient way to handle the remaining frames with the same CTA size that was booked for the priority frames. Dynamic fragmentation techniques therefore, need to be investigated to solve this issue. The FTS and PTS based algorithms proposed in chapter 4 can come in handy to solve such problems along with suitable dynamic fragmentation schemes in which the fragment size can be negotiated within the same superframe between the source and destination devices using the Implied –Acknowledgement (Imp-ACK) scheme [12].
2. To evaluate the STS and STS-R metrics, static routing is considered in the simulation model developed in the OPNET Simulator. Further insights and a more thorough performance evaluation can be obtained if a routing protocol is designed for meshed

WPANs which can make use of STS and STS-R and any possible upgrades or extensions to these two metrics. The routing protocol can be tree based or distributed based.

3. At present the STS and STS-R metrics take into account the EED and capacity utilization of a flow on a particular path. Since the queuing delay is the most significant factor contributing to the EED, those factors that help in reducing the queuing delay and signify the potential of a path to aid in further improvements need to be identified and possibly incorporated in the two metrics. Some of the criteria that define the desirable characteristics of a path that can prove to be helpful are

- Select the path with the lowest superframe utilization for each link
- Select the path with the lowest arrival rate of flows for each link
- Select the path, the final hop of which has the greatest probability to be moved to the left as much as possible after a single (STS) or multiple (STS-R) iterations of CTA switching procedure

The feasibility of the points mentioned above can be investigated and if appropriate can be used with the STS and STS-R metrics as an extension.

6.4 Capacity Improvements in WPANs with Power Control

In a wireless medium which is shared usually by more than one device, the capacity of a channel usually proves to be one of the major bottlenecks. The two approaches that have emerged as the most popular ones to improve the channel capacity in ad-hoc networks are spatial reuse through power control and by using multiple channels. In the thesis, the former approach is preferred over the later because there is a high possibility of reduced energy consumption when a single radio interface is used and power control procedures are adopted to reduce the power for a flow if possible in such a way, that another flow can initiate in the same time space. The disadvantage of this approach is that although reduced energy consumption is possible with a certain probability of supporting additional flows via power control, the multiple-channel/multi-radio approach ensures that the capacity of the channel is increased either two fold or multifold based on the number of radios used per device. A comparison between the power control approach and the multiple-channel

approach therefore can prove to be useful. A brief description of the related future work is as follows:

1. The WPAN family of standards has a common goal of utilizing low cost devices to achieve low power communications due to their limited battery life. Therefore, the use of multiple radios per devices raises questions on its applicability because of increased energy consumption which can drain the battery life even quicker. For similar reasons, the details and specifications for multi-channel communication by devices equipped with multiple radios are given for the mesh extensions of WLAN i.e. IEEE 802.11s and for WiMAX i.e. IEEE 802.16e. However, for the WPAN mesh extension (IEEE 802.15.5), the specifications for multi-channel usage are not included. The performance evaluation of multi-radio devices in single hop and meshed WPANs can substantiate the use of such approaches. The energy consumption can be compared with the power control approach. If there is not much difference in the energy consumption but there are significant improvements in the capacity i.e. more devices can be accommodated, the use of second approach can then be considered in the future.

The evolution of handheld devices regarding the processing power and memory along with the supported applications have been nothing less than phenomenal over the last decade. With the capability of handling rich multi-media content, the use of handheld devices for file sharing, multi-media streaming, multi-player gaming, HD Audio and Video on demand etc have increased. HDR WPANs can prove to be one of the most likely candidates to support such applications along with many more in a limited operating space. With more and more people starting to use their mobile phones or other portable devices to serve most of their computing needs as well as to communicate, the potential of HDR WPANs to provide the necessary Air Interface to handle such communication depends on the resolution of common issues which are either specific to WPANs or generally found in all wireless communication systems. The work done in this thesis is one step forward to deal with common issues like limited superframe capacity, resource reservation and allocation, capacity efficiency and reducing EED for meshed WPANs. The possible extensions mentioned in the future work can further aid in making WPANs a highly likely candidate for networking in a limited space between devices.

6.5 References

- [1] Yi-Hsien Tseng; Wu, E.H.-K.; Gen-Huey Chen, "Scene-Change Aware Dynamic Bandwidth Allocation for Real-Time VBR Video Transmission Over IEEE 802.15.3 Wireless Home Networks," *Multimedia, IEEE Transactions on* , vol.9, no.3, pp.642-654, April 2007.
- [2] Doulamis, A.D.; Doulamis, N.D.; Kollias, S.D., "An adaptable neural-network model for recursive nonlinear traffic prediction and modeling of MPEG video sources," *Neural Networks, IEEE Transactions on* , vol.14, no.1, pp. 150-166, Jan 2003.
- [3] Adas, A.M., "Using adaptive linear prediction to support real-time VBR video under RCBR network service model," *Networking, IEEE/ACM Transactions on* , vol.6, no.5, pp.635-644, Oct 1998
- [4] Kwong, R.H.; Johnston, E.W., "A variable step size LMS algorithm," *Signal Processing, IEEE Transactions on* , vol.40, no.7, pp.1633-1642, Jul 1992.
- [5] Kuo, W.-K.; Lien, S.-Y., "Dynamic resource allocation for supporting real-time multimedia applications in IEEE 802.15.3 WPANs," *Communications, IET* , vol.3, no.1, pp.1-9, January 2009.
- [6] G. Boggia; P. Camarda; L. A. Grieco, "Scheduling channel time allocations in 802.15.3 WPANs for supporting multimedia applications," *Wireless Communications and Mobile Computing*, pp. 1530-8669, 2009.
- [7] Sun-Myeng Kim; Young-Jong Cho, "Scheduling scheme for providing QoS to real-time multimedia traffics in high-rate wireless PANs," *Consumer Electronics, IEEE Transactions on* , vol.51, no.4, pp. 1159-1168, Nov 2005.
- [8] Liu, X., Dai, Q. AND Wu, Q. 2004. Scheduling algorithms analysis for MPEG-4 traffic in UWB. In *2004 IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall*.
- [9] B.-S. Kim, S. Kim, Y. Fang, and T. Wong, "Feedback-assisted mac protocol for real time traffic in high rate wireless personal area networks," *Wireless Networks*.
- [10] D. H. Johnson and S. Sinanovic, "Symmetrizing the Kullback-Leibler distance," *Technical report, Rice University*, 2001.
- [11] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065-1076, 1962.
- [12] "IEEE Standard for Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Specific requirements Part 15.3: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless

Chapter 6

Personal Area Networks (WPANs) Amendment 1: MAC Sublayer," *IEEE Std 802.15.3b-2005* (*Amendment to IEEE Std 802.15.3-2003*) , vol., no., pp. 0_1-146, 2006.