AN INVESTIGATION INTO THE SCALE-FREE NATURE OF HETEROGENEOUS NETWORKS

Thesis submitted for the
Degree of Doctor of Philosophy of
UCL, University College London.

Jason J. L. Spencer

Department of Electronic and Electrical Engineering
University College London
September 2008

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Jason Spencer, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

To my family;

especially Mary and Joseph,

my parents.

ACKNOWLEDGEMENTS

During the course of this research many people have helped and supported me, academically, technically, administratively, emotionally and financially. I would like to list as many as I can, for I am truly grateful. Thank you.

This research and thesis would not have become a reality were it not for the help, support, foresight, advice and knowledge of my supervisors, Prof. Polina Bayvel, Dr. Miguel Rio and Prof. John J. O'Reilly, and my friends and colleagues, Mick Flanagan, David Griffin, Theodoros Michalareas, Jonas Griem, Eduarda Mendes Rodriguez, Hamed Haddadi, Eleni Mykoniati, Aleks Lazarevic, Ioannis Liabotis, Ognjen Prnjat, Wayne Jarrett, Andreas Antonopoulos, Will Jenkins, Lawrence Latif, Tammy Saunders, Tomasz Paszkowski, Ewelina Szkutnik, John Mitchell, Panos Georgatsos, Barbara Głowala, Mohammed Boucadair, Anna Słupska, Michael Dueser, Alejandra Zapata, Ignacio de Miguel, Chris Matrakidis, Mark Wilby, Chris Todd, Walter Eaves, Natalie May, Tom Crummey, Lee Heaghney, Scott Landers, Sara Collins, Sharon Edwards, Michael Gladstone, Daniel Nicholls and many many more.

I would like to thank my colleagues at BT for their support, Ian Hawker, Dave Johnson, Dominic Marcenac, Paul France and Andrew Hastie.

Of course, none of this would have happened without the support of my mother, sister and the rest of my family.

AN INVESTIGATION INTO THE SCALE-FREE NATURE OF HETEROGENEOUS NETWORKS

JASON SPENCER

Department of Electronic and Electrical Engineering
University College London
September 2008
Supervisors: Prof. Polina Bayvel, Dr. Miguel Rio

ABSTRACT

In order to support a wide variety of services, to different user types, and under a variety of geographic situations, telecommunications networks are typically composed of a variety of layers and heterogeneous technologies. Layers (in terms of the OSI 7 layer model) such as the transmission layer (e.g. WDM), the data link layer (also known as the transport network e.g. SDH, Ethernet) and the network layer (e.g. IP). These layers may also contain logical layers within them such as virtual paths, as well as overlay networks such as a peer-to-peer system. No single layer is independent of the adjacent layer and the provisioning requirements of one layer become the demand on the layer below. Similarly the available resources become the delivered quality of service to the layer above.

This thesis is concerned with the design aspects of various layers and how they affect each other's topology. The thesis' main focus is topological analysis and modelling of layers, and it presents a detailed analysis of a deployed national SDH network, examining bandwidth distribution, topology, geography and the demand pattern. The thesis finds that even the strictly planned and provisioned SDH network, whose architecture contains explicit structures and hierarchy, has notable power-law traits in various metrics of the topology; traits similar to those which have been shown to exist in the Internet, as well as non-technological networks such as social graphs. There is also an examination of the protocols and architectures of the IP and SDH standards for features that affect topological development.

With a better understanding of the layers, design goals and assumptions are deduced and implemented in a new topology simulator called MITIE. MITIE (Modular Inter-layer feedback Topology InvEstigation tool and simulator) is a tool designed to investigate inter-layer feedback and differs from existing topology generators in that it considers the effect of serviced demands and allows the capacity usage to affect the further development of the topology. The thesis presents results from a series of experiments with MITIE and demonstrates that as the network is re-designed to accommodate demand, it can tend to power-law compliant topologies under the correct circumstances.

Such a reactive topology model could also be used to investigate the effect of topological change and the effect of increasing the number of layers (such as adding MPLS), or the use of peer-to-peer overlay networks, or the decrease of the number of layers (IP over WDM). The model could also be used to investigate link and node failure/addition and the real effect which will propagate through the rest of the multi-layer network.

Glossary of terms

{,,,}	MITIE {G,R,A,B} (Growth,Removal,addition A, addition B) tuple
ABR	Area Border Router
ADM	Add-Drop Multiplexor
AE	Absolute Euclidean – part of GRAB tuple
AL	Average Load – part of GRAB tuple
APNIC	Asia Pacific Network Information Centre
APS	Automatic Protection Switching
AR	Average Routes – part of GRAB tuple
ARDP	Absolute Route-Distance Product – part of GRAB tuple
ARIN	American Registry for Internet Numbers
ARPANET	Advanced Research Projects Agency Network
ARPL	Absolute Route Per Link – part of GRAB tuple
AS	Autonomous System
ASBR	Autonomous System Border Router
ASN	Autonomous System Number
ASON	Automatically Switched Optical Network
ASTN	Automatically Switched Transport Network
ATM	Asynchronous Transfer Mode
AUG	Administrative Unit Group
BA	Albert-Barasbasi (topology model)
BGP	Border Gateway Protocol
BT	British Telecom
CCR	Clustering Coefficient Rank
CI	Confidence Interval
CIDR	Classless Inter-Domain Routing
DFD	Degree Frequency Distribution
DPP	Dedicated Protection Path
DR	Degree Rank
DSxxx	Digital Signal xxx
DXC	Digital Cross Connect
E-BGP	Exterior-BGP
EGP	Exterior Gateway Protocol
ER	Erdös-Rényi (topology model)
FIB	Forwarding Information Based
FIT	Failutes In Time
FKP	Fabrikant, Koutsoupias, Papadimitriou (topology model)
FR	Frame Relay
FTP	File Transfer Protocol
GFP	Generic Framing Protocol
GMPLS	Generalised Multi-Protocol Label Switching
GRAB	Growth, Removal, addition A, addition B tuple from MITIE
HOP	High Order Path
HTTP	HyperText Transfer Protocol
I-BGP	Interior-BGP
ICMP	Internet Control Message Protocol
IG	Incremental Growth
IGP	Interior Gateway Protocol
IGRP	Interior Gateway Routing Protocol
IP	Internet Protocol

IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
IS-IS	Intermediate System to Intermediate System
ISP	Internet Service Provider
ITU	International Telecommunications Union
IXP	Internet eXchange Point
JANET	Joint Academic Network
LACNIC	Latin American and Caribbean Internet Addresses Registry
LAN	Local Area Network
LCAS	Link Capacity Adjustment Scheme
LOP	Low-Order Path (SDH)
LPL	Load Per Link
LPN	Load Per Node
LSA	Link State Advertisement
MAN	Metropolitain Area Network
MED	Multi-Exit Discriminator
MIB	Management Information Base
MITIE	Modular Inter-layer feedback Topology InvEstigation tool and simulator
MS	Multiplex Section (SDH)
MS-DPRing	Multiplex Section Dedication Protection Ring
MSLP	Multiple Section Linear Protection
MSOH	Multiplex Section Overhead
MS-SPRing	Multiplex Section Shared Protection Ring
MTTF	Mean Time To Failure
NAP	Network Access Point
OA&M	Operations, Administration and Management
OSPF	Open Shortest Path First
PCM	Pulse Code Modulation
PDH	Plesiosynchronous Digital Hierarchy
PFP	Positive Feedback Preference (topology model)
POS	Packet Over SONET/SDH
PPPoA	Point-to-Point Protocol over ATM
PPPoE	Point-to-Point Protocol over Ethernet
PVC	Permanent Virtual Circuit
QoS	Quality of Service
RDP	Routes-Distance Product – part of GRAB tuple
RIB	Routing Information Base
RIP	Routing Information Protocol
RIPE	Réseaux IP Européens
RPL	Routes Per Link – part of GRAB tuple
RPR	Resilient Packet Ring
RSOH	Regenerator Section Overhead
SDH	Synchronous Digital Hierarchy
SKA	Sender-Keeps-All
SNCP	Sub-Network Connection Protection
SNMP	Simple Network Management Protocol
STL	Standard Template Library
STM-x	Synchronous Transport Module level - x
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TL	Total Load – part of GRAB tuple

TOS	Type Of Service
TR	Total Routes – part of GRAB tuple
TTL	Time To Live
TUG	Tributary Units Group
UDP	User Datagram Protocol
VC	Virtual Container (SDH)
VC	Virtual Circuit (ATM)
PVC	Permanent Virtual Circuit
VPN	Virtual Private Network
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
xDSL	x-Digital Subscriber Line (where x is V(ery high),A(synchronous),S(ynchronous))

TABLE OF CONTENTS

CHAPTER 1 MULTI-LAYER NETWORK PROVISIONING AND INTER-LAYER FEEDBACK 22

1.1 INTRODUCTION 22

1.2 NETWORK PROVISIONING 22

1.2.1 A Typical Multi-layer Network 25

1.2.2 Network Events, Timescales and Inter-Layer Feedback 26

1.3 NETWORK DESIGN AND COST FUNCTIONS 30

1.4 SCOPE FOR RESEARCH: INTER-LAYER FEEDBACK AND NETWORK TOPOLOGY 35

1.5 THESIS STRUCTURE 36

1.6 RESEARCH CONTRIBUTIONS 37

1.7 CONCLUSION 39

CHAPTER 2 TELECOMMUNICATIONS NETWORK TOPOLOGY MODELLING 40

2.1 INTRODUCTION 40

2.2 NETWORK MODELLING 41

2.2.1 Power-laws in Telecommunications Networks 41

2.3 TOPOLOGY MODELS 42

2.3.1 Structured models 43

2.3.1.1 Transit-Stub 44

2.3.1.2 GHITLE: Generator of Hierarchical Internet Topologies using Levels 44

2.3.1.3 IGen: Topology generation through network design heuristics 44

2.3.1.4 AStop: A New Topology Generator at the AS Level 44

2.3.1.5 Inet: Internet Topology Generator 45

2.3.2 Random Models 47

2.3.2.1 Erdős-Rényi (ER) random graph model 47

2.3.2.2 Albert-Barábasi (BA) model 48

2.3.2.3 GLP (Generalized Linear Preference) model 50

2.3.2.4 IG (Interactive Growth) model 50

2.3.2.5 PFP (Positive Feedback Preference) model 51

2.3.2.6 PLOD (Power-Law Out-Degree) model 51

2.3.2.7 Waxman Topology model 52

2.3.2.8 Doar-Leslie Topology Model 52

2.3.2.9 Watts-Strogatz Small-World model 53

2.3.2.10 FKP (Fabrikant, Koutsoupias, Papadimitriou) model 53

2.4 MODEL IMPLEMENTATIONS 54

2.4.1 GT-ITM: The Georgia Tech Internetwork Topology Models 54

2.4.2 BRITE: the Boston university Representative Internet Topology gEnerator 55

2.5 THE DEVELOPMENT AND APPLICABILITY OF NETWORK MODELS 55
2.6 SCOPE FOR MODEL RESEARCH 57
2.7 CONCLUSIONS 58

CHAPTER 3 THE NETWORK LAYER: IP – THE INTERNET PROTOCOL 60

3.1 CHAPTER INTRODUCTION 60
3.2 IP (INTERNET PROTOCOL) NETWORKS 60
3.3 IP NETWORK PROTOCOLS AND ROUTING 62
3.3.1 Routing Hierarchies: Nodes, networks and areas 63
3.3.2 IP Addressing 64
3.3.3 IP Routing 65
3.3.3.1 Intra-Domain routing: Interior Gateway Protocol, IGP 66
3.3.3.2 Inter-Domain routing: Exterior Gateway Protocol, EGP 68
3.4 THE INTERNET 72
3.4.1 ISP Peering, agreements and BGP Policies 73
3.4.1.1 Interconnection Business models 73
3.4.1.2 Network Interconnection 74
3.4.2 Commercial IP Networks and Routers 75
3.5 INTERNET MAPPING 76
3.5.1 Internet Mapping Techniques 76
3.5.1.1 SNMP, the Simple Network Management Protocol 76
3.5.1.2 ICMP Ping and the IP Record Route option 76
3.5.1.3 Traceroute probing 77
3.5.2 Current Mapping Software and Methods 79
3.6 THE INTERNET TOPOLOGY 81
3.6.1 Traceroute Experiment: IP router topology in Europe 81
3.6.1.1 Experiment Results 82
3.6.1.2 Experiment Accuracy 85
3.6.2 The Faloutsos' Topology Observations 89
3.6.2.1 The rank exponent 90
3.6.2.2 The outdegree exponent 91
3.6.2.3 The hop-plot exponent 93
3.6.2.4 The eigenvalue exponent 94
3.6.2.5 Comments on Faloutsos' Empirical Study and other studies 95
3.6.3 Observations on the geographic location of Internet resources 95
3.7 INFLUENCES OF IP NETWORK DESIGN ON TOPOLOGY MODELLING 97
3.8 CHAPTER OVERVIEW AND CONCLUSIONS 98

CHAPTER 4 TRANSPORT NETWORKS: SDH – THE SYNCHRONOUS DIGITAL HIERARCHY 99

4.1 CHAPTER INTRODUCTION 99
4.2 THE SYNCHRONOUS DIGITAL HIERARCHY 99
4.2.1 SDH Overview *99*
4.2.1.1 SDH Network Elements 100
4.2.1.2 SDH Transport Hierarchy 101
4.2.1.3 SDH Planning 104
4.2.1.4 Related Tranport Technologies: PDH, GFP, VCat, LCAS 110
4.3 BRITISH TELECOM'S SDH NETWORK 112
4.3.1 Planned Topology *112*
4.3.2 The Data Set *113*
4.3.3 Topology Analysis *116*
4.3.3.1 Circuit Topology 116
4.3.3.2 SDH Network Geography 125
4.3.3.3 The spatial distribution of topology 130
4.3.3.4 SDH Bandwidth 133
4.3.3.5 Network Hierarchy 138
4.3.3.6 Network Routing Efficiency 141
4.3.3.7 Demand Topology 146
4.4 CONCLUSIONS 153

CHAPTER 5 SINGLE LAYER AND MULTI-LAYER NETWORK TOPOLOGY MODELLING 155

5.1 INTRODUCTION 155
5.2 THE MITIE TOPOLOGY SIMULATION TOOL 156
5.3 EXPERIMENTAL METHOD 158
5.3.1 On the use of line-fitting, least-squares and R^2 in network metrification *159*
5.3.2 Initial network conditions *159*
5.3.3 On Deterministic and Stochastic Node and Link Selection *161*
5.4 NETWORK GROWTH MODELLING 162
5.4.1 Single Layer Network and Connectivity Growth *162*
5.4.2 Multi-layer Network Growth *164*
5.5 NETWORK EROSION MODELLING 169
5.5.1 Single Layer Network Erosion *170*
5.5.2 Multi-layer Network Erosion *172*
5.5.3 Comparison of single and multi-layer erosion *174*
5.6 NETWORK REWIRING MODELLING 177
5.6.1 Single layer rewiring *177*
5.6.2 Multi-Layer: Load based rewiring *180*
5.6.2.1 Random link removal and load based reconnection 181
5.6.2.2 Load versus Load rewiring 182

5.6.3 Multi-Layer: Geography and Load based rewiring *186*
5.6.3.1 Geography versus Load rewiring 186
5.6.3.2 Load-Geography hybrid versus Load rewiring 189
5.7 COMPARISON OF SINGLE AND MULTI-LAYER MODELLING METHODS 191
5.8 CONCLUSION 192

CHAPTER 6 CONCLUSIONS 194

6.1 INTER-LAYER FEEDBACK CONDITIONS AND POWER-LAWS 197
6.2 FUTURE WORK 200
6.2.1 Types of Feedback *200*
6.2.2 Heterogeneous demands *201*
6.2.3 Hybrid node and link selection processes *201*
6.2.4 Multi-layer, not just dual layer *201*
6.2.5 Realistic network architectures and structures *202*

APPENDIX A THE MODULAR INTER-LAYER FEEDBACK TOPOLOGY INVESTIGATION TOOL AND SIMULATOR 203

MITIE DESIGN 203
MITIE CONNECTIVITY STRATEGIES 205
MITIE IMPLEMENTATION 207
LINK AND NODE SELECTION INVERSION 212
MITIE COMMAND LINE ARGUMENTS 214
OUTPUT DATA CONTROL OPTIONS 216
LINK DIMENSIONING OPTIONS 217
EXAMPLE USAGE 219

APPENDIX B SINGLE LAYER MODELLING RESULTS 222

BIBLIOGRAPHY 232

LIST OF FIGURES

FIGURE 1 NETWORK CAPACITY AND DEMAND OVER TIME, SHOWING HOW AVAILABLE CAPACITY MUST ALWAYS BE GREATER THAN DEMAND, WITH NETWORK UPGRADES KEEPING THE SUPPLY AHEAD OF DEMAND 23
FIGURE 2 THE GROWTH OF CAPACITY AND RANGE IN COMMERCIALLY AVAILABLE OPTICAL TRANSMISSION TECHNOLOGY 24
FIGURE 3 A TYPICAL MULTI-LAYER NETWORK 26
FIGURE 4 THE EVENTS AND PROCESSES IN A PACKET BASED TELECOMMUNICATIONS NETWORK 27
FIGURE 5 A MULTI-LAYER NETWORK, CONTAINING IP AND SDH LAYERS, AND THE FEEDBACK BETWEEN LAYERS 29
FIGURE 6 IDEAL TOPOLOGIES FOR 12 CUSTOMER NODE NETWORKS 31
FIGURE 7 A TWO LEVEL HIERARCHICAL NETWORK WITH TRANSIT NODES POSITIONED TO MINIMISE LINK LENGTH 31
FIGURE 8 A TWO LEVEL HIERARCHY NETWORK WITH TRANSIT NODES POSITIONED TO MINIMISE LINK DISTANCE-BANDWIDTH PRODUCT WHERE THE DEMAND MATRIX IS A FULL MESH OF CUSTOMERS NODES 32
FIGURE 9 A TWO LEVEL HIERARCHY NETWORK WITH TRANSIT NODES POSITIONED TO MINIMISE LINK DISTANCE-BANDWIDTH PRODUCT WHERE THE DEMAND MATRIX IS FROM EACH NODE TO FOUR ADJACENT NEIGHBOURS 33
FIGURE 10 DOMAINS AND THEIR INTERCONNECTION IN IP NETWORKS 64
FIGURE 11 A TYPICAL MULTI-AREA OSPF CONFIGURATION WITHIN AN AUTONOMOUS SYSTEM. 67
FIGURE 12 THE USE OF ROUTE REFLECTORS TO IMPROVE I-BGP SCALABILITY 70
FIGURE 13 THE USE OF CONFEDERATIONS TO IMPROVE I-BGP SCALABILITY 71
FIGURE 14 THE THREE TIER INTERNET MODEL WITH PEERING AGREEMENTS AND MULTI-HOMING 73
FIGURE 15 THE PROBLEM OF UNDISCOVERED LINKS FROM SINGLE NETWORK PROJECTIONS 78
FIGURE 16 THE PROBABILITY DISTRIBUTION OF OUTDEGREE, K, FOR AN IP INTERFACE TOPOLOGY MEASURED FROM A SINGLE SOURCE. 83
FIGURE 17 THE DISTRIBUTION OF ROUTE LENGTHS TO ALL RIPE PREFIXES FROM A SINGLE PROJECTION POINT 83
FIGURE 18 THE DISTRIBUTION OF THE NUMBER OF ROUTES PER LINK FROM A SINGLE SOURCE TO ALL DESTINATIONS 84
FIGURE 19 THE DISTRIBUTION OF CIDR SUBNET SIZES IN THE LIST OF SUBNETS AVAILABLE FROM RIPE 85
FIGURE 20 THE FRACTION OF LINKS DISCOVERED WITH FURTHER PROJECTIONS INTO THE NETWORK, WHETHER THE UNDERLYING TOPOLOGY IS ERDOS-RENYI LIKE OR BARABASI-ALBERT LIKE 86
FIGURE 21 THE R^2 CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DISCOVERED TOPOLOGY AS THE NUMBER OF PROJECTION SOURCES IS INCREASED 87

FIGURE 22 THE POWER-LAW FIT EXPONENT OF THE MEASURED TOPOLOGY AS THE NUMBER OF PROJECTION SOURCES INCREASES 88
FIGURE 23 THE RANK PLOT FOR AS TOPOLOGY INT-11-97 FROM FALOUTSOS ET AL. [FAL99]. THE RANK APPEARS ON THE ABSCISSA, AND THE OUTDEGREE ON THE ORDINATE AXIS 90
FIGURE 24 THE RANK PLOT FOR ROUTER TOPOLOGY ROUT-95 FROM FALOUTSOS ET AL. [FAL99]. THE RANK APPEARS ON THE ABSCISSA, AND THE OUTDEGREE ON THE ORDINATE AXIS 90
FIGURE 25 THE OUTDEGREE PLOT FOR AS TOPOLOGY INT-11-97 FROM FALOUTSOS ET AL. [FAL99]. THE OUTDEGREE APPEARS ON THE ABSCISSA, AND THE FREQUENCY ON THE ORDINATE AXIS. 92
FIGURE 26 THE OUTDEGREE PLOT FOR ROUTER TOPOLOGY ROUT-95 FROM FALOUTSOS ET AL. [FAL99]. THE OUTDEGREE APPEARS ON THE ABSCISSA, AND THE FREQUENCY ON THE ORDINATE AXIS. 92
FIGURE 27 THE HOP PLOT FOR AS TOPOLOGY INT-11-97 FROM FALOUTSOS ET AL. [FAL99]. THE HOP DISTANCE APPEARS ON THE ABSCISSA, AND THE NUMBER OF PAIRS ON THE ORDINATE AXIS. 93
FIGURE 28 THE HOP PLOT FOR ROUTER TOPOLOGY ROUT-95 FROM FALOUTSOS ET AL. [FAL99]. THE HOP DISTANCE APPEARS ON THE ABSCISSA, AND THE NUMBER OF PAIRS ON THE ORDINATE AXIS. 93
FIGURE 29 THE SORTED EIGENVALUE PLOT FOR AS TOPOLOGY INT-11-97 FROM FALOUTSOS ET AL. [FAL99]. THE INDEX INTO THE SORTED LIST OF EIGENVALUES APPEARS ON THE ABSCISSA, AND THE EIGENVALUE ON THE ORDINATE AXIS. 94
FIGURE 30 THE SORTED EIGENVALUE PLOT FOR ROUTER TOPOLOGY ROUT-95 FROM FALOUTSOS ET AL. [FAL99]. THE INDEX INTO THE SORTED LIST OF EIGENVALUES APPEARS ON THE ABSCISSA, AND THE EIGENVALUE ON THE ORDINATE AXIS. 94
FIGURE 31 SDH NETWORK LAYERS 100
FIGURE 32 SDH NETWORK LAYERS ACROSS VARIOUS SDH NETWORK ELEMENTS 101
FIGURE 33 THE SDH TRANSPORT HIERARCHY 103
FIGURE 34 THE DEDICATED PATH PROTECTION SCHEME IN OPERATION 105
FIGURE 35 THE SUB-NETWORK CONNECTION PROTECTION (SNCP) SCHEME 106
FIGURE 36THE MAIN AND BACKUP MULTIPLEX SECTIONS IN MSLP 106
FIGURE 37 AN SDH MS-DPRING DEDICATED PROTECTION RING WHERE NO FAULT HAS OCCURED AND THE CIRCUITS PASS IN THE SAME DIRECTION AROUND THE RING 107
FIGURE 38 AN SDH MS-DPRING DEDICATED PROTECTION RING WHERE A FIBRE HAS BEEN CUT AND THE ADJACENT ADMS HAVE SWITCHED THE TRAFFIC INTO THE BACKUP FIBRE 107
FIGURE 39 AN SDH MS-SPRING SHARED PROTECTION RING WHERE NO FAULT HAS OCCURRED AND EACH FIBRE CONTAINS MAIN AND BACKUP CIRCUITS AND, THEREFORE, EACH FIBRE IS USED TO SHARE MAIN AND BACKUP TRAFFIC 108
FIGURE 40 AN SDH MS-SPRING SHARED PROTECTION RING WHERE A FIBRE BREAK HAS OCCURED AND THE CIRCUITS HAVE BEEN SWITCHED INTO THE PROTECTION CIRCUITS OF THE OTHER FIBRES 108
FIGURE 41 THE ARCHITECTURE OF THE BT SDH NETWORK (DIAGRAM COURTESY OF BT) 113
FIGURE 42 THE RANK OF THE SDH NODE DEGREES SORTED IN DECREASING ORDER 118

FIGURE 43 THE DEGREE FREQUENCY DISTRIBUTION OF THE SDH CIRCUIT HOP TOPOLOGY 119
FIGURE 44 THE FIRST FORTY SORTED EIGENVALUES OF THE ADJACENCY MATRIX OF THE SDH NODE TOPOLOGY 120
FIGURE 45 A PLOT OF THE NUMBER OF PAIRS OF NODES WHICH ARE WITHIN A GIVEN NUMBER OF HOPS IN THE SDH NODE TOPOLOGY 121
FIGURE 46 THE NEIGHBOURHOOD AROUND VERTEX V, WITH VERTEX V IN RED AND THE NEIGHBOUHOOD NODES IN GREEN. 122
FIGURE 47 THE CLUSTERING COEFFICIENT OF SDH NODES AGAINST THEIR RANK WHEN SORTED IN ASCENDING ORDER. 123
FIGURE 48 SDH NODE ASSORTATIVITY: A PLOT OF THE PROBABILITY THAT THE NODE AT ONE END OF A LINK HAS A CERTAIN DEGREE GIVEN THE DEGREE OF THE NODE AT THE OTHER END. (DEGREES ARE IN BINS OF 10) 124
FIGURE 49 THE NUMBER OF NODE PAIRS WHICH ARE A GIVEN DISTANCE APART (WHETHER CONNECTED OR NOT)(IN 500 DISTANCE UNIT BINS). 125
FIGURE 50 THE FREQUENCY DISTRIBUTION OF THE DISTANCES BETWEEN CONNECTED NODES (IN BINS OF 500 DISTANCE UNITS) 126
FIGURE 51 ESTIMATING THE PARAMETERS FOR THE WAXMAN TOPOLOGY MODEL: THE PROBABILITY THAT TWO NODES ARE LINKED (P) VS THE NORMALISED DISTANCE BETWEEN THEM (D/L, WHICH IS IN BINS OF 0.02) 127
FIGURE 52 THE PROBABILITY THAT TWO NODES ARE LINKED AGAINST THE NORMALISED DISTANCE BETWEEN THEM (D/L IS IN BINS OF 0.02). 128
FIGURE 53 THE BT NATIONAL SDH CIRCUIT TOPOLOGY AT THE FINEST AVAILABLE GEOGRAPHIC GRANULARITY 130
FIGURE 54 THE BT NATIONAL SDH CIRCUIT TOPOLOGY WITH SITES MERGED TO THE NEAREST 4000 DISTANCE UNITS (ABOUT 40KM) 130
FIGURE 55 THE POWER-LAW EXPONENT AND CORRELATION COEFFICIENT FOR DIFFERENT GRID GRANULARITIES, M, WHICH IS IN DISTANCE UNITS. 131
FIGURE 56 THE CORRELATION COEFFICIENT OF THE POWER-LAW FIT TO THE DEGREE DISTRIBUTION FOR DIFFERENT GRANULARITY GRID SIZES, M, IN DISTANCE UNITS, AS WELL AS THE NUMBER OF NODES IN THE COLLAPSED TOPOLOGY 132
FIGURE 57 THE FREQUENCY DISTRIBUTION OF THE TOTAL CIRCUIT BANDWIDTHS ON EACH LINK (IN BINS OF 500 MBIT/S) 134
FIGURE 58 THE BANDWIDTH VALUES OF ALL SDH HOPS (INCLUDING LOOPS) RANKED IN INCREASING ORDER 135
FIGURE 59 THE DEGREE OF AN SDH NODE COMPARED TO THE TOTAL BANDWIDTH OF ITS LINKS. 136
FIGURE 60 A PLOT OF THE AVERAGE BANDWIDTHS OF LINKS GIVEN THE DEGREE VALUES OF THEIR END NODES (DEGREE VALUES IN BINS OF 20). 137

FIGURE 61 A SCATTER PLOT OF THE TOTAL BANDWIDTH BETWEEN EVERY CONNECTED NODE PAIR AGAINST THE GEOGRAPHIC DISTANCE BETWEEN THEM. 138
FIGURE 62 THE AVERAGE DISTANCE TO THE BEGINNING OR END OF A CIRCUIT, IN HOPS, THAT A NODE APPEARS AT FOR ALL CIRCUITS 139
FIGURE 63 THE AVERAGE DISTANCE IN HOPS THAT A LINK APPEARS IN A CIRCUIT AGAINST THE TOTAL BANDWIDTH OF THAT LINK, AVERAGED FROM ALL CIRCUITS. 140
FIGURE 64 THE FREQUENCY DISTRIBUTION OF THE ACTUAL NUMBER OF HOPS IN AN SDH CIRCUIT VS THE NUMBER OF HOPS IN THE SHORTEST PATH BETWEEN THE END-POINTS (NOTE DIFFERENT SCALES ON HOPS AXES) 141
FIGURE 65 THE CUMULATIVE FRACTION OF ALL CIRCUITS THAT HAVE A CERTAIN RATIO OF ACTUAL HOPS TO THE SHORTEST PATH HOPS OR LESS 143
FIGURE 66 A PLOT OF THE AVERAGE TOTAL CIRCUIT DISTANCE TRAVELLED BY CIRCUITS GIVEN THE DEMAND DISTANCE. ERROR BARS SHOW LOWEST AND HIGHEST CIRCUIT DISTANCES FOR THE GIVEN DEMAND DISTANCE. (DEMAND DISTANCE BINNED IN 500 DISTANCE UNIT BINS) 144
FIGURE 67 THE CUMULATIVE FRACTION OF ALL CIRCUITS THAT HAVE A CERTAIN CIRCUIT DISTANCE TO DEMAND DISTANCE RATIO 144
FIGURE 68 THE RANKED DEGREE VALUES OF THE SDH DEMAND TOPOLOGY 147
FIGURE 69 THE FREQUENCY DISTRIBUTION OF DEMAND TOPOLOGY NODE DEGREES 147
FIGURE 70 THE 40 LARGEST EIGENVALUES OF THE DEMAND TOPOLOGY 149
FIGURE 71 THE TOTAL NUMBER OF NODE PAIRS A GIVEN DISTANCE (IN HOPS) APART 149
FIGURE 72 THE CLUSTERING COEFFICIENT OF THE DEMAND TOPOLOGY 150
FIGURE 73 DISTRIBUTION OF GEOGRAPHIC DISTANCE BETWEEN DEMAND EDGE NODES 151
FIGURE 74 TOTAL BANDWIDTH BETWEEN EDGE NODES PAIRS (IN BINS OF 500MBIT/S) 152
FIGURE 75 DISTRIBUTION OF GEOGRAPHIC DISTANCE BETWEEN CIRCUIT EDGE NODES FOR VARIOUS CAPACITIES 153
FIGURE 76 THE CORRELATION COEFFICIENT, R^2, OF THE POWER-LAW FIT OF NODE DEGREE DISTRIBUTION, AS LOAD BASED NETWORK GROWTH PROGRESSES WITH {TR,0,0,0}, {AR,0,0,0}, {TR,0,G,TR}, {AR,0,G,AR} 165
FIGURE 77 THE DEGREE FREQUENCY DISTRIBUTION OF A 3000 NODE TOPOLOGY GENERATED WITH {TR,0,0,0} 166
FIGURE 78 THE DEGREE FREQUENCY DISTRIBUTION OF A 3000 NODE TOPOLOGY GENERATED WITH {AR,0,0,0} 166
FIGURE 79 THE CORRELATION COEFFICIENT, R^2, OF THE POWER-LAW FIT TO DEGREE RANK, AS LOAD BASED NETWORK GROWTH PROGRESSES WITH {TR,0,0,0}, {AR,0,0,0}, {TR,0,G,TR}, {AR,0,G,AR}. 167
FIGURE 80 THE R^2 CORRELATION COEFFICIENT OF THE POWER-LAW FIT TO CLUSTERING COEFFICIENT RANK, AS A FUNCTION OF THE NETWORK SIZE AS THE NETWORK IS GROWN ACCORDING TO THE LOAD BASED CONNECTION ALGORITHMS 168

FIGURE 81 THE R^2 CORRELATION COEFFICIENT OF POWER-LAW FIT TO NODE DEGREE FREQUENCY DISTRIBUTION, AS A FUNCTION OF THE LINKS-TO-NODE RATIO AS {0,R,0,0} LINK EROSION PROGRESSES. VALUES PLOTTED FOR ER AND BA INITAL GRAPHS, WITH GRAPH PARTITIONING PERMITTED AND NOT PERMITTED 170
FIGURE 82 THE R^2 CORRELATION COEFFICIENT PLOTTED AS A FUNCTION OF THE LINKS-TO-NODE RATIO AS {0,AE,0,0} LINK EROSION PROGRESSES. VALUES PLOTTED FOR ER AND BA INITIAL GRAPHS CONNECTED WITH RANDOM NODE PLACEMENTS, AND NODE PLACEMENTS EXTRACTED FROM THE SDH NETWORK. 172
FIGURE 83 THE R^2 CORRELATION COEFFICIENT OF POWER-LAW FIT TO NODE DEGREE FREQUENCY DISTRIBUTION, AS A FUNCTION OF THE LINKS-TO-NODE RATIO AS {0,ARPL,0,0} LINK EROSION PROGRESSES. VALUES PLOTTED FOR ER AND BA INITIAL GRAPHS. 173
FIGURE 84 THE R^2 CORRELATION COEFFICIENT OF POWER-LAW FIT OF NODE DEGREE DISTRIBUTION, AS A FUNCTION OF THE LINKS-TO-NODE RATIO AS {0,ARDP,0,0} LINK EROSION PROGRESSES. VALUES PLOTTED FOR ER AND BA INITIAL GRAPHS. 174
FIGURE 85 THE PERCENTAGE IMPROVEMENT IN THE CORRELATION COEFFICIENT OF POWER-LAW FIT TO DEGREE FREQUENCY DISTRIBUTION FOR VARIOUS EROSION SCENARIOS COMPARED TO THE {0,R,0,0} EROSION, FOR AN INITIALLY BA TOPOLOGY 175
FIGURE 86 THE PERCENTAGE IMPROVEMENT IN THE CORRELATION COEFFICIENT OF POWER-LAW FIT TO DEGREE FREQUENCY DISTRIBUTION FOR VARIOUS EROSION SCENARIOS COMPARED TO THE {0,R,0,0} EROSION, FOR AN INITIALLY ER TOPOLOGY 176
FIGURE 87 {0,R,R,R} AND {0,R,D,D} REWIRINGS ON INITIAL ER AND BA GRAPHS. 178
FIGURE 88 THE DEGREE FREQUENCY DISTRIBUTION PLOT FOR A SINGLE INSTANCE OF A BA TOPOLOGY REWIRED 3000 TIMES BY {0,R,D,D} 179
FIGURE 89 THE PROGRESSION OF CORRELATION COEFFICIENT TO POWER-LAW FIT OF DEGREE FREQUENCY DISTRIBUTION AS 3000 LINKS ARE ADDED TO 500 NODE, 1000 LINK, ER AND BA SEED TOPOLOGIES WITH {0,0,D,D} 179
FIGURE 90 {0,AE,R,R} AND {0,AE,D,D} REWIRINGS ON INTIAL ER AND BA GRAPHS TO EXAMINE THE POTENTIAL EFFECT OF GEOGRAPHY ON REWIRING. 180
FIGURE 91 {0,R,AR,AR} AND {0,R,TR,TR} REWIRINGS - BOTH MULTI-LAYER SCENARIOS AS THEY CONSIDER THE TRAFFIC LOAD TRAVERSING NODES 181
FIGURE 92 THE NODE DEGREE FREQUENCY DISTRIBUTION AFTER 3000 CYCLES OF {0,R,TR,TR} REWIRING 182
FIGURE 93 THE NODE DEGREE FREQUENCY DISTRIBUTION AFTER 3000 CYCLES OF {0,R,AR,AR} REWIRING 182
FIGURE 94 {0,ARPL,AR,AR} AND {0,ARPL,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES 183

FIGURE 95 THE DEGREE FREQUENCY DISTRIBUTION OF A SINGLE INSTANCE AFTER 3000 CYCLES OF {0,ARPL,TR,TR} REWIRING STARTING WITH A BA TOPOLOGY INSTANCE THAT HAS THE HIGHEST R^2 AFTER THE 3000 CYCLES OF REWIRING 184
FIGURE 96 THE NUMBER OF NODES WITH A DEGREE OF 1 AS THE {0,ARPL,TR,TR} REWIRING PROGRESSES 184
FIGURE 97 {0,ARPL,AR,AR} AND {0,ARPL,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES 185
FIGURE 98 {0,RPL,AR,AR} AND {0,RPL,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES 185
FIGURE 99 {0,RPL,MAXAR,MAXAR} AND {0,ARPL,MAXAR,MAXAR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES 186
FIGURE 100 {0,RPL,MAXTR,MAXTR} AND {0,ARPL,MAXTR,MAXTR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES 186
FIGURE 101 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DEGREE FREQUENCY DISTRIBUTION OF {0,AE,AR,AR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 187
FIGURE 102 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DEGREE FREQUENCY DISTRIBUTION OF {0,AE,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 187
FIGURE 103 THE DEGREE FREQUENCY DISTRIBUTION FOR THE INSTANCE OF {0,AE,TR,TR} REWIRING WITH THE LOWEST R^2 CORRELATION COEFFICIENT 187
FIGURE 104 THE DEGREE FREQUENCY DISTRIBUTION FOR THE INSTANCE OF {0,AE,TR,TR} REWIRING WITH THE HIGHEST R^2 CORRELATION COEFFICIENT 187
FIGURE 105 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DEGREE RANK OF {0,AE,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 189
FIGURE 106 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE CLUSTERING COEFFICIENT RANK OF {0,AE,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 189
FIGURE 107 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DEGREE FREQUENCY DISTRIBUTION OF {0,ARDP,AR,AR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 190
FIGURE 108 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DEGREE FREQUENCY DISTRIBUTION OF {0,ARDP,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 190

FIGURE 109 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE DEGREE RANK OF {0,ARDP,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTION 191
FIGURE 110 THE CORRELATION COEFFICIENT OF A POWER-LAW FIT TO THE CLUSTERING COEFFICIENT RANK OF {0,ARDP,TR,TR} REWIRING SCENARIOS FOR ER AND BA INITIAL TOPOLOGY TYPES AND UNIFORMLY RANDOM AND BT NODE DISTRIBUTIONS 191
FIGURE 111 THE MAIN EXECUTION PATH FOR MITIE 209
FIGURE 112 ROULETTE WHEEL NODE OR LINK SELECTION IN THE GROWTH PROCESS OF MITIE 210
FIGURE 113 MITIE ORPHAN LINK BRIDGE DETECTION ALGORITHM 212

LIST OF TABLES

TABLE 1 COMMERCIAL OPTICAL TRANSMISSION SYSTEMS (DERIVED FROM A PRESENTATION BY HERWIG KOGELNIK, ECOC 2004) 23
TABLE 2 A COMPARISON OF LINK DISTANCES AND DISTANCE-BANDWIDTH PRODUCTS IN THE IDEAL TOPOLOGIES 34
TABLE 3 THE RANK POWER-LAW EXPONENTS AND ABSOLUTE CORRELATION COEFFICIENTS FOR THE VARIOUS AS AND ROUTER TOPOLOGY DATA SETS AS FOUND BY FALOUTSOS ET AL. [FAL99] 91
TABLE 4 THE OUTDEGREE POWER-LAW EXPONENTS AND ABSOLUTE CORRELATION COEFFICIENTS FOR THE VARIOUS AS AND ROUTER TOPOLOGY DATA SETS AS FOUND BY FALOUTSOS ET AL. [FAL99] 92
TABLE 5 THE HOP-PLOT EXPONENTS AND ABSOLUTE CORRELATION COEFFICIENTS FOR THE VARIOUS AS AND ROUTER TOPOLOGY DATA SETS AS FOUND BY FALOUTSOS ET AL. [FAL99] 94
TABLE 6 THE EIGENVALUE POWER-LAW EXPONENTS AND ABSOLUTE CORRELATION COEFFICIENTS FOR THE VARIOUS AS AND ROUTER TOPOLOGY DATA SETS AS FOUND BY FALOUTSOS ET AL. [FAL99] 95
TABLE 7 THE BIT RATES OF VARIOUS SDH TRANSPORT LEVELS 102
TABLE 8 SDH TRANSPORT EFFICIENCY WHEN CARRYING ETHERNET TRAFFIC 110
TABLE 9 AN EXCERPT OF THE SDH CIRCUITS DATA SET 114
TABLE 10 THE BANDWIDTH AND PROTECTION TYPES OF SDH DATA SET. 114
TABLE 11 THE PARAMETERS OF THE TOPOLOGIES AVAILABLE FROM THE BT SDH NETWORK DATA 116
TABLE 12 THE SINGLE LAYER MODELLING SCENARIOS THAT PRODUCE NETWORKS WITH THE BEST POWER-LAW FITS 164

LIST OF EQUATIONS

EQUATION 1 THE DEGREE DISTRIBUTION MODEL IN THE INET TOPOLOGY GENERATOR 45
EQUATION 2 THE WEIGHT OF A NODE IN THE PREFERENTIAL CONNECTIVITY IN INET-3.0 46
EQUATION 3 THE PROBABILITY THAT NODE I CONNECTS TO NODE J IN THE INET-3.0 PREFERENTIAL ATTACHMENT MODEL 46
EQUATION 4 THE DEGREE DISTRIBUTION OF THE ERDÖS-RÉNYI RANDOM GRAPH MODEL 48
EQUATION 5 THE AVERAGE DIAMETER OF AN ERDOS-RENYI RANDOM GRAPH 48
EQUATION 6 THE PROBABILITY THAT A NEWLY ADDED NODE IS CONNECTED TO AN EXISTING NODE IN THE ALBERT-BARABASI RANDOM GRAPH MODEL 49
EQUATION 7 THE DEGREE DISTRIBUTION OF THE ALBERT-BARABASI RANDOM GRAPH MODEL 49
EQUATION 8 THE AVERAGE PATH LENGTH OF THE ALBERT-BARABASI RANDOM GRAPH MODEL 49
EQUATION 9 THE PROBABILITY THAT A NODE *I* IS SELECTED FOR CONNECTION WHERE *KI* IS THE DEGREE OF THE NODE *I* AND B IS A PARAMETER BETWEEN $-\infty$ AND 1 50
EQUATION 10 THE PROBABILITY THAT A NODE IS SELECTED FOR CONNECTION, WHERE K_I IS THE DEGREE OF NODE *I* AND Δ IS A PARAMETER IN THE RANGE [0,1] 51

EQUATION 11 THE DEGREE ASSIGNED TO A NODE IN THE PLOD MODEL WHERE $X \sim U(1,N)$ 51
EQUATION 12 THE PROBABILITY THAT TWO NODES ARE CONNECTED IN THE WAXMAN TOPOLOGY MODEL .. 52
EQUATION 13 THE PROBABILITY THAT TWO NODES ARE CONNECTED IN THE DOAR-LESLIE TOPOLOGY MODEL 52
EQUATION 14 THE CONDITION THAT MUST BE SATISIFIED WHEN SELECTING THE CONNECTING NODE IN FABRIKANT ET AL.'S HOT TOPOLOGY MODEL 53
EQUATION 15 THE PROBABILITY THAT TWO NODES ARE CONNECTED IN THE EXPONENTIAL MODEL OF GT-ITM 54
EQUATION 16 THE AVAILABILITY OF A SYSTEM CONSISTING OF ELEMENTS CONNECTED IN SERIES 109
EQUATION 17 THE AVAILABILITY OF A SYSTEM CONSISTING OF ELEMENTS CONNECTED IN PARALLEL 109
EQUATION 18THE CLUSTERING COEFFICIENT 122
EQUATION 19 THE WAXMAN TOPOLOGY CONNECTIVITY EQUATION 127
EQUATION 20 THE WAXMAN CONNECTIVITY EQUATION RE-ARRANGED TO THE STRAIGHT LINE EQUATION 127
EQUATION 21 THE CONNECTION PROBABILITY OF SDH NODES GIVEN THE EUCLIDEAN DISTANCE BETWEEN THEM 128
EQUATION 22 THE TWO BODY GRAVITATIONAL EQUATION 129
EQUATION 23 THE PROBABILITY THAT A LINK I OF L LINKS IS REMOVED IN THE RDP REMOVAL SCHEME ... 190

Chapter 1
Multi-layer Network Provisioning and Inter-Layer Feedback

1.1 Introduction

In order to support a wide variety of services, to different user types, and under a variety of geographic situations, telecommunications networks are typically composed of a variety of layers and heterogeneous technologies. Layers (in terms of the OSI 7 layer model) such as the transmission layer (e.g. WDM (Wavelength Division Multiplexing)), the data link layer (also known as the transport network e.g. SDH (Synchronous Digital Hierarchy), Ethernet) and the network layer (e.g. IP (Internet Protocol)). No single layer is independent of the adjacent layer and the provisioning requirements of one layer become the demand on the layer below. Similarly, the available resources become the delivered quality of service to the layer above. This thesis is concerned with the design aspects of various layers and how they affect each other's topology.

This chapter will examine the processes present in operating and maintaining a telecommunications network and propose how these processes lead to effects between network layers that can exert influence on one another in a feedback system. The chapter examines the forces as applied to the geographic location of network resources and goes on to suggest that similar forces influence network layer topology.

1.2 Network Provisioning

Telecommunications network operators are businesses and like most businesses offer goods or services, in this case the service of communication, and their aim is to make a profit while providing this service; to this end they try to maximise their revenue while minimising costs. Network provisioning, therefore, is an optimisation of network utility versus cost. To ensure the network operator provides an acceptable quality of service, as well as maximising return on investment, the network must be provisioned to support the demands of customers as long as possible. Thus the operator makes predictions of offered traffic and dimensions the network accordingly. This usually happens on a cyclic basis with the network having to be upgraded to support current and future demand. This can be seen in Figure 1 (from a similar diagram in [VAL07]) where a network design event is periodically invoked to upgrade the network to accommodate increasing demand. This is, in effect, a reaction (network redesign) to an action (increase in demand).

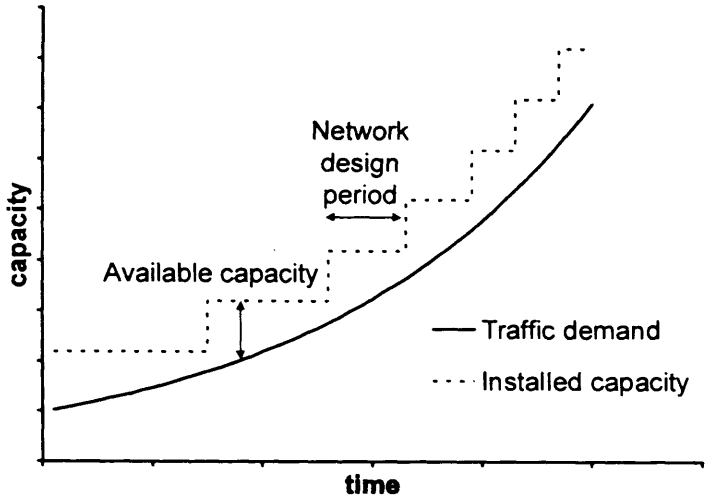


Figure 1 Network capacity and demand over time, showing how available capacity must always be greater than demand, with network upgrades keeping the supply ahead of demand

In practical terms network design is about selecting network elements, their location and connectivity – more simply, it is a trade-off between nodes and links; nodes to provide switching facility in the core, and links to extend the span of the network without the requirement of further switching.

While demand steadily increases, so does innovation and improved transmission and transport technology. Technological growth has increased the capacity and reach of transmission systems, for example, in Figure 2 we can see the rate of growth of the bandwidth-distance product (the maximum distance a transmission system can support multiplied by the capacity of the system) for commercially available optical systems, as described in Table 1.

System	Year	WDM channels	Bit rate/ Channel Gb/s	Bit rate/ Fibre Gb/s	Regenerator spans/ km	Bandwidth-distance product
FT3	1980	1	0.045	0.045	7	0.315
FTG-1.7	1987	1	1.7	1.7	50	85
FT-2000	1992	1	2.5	2.5	50	125
NGLN	1995	8	2.5	20	360	7200
Wavestar 400G	1999	80	2.5	200	640	128000
		40	10	400	640	256000
WaveStar 1.6T	2001	160	10	1600	640	1024000
LambdaXtreme	2003	128	10	1280	4000	5120000
		64	40	2560	1000	2560000

Table 1 Commercial optical transmission systems (derived from a presentation by Herwig Kogelnik, ECOC 2004)

Although transmission technology growth leads to a better ability to support demands, similar technological innovation in customer equipment leads to new uses (e.g. personal mobile telephone, broadband to the home, better video codecs to make video over data feasible, voice over data), new demands (e.g. always-on and pervasive Internet) and surges in traffic (e.g. increase in popularity of video streaming, peer-to-peer systems), which is caused by potentially unrelated technological advances (cameras, efficient video encoders, increased microprocessor ability), and especially the market forces which influence the affordability of end-user devices (i.e. cheaper mobile phones, cheap laptops). Technological innovation is, however, beyond the scope of this thesis; suffice to say that demand increase is almost inevitable, and often accompanied by paradigm shifts (fixed to mobile telephony switch, terrestrial television switching to cable television, client-server to peer-to-peer).

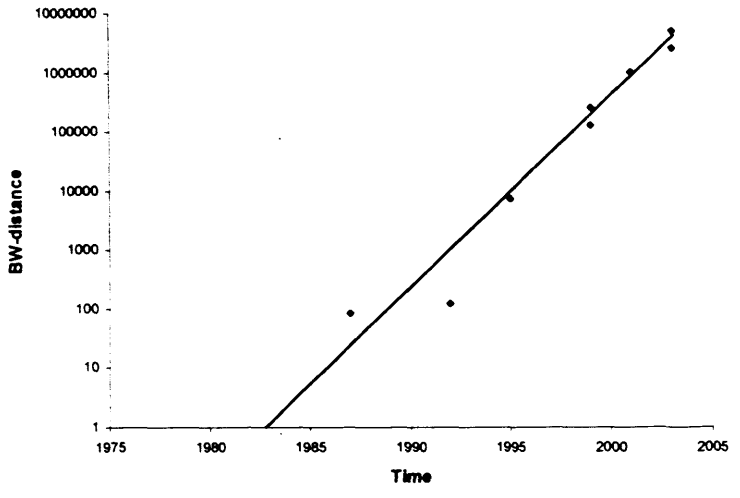


Figure 2 The growth of capacity and range in commercially available optical transmission technology

The changing uses and the resulting services offered by telecommunications providers were often matched by a technology designed to specifically support that service, for example telephony was originally provided over analogue phone circuits, and later migrated to digital PCM (Pulse Code Modulation) systems, which were on a separate network to X.25 data services.

With increased uptake (more users) and more demanding applications older technologies had difficulty scaling to the capacities, and at the same time telecommunications network operators tried to lower overheads by decreasing management costs and pushed for a multi-service network – in a typical network evolution PCM voice as well as data services were migrated to ATM networks to be carried by SDH transport networks. ATM interfaces, however, had difficulties scaling to gigabit speeds, even though SDH could support such speeds. Another

example of convergence, increased demand and technological advancement is the fact that in 1998 in the UK most residential Internet access was through analogue V90 modems at 56 kbit/s, while in 2008 the most common Internet connectivity technology is xDSL (x-Digital Subscriber Line) and is provided over ATM networks (PPPoA – point-to-point protocol over ATM), carried by SDH networks, although Ethernet based xDSL systems (PPPoE - point-to-point protocol over Ethernet) are also becoming increasingly common. The result is a wide variety of old and new network technologies being carried by various transport networks, being connected by various transmission technologies.

1.2.1 A Typical Multi-layer Network

The result of ever changing networking and transport technologies is something similar to the multi-layer network in Figure 3, which provides support for new and legacy technologies. Figure 3 also gives us an idea of the inter-dependence of network technologies: At the bottom layer ducts can only be dug between geographically nearby locations, whereas multiple fibres can then span multiple ducts to reach further sites. Each fibre can then support one or more demands by multiplexing wavelengths in a WDM system. Each wavelength can then carry one of many different transport and network technologies such as IP, SDH or Ethernet. These technologies all have different architectures, characteristics and protocols (we will be examining IP and SDH closely in Chapters 3 and 4) all of which make different demands of the WDM and fibre layers. Some of the technologies, like SDH and ATM also have different layers within them, for example the low and high order paths of SDH and the Virtual Paths and Virtual Circuits of ATM. The main purpose of this diagram is to demonstrate some of the technologies involved and the extent of heterogeneity in a typical network. Many networks today also operate GMPLS (Generalized Multi Protocol Label Switching) [RFC3945], which is a family of protocols (for signalling, routing, resource management and traffic engineering) designed to provide services over packet and circuit switched, optical and non-optical networks. ASTN (Automatic Switched Transport Network) [G.807] (and ASON (Automatic Switched Optical Network) [G8080]) further proposes a network architecture to simplify management, through the use of GMPLS protocols, and provides a unified view of paths and network elements throughout the network.

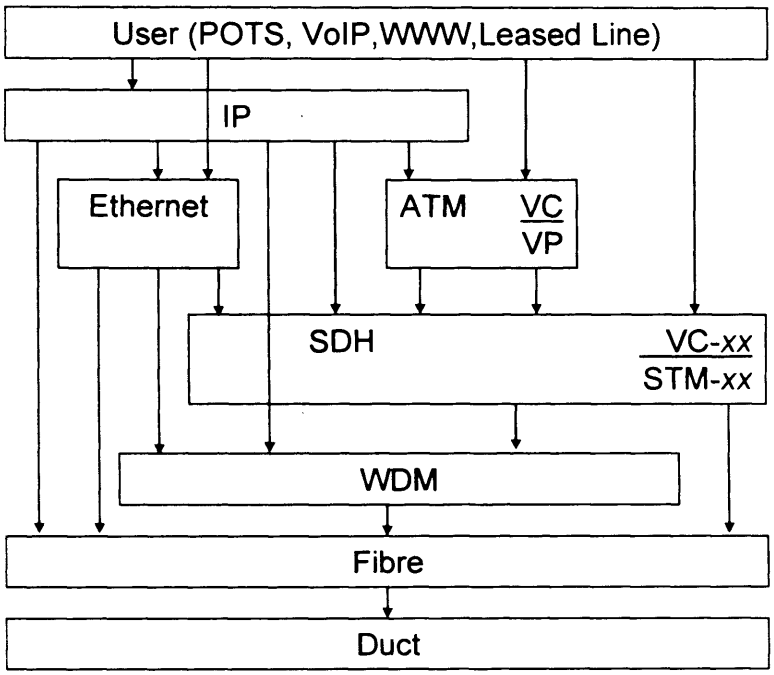


Figure 3 A typical multi-layer network

1.2.2 Network Events, Timescales and Inter-Layer Feedback

We have seen that in telecommunications networks there is always evolution, adaptation, expansion and change, not only in the demands, but also in the technologies. These changes are, however, different at differing timescale; in terms of the installed network the changes in demand can be characterised by the events in Figure 4 (for the case of a packet switched network), where, depending on the time scale, different events can occur which are handled by different control processes and implemented by different means. At the shortest scale packets are handled by the routing protocol, but as we move to changes in flows we then have to consider traffic engineering, and network engineering. When we experience shifts in demand, which are events that occur over periods of greater than a week then we are considering our design and configuration, and adapt the network through network planning – and this design and planning is what invokes a network design period from Figure 1.

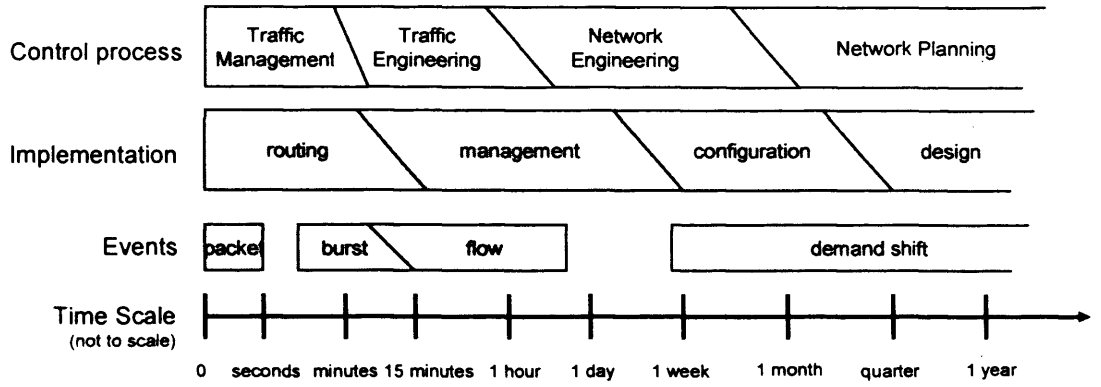


Figure 4 The events and processes in a packet based telecommunications network

In practical terms though this is not limited to a single layer, since a shift in demand of an IP network may require increased transport capacity from an SDH network, which is actually a shift in the demand pattern of the SDH network, which may in turn require the allocation of additional wavelengths from WDM, which may require more fibre to be installed.

The effects between adjacent layers can be seen in Figure 5, where an IP network is partly carried by an SDH network which is mostly co-located with the IP nodes. The IP network follows its own topology, with certain architectural patterns, such as the hub around the IXP (Internet Exchange Point), and is transported by an SDH network and an Ethernet network (those nodes labelled with an E) (Ethernet network not shown in the diagram). The SDH network follows a different topology, with architectural patterns such as rings, meshes and hubs, and also contains nodes (protection switches, repeaters and other SDH network elements) that do not co-locate or appear in the IP network (labelled with an S). The demand that the IP networks makes in terms of end-to-end circuits of the SDH network is the downward force (on the right of Figure 5) between the IP and SDH layers, and the available capacity should be enough to satisfy the demand. If it is not then packets are delayed or dropped and the delivered quality of service that the IP layer receives will then influence its design, through traffic and network engineering (Figure 4), which is the upward force on the left of Figure 5. As the SDH network operators notice an increase in demand they will invoke a network design period and increase the available capacity in the bottlenecked areas (causing an increased demand on the physical layer (WDM, fibre, ducts)) as per Figure 1. This increase in available capacity improves the delivered quality of service, potentially causing a change in traffic or network engineering and a change in the IP layer. We will see that a simple increase in the capacity of a link can have a direct impact on the routing as some IP routing protocols, such as OSPF, actually use an inverse function of capacity

to set link weights (see Chapter 3 for more details). Depending on the configuration and resource location of the rest of the network there may be potential for Braess' Paradox [BRA05] to occur and affect the rest of the network, finally settling either at equilibrium or causing more demands changes to adjacent network layers. Braess' paradox is where extra capacity is added to a transport network (road network or packet based computer network - any flow network) and the routing decision is made to favour the high capacity links, expecting them to provide a higher quality-of-service because of their high capacity, and instead the re-routing of large amounts of traffic overwhelms the extra capacity and the average throughput for all demands drops. This was seen, for example, in [GRI07] which proposed a quality-of-service aware inter-domain routing protocol, and under the wrong conditions the search for available resources overwhelmed parts of the network and brought down overall quality-of-service.

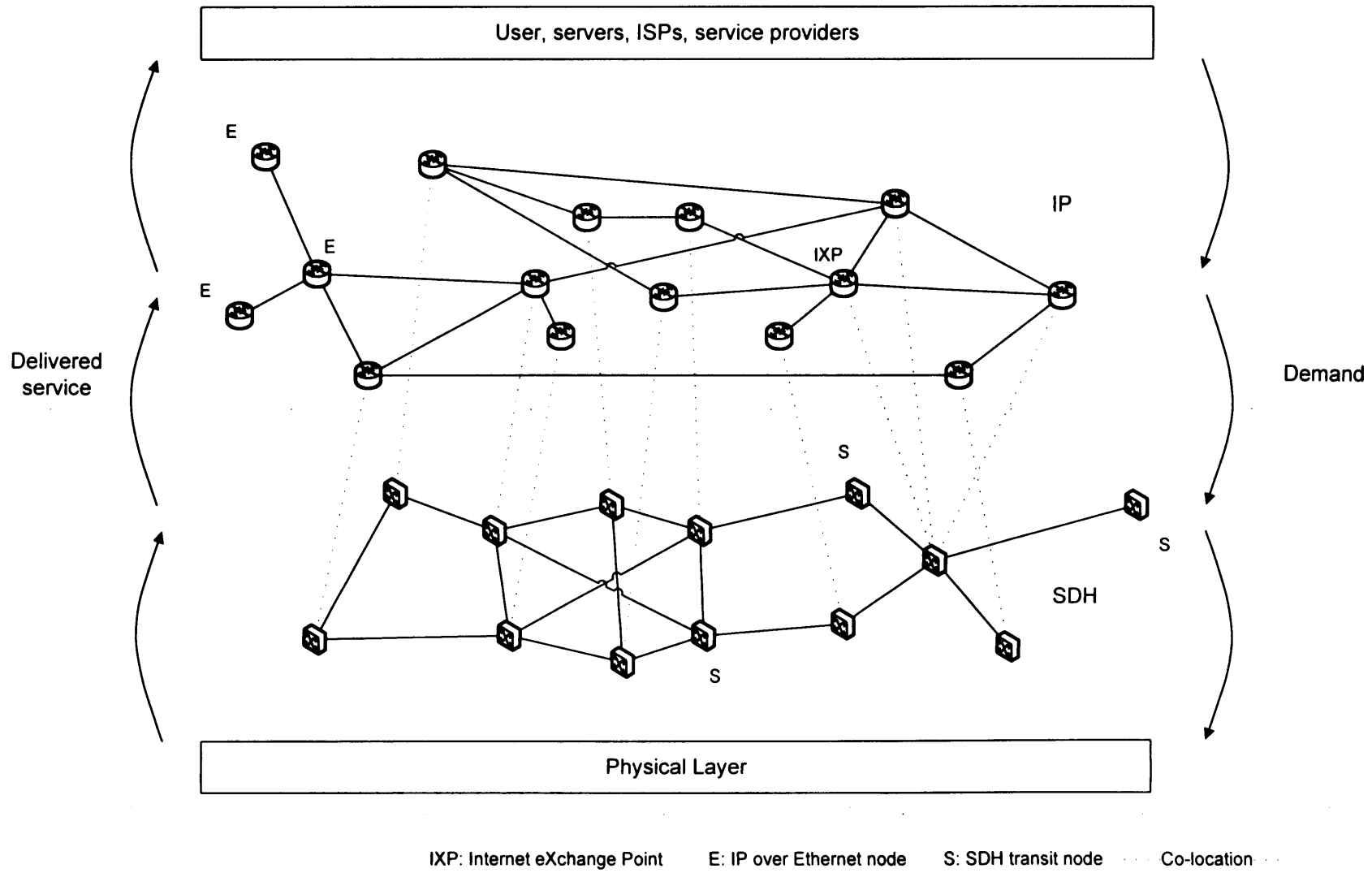


Figure 5 A multi-layer network, containing IP and SDH layers, and the feedback between layers

1.3 Network Design and Cost Functions

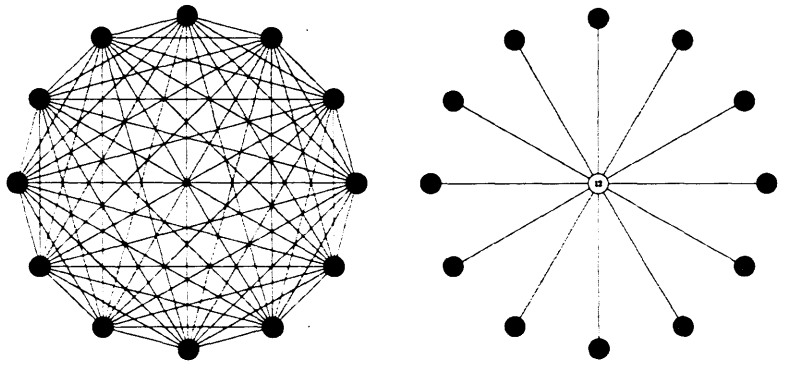
We have now seen that in theory each layer can affect adjacent layers. The exact nature of the effect is of course dependant on the planning and routing algorithms of the layers and the cost functions that affect the layers. This cost function is highly dependant on the layer in question – take the following two scenarios as an example:

- An IP over IP VPN (Virtual Private Network) tunnel is created from a customer network connected to the Internet, to another customer network connected elsewhere in the Internet. Assuming that both customer networks pay a fee based only on the amount of data transferred, and not the actual destinations of the traffic, then the cost to the user of establishing the link is only counted as increased traffic, and not a function of geographical distance.
- An SDH circuit is to be created between two sites which are not connected by fibre. The cost of installing the fibre is a function of the length of the fibre, the terminating equipment and any regenerating equipment required along the route of the fibre. The cost of the new SDH link is now directly a function of geography and the Euclidean distance between the two sites.

To demonstrate the effect of adjacent layers let us consider a simple network of twelve customer nodes, arranged in a ring on a Euclidean plane with a radius of 1 distance unit. All customer nodes wish to send one unit of symmetric traffic to all other nodes (that is a full mesh of demands). The network supporting the demands is closely coupled to the physical layer and, therefore, the cost of links is a function of distance. As mentioned before the design of a network is really a trade-off between links and nodes – and by implication cost. There are of course a very large number of possible configurations, especially if additional transit nodes are added, but we will concentrate on a few ideal topologies to demonstrate the effect of cost functions. The relative merits of a selection of network configurations can be seen in Table 2.

The simplest design would be to connect all the nodes in a full mesh (Figure 6a). This would require each of the N customer nodes to have N-1 links – and a new link per node would have to be installed every time a new customer is added. This would require very large and expensive equipment at every customer site and 0.5*N*(N-1) links. Alternatively, rather than have so many links, the addition of a transit hub (Figure 6b) in the centre would mean that only one link would need to be added to the hub every time a new customer is added (although the load on other links would increase), and the number of links scales with N. The position of the

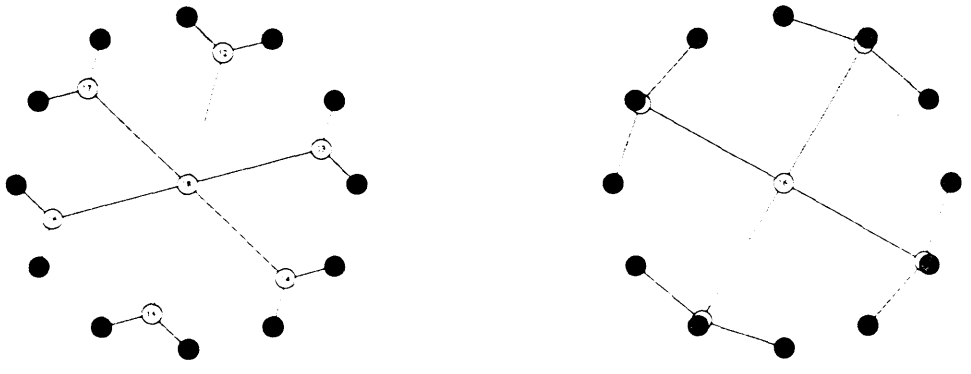
hub is in the centre to minimise the lengths of the links. The capacity of this hub node, however, has to be very large (Table 2, capacity scales with N*N) and creates a single point of failure.



(a) a full mesh topology (b) a hub topology

Figure 6 Ideal topologies for 12 customer node networks

To alleviate the load on the hub and unnecessarily forward traffic from nodes which are actually quite proximal (adjacent nodes in the perimeter ring), a second set of hubs are added, which we will call the transit layer. This is a series of nodes which connect to the customer nodes, and for demands which travel further to more distant nodes, a connection to the central hub is made. This configuration can be seen in Figure 7a where each transit node connects to two customer nodes, and in Figure 7b where they connect to three customer nodes. The line lengths are calculated to optimise solely on link length. Such a transit node may also be added if the maximum length of a link is limited by technology.



(a) a two level hierarchical topology with a transit fan-out of 2, nodes positioned to minimise link distance.

(b) a two level hierarchical topology with a transit fan-out of 3, nodes positioned to minimise link distance.

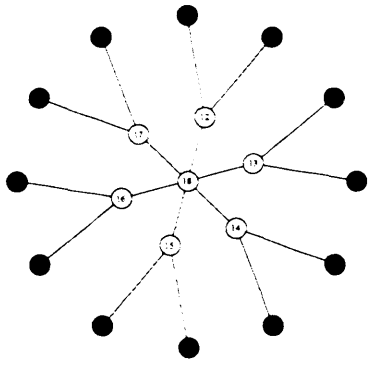
Figure 7 a two level hierarchical network with transit nodes positioned to minimise link length

The cost function is not, however, just a function of distance, it is also a function of capacity, either by having a single physical link with a higher capacity, or multiple bundled

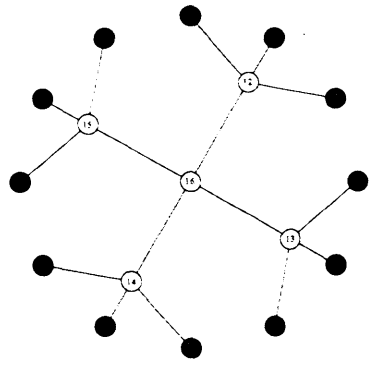
physical links to scale capacity. If we now consider the bandwidth-distance product (the product of link capacity and link length) and attempt to optimise for that the network nodes are repositioned as can be seen in Figure 8a and Figure 8b (optimisation based on bandwidth-distance product makes sense if you consider that capacity scaling is done through bundling multiple fibres together). The high load between the hub and the transit nodes causes the optimal location of the transit nodes to be moved geographically closer to the hub.

So far the demands were a full mesh, but if the pricing of demand delivery were based loosely on the distance connected there may be a tendency to shift from a full mesh of unity capacity demands, to, say, a demand pattern where a node only communicates with the four most immediate neighbours (referred to here as the Na4N demand topology), that is, for node 0, the demands reach nodes 1, 2, 10 and 11. Each demand is now 2.75 (11/4) capacity units to maintain the same total load in the network. Note that by changing the demand pattern to the network the utility to the end user may not necessarily decrease as there may be an additional layer at the nodes forwarding traffic between the end-points of demands that are presented to this optimisation process which will extend the reach of services.

The result can be seen in Figure 9a and Figure 9b where the transit nodes have again shifted towards the customer nodes, as that it where the demands are concentrated.

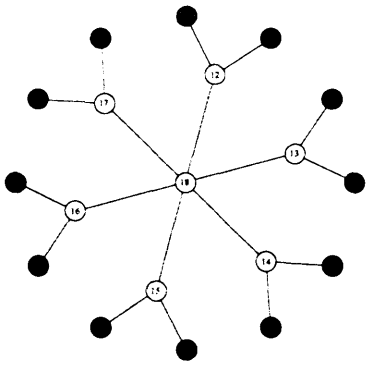


(a) a two level hierarchical topology with a transit fan-out of 2, nodes positioned to minimise link bandwidth-distance product. The demand pattern is a full mesh.

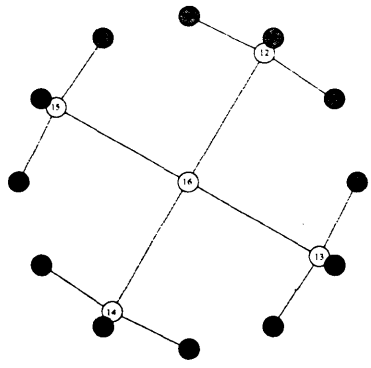


(b) a two level hierarchical topology with a transit fan-out of 3, nodes positioned to minimise link bandwidth-distance product. The demand pattern is a full mesh.

Figure 8 a two level hierarchy network with transit nodes positioned to minimise link distance-bandwidth product where the demand matrix is a full mesh of customers nodes



(a) a two level hierarchical topology with a transit fan-out of 2, nodes positioned to minimise link bandwidth-distance product. The demand pattern from node to four adjacent neighbours, for all nodes.



(b) a two level hierarchical topology with a transit fan-out of 3, nodes positioned to minimise link bandwidth-distance product. The demand pattern from node to four adjacent neighbours, for all nodes.

Figure 9 a two level hierarchy network with transit nodes positioned to minimise link distance-bandwidth product where the demand matrix is from each node to four adjacent neighbours

Table 2 makes further comparisons of the different network configurations – we can see that the two-level 3-fan-out topology with Na4N demands achieves the lowest link-distance product, which is important to our costs when installing physical connectivity, has the lowest hub throughput, which decreases costs of the equipment and potentially avoids scalability issues, and, apart from the full mesh topology, has the joint lowest maximum link capacity, which again improves scalability and cost. [SPE99] made a similar examination of a ring network but examined more closely the effect of demand patterns on resource allocation in WDM rings as an introduction into the effects of demand pattern on cost.

	Number of nodes	Number of links	Hub throughput	Max link capacity	Sum of link distances	Sum of link distance-bandwidth products.
Mesh	12	66	0	1	91.10	91.10
Hub	13	12	132	11	12.00	132.00
2-level 2-fan-out Minimise link distance	19	18	120	20	8.49	137.43
2-level 3-fan-out Minimise link distance	19	16	108	27	5.93	127.10
2-level 2-fan-out Minimise link distance-BW Full-mesh demands	19	18	120	20	9.86	130.14
2-level 3-fan-out Minimise link distance-BW Full-mesh demands	19	16	108	27	6.88	119.95
2-level 2-fan-out Minimise link distance-BW Na4N demands	19	18	99	16.5	8.73	118.22
2-level 3-fan-out Minimise link distance-BW Na4N demands	19	16	66	16.5	6.00	85.80

Table 2 a comparison of link distances and distance-bandwidth products in the ideal topologies

1.4 Scope for research: Inter-layer feedback and network topology

Now that we have examined the timescales, drivers, enablers and processes behind network provisioning, and have seen that there is potential for feedback between layers, we will concentrate on a more specific aspect of the network, namely, topology. It is very expensive to relocate node sites since this requires significant capital expenditure, mainly in the purchase of premises, and it is significantly cheaper to upgrade, replace and rewire links, either logical (IP routing, SDH circuits) or physical (fibres, ducts), so topology is potentially more prone to be affected by layer interaction, and this is, therefore, the main focus of this thesis.

Current literature takes a very isolationist approach to modelling networks, and specifically topology, as will be seen in Chapter 2, and concentrates on each layer individually, and at most considers a static demand pattern, or a static underlying resource configuration. This thesis addresses the issue of network topology modelling from the approach of dynamic, reactive models, which will develop their topology in response to changes in adjacent layers. The main issues to be addressed are:

- Modelling network layer development – what are the processes, design targets, and environment under which network layers develop? Are there explicit rules? For example the explicit creation of hierarchies, the deliberate use of diverse routing and so on.. and if there are explicit rules then under what conditions do they operate? If there are not, then what are the emergent development tendencies and how are they expressed in the topology of layers?
- Deployed network analysis and metrification – to be able to successfully model network layers accurately a modelling target is required with some metric to express the level of respresentativity of models, that captures topological traits, whether explicit or the emergent result of local design decisions. What are the traits present? and do any of them indicate any local design behaviour which is not explicit in the design principles?
- Multi-layer network development – How can knowledge of the development processes of network layers and the environment in which they develop be used to generate topologies that are similar to those found in deployed networks? Are there any circumstances under which modelled topologies become similar to measured topologies and do any of these circumstances involve inter-layer feedback?

This thesis will approach the above issues by first surveying current network models. Then, two types of network layers, the IP layer and the SDH layer will be examined, with a close examination of protocols, including routing protocols, planning methods and a discussion of how these influence topology. These influences will be distilled into development processes which are captured in a new tool called the "Modular Inter-layer feedback Topology InvEstigation tool and simulator", or MITIE. MITIE is then used to simulate a series of network development scenarios to investigate under what conditions networks develop to form topologies similar to those analysed and whether any of these scenarios require inter-layer feedback to be representative.

1.5 Thesis Structure

This thesis is organised into the following chapters:

Chapter 1: Multi-layer Network Provisioning and Inter-Layer Feedback – this chapter examines network provisioning and the processes behind changes in the configuration of the network, as well as time scales and implementation types and examines how the objective functions of one layer, and the configuration of another lead to changes in the network's configuration.

Chapter 2: Telecommunications Network Topology Modelling – this chapter is a review of current state-of-the-art of the models available to explicitly capture a layers' topology, as well as more general models which may capture certain other traits of theirs. We will see that while many are representative of certain types of network layer none consider in any way the demand being carried, nor the modelled layer's reaction to that demand. The current models are non-reactive and simply attempt to imitate the result of actual network growth, rather than capture the dynamic development of these layers.

Chapter 3: The Network Layer: IP Networks – this chapter examines the IP (Internet Protocol) protocol, and architectural aspects of deployed IP networks and specifically the architecture of the Internet. The chapter continues to describe network topology measurement techniques and presents results from such a measurement, as well as describing other measurement work.

Chapter 4: Transport Networks: SDH – this chapter examines the protocols and design behind SDH (Synchronous Digital Hierarchy) and presents an analysis of a deployed national network from a complete data set (not passively measured or inferred like the IP network), including that of SDH node topology and the demand pattern.

Chapter 5: Single Layer and Multi-Layer Network Topology Modelling – this chapter presents the results of experiments performed using MITIE (see Appendix A), which is a new topology

generation tool created based on observations made in Chapters 3 and 4, and through these results by inference and deduction suggests the existence of inter-layer feedback effects, and their emergent topological properties.

Chapter 6: Conclusions – presents conclusions based on the research in this thesis and proposes future directions in inter-layer effects research and the development of MITIE.

1.6 Research contributions

The following peer reviewed articles, book chapter and papers have been published during the course of this research:

- Spencer J., Antonopoulos A., Sacks L., O'Reilly J.J., "Resource Allocation in WDM OMS-SPRing Architectures with Arbitrary Demand Patterns"; Proceedings of European Conference on Networks and Optical Communications, NOC '99, Delft, The Netherlands, 22-24 June 1999.
- Spencer J. and Sacks L., "IP Network Topology and the Impact of Underlying Transport Networks", in Proceedings of London Communication Symposium 2001, 10-11 September 2001.
- Spencer J. and Sacks L., "Modelling IP Network Topologies by Emulating Network Development Processes", Proceedings of IEEE Softcom 2002, Split, Croatia, October 2002.
- Spencer J., Johnson D., Hastie A. and Sacks L., "Emergent Properties of the BT SDH network", BT Technology Journal, April 2003.
- Spencer J. and Sacks L., "On Power-Laws in SDH Transport Networks", in Proceedings of IEEE ICC 2003, May 2003, Anchorage, Alaska, USA.
- Chapter 5 by Spencer J., "Telecommunications Network Modelling. Planning and Design", S. Evans (Ed.), IEE Press 2003. ISBN 0863413234.
- Spencer J., Dueser M., Zapata A., I. de Miguel, Bayvel P., Breuer D., Hanik N., Gladish A., "Design considerations for 100-Gigabit Metro Ethernet (100GbME)" in Proceedings of ECOC 2003, September 2003, Rimini, Italy.
- Zapata A., I. de Miguel, Dueser M., Spencer J., Bayvel P., Breuer D., Hanik N., Gladish A., "Performance comparison of static and dynamic optical metro ring network architectures", in Proceedings of ECOC 2003, September 2003, Rimini, Italy.

- Zapata A., I. de Miguel, Dueser M., Spencer J., Bayvel P., Breuer D., Hanik N., Gladish A., "Impact of non-uniform traffic on performance and resource requirements in optical ring networks", in Proceedings of ECOC 2003, September 2003, Rimini, Italy.
- P. Georgatsos, J. Spencer, D. Griffin, P. Damilatis, H. Asgari, J. Griem, G. Pavlou and P. Morand, "Provider-level Service Agreements for Inter-domain QoS delivery", Proceedings of Fourth Int. Workshop on Advanced Internet Charging and QoS Technologies (ICQT04), Springer, September 2004.
- Zapata A., Dueser M., Spencer J., Bayvel P., I. de Miguel, Breuer D., Hanik N., Gladisch A., "Next-Generation 100-Gigabit Metro Ethernet (100 GbME) Using Multiwavelength Optical Rings", IEEE/OSA Journal of Lightwave Technology November 2004, Vol. 22, No.11. pp 2420-2434.
- H. Asgari, M. Boucadair, R. Egan, P. Morand, D. Griffin, J. Griem, P. Georgatsos, J. Spencer, G. Pavlou, M.P. Howarth, "Inter-Provider QoS Peering for IP Service Offering Across Multiple Domains", in Proceedings of 2nd International Workshop on Next Generation Networking Middleware (NGNM05), IFIP Networking Conference 2005, Waterloo, Canada, 2-6 May 2005.
- P. Levis, M. Boucadair, P. Morand, J. Spencer, D. Griffin, G. Pavlou, P. Trimintzios, "A New Perspective for a Global QoS-based Internet", Journal of Communications Software and Systems, 4th quarter 2005.
- D. Griffin, J. Spencer, J. Griem, M. Boucadair, P. Morrand, M. Howarth, N. Wang, G. Pavlou, H. Asgari, P. Georgatsos, "Inter-domain Routing through QoS Class Planes", IEEE Communications, special issue on Quality of Service Routing Algorithms for Heterogeneous Networks, Vol. 45, No. 2, pp. 88-95, IEEE, February 2007.
- N. Wang, D. Griffin, J. Spencer, J. Griem, J. Rodríguez Sánchez, M. Boucadair, E. Mykoniati, B. Quoitin, M. Howarth, G. Pavlou, A. J. Elizondo, M. L. García Osma and P. Georgatsos, "A Framework for Lightweight QoS Provisioning: Network Planes and Parallel Internets", poster, in Proceedings of the IEEE/IFIP International Symposium on Integrated Management (IM'2007), Munich, Germany, IEEE, May 2007

The topology generator, MITIE, which was the result of part of this research, is available at: **http://mitie.jasonspencer.org** or **http://www.ee.ucl.ac.uk/~jsp/mitie/**

MITIE also formed the basis (especially the load based reactive elements) for another simulation tool, qBGPSim which investigated how a QoS (Quality of Service) aware BGP (Border Gateway

Protocol), qBGP [QBGP], behaved on a large scale, which was research carried out as part of the IST MESCAL (**http://www.mescal.org/**) and IST AGAVE (**http://www.ist-agave.org/**) projects.

1.7 Conclusion

This chapter has examined the processes behind network design and the actions and reactions between layers, and gave an example of this in action, while laying the foundation of the rest of the thesis and the incentive for investigating the inter-layer effect on a layer's topology.

Chapter 2
Telecommunications Network Topology Modelling

"Essentially, all models are wrong, but some are useful" – George Box, economist, 1987

2.1 Introduction

When examining networks, it is often necessary to have a mathematical or analytical representation of the network to be able to examine behaviour which would not be feasible to examine in a real network. Network topology models allow us to examine many aspects of the network such as scalability and tolerance to failure as well as being a necessary foundation to the modelling of other network behaviour such as traffic congestion (e.g. traffic models and buffer design) and protocol behaviour (e.g. TCP behaviour). Since telecommunications networks are man made they are planned and dimensioned according to an algorithm or some kind of design paradigm. This algorithm is typically explicit and created to optimise or somehow influence a trait of the network (e.g. protection in SDH networks [G.841]), in which case it is easier to model the network because the decision processes are known; the only unknowns are external influences or conditions such as geography (land types, customer distribution etc), demand pattern dynamics and so on. It is also possible that networks are designed without a specific master algorithm but rather subject to local decisions, by reactive processes and by qualitative decisions (often the case with IP and Ethernet networks, depending on their use, LAN, MAN etc.). These networks are more difficult to model because they do not have well known processes and can rely on many external influences.

Whether the design is local or global, whether explicit or implied all depends on the application and the environment – thus a TDM network is often globally planned with an explicit algorithm because explicit resource reservation is required and the costs of provisioning are very high – in contrast to an Ethernet Local Area Network whose performance is usually best effort and which is not designed according to some kind of global requirement and tends to have a more adaptive design. The design can also contain significant heterogeneity, and a model would have to capture that – for example the Internet is collection of independently designed networks – some may be closely planned, some less so, and all with different strategies.

In this chapter we will examine the current state of the art of topological models, focusing particularly on models based on statistical physics. This chapter will look at the available implementations and then an examination of the models' applicability to real known networks is made.

2.2 *Network Modelling*

A typical application of network models would be to perform hypothetical operations on the network, such as extrapolating network size to investigate scalability [ADA01] or the effect of failure could be investigated [ALB00a]. The usefulness of the model, however, goes only as far as the model is representative of the network. While the model may be generated from known network planning rules, such as may be the case for SDH [SOR99][WAS94], it is often the case that the precise network creation laws are not known or differ by application (e.g. IP networks) and the models are selected, or designed, to fit to empirical network measurements. It is also common that the model does not capture all of the properties of the network and may be similar in one metric while failing to resemble the network in other metrics. For example, the model used in Inet 2.0 (see section 2.3.1.5) produced topologies which were quite similar in terms of degree distribution, rank and so on, but not in vertex cover [PAR01].

For a model to be useful it must be accurate and representative of one or more of the traits that are to be captured. Further, the model should be scalable and efficient to implement. Ideally the model should also be easily programmable and simple, while maintaining flexibility.

It is important to capture as closely as possible all features of a network in the model as experimentation on topologies generated by models can give significantly different results between models, for example: use of ideal topologies in congestion analysis [ANA02], resource consumption in multicast over regular topologies [MIT94], multicast tree efficiency in random vs Doar-Leslie topology [DOA93], attack and error tolerance in random and scale-free networks [ALB00a].

Another useful aspect of modelling networks is the ability to capture topological properties in metrics or parameters to the model. We will use these metrics, as well as a measure of correlation to fit to the empirical data as in Chapters 3 and 4 to validate models proposed in Chapter 5.

2.2.1 *Power-laws in Telecommunications Networks*

While there has been no intentional incorporation of power-laws into telecommunications networks it does seem like it is a common emergent property. We will see in Chapters 3 and 4

that when examining the deployed topology of the Internet and a national SDH network that a number of metrics are distributed with a power-law. The four best known ones, discovered by Faloutsos et al. in the Internet [FAL99] are based on node degree (the number of links connected to the node), the node's ranking in order of degree, the eigenvalues of the adjacency matrix, and the shortest paths through the network (for more details see section 3.6.2).

These are important because they are the basis of many of the models described in this chapter and are the result of not only many local design decisions in many autonomous systems, but also the result of global peering agreements and routing policies, user type and distribution and the technologies involved in implementing the Internet. Magoni and Pansiot [MAG01] also found four further power-laws. It should be noted that Magoni and Pansiot explicitly state that whenever they came across a distribution governed by a power-law that they would state it. This has been a general comment on the discovery of so many power-laws – that it is the process of searching for them that is causing so many to be found, while neglecting other possible relationships. Chen et al [CHE02] argued that the power-laws discovered by Faloutsos et al could be the result of their AS topology sampling method, and that they could be missing as much as 30% of the links in the topology; they also suggested that with their extended data set the degree distribution is less conforming to the power-law. Faloutsos et al returned with Siganos to revisit the power-laws in [SIG03] and found they still hold with newer and more complete data sets.

While in Chapter 3 we see that there is potential for the power-law to be an artefact of the Internet sampling method, it is not the case for the SDH network studied in Chapter 4, since it was a complete topology that was examined, and as such power-law compliance is a significant goal for the network models we examine below.

2.3 Topology Models

Topology models are varied and selected closely to the network under investigation and the types of experiments that will be performed. When modelling the topologies of layers in a telecommunications network the models could be classed into structured and random. Structured networks tend to be more regular, based on some kind of topological primitives or hierarchy, while random networks are more stochastic in nature and typically less structured and more irregular. The two are not, however, entirely exclusive – it would be possible to generate a model

which considers hierarchy, but uses stochastic processes within each hierarchical layer to govern connectivity.

2.3.1 Structured models

These models typically have a complete view of the network and create explicit structures which consist of more than one link (ring, triangles, meshes, as well as entire end-to-end paths) to construct the network. These models can capture some element of hierarchy or heterogeneity in the network and have assigned a specific purpose to certain constituent nodes, such as transit vs. stub ASes or ring interconnecting DXCs (SDH Digital Cross-connect, see Chapter 4), which influences their connectivity affinity and properties. These models can make decisions based on additional global information such as traffic matrices or traffic forecasts. Since explicit structure is imposed on the network, global properties can be specifically added like the existence of N:M protection paths (N number of circuits share their protection with M backup circuits - see Chapter 4 for a detailed description) and the network can be optimised for global properties such as availability, route diversity and so on.

The models could be based on the explicit design process of the network, if it is known, such as the algorithms in SDH network planning algorithms, but these are often proprietary information. There are also often parameters in the models, such as the geographic distribution of nodes, which can drastically change the model output and these again are often unknowns and require models themselves.

In Chapter 4 we will, however, see how the external influences such as geography and demand pattern cause emergent properties in a deployed SDH network and examine the explicit structures that exist in SDH networks. While the network is strictly planned and structured, many of the properties seen in Chapter 4 can be also modelled by a power-law random model, examples of which will see in the next section. Similarly Chapter 3 looks at the design principles of IP networks and the Internet, as well as topology measurement techniques and actual measurements, with the intention of extracting models, which again may be modelled either by structured models, such as transit-stub below, or by power-law compliant random models.

Currently available structured network topology models include:

2.3.1.1 Transit-Stub

This is a simple model [ZEG96] of Autonomous Systems and their inter- and intra- domain connectivity. The transit-stub architecture consists of a set of nodes, connected using some random link placement algorithm, which forms the set of transit domains (the Transit-Stub model is actually implemented as part of the GT-ITM topology tool and the link placement methods are better described in section 2.4.1 where GT-ITM is examined). Then, each node in the transit domain topology is replaced by another random graph, which represents the internal topology of the transit domain. Finally for each transit network one or more stub networks are created and attached to one or more transit networks.

2.3.1.2 GHITLE: Generator of Hierarchical Internet Topologies using Levels

GHITLE is a model [GHITLE] which attempts to generate AS level topologies by considering peer and customer-provider relationships, building a highly connected tier-1 core, a number of intermediate levels, and then stub networks. Parameters in the model include the ability to specify the number of levels in the hierarchy (how many additional levels of hierarchy exist between the stub level and the core), the random node inter-connection strategies (add new links based on uniform randomness, or use Albert-Barabasi style node preference (see later in this chapter)) and the ability to control the number of links that are generated between levels of the hierarchy.

2.3.1.3 IGen: Topology generation through network design heuristics

IGen [IGEN] implements various network design heuristics such as MENTOR [CAH98], MENTour [CAH98], Delaunay triangulation and Two Trees [GRO04] for the purpose of building network topologies. IGen works by first creating a set of nodes, these are grouped into PoPs (Points of Presence – see Chapter 3 for an overview of PoPs) and then these PoPs are interconnected with a backbone. This backbone is generated based on the MENTOR, MENTour, Delaunay triangulation or Two Trees heuristics. [QUO07] presents a more in-depth description of the heuristics and a comparison to known network topologies.

2.3.1.4 AStop: A New Topology Generator at the AS Level

AStop is a topology generator [NIE06] based on empirical measurements from BGP tables (taken from RouteViews databases) and business relationships that exist between ASes. The model is based on the observation that there is a highly connected core of the Internet and, therefore, has two parts, the first builds a core network of ASes and the second attaches edge/customer ASes. The method used to connect the core is similar to the PLOD model (see Section 2.3.2.6) except that the sum of the expected degrees is the number of links in the core added to those that connect to stub nodes. Then, only the number of links expected in the core (that is the total number of

links less the number of links to the edge ASes) are added using the PLOD model. In the final step edge ASes are connected to the core nodes.

2.3.1.5 Inet: Internet Topology Generator

Inet [INET] has gone through a number of different versions and changed its method of topology generation. In version 1.0 (released in 2000) the topology was formed by first assigning an expected degree to each node according to Faloutsos et al's [FAL99] degree rank power-law and connecting the τ highest degree nodes (where τ is a parameter) with a full mesh, and then a fraction of the remaining nodes (outside the set constituting the full mesh) are randomly connected to two of the nodes in the full mesh, as long as they have not achieved their expected degree yet. The remaining nodes are then connected to either a node in the τ set or to a node that is directly connected to a member of the τ set, again within the constraints of the expected degree. There is a second optional phase where the κ highest degree nodes are expanded into networks of η nodes each.

Inet 2.0 (released also in 2000) [INET20] takes a very empirical approach to Internet modelling by looking at various measured Internet topologies (NLANR [NLANR] datasets at different points in time between 1997 and 2000) and attempts to match various analytical metrics and distributions as closely as possible. The model also incorporates the observation that the Internet exhibits exponential growth, and took the Faloutsos et al. [FAL99] degree distribution observation further and specified the distribution as

$$f_d = e^{at+b} d^O;$$

Equation 1 The degree distribution model in the Inet topology generator

By fitting this distribution to their data sets taken at different times, they found values for a,b and O. They made similar changes to the rank power-law of Faloutsos et al. and found values for those coefficients also. The only parameters their topology generator takes is N, the size of the network, and k, the fraction of the nodes that have a degree of 1. From N they can extrapolate the time t at which the Internet would be that size, and using the known values for a, b and O from Equation 1 and the rank-degree equivalent they find the hypothetical distribution for node degree and rank for that network size. The degree rank equation is then used to find the degrees of 2% of the nodes, while k.N nodes are assigned a degree of 1, while the rest have degrees assigned by the

degree frequency distribution extrapolation. The nodes are then connected using a preferential attachment similar to the BA model, but with caveats, as described in [INET20].

Inet 3.0 [INET30] then added tweaks to the Inet 2.0 model by being more accurate about the treatment of the three highest degree nodes, as well as substantially changing the preferential connectivity function. The function remains as a normalised weight, but rather than the weight being degree it is a function of degree and the frequency of the degrees in the rest of the network, thus:

$$w_i^j = \max\left(1, \sqrt{\left(\log\frac{d_i}{d_j}\right)^2 + \left(\log\frac{f(d_i)}{f(d_j)}\right)^2}\right).d_j$$

Equation 2 The weight of a node in the preferential connectivity in Inet-3.0

where w^j_i is the weighted value of d_j with respect to a node i, d_i is the degree of node i and $f(d_i)$ is the frequency of degree d_i.

The preferential connectivity function then becomes:

$$p(i,j) = \frac{w_i^j}{\sum_{k \in G} w_i^k}$$

Equation 3 The probability that node i connects to node j in the Inet-3.0 preferential attachment model

where $p(i,j)$ is the probability that node i is connected to node j.

This change in the weights in node selection is a significant change because it makes the relative frequencies of degrees a factor in selection and, therefore, captures an element of assortativity [NEW03][ZHO04][VAZ02] (assortativity is the tendency for nodes of a similar degree to link to each other, while disassortativity is the tendency for links to occur between nodes with more different degree values), although the original purpose is to improve the vertex cover accuracy of

the model (vertex cover is a minimal set of nodes which, together with their direct neighbours constitute all of the nodes).

2.3.2 Random Models

These models have no initial explicit design goal, and are created using local connectivity decisions based on existing metrics of the nodes, often capturing an element of growth, and then exhibit macroscopic properties which are not inherently present in the local decisions. It is this mapping of local decisions to global emergent properties that is the main focus of Chapter 5. These models could also be described as unstructured, implicit, statistical physics or stochastic models.

There are a very large number of random graph models all with interesting and potentially applicable properties but this chapter will focus as much as possible on random models that may be useful in modelling IP and SDH networks, as these are the ones that are more closely examined in later chapters. The main focus is on models that produce power-law compliant networks (that is networks which have a power distribution with some metric of the topology) as these are similar to the IP and SDH networks, as we shall see, and are a well researched area. Power-laws also appear in the topologies of other natural and man-made networks such as the world-wide-web [ALB99], the network of chemical reactions in a cell [JEO00] and social networks (research paper co-authorship) [BAR02], human sexual contact [LIL01] and actors [ALB00b]), the telephone call graph [ABE99] and the power grid [BAR99] among others.

2.3.2.1 Erdös-Rényi (ER) random graph model

The Erdös-Rényi random graph model [ERD60] is one generated by either of two variants. G(N,E) is a graph which is chosen uniformly at random of the set of graphs that have E edges and N vertices. The second variant, G(N,p) is a graph that consists of N nodes and is constructed by connecting every node pair with a probability p. Therefore, a G(N,p) has an average of 0.5N(N-1)p edges. Other properties of the model include:

The degree distribution of a vertex v is binomial, that is:

$$P(d_v = k) = C_k^{N-1} p^k (1-p)^{N-1-k}$$

Equation 4 The degree distribution of the Erdös-Rényi random graph model

where d_v is the degree of vertex v, N and p are from the graph definition, G(N,p), while $C^{N-1}{}_k$ is the number of combinations k elements can be chosen from a set of N-1 elements, without repeating elements.

For large values of N this tends to a Poisson distribution, and, therefore, ER graphs are often synonymous with exponential degree distribution graphs.

E-R graph have been shown [CHU01] to have a diameter of

$$d = \frac{\ln(N)}{\ln(\langle k \rangle)};$$

Equation 5 the average diameter of an Erdos-Renyi random graph

Where N is the size of the network and $<k>$ is the average degree.

E-R graphs have another interesting property, that as vertices are added the graphs have a tendency to reach a critical point in the vertex to edge ratio where the graph suddenly consists of one very large cluster [BOL84] rather than many small ones. That is, for values of N/E less than 0.5 (that is there is less than one link per node) the network tends to be a large collection of small clusters of nodes and most likely a disconnected network, but if N/E increases past 0.5 the network tends quickly toward a single large cluster with additional small clusters.

2.3.2.2 Albert-Barábasi (BA) model

This graph model [BAR99] incorporates two processes that widely exist in real networks, that is growth, where the number of edges increases over time, and preferential attachment, where during the addition of vertices to the network the decision of which edges to connect to is not uniformly random and exhibits a preference to some kind of property of an edge, in this case the degree of that edge.

Specifically a BA graph is initialised with a seed network where there the network is a single cluster and edges are added to it and a new vertex is added from the new edge to an existing edge with the probability:

$$p_i = \frac{k_i}{\sum_j k_j}$$

Equation 6 The probability that a newly added node is connected to an existing node in the Albert-Barabasi random graph model

where p_i is the probability of connecting the new edge to edge i, and k_j and k_i are the degrees of edges j and i respectively.

During growth the tendency is for the high degree (relative to the other node degrees) nodes to accumulate more links and increase their degree, which is often referred to as the "rich get richer" effect.

As the network grows in size, the degree distribution can be shown to approach:

$$P(k) \propto k^{-3}$$

Equation 7 The degree distribution of the Albert-Barabasi random graph model

where P(k) is the probability that a node has degree k.

Due to the power-law nature of the degree distribution these networks are often called power-law networks or scale-free networks. That is, the shape of the graph will not reveal where on the graph we are looking – the left side is similar to the right side (subject to translation) – in comparison to, say, the ER model earlier which had a degree distribution (Equation 4) which was a binomial and therefore each section of the graph looked unlike other sections.

The model also has an average path length of:

$$l \sim \frac{\ln(N)}{\ln(\ln(N))};$$

Equation 8 The average path length of the Albert-Barabasi random graph model

where N is the network size. This is significantly shorter than the E-R model when considering sparse networks (k<<N).

The BA model also has two variants which are the limiting cases – known as BA model A, where there is no preferential attachment, just growth, and does not exhibit the power-law properties, and BA model B which has growth, but no preferential attachment and adds links to a graph of fixed size (which is equivalent to the ER model) and while early on in the growth may look power-law like, later tends away towards a Gaussian degree distribution as more vertices are added. Albert and Barabasi [ALB00b] also considered the requirements of node and link growth, as well as rewiring, and the conditions under which power-law networks occur. Barabasi, Albert and Jeong examined how the networks evolved as they grew and examined the dynamics of the degree of nodes over time [BAR99b].

2.3.2.3 GLP (Generalized Linear Preference) model

The GLP model [BU02] is an extension of the BA model and adds a parameter to the probability function used in selecting the nodes to attach to, and in an additional link growth stage. The new node selection probability function is:

$$p_i = \frac{(k_i - \beta)}{\sum_j (k_j - \beta)}$$

Equation 9 The probability that a node *i* is selected for connection where *ki* is the degree of the node *i* and β is a parameter between –∞ and 1

The parameter β (which is between –∞ and 1) can now be used to influence to what extent the selection is biased towards the higher degree nodes. The initial network of m_0 nodes is connected with m_0-*1* edges, then with a probability p, m links are added (where $m \leq m_0$) (where p is a probability parameter) between nodes which are selected based on Equation 9. Then, with a probability *1*-p a new node is added and connected to m existing node, again selected by the above equation.

2.3.2.4 IG (Interactive Growth) model

Zhou and Mondragon [ZHO03] proposed a model similar to the GLP model but instead they first add the new node, and then make the link addition process a related event: they fixed one of the endpoints of the link growth link to be the existing node that was connected to in the node growth stage (that is the node to which the new node was connected to).

2.3.2.5 *PFP (Positive Feedback Preference) model*

Zhou and Mondragon also proposed a second model, the PFP model [ZHO05] which extended the IG model to add a third growth possibility. The PFP model node has three growth scenarios:

- Add a new node and link to an existing node in the network, and add a second link from the existing node to another node in the existing network.
- Or, add a new node and link to an existing node in the network, and add two further links from the existing node to two other existing nodes in the network.
- Or, add a new node and add two links from it to two other existing nodes in the network.

The probability of selecting a node for linking was also changed to:

$$\prod(i) = \frac{k_i^{1+\delta \log_{10} k_i}}{\sum_j k_j^{1+\delta \log_{10} k_j}}$$

Equation 10 The probability that a node is selected for connection, where k_i is the degree of node i and δ is a parameter in the range [0,1].

2.3.2.6 *PLOD (Power-Law Out-Degree) model*

The PLOD topology model [PAL00] explicitly selects node pairs to be connected such that the overall degree distribution is a power-law. The model initially creates a set of disconnected nodes and assigns "degree credits" to each node at random. These credits are chosen such that each node is assigned a degree of

$$\beta x^{-\alpha}$$

Equation 11 The degree assigned to a node in the PLOD model where $x \sim U(1,N)$

where α and β are parameters and x is a uniformly randomly chosen number between 1 and N inclusive, where N is the number of nodes in the graph. Then, pairs of nodes with degree credits greater than 0, and which are not already connected, are selected uniformly at random and connected, and the degree credit is decremented for both of the newly connected nodes. [PAL00] shows that this generates topologies that are compliant in the three power-laws found by Faloutsos et al [FAL99] in Internet topologies, as well as being similar to their approximation of hop-count.

2.3.2.7 *Waxman Topology model*

This model [WAX88] uses the position of nodes on a Euclidean plane and the distances between them to decide their connectivity. The model states that the probability of two nodes being connected drops exponentially as the Euclidean distance between them increases. Therefore, for a plane containing disconnected nodes the probability of two nodes, u and v being connected is:

$$P(u,v) = \beta e^{\frac{-d(u,v)}{L\alpha}};$$

Equation 12 The probability that two nodes are connected in the Waxman topology model

where α and β are in the range (0,1] and are both parameters, $d(u,v)$ is the Euclidean distance between nodes u and v and L is the largest distance between any two nodes in the network. The model was first proposed to generate random topologies for use in experiments on building multicast trees.

The exact topological properties, such as average network diameter and degree distribution, are difficult to specify because they are so dependant on the node distribution and the α and β parameters. Zegura [ZEG97] when quantitatively comparing a selection of models did, however, plot the number of links given different value for α and β, as well as offering some results for average degree and topological diameter, assuming a uniformly random distribution of nodes.

2.3.2.8 *Doar-Leslie Topology Model*

The Doar-Leslie model link placement [DOA93] is a modification of the Waxman model, [WAX88] where the probability of two nodes being linked is proportional to the Waxman probability, but also scaled by a function of inverse of the network size and the average degree of nodes, thus

$$P(u,v) = \frac{k < d_i >}{N} \beta e^{\frac{-d(u,v)}{L\alpha}};$$

Equation 13 The probability that two nodes are connected in the Doar-Leslie topology model

Where N is the number of nodes in the network and $<d_i>$ is the average degree of all nodes, while k is a parameter which Doar and Leslie attempted to match to real topologies.

2.3.2.9 *Watts-Strogatz Small-World model*

This model attempts to capture small-world properties, that is properties of a graph which make its diameter small. The Watts-Strogatz model [WAT98] proposes an initial regular ring lattice of N vertices, connected to K neighbours, then, with a probability β each edge has one of the endpoints rewired to another edge which is chosen uniformly at random.

This kind of a network approaches an Erdös-Rényi network (equivalent to β of 1, i.e. fully randomly rewired) as β increases but before that happens it demonstrates a number of interesting properties. The networks tend to quickly drop in characteristic path length (the average diameter between all node pairs) after β of 0.001, as well as maintaining their high clustering coefficient until a relatively high β of 0.1.

2.3.2.10 *FKP (Fabrikant, Koutsoupias, Papadimitriou) model*

The FKP model [FAB02] is significantly different from the other models in that instead of attempting to model simple connectivity processes, which then potentially have an implicit feedback process, like the rich-get-richer of the BA model, or the complete lack of feedback in the Waxman model, the FKP model explicitly has two components to its link selection function that would work against each other. FKP's model was inspired by work on HOT (Highly Optimized Tolerance) [CAR99] and is often referred to as the HOT model (but referred to here as FKP to avoid confusion). The FKP model's node selection function has, on the one hand a component that prefers lower Euclidean distances, but on the other hand has a component that prefers more topologically central nodes.

In the FKP model the network experiences node growth, like in other models such as the BA model, but places them uniformly at random on a Euclidean plane and starts by connecting newly added nodes with a single link to an existing node selected uniformly at random to build an initial tree. Then every newly added node is linked to a single old node selected to satisfy the condition following condition:

$$\min_{j<i} \alpha.d(i,j) + h_j;$$

Equation 14 the condition that must be satisified when selecting the connecting node in Fabrikant et al.'s HOT topology model

where i is the node being added, and node j is one of the previous nodes, while α is a parameter, $d(i,j)$ is the Euclidean distance between nodes i and j, and h_j is some measure of the centrality of node j. h_j can take a number of forms as long as it expresses the node's topological centrality in

the existing graph; Fabrikant et al gave examples such as (a) the average number of hops from other nodes, (b) the maximum number of hops from another node or (c) the number of hops from a fixed centre of the tree.

The parameter α is used to control the relative influence the distance or the centrality plays. Fabrikant et al [FAB02] go on to prove that for the (c) measure of centrality above a power-law degree distribution occurs in the created topology and showed the same experimentally for (a) and (b), as long as α is within a given bound. They showed that if α was below $1/\sqrt{2}$ the graph became a star topology, if $\alpha = \Omega(\sqrt{n})$ (that is $\sqrt{n}$ is an asymptotic upper bound), where n is the number of nodes in the graph, then the degree frequency distribution is exponential, while if $\alpha \geq 4$ and $\alpha = o(\sqrt{n})$ (that is α is dominated by $\sqrt{n}$) then the degree distribution is a power law. These results are highly significant because similar push-pull forces will be seen in the multi-layer network models proposed later in this thesis.

2.4 Model Implementations

We have now examined the models, but there are also aspects of models that are specific to their implementations and this is what we will now consider. GHITLE, IGen and ASTop and FKP are already directly implemented and do not differ from their descriptions above, so in this section we will look at the remaining better known topology generator tools, which models they implement and their specifics.

2.4.1 GT-ITM: The Georgia Tech Internetwork Topology Models

This is a collection of tools to generate and analyse a wide variety of internetwork topologies, including routines to generate topologies based on geography, with a flat, N-level hierarchical or transit-stub architecture. GT-ITM [GTITM] supports a number of link placement algorithms, including an implementation of a uniformly random model (equivalent to E-R), the Waxman model, the Doar-Leslie model, and a model [ZEG96] based on the exponential of the distance between nodes.

$$P(u,v) = \beta e^{\frac{-d(u,v)}{L-d(u,v)}};$$

Equation 15 The probability that two nodes are connected in the Exponential model of GT-ITM

where the variables are as in Equation 12.

2.4.2 *BRITE: the Boston university Representative Internet Topology gEnerator*

This topology generator is one of the most used topology generators in recent years for creating Internet-like topologies at the router and AS level. BRITE [MED01] includes implementations of the BA model, including the two variants, as well as the Waxman model. BRITE also has support for a number of input file formats for loading topologies from other topology generators.

For use with the Waxman model BRITE is also capable of Euclidean random node distribution including the ability to place nodes in clusters and then distribute those clusters around the Euclidean space, emulating a heavy tailed node distribution (specifically a pareto distribution).

BRITE is also capable of assigning bandwidths to links according to a range of distributions (constant, uniformly random, exponential distribution, pareto distribution), but the bandwidth values are randomly generated according to the distribution and not derived from any network property. The link bandwidths are also not used at all in any part of the topology generator, just as part of the output.

Because BRITE is intended to model AS and router topology it also includes support for hierarchical topologies. This is where a two level topology is created, one of ASes, and contained within these ASes are router topologies. The connectivity at each level of hierarchy is generated by one of the previously mentioned models, that is BA, Waxman, or imported topology files.

2.5 *The Development and Applicability of Network Models*

In the last 20 years the evolution of models and topology generators has closely followed what is known about the networks being modelled. In the case of the Internet it was known originally that ASes have different roles and different connectivity, hence the creation of transit-stub/GT-ITM, and that there is some correlation between connectivity and Euclidean distance, hence the creation of Waxman and Doar-Leslie. These were created based on design principles, but without any empirical knowledge of the final resulting deployed network.

With the discovery of power-laws by Faloutsos et al [FAL99] and the other analyses [MAG01][CHE02][SIG03] the old models were found not to be topologically similar [TAN02]. The focus in network modelling then moved to the creation of power-law compliant networks - initially the BA model was popular because it was simple to implement and created a power-law distribution – but the specific metrics, such as the exponent values were significantly different (the BA model had an exponent fixed at 3, while the AS topologies from [FAL99] had exponents of 2.15, 2.16 and 2.20). The GLP, PLRG and PLOD models were then proposed to create power-law degree distribution topologies with arbitrary exponents.

The BA model also continued the concept of modelling local design processes (e.g. previously seen in distance cost in Waxman) and it was posited that newly added nodes would prefer to connect to better connected nodes as they had better access to the rest of the network. The observation of assortativity in Internet topologies [NEW03][ZHO04][VAZ02] then demonstrated a weaknesses of the BA model, which was not representative of the assortative mixing in the internet, and then the GLP model was proposed which added the second link addition stage and could control the extent to which high degree nodes were favoured.

GLP, although power-law compliant, was not, however, representative of the AS graph, which featured a less than power-law predicted frequency for one, two and three degree nodes [INET20][ZHO03]. This lower than power-law predicted frequency of low degree nodes was captured in Inet 2.0, as was a decreased degree in the very highest degree nodes [INET20]. Inet, through various versions, initially worked by attempting to mimic the construction of the core and edge networks, and then by direct curve fitting in versions 2.x and 3.0. Inet 2.0 was criticised [PAR01] for not representing correctly the AS graph's vertex cover and, therefore, Inet 3.0 was an improvement over Inet 2.0 in terms of assortativity conformance and vertex cover.

Research into node and link growth and preferential attachment was taken further by Zhou and Mondragon with the Interactive Growth (IG) [ZHO03] and Positive Feedback Preference (PFP) [ZHO05] models. The IG model used a more intuitive method of adding a second link - rather than adding it somewhere deep in the network disjoint to any other growth, the new link was added to the node that was the host of the newly added node, since that would be where the increased load would be coming from. They showed that IG performed better than GLP for the low degree nodes. PFP took the preferential attachment class of models yet further by adding a dual homing scenario for newly added nodes and this created a model that was in many metrics very representative of the AS graph. The model captured [ZHO05] the rich-club (assortative mixing among high-degree nodes) component of the AS graph, shortest path lengths and betweeness centrality more accurately than the models before it.

Fabrikant et al's HOT model (the FKP model) is a paradigm shift away from the others – while it is still about preferential attachment it introduced self-organising criticality into modelling and two competing forces that together co-adapt during network growth to create what is surprisingly a power-law compliant network. It has been shown [TAN02] that the Waxman and Doar-Leslie models do not create topologies that conform to the power-laws but when used together with a second component that instead of wanting to connect to nodes that are nearer by (on the Euclidean plane) exerts the need to connect more to the core of the network (as expressed

by the measure of centrality), then power-law networks are produced. The relative influence of these components is important as if too much of one or too much of the other dominates the node selection equation then the network is no longer a power-law. The concept of this push-pull was proposed by Carlson and Doyle [CAR99] and is based on the concepts of self-organising criticality, such as that found in sand-piles (where increasing pile size is balanced by gravity).

This is not to say, however, that the other models are without merit – they share various features and properties that relate to various features we will see exist in telecommunications networks – for example, the Watts Strogatz model is ring based, like SDH networks often are (and the one analysed in Chapter 4 definitely is), and has the feature of having a much smaller diameter than a random network – and this is a feature found in the SDH network in Chapter 4. Chapter 4 will show that there is a clear relationship between the probability of two nodes being connected and the Euclidean distance between them – and the Waxman, Doar-Leslie and GT-ITM exponential Euclidean models are all examples of this.

2.6 Scope for model research

There is a common failing of all the models described in this section, and that is a disregard for the actual (not estimated or generalised) processes behind the formation of the networks, as well as a complete lack of any reactive elements or dynamism in the models.

While models like PFP are very similar to the Internet, and have shown it in a wide selection of metrics, they will only generate static clones of the Internet, and will not capture the dynamic nature and the response to stimuli that is required to properly test and investigate "what-if" scenarios (the investigation of link failure for example). There is an implicit response mechanism present in some models – for example with PFP and the link added to the node that was most recently attached to but that is not actual feedback to a change, only an estimated possible response. What if the new host node actually had sufficient capacity and it was elsewhere in the network that there was an actual bottleneck? There is the co-adaptive nature of FKP, but there is still much scope in examining the behaviour of a network layer and the feedback it provides through demand matrix changes and through delivered quality of service, as was described in Chapter 1. The behaviour of routing protocols should be an important element in the modelling of the network, since that decides the demand matrix, as should the cost function of the transport network.

We will see in Chapter 5 a broad investigation into the forces a network layer's technology has on the traffic, the methods of feedback and how this comes together to affect the network's topology.

The most interesting concepts in the current models that should be kept in mind when creating future models are:

- The use of hierarchy - definitely present in the Internet and potentially in SDH networks – as seen in GHITLE and AStop, and to a lesser extent in transit-stub and Inet.
- The use of implicit or explicit feedback – BA, IG, PFP, FKP.
- The use of actual network planning structural components - as seen in Igen and in part in WS.

Another aspect that is also missing in previous work is the pervasiveness of heterogeneity in the network and that depending on legacy (whether this is an incrementally grown network with previous generation technology, or a complete redesign), purpose (best effort vs. 99.9999% uptime, LAN vs. MAN, business vs. residential customers), extent (the geographic area covered), business and administrative issues (who controls which networks), and so on - there is potential for a lot of heterogeneity in the network – something which has not been captured by any of these models.

Realistically it is not possible to create a universal topology generator for all types of network layer, nor is it possible to capture all the heterogeneous elements, but the intention of the research described in this thesis (particularly Chapter 5) will be to investigate the multi-layer feedback and to see if it is capable of producing power-law compliant networks, since that is what would be expected in a number of network layers (as will be seen in Chapter 3 and Chapter 4).

2.7 Conclusions

We have seen in this chapter a number of models which over time have adapted to the better understanding of the Internet topology that we have, while others are simple and elegant ways of modelling components (WS) and connectivity factors (Waxman) of SDH networks. We have also seen that feedback systems in topology generation can create topologies similar to those that we will see for the AS graph and an SDH network.

Current models are based on existing knowledge of structure and on empirical data but rarely capture a dynamic element, nor do they capture a "holistic" approach to network modelling. We have shown that there are many models intended for various applications and redefined for new applications and while they can be useful for many scenarios none of them attempt to capture any kind of second order design influences, such as those from carried traffic flows or heterogeneity in the layers being modelled. Next we will see the technologies, protocols and design processes behind IP networks and an analysis of the Internet, in Chapter 3. In Chapter 4 we will examine SDH network design process and look in depth at the analysis of an actual deployed nation-wide SDH network. Chapter 5, inspired by the models here, by the concept of multi-layer feedback, and the knowledge gained from the analysis of deployed networks, will propose and examine new network models in search of the effect of the inter-layer feedback.

Chapter 3
The Network Layer:
IP – the Internet Protocol

3.1 Chapter Introduction

In this chapter we examine IP networks and specifically their topologies for the purposes of extracting design rules, as well as modelling targets for use in the modelling work in Chapter 5. We start by surveying the actual protocol and then the various associated protocols that are responsible for route discovery and distribution. These protocols, the algorithms on which they are based and their specific implementations are responsible for the creation of the IP route topology, as well as being the source of design constraints on the network planners, and are, therefore, closely examined with a special focus on topology influencing features. We also examine the macroscopic Internet architecture, ISP business models and policies, as well as common equipment in networks that comprise the Internet. As there are no strict universal planning rules for the Internet, later in the chapter we will examine the techniques available to create maps of real world deployed IP networks. Original results from a mapping of the European part of the Internet are also presented, as well as an examination of other mapping results in literature.

3.2 IP (Internet Protocol) Networks

An IP network is one that supports the Internet Protocol. The most common version in use is IPv4 [RFC0791], but there is a slow movement to IPv6 [RFC2460]. The IP protocol, first developed circa 1980 provides for the transport of variable size datagrams and provides a connectionless and unguaranteed (there is no end-to-end concept and no guarantee that packets arrive at their destination) packet delivery system. Its main functions are that of payload fragmentation and packet addressing. In the OSI seven-layer model it appears in Layer 3, the network layer. IP was developed as a protocol to provide a standard interface to transfer variable sized datagrams across any supported type of transport network. While IP does have an optional TOS (Type of Service) field in the header IP provides no guarantee of delivery, order of delivery, or any other QoS (Quality of Service) parameter.

To provide more useful services to applications, protocols such as UDP [RFC0768] or TCP [RFC0793] are used and carried over IP. These provide the ability to distinguish between

services on a host through the use of port numbers, some of which are standardised, such as the http (the HyperText Transfer Protocol [RFC2616]) protocol on TCP port 80. TCP also has further features such as reliability (lost packets will be resent), flow control (packet transmission rate adapts to minimise packet loss and delay) and sequencing (packets which are delivered out of order are re-arranged before being sent to the application). A further protocol that is implemented over IP is ICMP (The IP Control Message Protocol [RFC0792]) and is responsible for in-band error and control communication.
From the perspective of this investigation into network layers and demands it is also interesting to note that protocols such as UDP do not adapt to network conditions and the sending host will send such packets at the highest speed it can, as soon as it can, with no concern for congestion; such traffic is commonly called inelastic. TCP on the other hand monitors packet loss and retransmits and adapts the sending rate accordingly; this type of traffic is usually called elastic.

The aspects of IP, related routing protocols, network design and available equipment that most impact delivered quality of services, and, therefore, the resources available for the layer above can be divided into the following categories:

- **IP Addressing**: The use of address and, therefore, route aggregation has a direct effect on the aggregation of traffic.
- **Routing Protocols**: The routing protocols commonly used today at every level of IP networks and the Internet. By examining these we will see if there are any suggestions of structure that are introduced into router and network topologies.
- **Routing Hierarchies**: The division of networks into smaller networks interconnected by border routers creates gateways, or hubs, toward which non-local traffic is focussed. This explicit structure is examined together with its impact on topology.
- **The Design of the Internet**: With no deliberate and explicit global design we examine the post-hoc architecture that has formed in the Internet and the policies of Internet Service Providers, interconnection business models and an example of a real Internet Service Provider to understand what influences the Internet to be the topology that it is.

With an understanding of the factors that influence IP network topology we examine measurements of the actual Internet:

- **Internet Mapping**: The techniques used and problems associated with various approaches to mapping the Internet.
- **A measurement of the European part of the Internet**: A brief look at an experiment that was performed to trace the router topology in Europe and the problems associated with the measurement method used.
- **Faloutsos' empirical Internet Topology study**: A look at the emergent traits that were found by Faloutsos et al. [FAL99] and a survey of commentary on their study.

The aim of this chapter is to search for sources of structure and planning rules in IP networks and examine the emergent topology for the purpose of comparison to topology generator models which will be introduced in later chapters.

3.3 IP Network protocols and routing

The local design and planning of an IP network varies widely with its purpose. Often in the LAN (Local Area Network) IP networks are just incrementally grown and expanded as needed. In the core IP networks, WANs (Wide Area Networks) are planned and dimensioned through the use of offline traffic engineering algorithms but also have an element of dynamicity to them and are sometimes configured and developed on a reactive basis. In both cases (LAN and WAN) IP often contains a predominantly dynamic element to its routing with little or no pre-reserved resources (other than link capacity) and provides a best-effort service. In this section we will examine the route discovery processes, any implied structure they provide, and the structure introduced by the interconnection of IP networks.

Since IP is meant to provide the ability to interconnect many networks to extend the overall reach of the extended network there is no one single global control point and it is not globally planned. Due to this a system for topology discovery is required and also a method for route computation. While IP provides addressing functionality it is left to routing protocols to discover these routes and create routing and forwarding tables. A routing table is a table of subnetworks, or aggregations of subnetworks which are reachable from a particular interface in the router. Assigned to each routing table entry is a route metric that specifies the relative cost of this interface as the next hop. When choosing a next hop the metrics are compared and when there are alternatives the "longest match" (the entry where most network address bits match) is

chosen, otherwise the routing table contains a default next hop for all traffic not matching any of the routing table network address entries.

3.3.1 Routing Hierarchies: Nodes, networks and areas

For very large networks (1000s of nodes or more) it would be difficult to propagate routing information between every node pair and, therefore, networks are commonly partitioned into areas. These areas are themselves seen as nodes in a "higher level" network, as part of a hierarchy. The areas are then inter-connected via specially assigned border routers. The internal configuration of these areas is of no importance to the "higher level" network, only how the border routers of each area are externally connected to other border routers. These areas can cover logical areas (groups of inter-connected routers) or can follow areas defined by geography boundaries or administrative domains. The use of such areas is understandable, as the majority of an area's traffic will usually terminate within that area (e.g. a university campus LAN). An example of the use of logical areas is OSPF (Open Shortest Path First) which will be seen later. Using areas to define administrative domains is usually done in BGP (the Border Gateway Protocol) which uses Autonomous Systems (ASes) to define networks under independent control. In Figure 10 we can see three BGP Autonomous Systems (see section 3.3.3.2.1 for an examination of BGP). Internally an arbitrary topology runs an Interior Gateway Protocol (IGP), such as RIP (see section 3.3.3.1.1) or OSPF (see section 3.3.3.1.2) and the routing of traffic between domains is handled by an Exterior Gateway Protocol (EGP), such as BGP. Interconnection of domains is either through direct peering or through connection to a NAP (Network Access Point) (also referred to as an IXP, Internet Exchange Point).

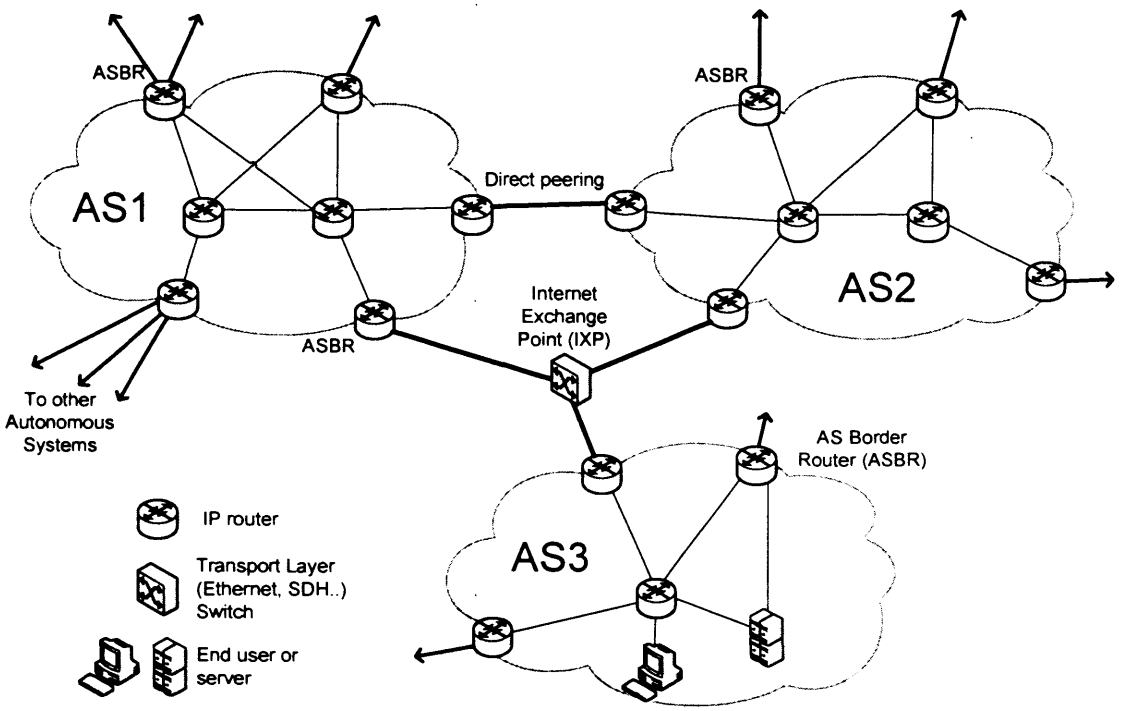


Figure 10 Domains and their interconnection in IP networks

Such a grouping of nodes into areas is a possible source of implicit topology. It also allows for greater heterogeneity in the network, with one area making different assumptions about routing metrics than another. To better understand the topology implications we must consider the actual implementations of routing protocols, but before we do we should consider how hosts in an IP network are addressed.

3.3.2 IP Addressing

In IP version 4 every IP speaking interface on a host is identified with a 32-bit address. This address comprises of two parts: the network (or subnet) address in the higher-significance bits, and the host address in the lower significance bits. The address is usually specified as a four byte tuple, with bytes delimited by dots, e.g. 128.40.40.82. A netmask is also specified which is a bitfield that is logically ANDed with the address to find the network address. Originally the IP address space [RFC0791] was split into four classes: Class A addresses, which were networks consisting of up to 2^{24} host addresses (where the most significant byte specified network, the other three bytes specify host), Class B addresses which had two bytes for the host address and Class C addresses which had one byte for the host address. Class D addresses were special in that they did not specify single hosts, but are a multicast group of hosts. The class of address and, therefore, subnet mask, could be found from the top three most significant bits. This regime for

partitioning the address space is, however, wasteful as the finest granularity of address block allocations is 256 addresses (one Class C subnet). CIDR (Classless Inter-Domain Routing) [RFC1519] was, therefore, introduced to allow for the finer allocation of address space. Here the netmask would be specified separately and would express the size of the subnet and it would not be restricted to the 8, 16 or 24 highest order bits, it could be any value from 5 to 32 bits in length. The written notation used for CIDR is 128.40.0.0/17, which specifies that the first 17 bits specified the network address, i.e. addresses 128.40.0.0 to 128.40.127.255 belong in this subnet. The use of CIDR also makes it easier to aggregate network addresses, e.g. 128.40.0.0/17 and 128.40.128.0/17 could be aggregated into 128.40.0.0/16. From Figure 19 we can see that Class C networks (the peak at 24 bit network size) are the predominant granularity of network address allocation.

IP addresses are assigned to networks by organisations such as RIPE, ARIN, APNIC and LACNIC and are globally unique. Additionally, a few network addresses in classes A, B and C have been allocated to be a private address space [RFC1918] which can be used internally in an organisation but are not globally unique and, therefore, not routable from the Internet.

3.3.3 IP Routing

Routing can be achieved in a number of ways, from sending an incoming packet out on a random interface to sending it out on all interfaces. In real networks, however, topology discovery and route management must be done in a de-centralised and scaleable way, to allow for growth and robustness. In IP networks the most common routing protocols fall into two main categories: Link-State and Distance Vector.

Distance Vector

In this case a node collects information about its distance (could be some Euclidean distance metric but is usually just the smallest number of logical hops) to all other nodes in the network and distributes this to all of its adjacent nodes at regular intervals. A neighbouring node then takes this information together with its current routing table and re-calculates the distances; these are then broadcast to all logically adjacent neighbours. With no changes in network topology the routing tables should eventually settle. An early implementation of this was used in the original ARPANET network where "distance" was a function of the outgoing buffer occupancy on an interface and, therefore, the routing algorithm could respond to network congestion. This was, however, found to be slow to converge, slow to respond to network failure and prone to

oscillation. As the amount of information distributed is dependant on network size this is not a very scalable solution either. Since the distances are dependant on previous distances received any changes are slow to propagate and, therefore, routing messages must be sent quite often (a few times a second).

Link State

In this case a node would, rather than send distances-to-all-nodes to neighbours, it would send distances-to-neighbours (where distances are link states) to all nodes in the network as part of a Link State Advertisement (LSA). These LSAs are flooded throughout the network, together with sequence numbers to prevent multiple copies being unnecessarily propagated. This may seem like a lot of advertising but since the information is not a function of other received information it can be sent much less frequently (e.g. every few minutes). The amount of information sent is also much less as it is only a function of the number of links from a node rather than the number of nodes in the network. From the LSAs each node builds a map of the network and can then make a next hop decision based on that.

3.3.3.1 Intra-Domain routing: Interior Gateway Protocol, IGP

Within the domains seen in Figure 10 a method for routing traffic between domains is required, and this is the Interior Gateway Protocol or IGP. The two most commonly used IGPs today are the Routing Information Protocol (RIP) and Open Shortest Path First (OSPF), but others such as Integrated IS-IS [RFC1195] and IGRP (a distance vector routing protocol which does not support subnet masks; it does, however, support multiple routing metrics) also exist.

3.3.3.1.1 Routing Information Protocol, RIP

This is a routing protocol based on the distance vector algorithm which uses the number of logical hops as a distance metric, ignoring interface speed or link utilization. Being a distance vector protocol it is usually slow to converge, slow to respond to link failure and due to delays in information propagation can cause routing loops. RIP version 2 is also limited to 15 hops maximum and, therefore, restricts the network's diameter. RIP is still in common use in small networks where simple route discovery is required. RIP does not support the division of the network into a hierarchy or any kind of areas; it only supports flat networks, and does not support netmasks, and, therefore, only classful addressing. The RIP metric is also simple in that link

metrics are configured as fixed values (usually 1, i.e. sum of metrics across route is number of hops) rather than an automatic function of link metrics such as delay or bandwidth.

3.3.3.1.2 Open Shortest Path First, OSPF

OSPF is a routing protocol based on the link state algorithm with a number of scalability enhancements. For example OSPF is capable of only sending its full Link State Advertisement (LSA) very rarely, and then only send changes in link state rather than the whole table again. OSPF also allows for the use of hierarchical routing and areas within the network. These areas are collections of routers and within these areas LSAs are propagated, but not outside. The area's edge routers only propagate LSAs outside of the area which are summaries of internal LSAs. The area labelled Area 0, which always exists, is the backbone area and all areas must be connected to it (see Figure 11; here the network is the interior of an AS, and Autonomous System (see section 3.3.3.2.1). The link metrics that are broadcast as part of an LSA can be explicitly set by the network administrator, however, the default behaviour in routers such as those manufactured by Cisco is to create a metric which is inversely proportional to the link's line speed: metrics are relative to a 100 $Mbits^{-1}$ line speed, therefore, a $10Mbits^{-1}$ link has a metric of $100Mbits^{-1}/10Mbits^{-1} = 10$. This has the advantage of favouring higher capacity links that often lead into the backbone but may not create routes that have the smallest number of hops.

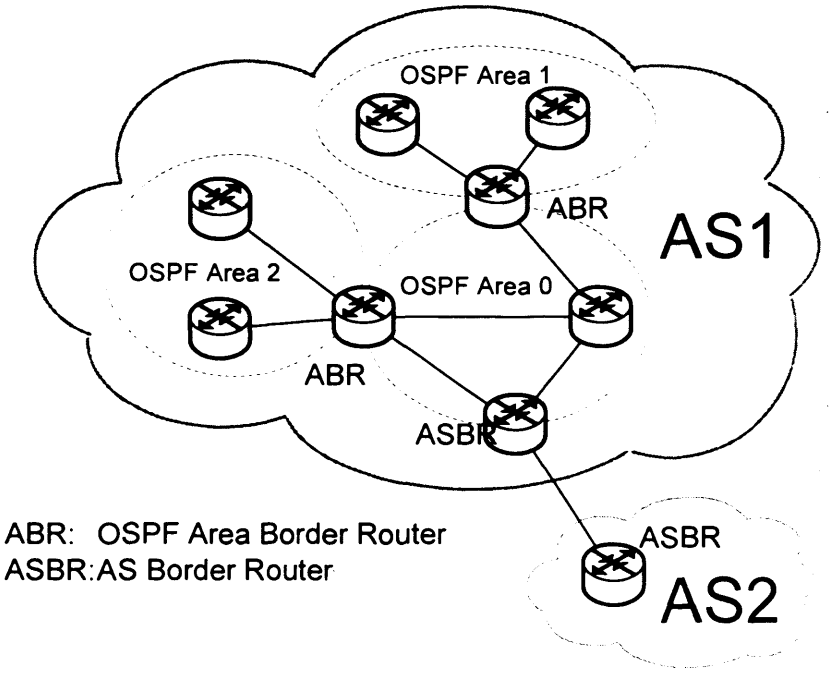


Figure 11 A typical multi-area OSPF configuration within an Autonomous System.

There are, however, special cases where it may not be possible for all OSPF areas to be physically connected to Area 0. In this case it is possible to introduce a virtual link to transit a common area. This is, however, rarely used and the majority of OSPF networks use the area topology shown above, if not a single area.

Since the link states of all routers must be stored on each router, and then a shortest path must be calculated OSPF is more demanding in terms of router memory and CPU requirements than other IGPs like RIP.

3.3.3.2 Inter-Domain routing: Exterior Gateway Protocol, EGP

EGPs are routing protocols used to distribute information about the connectivity of domains between the domains. The most widely used EGP is the Border Gateway Protocol. Other examples include EGP [RFC0827] (EGP, the Exterior Gateway Protocol, is also the name of a protocol, not just a class of protocols). EGPs are used because IGP metrics can differ between networks and also so as to prevent the need for full Internet routing tables in every router in every domain.

3.3.3.2.1 Border Gateway Protocol, BGP

BGP, the Border Gateway Protocol [RFC1771][RFC4271] is a routing protocol that is used to distribute routes between ASes (Autonomous Systems). BGP is based on the Distance Vector algorithm but is referred to as a Path Vector protocol [RFC1322] as it includes not only the destination prefix (a subnet address or aggregate of addresses) but also multiple attributes of the path and the path itself, rather than just the destination and a single metric. The path is specified by a list of ASN (AS Number) values. These ASN values are assigned to every transit network AS and have a globally unique 16-bit integer value. ASN values of 64512 to 65535 are assigned to a private range in a similar way to the private IPv4 address blocks [RFC1918]. The actual BGP communication is performed over TCP between BGP peers (routers running BGP), which are usually the gateway routers of ASes. Within these BGP sessions information about network prefixes (network addresses or aggregates of network addresses) is advertised together with a range of attributes. There are a number of attributes but the more significant ones are:

AS-PATH: Consists of a series of AS-SET and/or AS-SEQUENCE segments that contain the ASNs traversed by this advertisement. Which of these segments (AS-SET or AS-SEQUENCE) is used is dependant on whether this advertisement describes an aggregate prefix or not.

NEXT-HOP: The IP address of the router that is the next hop in the path back to the prefix. This is usually the IP address of the last BGP router the advertisement traversed.

LOCAL-PREF and **MULTI-EXIT-DESCRIMINATOR**: Used to express preference or weight to AS routes. These values are non-transitive and, therefore, only local to the AS (they are not propagated outside of the AS). Their actual value is set by a local policy, usually by some off-line traffic engineering algorithm driven by business based decisions (i.e. which networks to prefer based on peering agreements).

These attributes and others are then used by the local BGP route selection policies of the AS to decide on whether to use the route and whether to propagate it further.

To fully propagate routing information in an AS and to synchronise routing tables as quickly as possible BGP speaking routers within the AS are all programmed with the same policies and receive all the same advertisements (apart from those which are filtered out at the edge for being inappropriate) and together come to the same conclusions, therefore, having the same routing tables. To achieve this all BGP speakers must communicate with each other directly and this is in the form of long lasting BGP sessions running over TCP connections.

BGP actually consists of two sub-protocols: for communication between BGP speaking AS border routers in different ASes E-BGP (Exterior BGP) is used, but for communication within the AS between BGP border routers I-BGP (Interior BGP) is used. The two are similar but since E-BGP crosses an AS boundary and I-BGP does not their operation is slightly different. By default, on the inside of an AS a full mesh of I-BGP sessions is used to prevent routing loops (as no AS boundaries are crossed the AS-PATH attribute cannot be used to prevent loops) but this is an unscalable solution as it requires N(N-1)/2 TCP sessions to be configured (where N is number of BGP peers). With a small number of BGP peers this may be a feasible solution but in practise there are two solutions that are used to make I-BGP scale better:

Route Reflectors

These deal with scalability by creating a two level hierarchy. Certain I-BGP peers are configured as "route reflectors" which pass on I-BGP updates between ASBRs. The *route reflectors* are then configured as a full mesh (see Figure 12). To prevent route reflectors re-reflecting an already reflected advertisement, and thereby causing loops, the previously optional attributes such as CLUSTER-LIST and ORIGINATOR-ID are used: CLUSTER-LIST is a list of RR-ASBR pairs that the reflected advertisement has traversed.

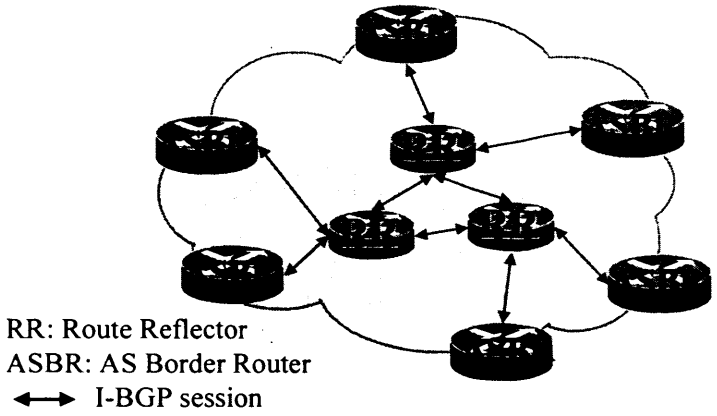


Figure 12 The use of route reflectors to improve I-BGP scalability

Confederations

Confederations attempt to solve scalability issues by dividing the AS into sub-ASes. These sub-ASes act similarly to ASes and are also labelled with an ASN value but from the private range mentioned earlier. To prevent loops across the external sub-AS topology the AS-PATH attribute of the advertisement is pre-pended with AS-CONFED-SET and AS-CONFED-SEGMENT segments similarly to AS-SET and AS-SEGMENT for outside the AS. These segments are removed when the advertisement leaves the AS. Attributes such as LOCAL-PREF, MULTI-EXIT-DESCRIMINATOR and NEXT-HOP are preserved across the sub-ASes. Internally, the I-BGP speakers of each sub-AS must still be fully meshed. Therefore, to the ASBRs of neighbouring ASes the sub-ASes look like a single AS: the internal configuration of the sub-ASes is transparent (see Figure 13). One notable requirement of confederations, however, is that all sub-ASes must all run a single IGP.

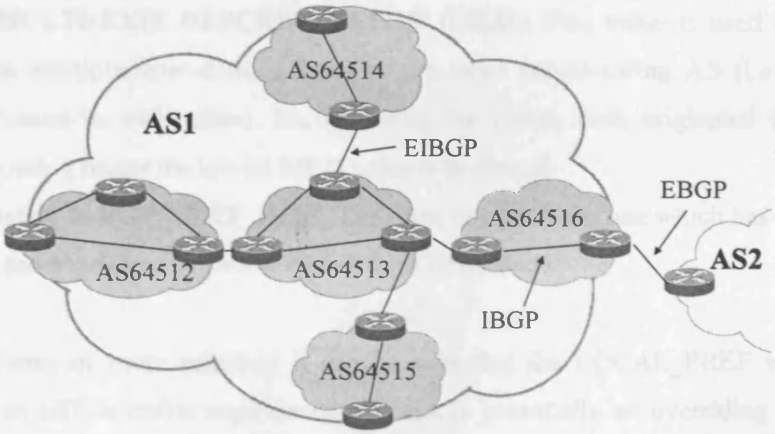


Figure 13 The use of confederations to improve I-BGP scalability

While all types of BGP sessions are simply TCP connections which can be carried over an arbitrary IP topology *core ISPs often implement their IP topology so as to follow the I-BGP topology* (see section 3.4.2). Therefore, depending on router configuration the routing protocols and their implementations can be implicit sources of topology, that is, they will not impose explicit structure, but the network designer may make design decisions with consideration for the BGP session topologies. Confederations introduce clusters and smaller cliques (sub-ASes), while route reflectors form a highly connected inner core with edge routers connecting to the route reflectors in a hub topology.

BGP Policies and routing decisions

BGP ASBRs will only propagate advertisements for routes which they themselves have decided to use and have already injected into their FIBs (Forwarding Information Bases, the list of forwarding rules). The standard route selection process for BGP routers is to compare various values and metrics in order of priority, and when found equal the next priority value is compared. The exact priority order is dependent on router manufacturer but a typical priority order for BGP routers is:

1. **Route source preference**: routes available from IGP are preferred to those from the EGP.
2. **BGP LOCAL_PREF**: The route with the highest local preference is chosen.
3. **BGP AS_PATH length**: The route with the smallest number of entries in the AS_PATH attribute is chosen.

4. **BGP MULTI-EXIT-DESCRIMINATOR (MED)**: This value is used to differentiate between multiple inter-domain links to the same neighbouring AS (i.e. the ASes are multi-homed to each other). So, assuming the routes both originated from the same neighbouring router the lowest MED value is preferred.
5. **IGP metric to BGP NEXT_HOP:** The route chosen is the one which has the lowest IGP cost to the IP address in the BGP NEXT_HOP attribute.

Therefore, in terms of route selection it can be seen that the LOCAL_PREF value, which is configured by an offline traffic engineering process, is potentially an overriding factor in route selection, and only once those are equal does the "shortest inter-domain path" decision come into play.

3.4 The Internet

We have now seen the logical topology implications of the various routing protocols, but there is also the macroscopic topology of the Internet that must be considered, which is the result of various independent decisions by network operators and planners.

The Internet was formed from the interconnection of a number of independent networks, some military or governmental, some research or academic and a few which were for public use. Originally these networks were interconnected as needed and connected directly: there was little concept of a transit network or hierarchy. To improve scalability dedicated transit networks were introduced and a hierarchy was created. The branching in this hierarchy was loosely defined by the allocation of Class A, B and C IP addresses (see section 3.3.2).

With the commercialisation and global acceptance of the Internet it has now tended towards the arrangement shown in Figure 14. The global Internet topology comprises of networks and ISPs (Internet Service Providers) covering different geographic areas and performing different functions:

- Tier 3 ISPs are smaller local and regional ISPs which connect often to the end user. These ISPs are usually considered the edge of the network and purchase their connectivity into the rest of the Internet from Tier 2 ISPs usually.

- Tier 2 ISPs are large regional ISPs or national ISPs that have a larger number of users, or large business users and can often provide transit services to Tier 3 ISPs. They do not have their own international connectivity and purchase international connectivity, or connectivity to other national ISPs, from Tier 1 providers.

- Tier 1 ISPs are the global network operators that cross the largest geographic span and are the operators of international links. Their main purpose is to provide transit services to Tier 2 ISPs, and they do not typically have edge users of their own.

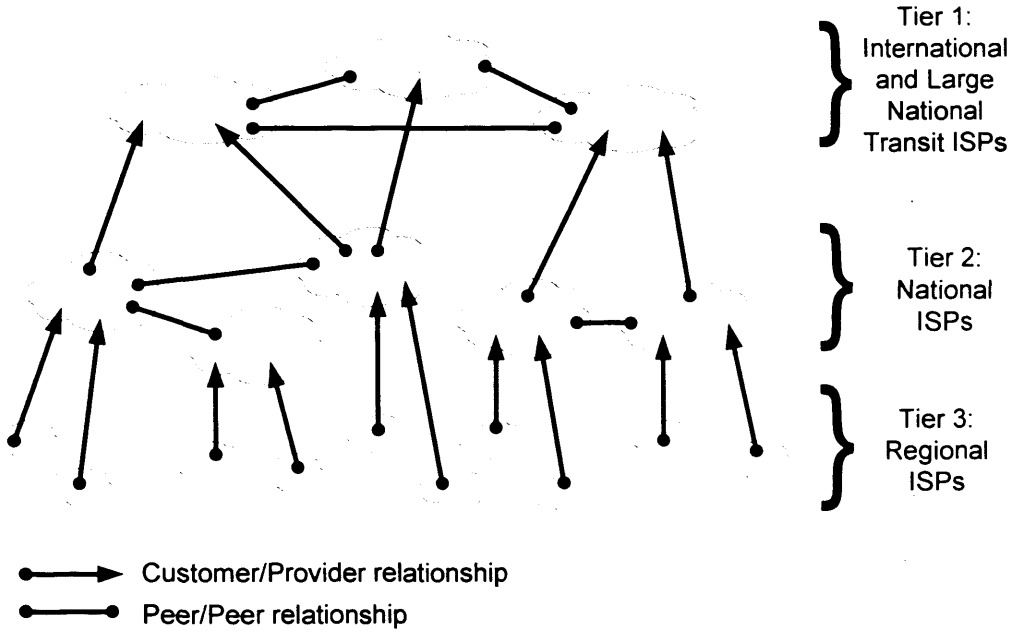


Figure 14 The Three Tier Internet Model with peering agreements and multi-homing

The relationship between ISPs generally falls into two categories, the customer-provider relationship and the peer-peer relationship, and this is described in the next section.

3.4.1 ISP Peering, agreements and BGP Policies

While BGP is a relatively simple and well defined protocol most of the topology and connectivity complexity comes from business relationships and the configuration of the forwarding policies.

3.4.1.1 Interconnection Business models

The decision to interconnect two ISPs, and the type of interconnection, is dependant on geography, the available physical layer, the ISP's demand matrix and the size of the two

interconnecting ISPs (and various other reasons, such as commercial issues and political influences). The next decision is the business relationship type:

The Customer-Provider relationship is when two ISPs, usually in different tiers interconnect and monies are paid either based on the amount of traffic sent, or sometimes on a flat rate per time interval (and dependent on the capacity of the link). The customer is in the tier below, while the provider, to whom monies are paid is in the tier above. The customer is typically then offered access to all network destinations offered by the provider, whether they are directly connected, or via customer-provider relationships with other ISPs, or the customers of ISPs with which the provider has a peer-peer relationship.

The Peer-Peer relationship is a relationship between two ISPs, usually in the same tier, which permits the transit of traffic to each other's customers. The business model used is usually the SKA (Sender Keeps All) (also known as the *settlement-free model*) model.
The Peer-Peer relationship does not typically allow one peer the use of the other peer's customer-provider interconnections where that peer is the customer, i.e. one peer in Tier 2 cannot gain access to a Tier 1 through the peer-peer relationship with another Tier 2 ISP.
This relationship is commonly used between Tier 2 ISPs so they do not need to send traffic to each other via a Tier 1 ISP, for which they would both have to pay. In this scenario this is only feasible, however, if direct peering is cheaper.
In SKA there is usually no monies exchanged between the peers, as long as the perceived value to each side is equivalent (which usually suggests a parity of traffic volume); the only costs are a sharing of the physical interconnection cost. Tier 1 ISPs typically always have SKA agreements with other Tier 1 ISPs.

The above is only a brief overview of the two main interconnection models as they strongly influence transport network traffic patterns; a more detailed analysis can be found in [MCK97].

3.4.1.2 Network Interconnection

Not only do networks interconnect directly but it is also common for large numbers of networks to terminate links at NAPs (Network Access Points), which are also known as Internet Exchanges or interconnect points (IXPs). These are premises, operated by a third party, that provide high physical security and protected electrical supplies to ISPs. At these NAPs ISPs situate border routers (or just transport network interconnection hardware) and link directly into the edge routers

of other ISPs with which they have an interconnection relationship. Logically these interconnection points can have various topologies, such as switched Ethernet or meshes of point-to-point connections. In the eventual logical router topology such connectivity may be seen as major hubs or highly clustered groups of routers. There was an example of an IXP in Figure 10.

3.4.2 Commercial IP Networks and Routers

Now with some knowledge of the logical configuration of networks we will briefly examine what capabilities and limitations current commercially available routers have, as well as the configuration of a real-world deployed national network to give a sense of scale to real IP networks.

The equipment comprising the router mesh of the Internet range in forwarding performance, interface count and routing functionality. One particular manufacturer, Cisco, offers routers for the small business, such as the 1700 series access router, as well as routers for Network Access Points, such as the 12000 series router. The 1700 series may only be capable of having three physical interfaces and forward approximately 12000 pps (packets per second). The physical interfaces include xDSL, Ethernet and asynchronous serial. A 12000 series router is capable of 375 million pps across a 320 Gbps switching fabric. The Cisco 12416 router can consist of up to 16 slots, each of which can be populated with interface cards of between one and sixteen ports. These ports are capable of a range of transmission technologies including FR, ATM, POS (up to OC-192c/STM-64c) and Ethernet (10Mbps, 100Mbps, 1000MBps, 10Gbps), transported over point-to-point links, MPLS, RPR or SDH rings. The router supports many protocols including IPv4, IPv6, IS-IS, OSPF, RIP and BGPv4. Such large routers would usually appear in Internet Exchange points.

It should be noted, however, there is not necessarily a link between the number of ports and the number of logical IP ports in the IP topology: ATM for example would require one physical port, however, IP links could be carried by a number of different PVCs; Ethernet could support multiple VLANs and so on.

To get an idea of what a real world deployed IP network looks like we consider a real national ISP. This ISP has one AS domain situated in Tier 2 and a second in Tier 2/3. The Tier 2/3 domain runs OSPF on the routers and is responsible for the source and sink of traffic from its end customers and privately owned networks, which do not run BGP. The Tier 2 domain is a transit network consisting of 10's of routers which all run BGP in a range of route reflector and confederation configurations. As all routers run BGP there is no need for an IGP and the physical

topology closely follows the BGP topology. The use of BGP with no IGP in the core is common practice as otherwise the IGP would need to contain routes for all prefixes in the Internet and this would be unscalable.

3.5 Internet Mapping

In previous sections we have seen the influences of technology and business on the router and AS topologies, however, there are still many unknown variables, from actual deployed network planning and bandwidth distribution, to inter-domain routing policies, to the interaction between IGP, EGP and the available transport networks. Since the Internet is distributed among many providers it is nearly impossible to find the actual deployed topology of each network and/or of the entire Internet. While some network operators do release general maps of their network, precise topology information is often considered confidential. To find the topology, or at least a good approximation to it, topology mapping techniques are required. This can be performed by a number of methods, some like SNMP (Simple Network Management Protocol) [RFC1157] interrogation are considered active probing and usually require the co-operation of network operators and, others like the use of traceroute style probes are more passive.

3.5.1 Internet Mapping Techniques

3.5.1.1 SNMP, the Simple Network Management Protocol

SNMP [RFC1157] is a UDP based service supported by most routers in the Internet and provides facilities to interrogate and configure the router state. The router provides a known hierarchical data structure through the MIB (Management Information Base). Provided in the MIB is everything from router name to interface types, link speeds, utilisation and routing tables. Information about the geographic location of the router is also often available. Through the interrogation of MIB routing table entries an accurate topology can be built, together with a map of capacity allocation, link utilisation and geographic distribution of routers. Unfortunately while some MIB variables are public most of the useful MIB entries listed here are not accessible without authorisation and, therefore, such active interrogation is not usually possible.

3.5.1.2 ICMP Ping and the IP Record Route option

It is possible to send ICMP ping requests with a "record route" option set in the header that causes routers supporting this function to append their incoming interface's IP address to an ordered list of IP addresses in the header of the packet, before it is forwarded. This is a very efficient method for discovering paths through a network since only a single packet is required to

find the route to a host. This method of discovering routes is, however, unreliable as many routers do not support this option for such packets coming from other domains and do not provide notification of the fact, leaving missing entries in the list, with no indication that a hop has been missed. The length of this list is also restricted to nine entries by the standard [RFC0791].

3.5.1.3 Traceroute probing

With active probing of routers not being possible a passive method is required, one which does not require the co-operation of network operators. Traceroute is a technique originally developed by Van Jacobsen, whereby a number of carefully crafted ICMP or UDP packets are sent towards a destination host. The packets start with a low TTL (Time-to-Live) field value and increment with every consecutive packet. The TTL field of an IP packet specifies the maximum number of hops this packet can traverse and is decremented at every hop (the field's purpose is to prevent packets traversing routing loops forever). When this value reaches zero without reaching the destination the router that decremented the packet's TTL to zero typically sends an ICMP Time Exceeded control packet to the original sender of the packet. The act of starting with a low TTL and increasing it in further packets should cause these ICMP packets to be returned from every router in the *forward path* towards the destination host. This technique can then be used repeatedly to trace the route to single hosts within stub networks to build a projection into the map of the core of the Internet. A single such projection was performed as part of this thesis and is described in section 3.6.1. There are, however, a number of problems associated with this method of topology measurement, described in the following sections.

3.5.1.3.1 ICMP throttling

An increasing number of routers do not respond with ICMP Time Exceeded packets, even though this is against the IP host behaviour standards. Some routers also limit the number of ICMP responses that can be sent within a certain time, thereby preventing some routers from appearing in the route. In contrast to the IP record route option described above it is obvious when a router ignores an ICMP Time Exceeded packet as a timeout expires.

3.5.1.3.2 Projections and incomplete maps

A single instance of tracerouting to a set of destination hosts only produces a single projection into the network from one viewpoint. This can result in many links not being discovered. For

this reason this experiment should be repeated a number of times from different networks and the multiple projections should be superimposed for a more complete map. The problem of undiscovered links from a single projection can be seen below in Figure 15.

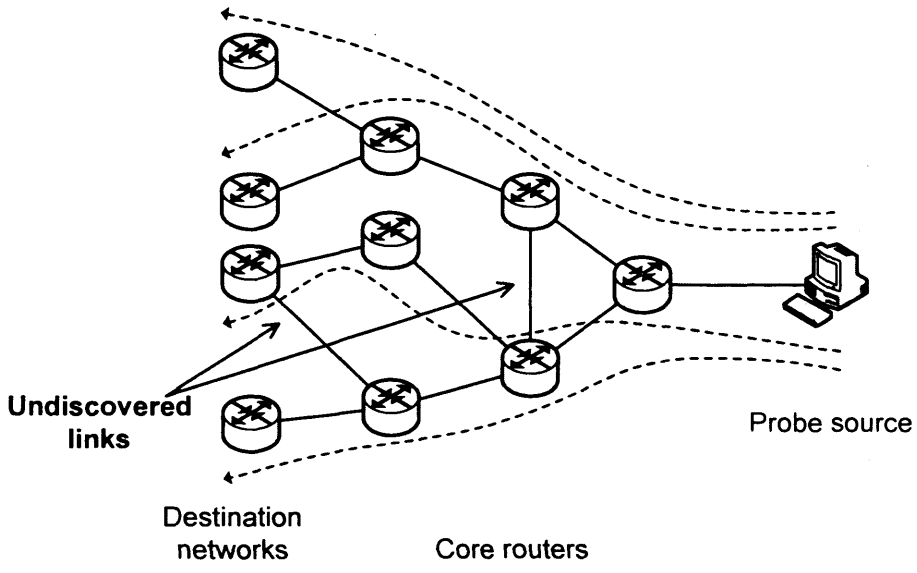


Figure 15 The problem of undiscovered links from single network projections

The problem of missing links can be partially alleviated by use loose source routed packets, however, many networks filter such packets.

For a further examination of the problems inherent with this method of topology measurement see section 3.6.1.2.

3.5.1.3.3 Router identification

It is not enough to use the IP address of an incoming interface on a router to identify the router in the measured topology as subsequent traceroutes may not follow the same path and arrive at a different interface on the same router, which will identify itself with a different IP address. There are two ways to discover which interfaces belong to which routers, the first is to interrogate the router with SNMP, the second is to ping the router: the resulting reply packet should be sent on the interface that is the best route back to the ping source, independent of which interface the ping arrived on at the router; this method is described in [PAN98].

3.5.1.3.4 Network dynamics

The mapping of routes to a large number of destination networks is not an instantaneous atomic operation and routing tables can change during the progress of the scan. This could cause more links to be discovered per projection, but it could also mean less are discovered. If the fluctuation occurs between the increments of TTL it could introduce non-existent links into the topology.

3.5.1.3.5 Router links with private or unnumbered addresses

With the increased scarcity of IP addresses network operators are more commonly using addresses from the private IP address space [RFC1918] within networks. For point-to-point links it is possible to have an interface unnumbered and just refer to it by the interface name rather than IP address. Routers with links in the private address space may respond with ICMP packets but since the source address is private these packets are often filtered at the network edge or by other networks along the route. Unnumbered interfaces on the other hand would send an ICMP response with the IP address of one of the numbered interfaces of the router.

3.5.2 Current Mapping Software and Methods

Since the above investigation into mapping techniques was conducted and the experiment in section 3.6.1 was performed the whole area of Internet mapping and tomography has seen a lot of research and a number of tools became available to perform various measurements, most notably:

- **Skitter** from CAIDA [SKITT] – a traceroute based topology measurement tool with sensors (projection sources) in about 20 locations thoughout the world, including some in backbone networks and near root DNS servers.
- The **DIMES** project [DIMES] - a traceroute and ping capable software application with a downloadable client for distributed topology and delay monitoring by anyone willing to participate.
- **Rocketfuel** [ROCKE] – a more holistic approach to topology measurement. Using traceroute and additional information like knowledge of BGP routing policies and other sources of information Rocketfuel attempts to minimise the required number of measurements to obtain accurate topological maps, and also measures of link bandwidth/latency and weights.
- **Route Views** [ROUTV] – this is a repository of global inter-domain routing tables, monitored at specific locations around the world, including a number of Internet

eXchange Points. This database is often used by router level mapping efforts to derive further information about their own measurements.

- **Mercator** [GOV00] – another traceroute based topology mapping effort, but which is also capable of sending source routed packets to alleviate problems with undiscovered links, and has a built in implementation of interface to router de-aliasing [PAN98].

For the purposes of this chapter we will concentrate on a tool developed as part of this thesis research and described in section 3.6.1, and the topology analysis performed by Faloutsos et al and described in section 3.6.2.

Murray and Klaffy [MUR01] have an overview of Internet measurement efforts including active and passive measurement, looking at topology, bandwidth and traffic. They state that as of April 2001 Skitter was performing scans similar to those described in section 3.5.1.3 with 22 probe sources; we will see later in section 3.6.1 the link discovery efficacy of more probe sources.

3.6 The Internet Topology

In this section we will first examine a brief experiment performed using the repetitive traceroute technique described in Section 3.5.1.3 and the resulting discovered topology. Our main source of Internet topology information will, however, be the work by Faloutsos et al. published in 1999 which will be reviewed later in this section.

Since the mapping work in Section 3.6.1 was performed in early 1999 great interest has been shown in Internet mapping, tomography, measurement and visualisation. Part of the tools available were described in section 3.5.2 and I direct the reader to also reference the sites of the CAIDA[1] and NLANR[2] projects.

3.6.1 Traceroute Experiment: IP router topology in Europe

To measure router and link topology at the IP packet level an experiment was created to repetitively perform the same function as the UNIX "traceroute" command. This utility takes advantage of the TTL (time to live) field in the IP header and finds the forward path towards a host by sending multiple packets. The operation of traceroute, originally written by Van Jacobsen, is described in Section 1.3.1.3.

Ideally we would trace the routes from all hosts to all hosts in the Internet, but since this is not realistic (in December 1998 there were approximately 60 million hosts in the Internet) a scan was performed from a single source from within the Department of Electronic and Electrical Engineering, UCL in the first quarter of 1999. UCL is provided with Internet connectivity by JANET, the Joint Academic NETwork (AS number 786) and, therefore, UCL would be seen as the customer of a Tier 2 provider (tier 2 is assumed since JANET has many customers and also peers directly with ISPs that provide international connectivity (Tier 1 ISPs)). The experiment performed traceroutes to a single host in every prefix listed by RIPE (http://www.ripe.net) (the host was chosen simply as the prefix with a 1 appended, such that the entry 83.16.128.8/28 (ASN 5617) would result in a scan to 83.16.128.9)(the prefix lists are available by anonymous FTP from RIPE). The destination host did not have to exist as we are not interested in the internal topology of the edge networks, just the topology of the core of the network. By taking the allocated network prefixes from RIPE the experiment avoided tracing to unallocated, reserved, or

[1] http://www.caida.org

[2] http://www.nlanr.net

private prefixes such as those described in [RFC1918]. Since subnet prefixes were being taken from RIPE, and RIPE's area of operations is Europe, only hosts of organisations registered in Europe were scanned to.

The experiment was performed from the single host in EE, UCL to 18311 subnets in Europe. The experiment ran over two days which meant that routes could have changed between the first subnet and last one i.e. it was not an instantaneous snapshot. This should not be a problem as we are not concerned about the relative path between routes, just the actual connectivity of routers. As this is traceroute and the IP addresses being captured are those of interfaces it should be noted that no attempt was made to map the interfaces into routers and, therefore, we are measuring a topology of interfaces, not routers.

3.6.1.1 Experiment Results

The result will inevitably be a tree like structure predominantly since it is a projection from a single source, but cycles are still possible. The total number of interfaces traversed was 17599 across 19416 links. The number of interfaces traversed may be questionable since it is a value close to (and less than!) the number of subnets (18311) that we are trying to trace to. There are a number of possible reasons for this. The main reason is probably that ICMP traffic is being filtered by certain ISPs and this is causing sections of the network not to appear; the severity of this problem is dependent on whether these ISPs are closer to our source network or further away as they would hide all the networks behind them too. Since there is no correlation between IP addresses on consecutive hops there is no way we can deduce how close we have come to our destination.

If an interface did not respond to the TTL expiration (see section 3.5.1.3) then it was removed from the data set; there were 4526 links to such interfaces (note this is not necessarily the same as saying 4526 such interfaces as here a link is an interface pair and one interface may respond and the other may not).

The results revealed interesting patterns in the topology of interfaces. In Figure 16 we can see the outdegree (outdegree is the number of egress links from and interface) distribution of interfaces.

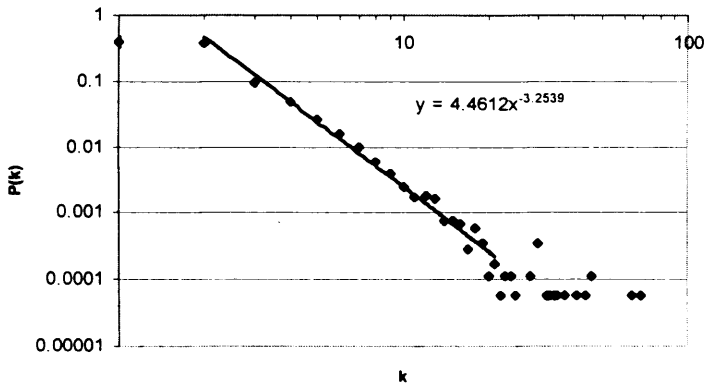


Figure 16 The probability distribution of outdegree, k, for an IP interface topology measured from a single source.

This shows a clear power distribution in the connectivity of the interfaces, save the extreme left where there is a point for one degree interfaces and the heavy tail on the right. The inability to find edge routers of the edge networks (the last hop) (this is because we are guessing a host at the target prefix to trace to, and, therefore, do not know how many hosts there are at the edge) could be to blame for the lower probability of one degree interfaces.

The next measurement of interest is the distribution of lengths of each route, below:

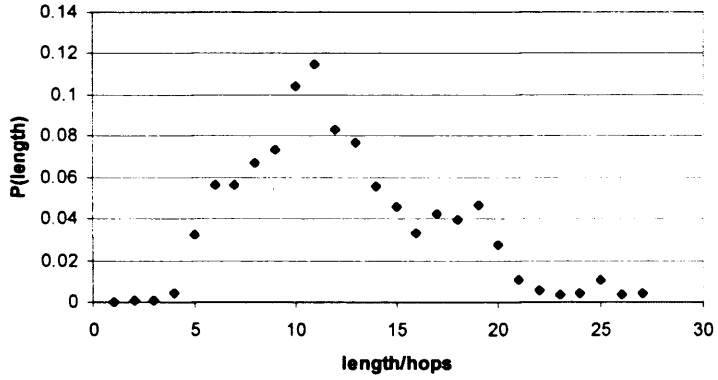


Figure 17 The distribution of route lengths to all RIPE prefixes from a single projection point

This shows a very small number of interfaces in the first few hops (as this is still within UCL) and a significant jump for interfaces within the UK (hops 5-10), the next jump and the peak are most probably hosts across Europe; they just seem to be close in terms of hops to UK hosts as the route passes through Tier 1 ISPs and this makes the route shorter in terms of router hops.

Since we are effectively looking at a distribution tree let us now consider how efficiently the load is spread across the outputs of the branching nodes. To do this we can find the distribution of the number of routes along each link.

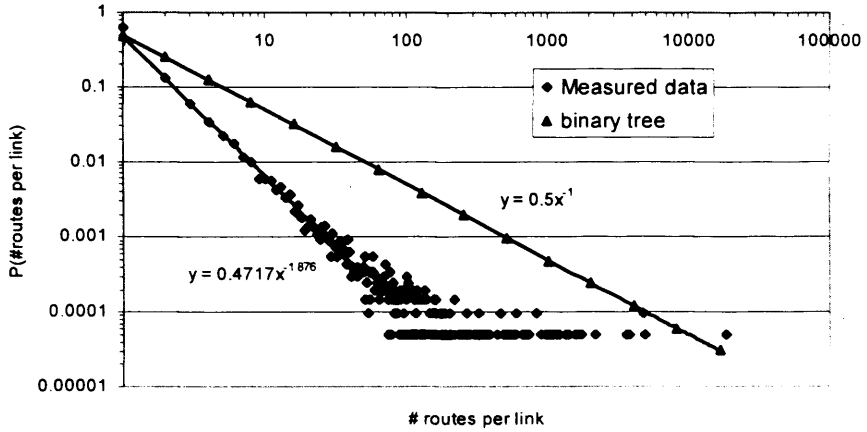


Figure 18 The distribution of the number of routes per link from a single source to all destinations

The distribution again follows a power law which is not surprising in this case because a normal binary tree would perform similarly. The binary tree case shown is where each incoming link is split into two outgoing links and the load distributed evenly. The difference between the binary tree and the measured data is that the route lengths to all destination nodes in a binary tree are the same, we have already seen this is not true of our real network. Real networks also very rarely split their load evenly between outgoing links; this is the effect of peering, routing policy and aggregation.

It should be noted that the above graph is not the same as examining load distribution as load is also a function of usage patterns and user distribution. To get an idea of how the users may be distributed throughout the network in Figure 19 we can see the distribution of prefix sizes.

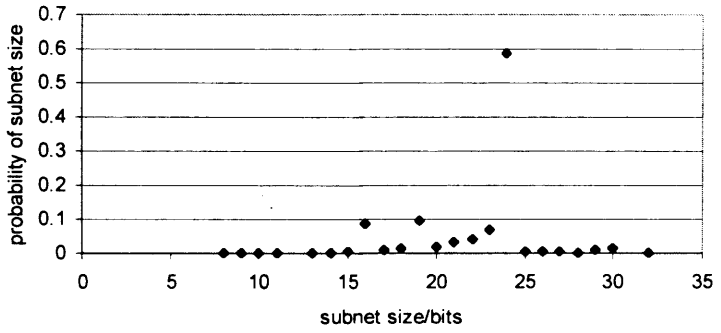


Figure 19 The distribution of CIDR subnet sizes in the list of subnets available from RIPE

The large peak at 24 is the equivalent of a Class C network address which is a subnet with 256 hosts. The peak at 16 is the old Class B address. So, although there has been a move to CIDR and there are no restrictions on subnet sizes it is still the case that an overwhelming number of legacy allocations mean that the most common sub-network size is /24, and, therefore, 256 hosts.

This experiment was performed before Faloutsos et al. presented their analysis and their work is a more thorough treatment of Internet topology, and is examined in section 3.6.2.

3.6.1.2 Experiment Accuracy

Given that we are only taking a sample of the entire European Internet graph from a single source it would be interesting to see how the measured topology would change with more projection sources. To examine this, six instances of two types of graph were created. One type of graph was an Erdős-Rényi (ER) graph where the probability of two nodes being connected is uniform, the degree distribution follows a binomial distribution, and the other was a Barabasi-Albert (BA) model where preferential attachment is used when growing the graph and the degree distribution tends to follow a power-law. See sections 2.3.2.1 and 2.3.2.2 for more details. For each graph type six instances were created with 1000 nodes and 2000 undirected links each, and six randomly ordered sets of 1000 projection sources were created. By making projections into the graphs from a varying number of sources to all destinations, and then superimposing the measured topologies, maps were constructed of each underlying network. It was assumed that routing through the underlying network follows the shortest path.

First, to see how much of the topology is actually being mapped we can see in Figure 20 the fraction of the total number of links being discovered as more and more sources are being used for projections and their maps are superimposed.

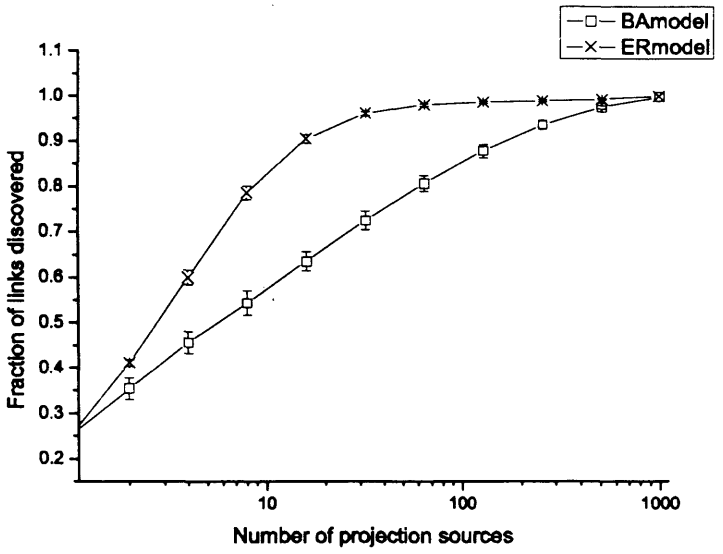


Figure 20 The fraction of links discovered with further projections into the network, whether the underlying topology is Erdos-Renyi like or Barabasi-Albert like

From the graph we can see that with a single source, whether the underlying topology is ER or BA only a small fraction of links is discovered. In an ER network this would improve faster than in a BA network, however, it looks like it would require nearly a third of the network to be projection sources before at least 70% of the links in a BA graph are discovered, but only 6% would need to be projection sources for an ER network to discover 70% of links.

Next, we consider whether the power-law in the measured topology is accurate or just an artefact of the measuring method. Since the measurement in Figure 16 exhibited a power law like distribution, in Figure 21 we can see the correlation coefficient to a power-law fit for both types of underlying networks, with an increasing number of projections.

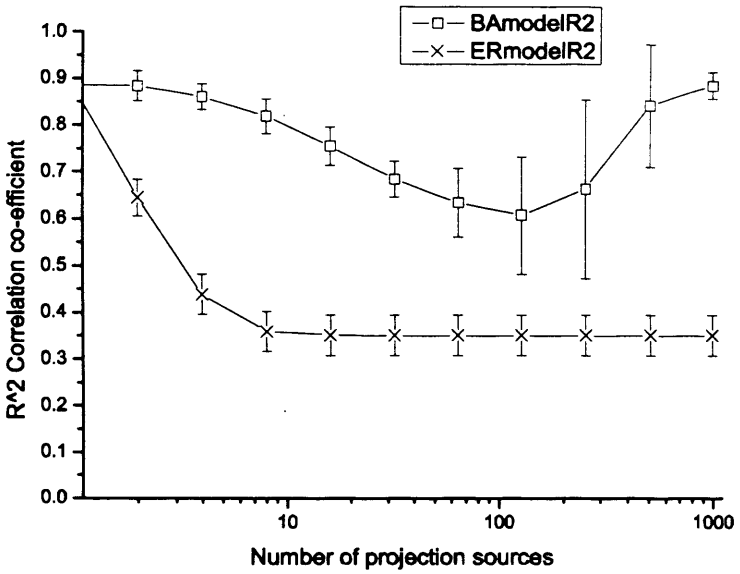


Figure 21 The R^2 correlation coefficient of a power-law fit to the discovered topology as the number of projection sources is increased

We can see from the plot that **irrespective** of underlying network type the measured topology from a single source has a high correlation (>0.85) to a power-law graph. Then, the ER graph moves away from that with only a few additional projections, while the BA graph slowly tends to shift away from being power-law compliant, until the end when all links are discovered (see Figure 20), where it then returns to a high correlation as the entire topology has been mapped.

Another observation that may be drawn from the graph is that it may not be possible to recognise the precise underlying topology until nearly half of the nodes are used as probe sources because it is not until ~500 projection sources does the mapped topology suddenly return to a power-law like network – the network could just as easily evolve into a different topology.

It should be noted that not only is the final superimposed map of use to finding the underlying topology but it may also be useful to examine how the map evolved as additional projections are added; the ER underlying model graph showed significant changes in the measured topology early on with only a few additional projections, while the BA model kept a high power-law correlation for at least the first 10 projections – by that time the ER model was very unlike a power-law topology.

The fact that the BA model in Figure 21 moves through a region of low correlation with a power-law as the number of projections increases may also raise the question of how accurate the actual power-law exponent of the measured topology is, if in fact the underlying topology we are trying to discover is a BA type topology. In Figure 22 we can see a plot of the power-law fit exponent and how that changes as the number of projection sources increases.

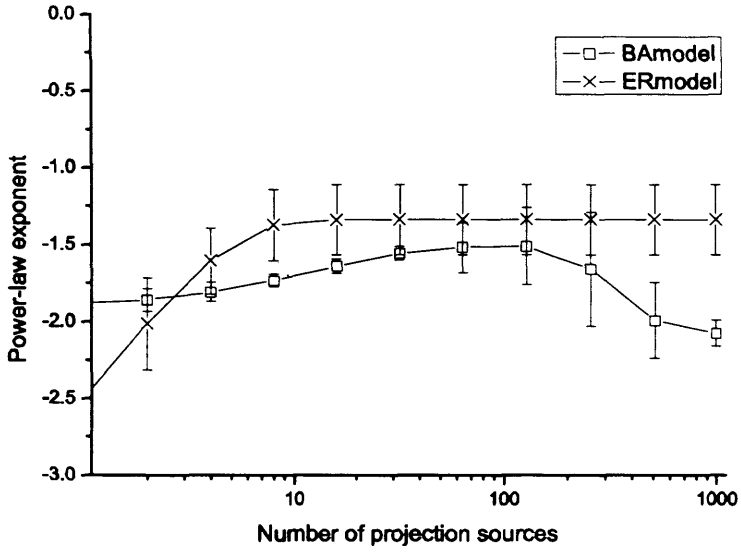


Figure 22 The power-law fit exponent of the measured topology as the number of projection sources increases

It can be seen from the above plot that the exponent for the measured topology of the BA model can vary significantly between -2.24 (at 512 projection sources) and -1.25 (at 128 projection sources).

Throughout the plots and conclusions above it should be noted that the graph size of 1000 is significantly smaller than the Internet and we may be seeing size-dependent properties and at other scales the above observations may not be accurate. However, observations like the fact that power-law degree distributions are inherent to the measuring method and independent of underlying topology cannot be ignored.

Recently it was shown [LAK03] by Lakhina et al. that the sampling bias is even more significant than shown here. Therefore, while there are very severe problems with the traceroute method of router topology mapping it is still better and more feasible than the other methods of mapping.

The experiment shown above was limited also in its reach, as it was only measuring the European part of the Internet, and it was only from a single source. In the next section we examine an empirical study by Faloutsos et al of the Internet at the AS level and router topology level, and while it will be said later in section 3.6.2.5 that their study too may suffer from these sampling biases it is still a very significant body of work and, as we shall see in Chapter 4 holds a remarkable similarity to an analysis of an SDH network which was performed as part of the research in this thesis; a network whose configuration was known exactly, and not subject to any sampling bias.

3.6.2 The Faloutsos' Topology Observations

In 1999 Faloutsos M., Faloutsos P. and Faloutsos C. published a paper [FAL99] which demonstrated for the first time a series of emergent properties of three instances of the Internet's AS inter-domain topology and one instance of router topology. They found four power-laws governing the topologies of their data sets. Treating the topologies as directed graphs they focussed on the outdegree of the nodes (where outdegree is the number of egress links from a node, where a node is either an AS in the case of the AS topologies, or a router in the router topology case), topological distances and eigenvalue spectrum.

The plots are reproduced in the next few sections, and the power-laws quoted verbatim as they are a very significant discovery in large-scale IP topology. Rather than just referencing them they are reproduced here for direct comparison to very similar laws discovered in SDH topologies, as part of this research, and described in Chapter 4. The power-laws also form the target of the modelling work in Chapter 5.

The data used in the study was three instances of the AS topology and one instance of the router topology:

Int-11-97, Int-04-98 and **Int-12-98** were three instances of the AS topology created by collecting data from geographically distributed BGP routers, although the paper does not specify where they were or how many there were. The captured topologies had 3015, 3530 and 4389 nodes respectively and 5156, 6432 and 8256 edges.

Rout-95 – this was a router level topology from 1995 which was originally collected by Pansiot and Grad [PAN98]. The topology data obtained by Pansiot and Grad was from a series of traceroutes, with source routing, from 12 sources to various known hosts in the Internet. They

used the "ping" technique described in Section 3.5.1.3.3 to map the interfaces into routers. They found 3888 nodes and 5012 edges in their router topology.

3.6.2.1 The rank exponent

Faloutsos et al. found that if the outdegree values are sorted in descending numerical order and plotted value against the ranking, then the distribution approximately follows a power law, such that (quoting from [FAL99]):

> **Power-Law 1 (rank exponent)**
>
> The outdegree, d_v, of a node v, is proportional to the rank of the node, r_v, to the power of a constant, r:
>
> $$d_v \propto r_v^R$$

In Figure 23 the actual plot for the AS topology from November 1997 can be seen, and in Figure 24 the plot for the router topology from 1995 can be seen. The paper has further plots for the other AS data sets.

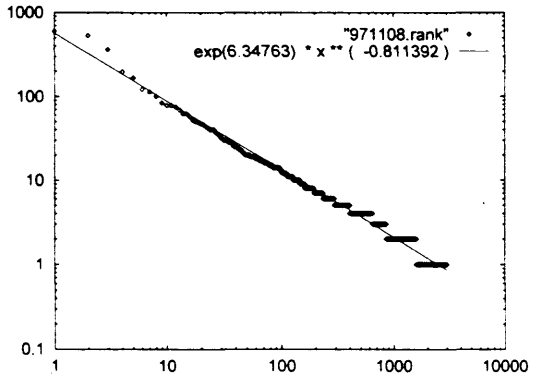


Figure 23 The rank plot for AS topology Int-11-97 from Faloutsos et al. [FAL99]. The rank appears on the abscissa, and the outdegree on the ordinate axis.

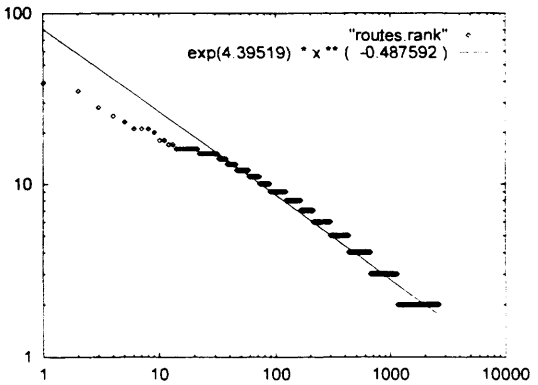


Figure 24 The rank plot for router topology Rout-95 from Faloutsos et al. [FAL99]. The rank appears on the abscissa, and the outdegree on the ordinate axis.

The rank exponents for the four data sets are shown below, together with approximate least-squares correlation coefficients to a power-law fit taken from the paper [FAL99] (correlation coefficient is a measure of conformance to a function, in this case the power-law fit, of a data set, where values closest to 1.0 signify the closest fit):

Data set	Power-law exponent	Correlation coefficient
Int-11-97	-0.81	0.981
Int-04-98	-0.82	0.979
Int-12-98	-0.74	0.974
Rout-95	-0.48	0.948

Table 3 The rank power-law exponents and absolute correlation coefficients for the various AS and router topology data sets as found by Faloutsos et al. [FAL99]

3.6.2.2 The outdegree exponent

The second power-law that is described in the paper [FAL99] is that formed when plotting the frequency distribution of outdegree (number of egress links from a node) values, such that (quoting from [FAL99]):

Power-Law 2 (outdegree exponent)
The frequency, f_d, of an outdegree, d, is proportional to the outdegree to the power of a constant, O:

$$f_d \propto d^O$$

In Figure 25 we can see the actual plot for the AS topology from November 1997, and in Figure 26 we can see the same plot for the 1995 router topology. The paper includes plots for the other AS data sets.

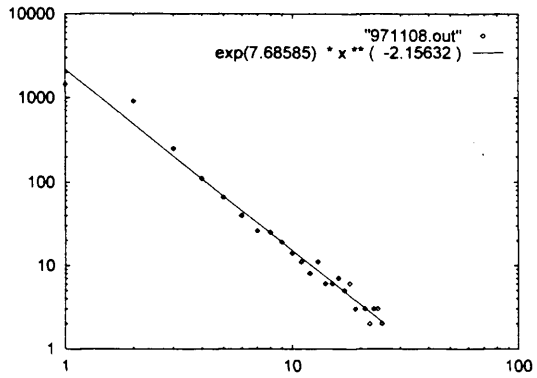


Figure 25 The Outdegree plot for AS topology Int-11-97 from Faloutsos et al. [FAL99]. The outdegree appears on the abscissa, and the frequency on the ordinate axis.

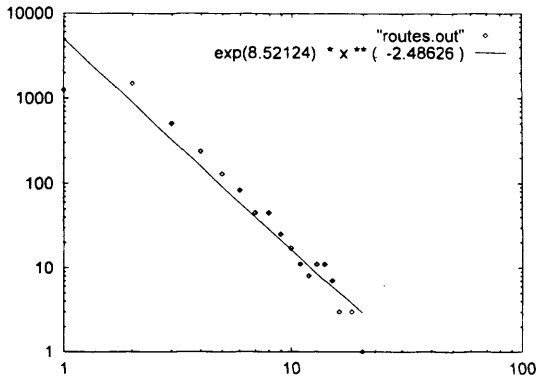


Figure 26 The Outdegree plot for router topology Rout-95 from Faloutsos et al. [FAL99]. The outdegree appears on the abscissa, and the frequency on the ordinate axis.

The outdegree exponents for the four data sets are shown below, together with approximate correlation coefficients to a power-law fit taken from the paper [FAL99]:

Data set	Power-law exponent	Correlation coefficient
Int-11-97	-2.15	0.991
Int-04-98	-2.16	0.979
Int-12-98	-2.20	0.968
Rout-95	-2.48	0.966

Table 4 The outdegree power-law exponents and absolute correlation coefficients for the various AS and router topology data sets as found by Faloutsos et al. [FAL99]

3.6.2.3 The hop-plot exponent

The third power law discovered [FAL99] is actually more of an approximation as it does not strictly follow a power law. The approximation centres around neighbourhood sizes and the total number of node pairs within a certain distance of a certain node, such that (quoting from [FAL99]):

> **Approximation 1 (hop-plot exponent)**
> The total number of pairs of nodes, P(h), within h hops, is proportional to the number of hops to the power of a constant, H (where h is smaller that the diameter, δ, of the graph):
>
> $$P(h) \propto h^H, h << \delta$$

In Figure 27 we can see the plot of the total number of node pairs which are at most h hops apart for the November 1997 AS topology, and in Figure 28 the same plot for the router topology from 1995. In both cases the plot points deviate from the power-law fit as *h* approaches the diameter of the graph and all the node pairs are counted in the total.

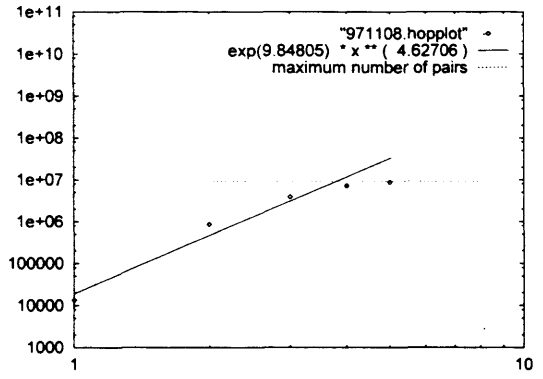


Figure 27 The hop plot for AS topology Int-11-97 from Faloutsos et al. [FAL99]. The hop distance appears on the abscissa, and the number of pairs on the ordinate axis.

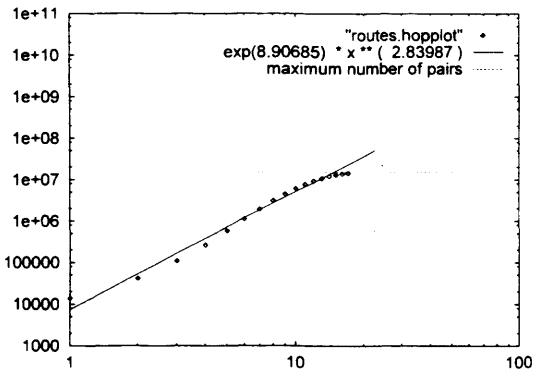


Figure 28 The hop plot for router topology Rout-95 from Faloutsos et al. [FAL99]. The hop distance appears on the abscissa, and the number of pairs on the ordinate axis.

The hop-plot exponents for the four data sets are shown below, together with approximate correlation coefficients to a power-law fit taken from the paper [FAL99]:

Data set	Power-law exponent	Correlation coefficient
Int-11-97	4.62	0.983
Int-04-98	4.71	0.981
Int-12-98	4.86	0.980
Rout-95	2.83	0.991

Table 5 The hop-plot exponents and absolute correlation coefficients for the various AS and router topology data sets as found by Faloutsos et al. [FAL99]

3.6.2.4 The eigenvalue exponent

The last power-law the authors discovered was that the first 20 entries in a sorted list of the eigenvalues of the adjacency matrix (this is a square matrix, a, where $a[i,j]$ is 1 if there is an edge from vertex i to j, and 0 otherwise) of the data sets follows a power law, such that (quoting from [FAL99]):

Power-Law 3 (eigen exponent)
The eigenvalues, λ_i, of a graph are proportional to the order, i, to the power of a constant ε:

$$\lambda_i \propto i^{\varepsilon}$$

In Figure 29 we can see a plot of the 20 highest eigenvalues of the adjacency matrix for the November 1997 AS topology, and in Figure 30 the same plot for router topology. The paper includes plots for the other two AS topologies.

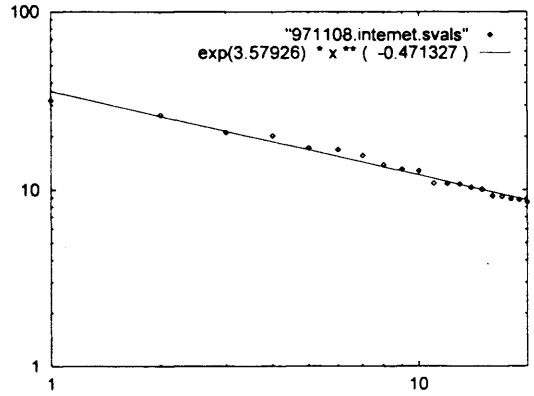


Figure 29 The sorted eigenvalue plot for AS topology Int-11-97 from Faloutsos et al. [FAL99]. The index into the sorted list of eigenvalues appears on the abscissa, and the eigenvalue on the ordinate axis.

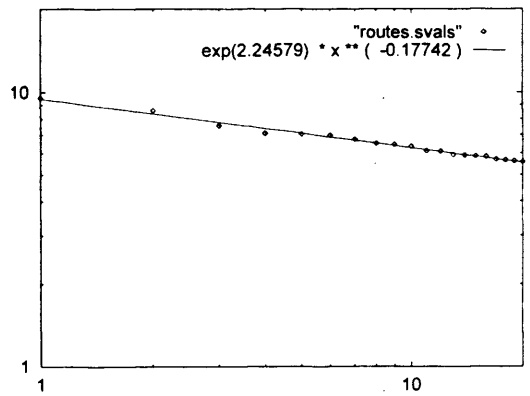


Figure 30 The sorted eigenvalue plot for router topology Rout-95 from Faloutsos et al. [FAL99]. The index into the sorted list of eigenvalues appears on the abscissa, and the eigenvalue on the ordinate axis.

The sorted eigenvalue exponents for the four data sets are shown below, together with approximate correlation coefficients to a power-law fit taken from the paper [FAL99]:

Data set	Power-law exponent	Correlation coefficient
Int-11-97	-0.471	0.990
Int-04-98	-0.502	0.989
Int-12-98	-0.487	0.991
Rout-95	-0.177	0.994

Table 6 The eigenvalue power-law exponents and absolute correlation coefficients for the various AS and router topology data sets as found by Faloutsos et al. [FAL99]

3.6.2.5 Comments on Faloutsos' Empirical Study and other studies

The results from Faloutsos' empirical study in 1999 have drawn a lot of scrutiny and have been criticised for their data sets; specifically the use of Pansiot and Grad's router data set and the fact that there is potential bias in the measurement method. We have already shown that this is possible in section 3.6.1.2. However, other more recent topology studies such as those by Mercator's authors have shown [GOV00] the underlying emergent property is still very similar to a power-law. So, while section 3.6.1.2 and [LAK03] have shown that there is inherent bias in the measurement method for router topologies Faloutsos' results are still significant for the AS topology analysis and that their route topology analysis is still one of the more complete examinations to date. [LAK03] also devised a method for detecting the sampling bias and found that data sets from Pansiot-Grad, Mercator and Skitter all suffer from this bias. The AS topology measurements used by Faloutsos et al. was also criticised by Chen et al. [CHE02] for possibly being incomplete, however, the authors of [FAL99] reinvestigated the AS topology in Siganos et al. [SIG03] which examined larger and more complete AS topology data sets and demonstrated that the power laws do exist over more than one data set captured at various dates.

It may be easy to dismiss all of Faloutsos' conclusions but as we shall see in the next chapter other network layers, such as a SDH network (transport layer), with a known complete map, rather than an inferred map, show the same power-law properties.

3.6.3 Observations on the geographic location of Internet resources

Given that in this thesis the link between various network layers is being examined, and potentially some layers are more significantly influenced by geography than others (especially the lower layers) it would be interesting to examine the geographic location of resources that comprise the Internet. This is unfortunately a rarely investigated topic, probably due to the

supposed agnostic nature of IP links and the fact that they are supposed to abstract away the features of the transport network. The two most significant contributions in this direction are by Yook et al [YOO02] and Lakhina et al. [LAK02].

Yook et al. showed [YOO02], by examining the geographic position of routers in the Internet showed that the distribution was not random as it was often modelled, but formed a fractal set with a well defined fractal dimension. Their main conclusions were:

1) **Self similar router distribution.** The physical distribution of routers in the Internet follows a self-similar pattern with a fractal dimension of about 1.5 (see section 4.3.3.3 for a comparison with SDH)
2) **Router distribution follows human population.** For all 1° by 1° boxes across North America (i.e. one degree of latitude by one degree of longitude box) there is a close correlation between the number of ASes and human population, and routers and human population. This strong correlation was also seen in Australia and Europe, but a much weaker correlation was seen in less developed areas like Africa of South America.
3) **Probability of two routers being linked is a function of distance.** The probability two routers are linked is approximately inversely proportional to the geographic distance between them.
4) **There is preferential attachment between ASes.** Yook et al also showed that over time the increase in connectivity of an AS increased as it's connectivity grew; therefore, for the first time this suggested quantitatively that new attachments were being formed between ASes with a preference to link to the more connected ASes. This is a significant result for Chapter 5 of this thesis where statistical modelling is examined.

In their 2002 paper Lakhina et al [LAK02] made a number of very interesting measurements and observations about the physical diversity of routers, the number of routers per region, and how this is impacted by the geographic regions economic status. Based on Skitter (their data was collected January 2001) and Mercator measurements (scan performed in August 1999) they concluded:

1) **Router distribution follows human population.** Similarly to [YOO02], Lakhina et al showed that there is a correlation between the number of interfaces or routers in a given area and the population size. They showed that there was a superlinear relationship in developed regions of the world.
2) **Probability of two routers being linked is a function of distance.** This was a more thorough analysis than Yook et al. and showed that the probability of two interfaces being linked decreased exponentially with distance until a certain cut-off point, after which the probability remained relatively independent of distance. They found that most links (from 75% to 95%) fell into this distance-dependent region.

3.7 Influences of IP network design on topology modelling

The examination of the protocols in this chapter have suggested a series of factors that impact on network topology, specifically:

- Clustering – in the cases of AS domains and OSPF areas there is a tendency to partition off parts of a topology and make it accessible to the outside only through border routers – this leads to the border routers being a hub-style transit point for all external traffic.
- Bandwidth dependence – The use of inverse capacity in the default configuration of OSPF on Cisco routers would suggest that routing is influenced by the available underlying capacity. This could potentially lead to a very good example of inter-layer feedback – traffic is drawn toward high capacity, increasing the utilisation of the link, either leading to Braess's paradox as described in Chapter 1, or lead to requiring a capacity upgrade.
- Heavy tailed node location distribution – as [YOO02] and [LAK02] showed, nodes are placed in a self-similar distribution on a Euclidean plane, and this provides a good model for the node placement in the experiments in Chapter 5.
- Heterogeneity and asymmetry – while the IGP protocols typically have a homogeneous view of the network and symmetrical routes (unless there are special cases like explicit non-symmetric weights or asymmetric bandwidth connected to Cisco OSPF routers), the EGP, especially the policies of BGP (with features such as the customer/provider and peer/peer relationships offering different prefix announcement policies) can be very complex and treat different sources and destination prefixes very differently, which often leads to non-shortest paths being used through the network.

- Connectivity affinity – Yook et al [YOO02]'s demonstration of the affinity to link to highly connected nodes lends credibility to the use of degree preference in models such as the BA model described in Chapter 1.
- Dependence on geography - Lakhina et al [LAK02]'s discovery that the probability of two nodes being connected is a function of the exponent of the distance between them is significant in that it lends credibility of the Waxman model described in Chapter 1.

There are also many other features that influence the topology, such as the use of non-optimal routes because of prefix aggregation policies, but these listed above are the features which have the most explicit and observed effect.

3.8 Chapter Overview and Conclusions

In this chapter we have seen the internal workings of routing protocols and the possible structure they may impose on IP networks, and we've seen the implications inter-domain policy and Internet structure may have on topology. Since it is not possible to instantaneously and accurately capture the Internet topology for study we have also examined the emergent topology through an original experiment as well as through the investigations of others like Faloutsos et al. [FAL99].

In the end the Internet has little global planning, only the planning within networks which is performed on a local basis with a thought for inter-domain traffic and peering decisions. Although given this apparent anarchy we have seen that there is possible emergent structure to the whole inter-network through the power-laws of Faloutsos'.

While not of the extent as Faloutsos' at al's analysis the work described in section 3.5.1 and section 3.6.1 was at the time novel work, albeit unpublished. The significance of Faloutsos' work is large and the power-laws will serve as a modelling target in Chapter 5.

Chapter 4
Transport Networks:
SDH – the Synchronous Digital Hierarchy

4.1 Chapter Introduction

This chapter serves to examine the SDH transport architecture and present an analysis of an actual deployed SDH network for the purposes of extracting traits and development processes that will aid the modelling of layers in later chapters.

The chapter begins by looking at the design of SDH and the SDH transport hierarchy, and the transport capabilities that it offers, as well as the requirements it places on the transmission layer. With an understanding of the design and inherent topology aspects of SDH the chapter continues with an in-depth analysis of an actual, deployed, country-wide SDH network. Topology, geography and capacity distribution of the SDH network are studied and correlations between metrics are sought. The chapter makes observations of the emergent properties of the network and attempts to build rules and policies which potentially exist in the explicit or implied planning of the networks, for example the correlation between geographic distance and bandwidth, or affinity of nodes to connect to similar nodes. These observations will be investigated further through the use of various models in following chapters of this thesis.

This chapter finds that even though the SDH network is very strictly planned and has a deliberately imposed architecture there are very clear power-law aspects to the topology, (which were not deliberately planned into the network) and these are not the result of the topology measurement method, as was potentially the case in Faloutsos' et al [FAL99] observations of the internet (Section 3.6.2) as this source data is a single and instantaneous snapshot extracted from the management database. It is also shown that some of the topological traits are maintained over a wide range of geographic scales i.e. the connectivity of towns, then cities, then entire regions.

The SDH results in this chapter are completely novel and currently the only known public study of such an SDH network.

4.2 The Synchronous Digital Hierarchy

4.2.1 SDH Overview

SDH is an ITU standard [G.707][G.708] that provides a generic transport service with circuit-based guaranteed delivery of synchronous data but also a specification for OA&M (Operations,

Administration and Management) functions, interface with the physical layer and the management and control communication between SDH elements.

SDH divides the functions of transport provisioning into a number of layers as can be seen in Figure 31 below.

At the top of the layer stack is the data to be transported, including IP, ATM or TDM based telephony. This is then packaged into either a Low-Order Path (LOP) with capacities up to 6.312 MBit/s (VC2) or a High-Order Path (HOP), starting at 49.536 Mbit/s and going up to many gigabit/s (highest standardised HOP is STM-256 at 38.338 Gbit/s), which is then packaged with the MSOH (Multiplex Section OverHead) into the Multiplex Section. The MSOH contains information about alarms, error correction, maintenance and signalling, including APS (Automatic Protection Switching) which is how circuit protection is implemented (described later). To this RSOH (Regenerator Section Overhead) signalling is added before it is passed to the physical layer. RSOH contains framing, channel identification and error correction.

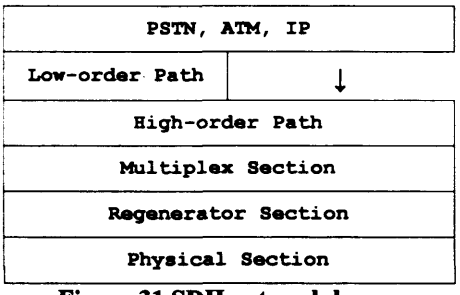


Figure 31 SDH network layers

4.2.1.1 SDH Network Elements

There exist in SDH networks four common types of network elements and their function is described below. The operation of each element covers one or more of the SDH network layers listed above:

4.2.1.1.1 SDH Multiplexer, MUX

Sometimes known as a Terminal Multiplexer this network element takes input signals of various types, including PDH, SDH, ATM and IP and packages them into higher bit rate STM-n signals.

4.2.1.1.2 Add/drop Multiplexer, ADM

An ADM performs the extraction and injection of low and high order path signals into the timeslots of high order path signals, that is, it can extract, say, a VC-12 tributary signal and replace it with another one while allowing all other traffic to pass through unchanged. The use of

ADMs can be seen in Figure 41. They are the basic element of SDH ring topologies as can be seen in Figure 37-Figure 40.

4.2.1.1.3 Digital Cross-Connect DXC

A DXC is similar in operation to the ADM in that it can extract and insert signals into timeslots but it is also capable of selectively moving signals between timeslots as the signal passes through the DXC and as such has switching capabilities. The use of DXCs can be seen in Figure 41.

4.2.1.1.4 Regenerator

Regenerators perform functions on the lowest layers of SDH including regenerating clock and amplitude of the incoming signal. Regenerators also supervise the physical layer channel properties and communicate with adjacent network elements through the regenerator section overhead (RSOH) of the SDH signal.

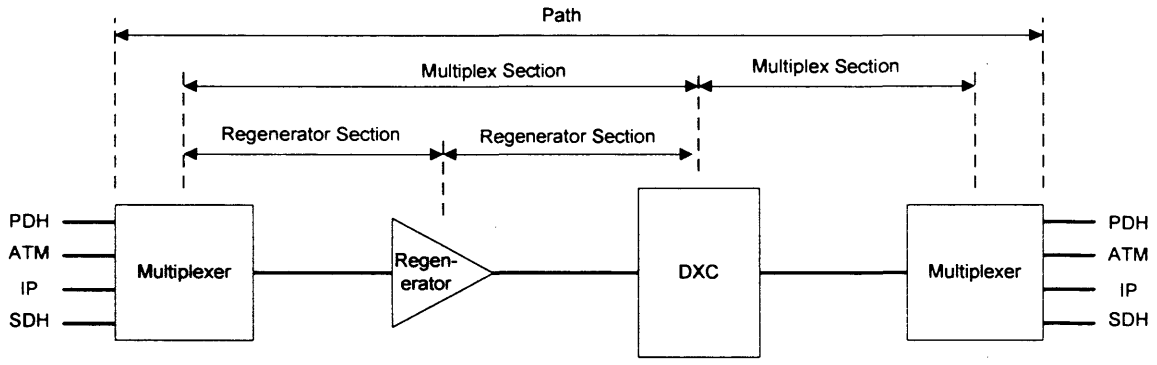


Figure 32 SDH network layers across various SDH network elements

4.2.1.2 SDH Transport Hierarchy

Due to the strict timing of SDH signals and the precise framing structure there is an exact hierarchy of payload types and of their multiplexing into higher-order signals, which can be seen in Figure 33. Each input stream is packaged into a container which is then mapped into a virtual container (VC). These are then processed into Administrative Units (AUs) or Tributary Units (TUs) and then either multiplexed into AUG or TUG (AU groups, TU groups). These groups (AUG,TUG) are then multiplexed into STM signals. The most significant point to note in this diagram is the granularity of input streams as well as multiplexed streams, and the hierarchy in which they are packaged.

The table below shows the most commonly used input streams, and the transport signals, together with their prospective speeds.

Circuit type	Data rate	Line rate
VC12	2.176 Mbit/s	
VC3	49.536 Mbit/s	
VC4	149.760 Mbit/s	
STM1		155.52 Mbit/s
STM4		622.08 Mbit/s
STM16		2488.32 Mbit/s

Table 7 the bit rates of various SDH transport levels

It should be noted that although this hierarchy exists there is no architectural requirement that the higher order streams be decomposed entirely into lower order streams before being operated on as was the case with earlier technologies such as PDH. The only requirement is that LOP tributaries must be multiplexed with other LOP tributaries before they can be packed into an STM-N stream.

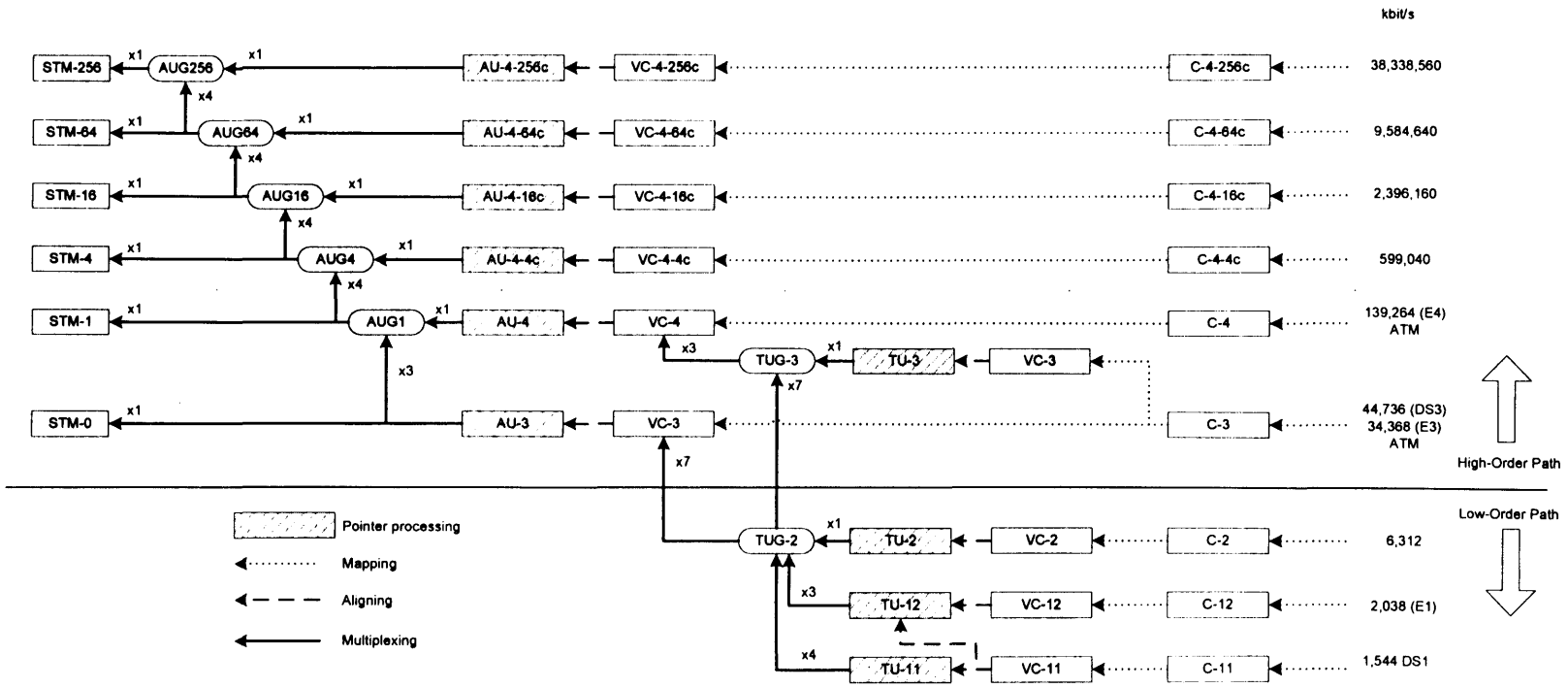


Figure 33 The SDH transport hierarchy

4.2.1.3 SDH Planning

SDH networks are very specifically planned and configured a priori with no dynamic reconfiguration through routing protocols as may be the case with network layer protocols such as IP. All end-to-end circuits are expressed as a set of traffic demands from which circuits are generated and then turned into an SDH network configuration through off-line planning algorithms. These algorithms attempt to minimise the cost of deployment of the network while providing support for the required demand pattern. Since SDH networks provide guaranteed levels of service and require the allocation of resources and pre-configuration of circuits the network configuration typically includes unallocated capacity to quickly and incrementally cater for new demands without interfering with existing circuits.

The configuration of the network can be any arbitrary topology with no explicit structures, but for ease of management and planning various structures are usually incorporated into the network design, such as rings and a hierarchy. To provide increased availability of circuits the circuits are often dimensioned as one or more alternative paths which can be switched between in case of a fault. These protection mechanisms are described in the next section and have an impact on the structures found in the SDH topology, such as the rings mentioned earlier.

The actual method of locating equipment and the topological structures within the network is an NP complete problem and usually solved by some kind of heuristic or global optimisation technique such as Simulated Annealing [WAS94][FOR03] [DAC00] [SOR99].

4.2.1.3.1 Circuit Protection

One common requirement of a transport network and one of SDH's strengths is its ability to quickly discover line or equipment faults and a loss of signal and to then perform either restoration (the post-hoc reconfiguration of spare capacity through a management system) or protection (the use of pre-configured alternative backup paths and fast circuit switching to minimise disruption to traffic). Restoration can typically take about two seconds to perform while protection is done in the order of 50 milliseconds. In SDH there are a number of ways that protection can be achieved and they all deliver different levels of protection and create different requirements of the physical layer. Protection can be for the entire path of the circuit, known as Path Protection, or protection of the Multiplex Section. The level of protection can also vary:

- **1+1 protection**: where a main path has a single backup path, but the traffic is duplicated and passed on both paths, and protection switching only occurs at the tail end.
- **1:n protection**: where a main path has *n* backup paths, each of which could be used when a failure in the main circuit occurs.
- **m:n protection**: where when any of *m* main paths could fail and be switched over to *n* backup paths.

The various protection schemes are described in the ITU-T G.841 standard [G.841] and briefly outlined below. A single scheme is not necessarily used in the entire network, and they can often be used individually, concatenated, or nested within each other [EDW98].

4.2.1.3.1.1 DPP (Dedicated Path Protection)

This is where there exists one main path between the SDH path termination nodes (typically the ingress and egress nodes) and one or more backup paths consisting of dedicated capacity. On the discovery of a fault the entire circuit is switched at the ingress from the main to the backup path (1+1 or 1:n where n=1 protection). More than one backup path can exist to provide increased protection (1:n or m:n protection).

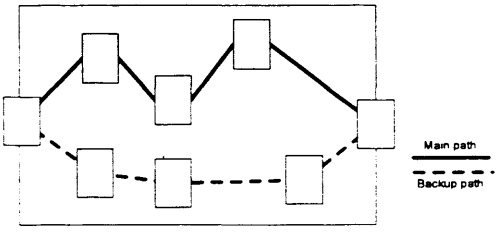


Figure 34 The Dedicated Path Protection scheme in operation

4.2.1.3.1.2 SNCP (Sub-Network Connection Protection)

This is where the whole network is configured into smaller networks (sub-networks), possibly different tiers in a hierarchy, or just an arbitrary division and there are main and standby paths for the circuit in each sub-network, with a common node in both paths as the interface between sub-networks. The protection switching in each sub-network can happen independently of the protection switching in the other sub-networks.

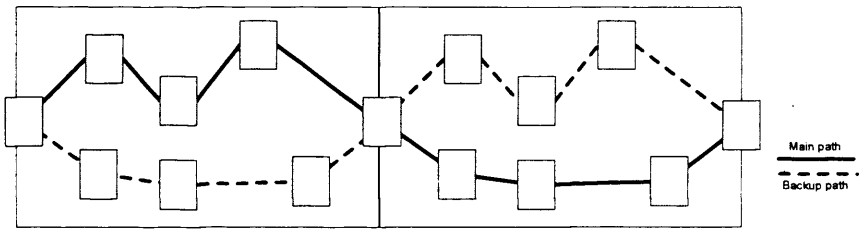


Figure 35 The Sub-Network Connection Protection (SNCP) scheme

4.2.1.3.1.3 MSLP (Multiple Section Linear Protection)

This is where a Multiplex Section (MS) (see Figure 32) is individually protected with one or more backup Multiplex Section spans between multiply connected ADMs/DXCs. On detecting failure of the main MS the ADMs or DXCs at the ends of the MS switch over to a backup MS.

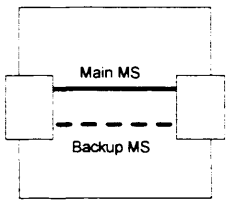


Figure 36The main and backup multiplex sections in MSLP

4.2.1.3.1.4 Multiplex Section - Dedicated Protection Ring: MS-DPRing

To simplify network design and protection management ring structures are commonly used in SDH networks and consist of a ring of ADMs which are directly connected. In the case of MS-DPRing each Multiplex Section has a main and a backup link and these form the clockwise and counter-clockwise ring; the main path always traverses the ring in the same direction and when there is a failure in one of the Multiplex Sections of the ring the adjacent ADMs then route the path in the reverse direction around the ring to the original path destination. This can be seen in Figure 37 and Figure 38.

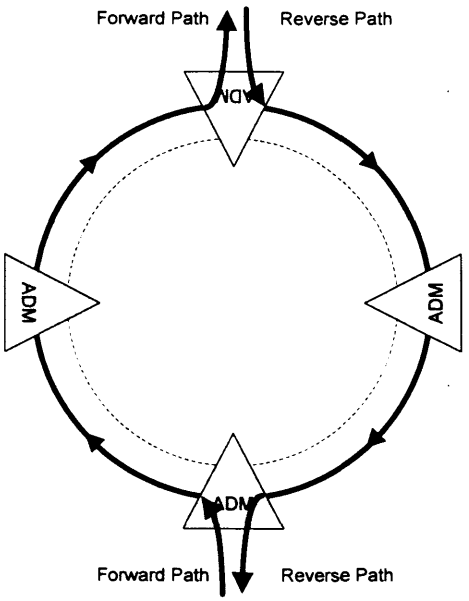


Figure 37 An SDH MS-DPRing dedicated protection ring where no fault has occured and the circuits pass in the same direction around the ring

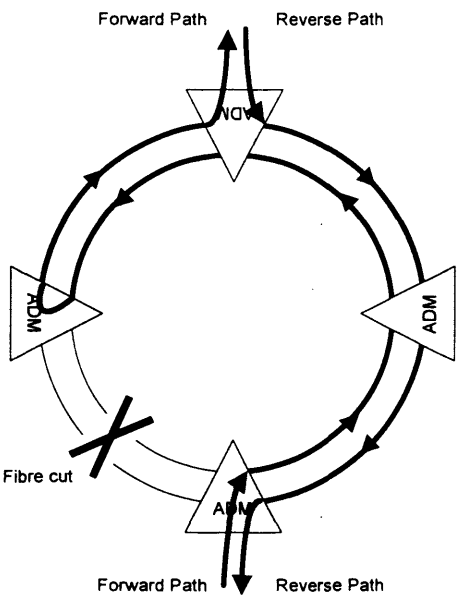


Figure 38 An SDH MS-DPRing dedicated protection ring where a fibre has been cut and the adjacent ADMs have switched the traffic into the backup fibre

4.2.1.3.1.5 Multiplex Section - Shared Protection Ring: MS-SPRing

An alternative ring protection is MS-SPRing (Multiplex Section – Shared Protection Ring) where the traffic is shared evenly between the clockwise and counter-clockwise rings up to half of their multiplex section capacity, and when a fault occurs the traffic arriving in the nodes adjacent to the fault is passed onto the other ring in the opposite direction in the empty timeslots. This can be seen in Figure 39 and Figure 40.

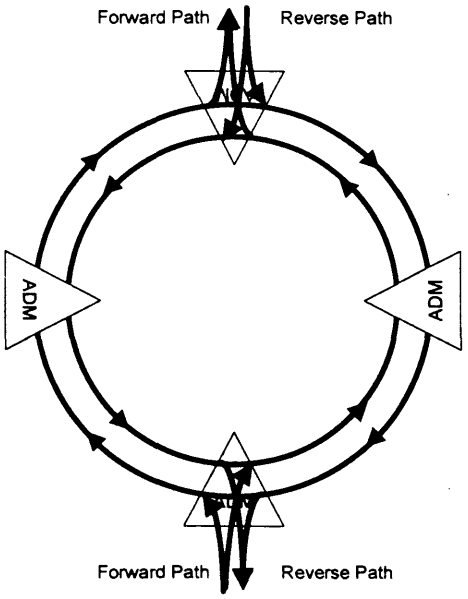


Figure 39 An SDH MS-SPRing shared protection ring where no fault has occurred and each fibre contains main and backup circuits and, therefore, each fibre is used to share main and backup traffic

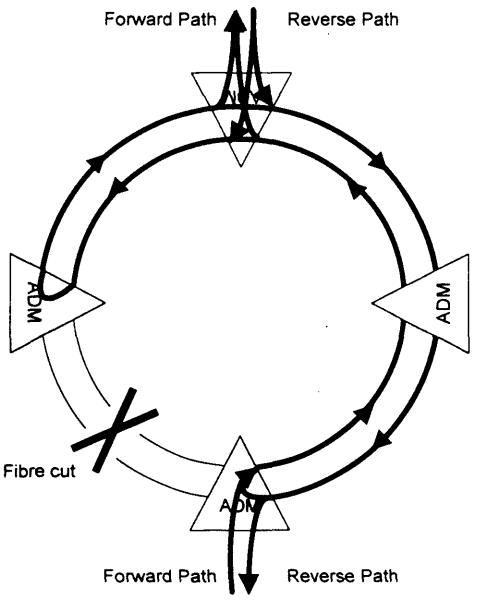


Figure 40 An SDH MS-SPRing shared protection ring where a fibre break has occured and the circuits have been switched into the protection circuits of the other fibres

4.2.1.3.2 Circuit Availability

An important specification of a circuit is its availability, which is the fraction of time that a circuit will be available for its purpose. The availability of a circuit is a function of the availability of each network element constituting that circuit and their configuration. Therefore, if a circuit consists of three elements, two end nodes and a single physical layer link, all connected in series, then all elements must be available for the circuit to be available. If there were two links between the end nodes (MSLP) then both end nodes must be available and at least one of the links. To find the availability of a circuit the configuration is decomposed into elements which are in series and those which are in parallel. The availability is a function of the MTTF (Mean Time To Failure) and MTTR (Mean Time To Repair). MTTF is sometimes expressed as a FIT rate (Failure In Time) such that MTTF is inversely proportional to FIT (Failure In Time, 1 FIT = 1 failure in 10^9 hours).

A system can now be decomposed into sub-systems of components in series and in parallel. It is possible to find the combined availability of these components to eventually obtain the availability of the system.

The availability of series elements can then be found with:

$$A_S = \prod_{i=1}^{n} A_i;$$

Equation 16 The availability of a system consisting of elements connected in series

That is, for *n* elements connected in series, all elements must be available for the system to be available, and, therefore, the overall system availability is the product of the individual elements' availabilities.

While the availability of parallel elements can be found with:

$$A_P = 1 - \prod_{i=1}^{n} (1 - A_i);$$

Equation 17 The availability of a system consisting of elements connected in parallel

That is, for *n* elements connected in parallel, all elements must be unavailable for the system to be unavailable, and, therefore, the overall system unavailability is the product of the individual elements' unavailabilities. To put this in terms of availability we simply subtract the unavailability probability from one, and vice versa, hence Equation 17.

The calculation of the entire system's availability is further complicated if the failure probability of an element is not entirely independent from the failure probability of other elements, for example two ADMs may be connected by what looks like two separate WDM links, configured, say, as an MSLP, and, therefore, in a parallel configuration, however, it may be the case that both wavelengths pass over the same physical fibre and, therefore, their failure of probability is actually a function of the availability of the line drivers and the fibre itself.

Further information about availability of network elements can be found here [RAD01] where the availabilities of ring structures are discussed.

SDH availability is, therefore, something that will influence WDM or physical layer planning and configuration (because of the requirement of physical diversity) and will be a potentially positive influence on the "value" of a circuit as it is provided to the layer above.

4.2.1.4 Related Tranport Technologies: PDH, GFP, VCat, LCAS.

SDH was created to deal with the limitations of the previously dominant transport technology called PDH (Plesio-synchronous Digital Hierarchy) which did not use such accurate timing as SDH. In PDH it was not possible to extract a single flow without de-multiplexing all the flows at each multiplex level, so to insert and extract a 2 Mbit/s tributary circuit into a 140 Mbit/s PDH circuit multiplex it was necessary to de-multiplex from 140 Mbit/s into multiple 34 Mbit/s multiplexes, then into multiple 8 Mbit/s multiplexes, then into multiple 2Mbit/s circuits, where the 2 Mbit/s circuit was removed/inserted and then re-multiplexed back up through the stages to 140 Mbit/s (this is basically the function of an SDH ADM although without the need to de-multiplex into stages – it is possible to extract a 2Mbit/s circuit directly from a 155 Mbit/s circuit). In PDH this required a large amount of equipment (so called the "multiplexer mountain") and made PDH much more difficult to manage and configure, as well as making it expensive and less reliable. While SDH does not suffer from the management and configuration problems that PDH did, SDH was introduced into networks such as British Telecom's as an upgrade, co-existing with PDH, with circuits being migrated in parts, and, therefore, there are many potential influences from the legacy PDH network topology.

While SDH was designed to support legacy telecommunication transport services it was not designed as a networking protocol for the asynchronous communication for packet-based systems, but as it is a transport protocol capable of traversing long distances it is often used to carry IP traffic directly in PoS (Packet over SDH/SONET) systems. This IP traffic is often delivered at the edge from Ethernet based networks which operate at 10Mbit/s [8023i], 100Mbit/s [8023u], 1Gbit/s [8023ab] and more recently 10Gbit/s [8023an] with standardisation efforts for 100Gbit/s [8023ba]. This leads to either a capacity bottleneck in the transport network or inefficiency since these capacities do not map directly to SDH Virtual Container capacities. The nearest VC capacity compared to Ethernet speeds can be seen in Table 8 below.

Ethernet Speed	Equivalent SDH level	Equivalent SDH VC speed	Transport efficiency
10 Mbit/s	VC-3	34 Mbit/s	29.4 %
100 Mbit/s	STM-1	155 Mbit/s	64.5 %
1 GBit/s	STM-16	2.48 Gbit/s	40.3 %

Table 8 SDH transport efficiency when carrying Ethernet traffic

ATM was used to carry IP over SDH networks, which allowed the capacity to be statistically multiplexed with other traffic in ATM virtual circuits, however, this required an additional network layer and increased deployment complexity. As a more lightweight approach GFP [G.7041] was introduced by the ITU to manage capacity and present it as either a constant bit-rate stream or variable-sized packet based traffic which could be multiplexed with other GFP traffic. Additionally VCat [G.707] was created to enable the bonding of multiple GFP circuits into any sized circuits. LCAS [G.7402] adds further flexibility by enabled the on-the-fly configuration and re-configuration of VCat GFP flows. The result is a dynamic, reconfigurable stream or packet based transport system which is still lightweight in its implementation (not requiring plenty of management overhead, additional equipment, topology changes and so on) and is capable is scaling to the current standardised SDH speeds.

4.3 *British Telecom's SDH Network*

As part of this research, British Telecommunications PLC (BT) supplied a map of their SDH network that they termed their "Narrowband SDH network", which carried their Frame Relay, PDH and other services, including some ATM and IP. The data provided included circuit routes, geographic SDH node location and details of the bandwidth assigned. The data was released under the condition that it was presented in a partially anonymised state and, therefore, this section may selectively normalise some values. In this section we will examine the data set and make observations on it and extrapolate the implicit processes that created it.

The network is used by both residential and business customers but is not the only transport network that BT operates. Their DSL services for example are provided over a different network – hence their naming of this network as the "Narrowband SDH network". The network configuration was captured in 2001.

4.3.1 *Planned Topology*

The BT SDH network is a nationwide transport network covering England, Wales, Scotland and Northern Ireland. It is a strictly planned topology which is expanded incrementally and undergoes an entire or partial reconfiguration of the SDH overlay sometimes as often as a couple of years.

The network is designed as a tiered hierarchy with four tiers as shown in Figure 41. Each tier contains rings of ADMs, where each ring covers a different regional expanse, depending on which tier it belongs to. The rings and tiers are inter-connected by ADMs and DXCs. The top tier is not a ring but a near full mesh of DXCs.

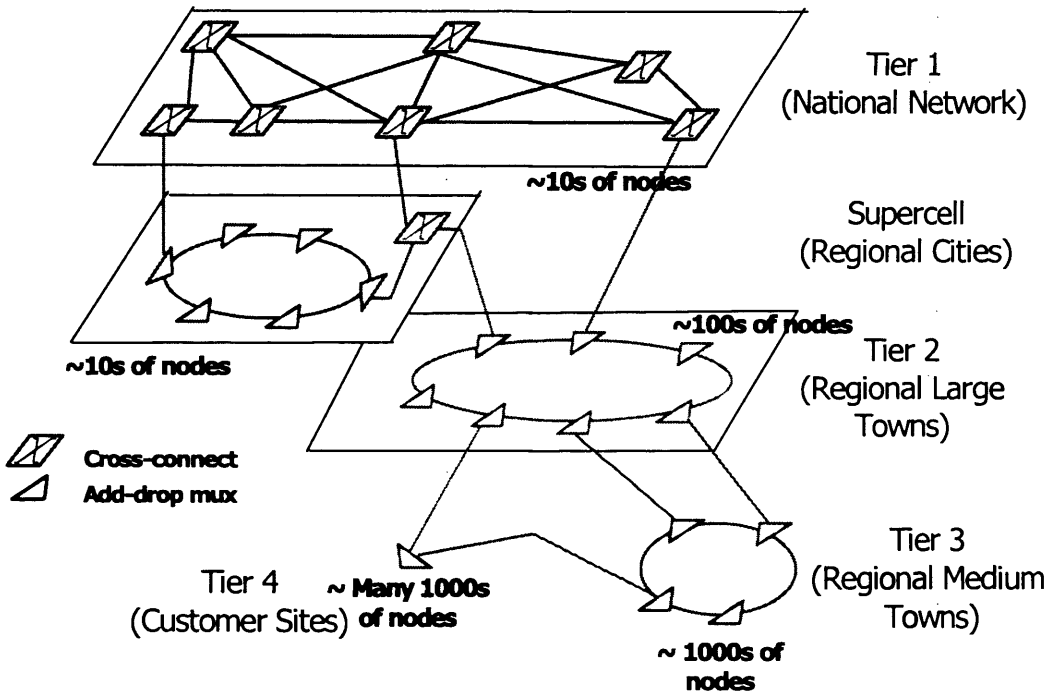


Figure 41 The architecture of the BT SDH network (diagram courtesy of BT)

The network is planned and dimensioned to support all current demands and with additional available capacity to be able to provision transport for new demands. The network provides circuits for point to point communication and provides various levels or capacity to demands and protection, as will be seen in the next section.

4.3.2 The Data Set

The data set consisted of a list of about 330000 circuits between at least 4500 customer sites. The data set was delivered in an anonymised state with one or both end points of each circuit removed from the data set. For this reason the following analysis is performed as if the circuits began and terminated at the edge node of the BT network: any customer sites that are listed in the data are not included in the analysis. In the circuit list there were separate entries for the main and backup circuits in the case where the circuits were protected. Since the data set is that of end-to-end circuits there was no indication of any kind of high-order paths which may have been used for aggregation or traffic marshalling.

The data set was presented as a list of circuits, each of which consisted of a list of SDH localities, an example of which can be seen in Table 9. The entries in each row were the circuit identification, whether the circuit listed was the main or standby circuit, the sequence or hop number that is being specified, and the locality were the hop reaches. Therefore, in the case of the

main circuit path of circuit CIRCUIT_10 there are 6 hops at the localities CUST_SITE_2820, CUST_SITE_2820, MI, MI, MI and DL, in that order.

```
"Circuit ID","MAIN/STANDBY","SEQ","LOCALITY"
"CIRCUIT_10","M",1,"CUST_SITE_2820"
"CIRCUIT_10","M",2,"CUST_SITE_2820"
"CIRCUIT_10","M",3,"MI"
"CIRCUIT_10","M",4,"MI"
"CIRCUIT_10","M",5,"MI"
"CIRCUIT_10","M",6,"DL"
"CIRCUIT_100","M",1,"AB"
"CIRCUIT_100","M",2,"AB"
"CIRCUIT_100","M",3,"AB"
"CIRCUIT_100","M",4,"EH/L"
"CIRCUIT_100","M",5,"EH/L"
"CIRCUIT_100","M",6,"EH/L"
"CIRCUIT_100","S",1,"AB"
"CIRCUIT_100","S",2,"AB"
"CIRCUIT_100","S",3,"AB"
"CIRCUIT_100","S",4,"AB"
"CIRCUIT_100","S",5,"AB/WEST"
"CIRCUIT_100","S",6,"FRM"
"CIRCUIT_100","S",7,"PH"
"CIRCUIT_100","S",8,"SXR/CA"
"CIRCUIT_100","S",9,"FK"
"CIRCUIT_100","S",10,"EH/L"
...
```

Table 9 An excerpt of the SDH circuits data set

Further details released included the geographic position of each "Locality" or site, as well as attributes of each circuit, which were only the bandwidth and protection type. The frequency of each property occurring can be seen in Table 10. The exact meaning of the circuit protection types was not known but those circuits which were not labelled as "Bronze" or "Bronze+" also had backup circuits associated with them in the circuit list.

Circuit protection type	Protection type frequency
Bronze	77.3%
Bronze+	0.04%
Steel	11.6%
Silver	3.3%
Tungsten	7.8%

Circuit bandwidth	Bandwidth frequency
DS2	95.8%
DS34	0.7%
DS45	0.1%
DS140	0.4%
STM-1	1.8%
VC-4	1.0%

Table 10 The bandwidth and protection types of SDH data set.

The list of circuit hops was then sanitised to remove any extraneous customer sites (since there was no visible heuristic as to when customer sites are included in the data set and when they were

not). There were instances when the customer site was in the middle of the circuit, in which case the circuit was spilt into two and treated as two separate circuits. Other strange anomalies include the circuit starting and ending at the same node, where the main circuit was this single hop, but the standby circuit was multiple hops, with the single hop of the main circuit in common, although such occurrences were very rare.

From this data it is, therefore, possible to extract four topologies:

1. **The SDH node adjacency topology**. This is where each node on the adjacency matrix is an SDH node and is formed by examining node connectivity as per the paths in the circuit data set.
2. **The SDH site adjacency topology**. This is like the node adjacency above but a node in this topology is one site, therefore, SDH nodes that are co-located according to the geographic positioning data set are considered as a single node.
3. **The demand node adjacency topology**. This is the topology formed when the intermediate hops in the circuit are ignored and the topology is simply formed by linking the ingress and egress edge nodes in the circuit. This is the topology that is seen to any layer above as the layer above does not see SDH circuit hops, only the endpoints.
4. **The demand site adjacency topology**. This is like the topology above, only nodes in the above topology that exist at the same physical site are collapsed into a single node in this topology, as with the SDH site topology above.

The geographic position of SDH nodes was based on a translated version of the Ordnance Survey Northings and Eastings format. The Ordnance Survey Northings and Eastings format is a grid encapsulating the United Kingdom within a 700km wide and 1300km long grid with a false origin off the coast of Cornwall, although in the BT data set case the grid also encapsulates Northern Ireland. The grid covers an area large enough that the size of each grid square is not uniform and tends to get smaller towards the north of the grid (due to the mercatorial projection used). For this reason this thesis maintains the use of the original units, however, to get a better feel for the numbers it was found that one distance unit was approximately 10 metres[3].

[3] To find the conversion factor two sites were taken (L/MUS and AB/NTH were chosen because the first is in London and the second is some distance away in Aberdeen) and matched to two in a public database (http://www.samknows.com/broadband/) to find their postcodes, and the distance between their postcodes in metres was found (http://www.dfes.gov.uk/cgi-bin/inyourarea/dist.pl). The conversion factor was 9.88.

In Table 11 we can see the details of the possible topologies mentioned above. Only 0.1% of the SDH nodes were co-located with other nodes and we will, therefore, not be considering the SDH sites topology further, unless implied, such as the geographic scaling analysis in Section 4.3.3.3. Similarly the demand node topology is considered rather than the demand site topology.

Topology	Size/nodes	Average degree	Maximum degree	Average diameter	Maximum diameter
SDH nodes	Approx 4000	1.840	About 85	5.88	16
SDH sites	Approx 4000	1.634	About 85	5.88	16
Demand nodes	Approx 2000	12.14	About 400	2.69	7
Demand sites	Approx 2000	11.78	About 400	2.69	7

Table 11 The parameters of the topologies available from the BT SDH network data

4.3.3 Topology Analysis

In this section we will consider the topology formed by the superposition of circuits between the SDH nodes, as well as the topology formed by the linking of the endpoints of the circuits, although without the customer sites as they were partially missing from the data set, as described in the previous section. The observations made here will be used to deduce and develop policies for the connection of random and structural models in the following two chapters. In this section the concluding observations are repeated in the page wide boxes for clarity.

4.3.3.1 Circuit Topology

Here we take the circuit list as can be partly seen in Table 9 and parse it to extract an adjacency matrix. For the purposes of this analysis each locality was considered to be an SDH node although it is possible that multiple equipments were present at each locality, however, the data set did not contain the necessary level of detail. Therefore, if there are two nodes appearing in sequence in the list of localities in the circuit then it is considered that these nodes are linked and that there is a vertex connecting the edges representing these nodes in the adjacency matrix. If the same two edges are linked through more than one circuit then there is only one vertex between the edges. The topology formed by the endpoints is examined separately in Section 4.3.3.7. The circuits were bidirectional and, therefore, the vertices were considered to be undirected and the adjacency matrix was symmetrical.

The first four plots considered below are those that were first discovered by Faloutsos et al in Internet router and AS topologies [FAL99] and were discussed earlier in Chapter 3 of this thesis.

In Figure 42 we can see a plot of the degree of each SDH node against the rank of the node when the list of nodes is sorted by order of decreasing degree. Nodes with the same degree are listed separately. Here degree is measured rather than in-degree (number of ingress) or out-degree (number of egress links) because the links are undirected.

The plot shows a very strong correlation to a power-law with an exponent of -0.759 and a correlation coefficient of 0.947. The degree values of the higher degree nodes are noticeably lower than the power-law fit. This could be caused by the removal of the customer nodes, assuming the edge nodes are the ones that have the highest degree (lowest rank).

The degree of a node is proportional to the rank in a list of degrees sorted in descending order, to the power of a constant.

It should be noted that making the assumption that circuits must end and terminate at a customer site and simply adding phantom customer sites to extrapolate the possible topology is not a good idea and could add possible skew since without a heuristic to tell which customer sites were removed it would not be possible to differentiate between multiple circuits terminating at the same customer, or multiple circuits terminating at multiple customers. It is also not possible to distinguish between circuits that terminate at a customer site which has been truncated and at a BT site, which is the case for many of the circuits.

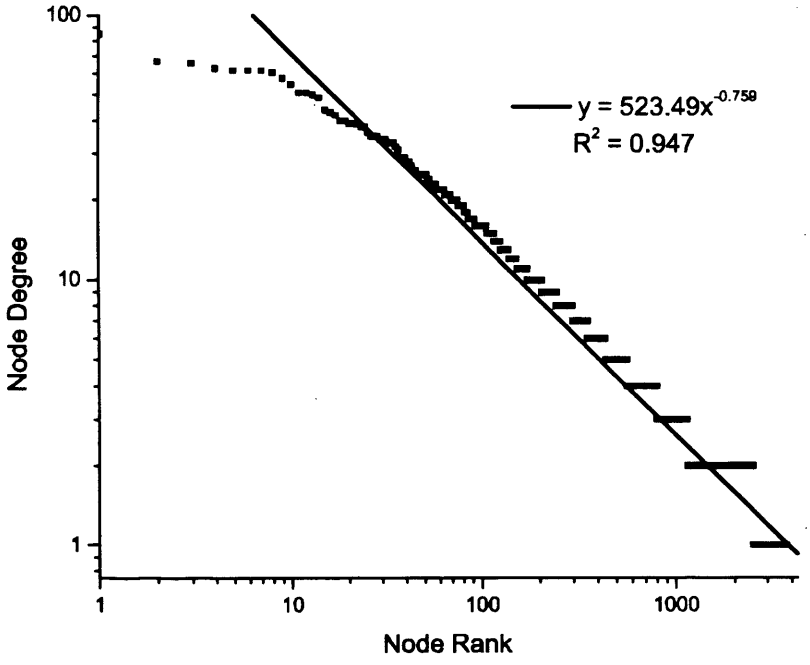


Figure 42 The rank of the SDH node degrees sorted in decreasing order.

Another way to examine the degrees is by looking at the frequency distribution of degrees. In Figure 43 we can see the probability that a node has a given degree. Again, there is very good correlation to a power law fit with a correlation coefficient of 0.9315. The exponent this time is -1.957. The power-law fit, therefore, suggests that there are a very large number of low degree nodes and a relatively small number of high degree nodes.

The frequency of a node degree is proportional to the node degree, to the power of a constant.

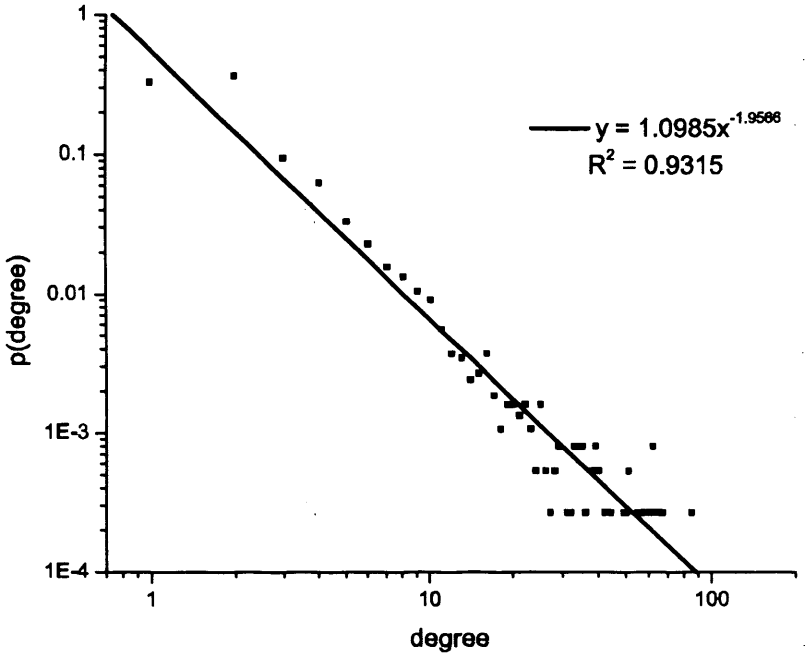


Figure 43 The degree frequency distribution of the SDH circuit hop topology

The next plot that Faloutsos et al. showed was that of the 30 largest eigenvalues of the adjacency matrix. This is known as the eigenvalue spectrum and is an illustration of the general connectivity and largest cliques in the topology. In Figure 44 we can see the first thirty eigenvalues of the adjacency matrix of the SDH node topology ranked in descending order. The adjacency matrix here is an N by N matrix, where N is the size of the network in nodes and the element at `(u, v)` is a `1` if nodes `u` and `v` and directly connected, and `0` otherwise. The matrix is symmetrical because the links are undirected. As before, the conformance to a power-law fit is very good with a correlation coefficient of 0.995. The exponent of the power-law fit is -0.299.

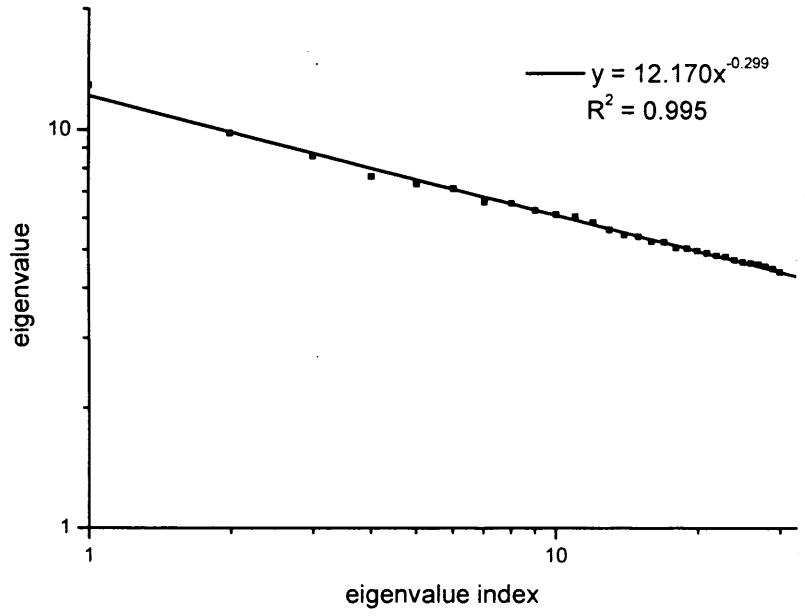


Figure 44 The first thirty sorted eigenvalues of the adjacency matrix of the SDH node topology

The first 30 largest eigenvalues of the adjacency matrix of the SDH topology are proportional to their rank (in decreasing order) to the power of a constant.

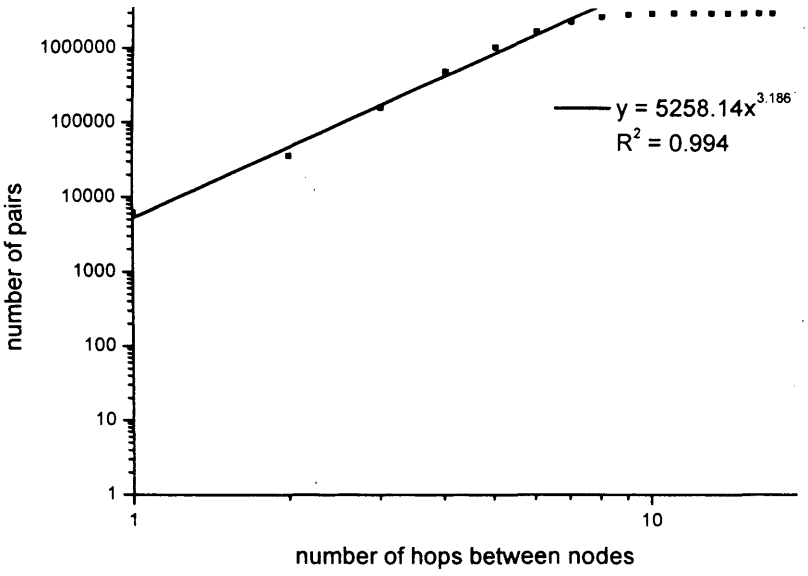


Figure 45 A plot of the number of pairs of nodes which are within a given number of hops in the SDH node topology

The last plot in common with Faloutsos et al. and the Internet is more of an approximation rather than a strict power-law. We can see in the Figure 45 that for distances much less than the network diameter the number of node pairs which are the given distance apart or less follow a power-law with the distance. In the case of the SDH node topology the cumulative number of nodes pairs up to a distance of six hops conforms to a power-law with a correlation of 0.994. The exponent of the power-law fit is 3.186.

The number of pairs within a given number of hops is proportional to the number of hops, to the power of a constant, assuming the number of hops is much less than the network diameter.

There are further power-laws existing in the SDH topology, however.

To get a better idea of the general connectivity around a node and the possible alternative routes in the immediate vicinity a metric is required to find the extent of clustering. Such a metric was proposed by Watts and Strogatz [WAT98] and they specified that for a graph of undirected links as in Figure 46, the clustering coefficient of a given node is the fraction of links that exist in the

immediate neighbourhood to the total number of possible links (where neighbourhood is the node and its immediate adjacent nodes).

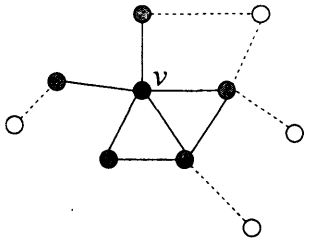


Figure 46 The neighbourhood around vertex v, with vertex v in red and the neighbouhood nodes in green.

Therefore, for the vertex v the neighbourhood consists of the directly connected nodes, coloured in green. The clustering coefficient can then be found as:

$$c_v = \frac{l_{neigh}}{0.5d_v(d_v + 1)};$$

Equation 18The clustering coefficient

Where l_{neigh} is the number of links in the neighbourhood and d_v is the degree of vertex v, so for this example l_{neigh} *is* 7, d_v is 5 and, therefore, the clustering coefficient, c_v is 0.467. Therefore, a value approaching 0 is signifies a relatively poorly connected neighbourhood, while a clustering coefficient of 1.0 signifies a full mesh of links in the neighbourhood. The clustering co-efficient has valid values in the boundary (0,1].

In Figure 47 we can see a plot of the rank of clustering coefficients when sorted in ascending order.

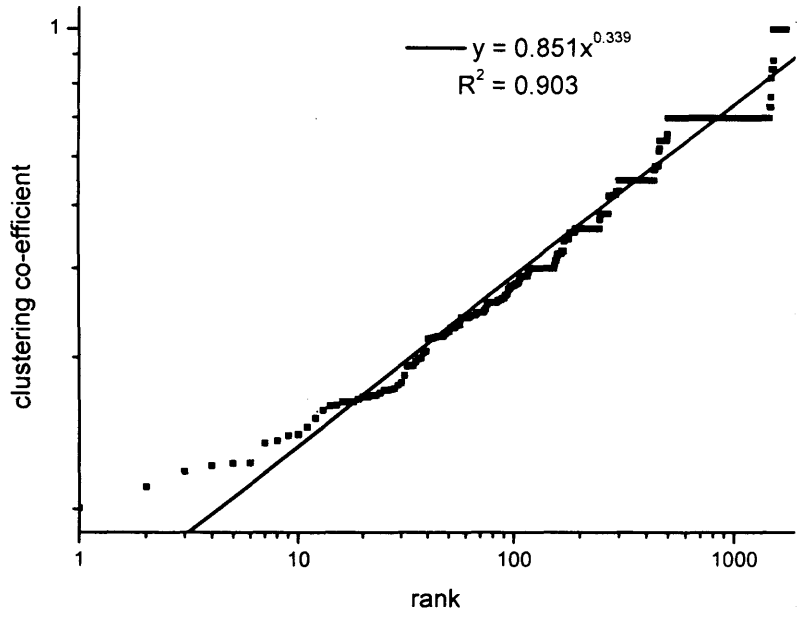


Figure 47 The clustering coefficient of SDH nodes against their rank when sorted in ascending order.

The plot is interesting due to the very large number of nodes with a clustering coefficient of 1.0 (14.0% of all nodes) and 0.66 (52.7% of all nodes), and other smaller plateaus at lower values. This would suggest that there may be some kind of ordered repeating structure in the topology. The 1.0 may be the very highly connected Tier 1 in Figure 41, while the 0.66 could be Tier 2 and Tier 3. The plot conforms relatively well to a power-law fit with a correlation coefficient of 0.903. The power-law exponent is 0.339.

> The correlation coefficient is proportional to the rank of the coefficient in a list of coefficients sorted in ascending order, to the power of a constant.

These five power-law observations are surprising as they are an unintended (or at least not explicitly enforced) emergent trait of the SDH network and it cannot be said that they are the result of the data collection method as may be the case for the Internet (see section 3.6.1.2). In the rest of this chapter we examine further traits and attempt to make observations that will help us understand better why these traits appear.

Now that we have seen metrics of individual nodes we can examine the connectivity between those metrics, specifically whether nodes have a tendency to connect to nodes of a similar degree

or not. This is known as assortativity [NEW03]. Here the connection of nodes with a similar degree is known as assortative mixing and connection of nodes with degrees that are significantly different is known as dissasortative mixing.

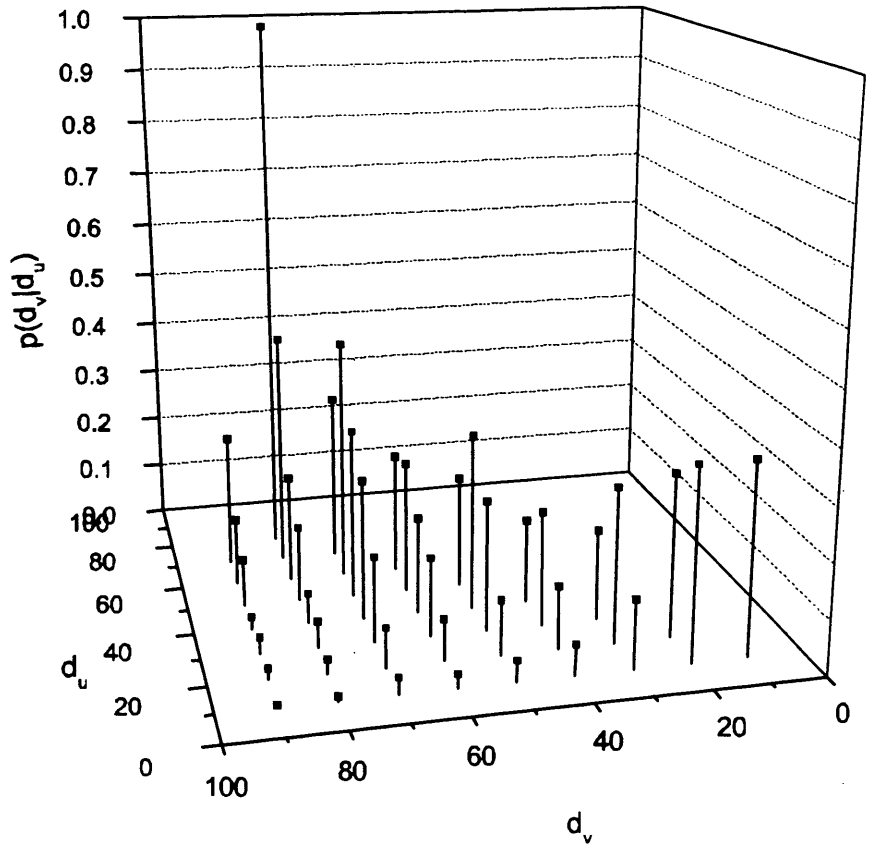


Figure 48 SDH node assortativity: A plot of the probability that the node at one end of a link has a certain degree given the degree of the node at the other end. (degrees are in bins of 10)

In Figure 48 we can see a plot of the probability that given a certain degree d_u of one endpoint node that the other endpoint node has a degree d_v. The plot has the degrees grouped in bins of 20 and d_u is the lower of the two node degrees.

The plot suggests that not much disassortativity occurs in the SDH network but there is a noticeable tendency for assortativity at both low degrees and high degrees.

SDH nodes exhibit assortative mixing at both high and low degrees and little disassortative mixing.

4.3.3.2 SDH Network Geography

In this section we examine aspects of the geography of the SDH network. Geography plays an essential role in the design of the network through a number of aspects including:

- **Transmission range limitations:** A single transmission span can be limited due to signal attenuation requiring either an SDH regenerator or either DXCs or ADMs (which would appear as nodes in the SDH topology).
- **Transmission cost:** transmission over longer distances require more expensive equipment changing the balance between cost and utility.
- **Circuit availability:** Longer transmission spans increase the probability of failure either through long physical layer sections or through increased equipment count.
- **Demand locality of reference:** Due to the fact that customers cluster in towns and cities and that there is an increased tendency to communicate with others which are closer geographically the demand topology may show an affiliation for demand links between geographically closer edge nodes.

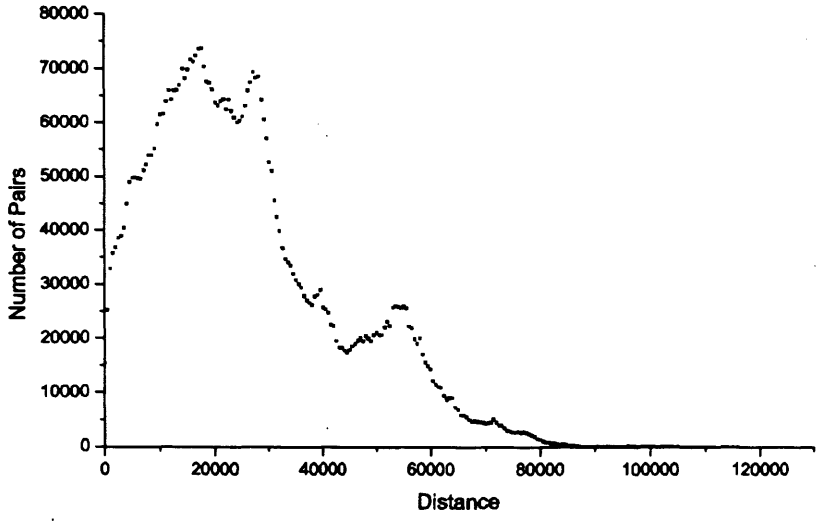


Figure 49 The number of node pairs which are a given distance apart (whether connected or not)(in 500 distance unit bins).

In Figure 49 we can see the count of numbers of pairs which are a given distance apart, in bins of width 500 distance units. The distances between nodes is a function of node density which is dictated by customer distribution, node purpose (edge customer facing node, or core Tier 1 node),

and cost. The transient peaks at approx. 20000, 30000, 40000 and 55000 are most likely caused by large groups of nodes in cities and this being the approximate distances between those cities. The lower values at the start are probably caused by the fact that to have the nodes so close together is not cost effective and avoided unless required for capacity or profit reasons. The largest distance between any two nodes is 116000 distance units (about 1160 kilometres).

Examining the distances between connected nodes we can see in Figure 50 the frequency distribution of the distances.

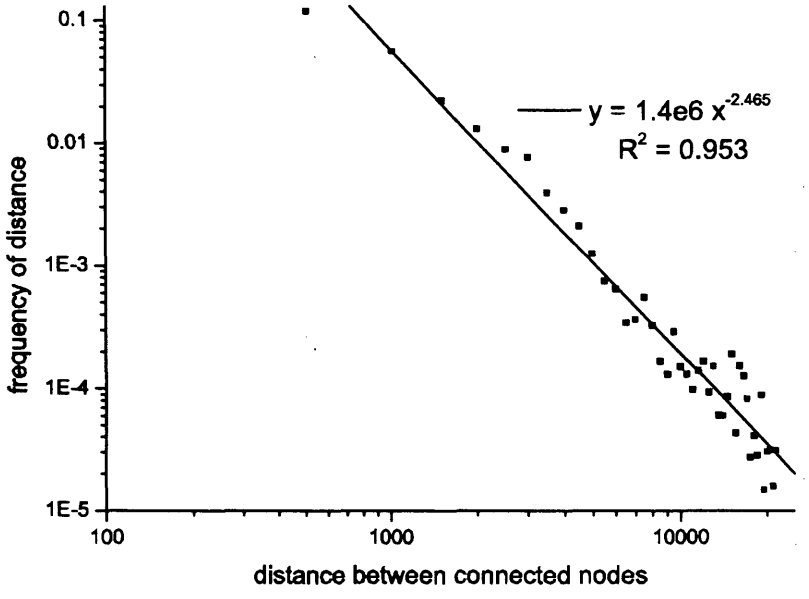


Figure 50 The frequency distribution of the distances between connected nodes (in bins of 500 distance units)

This is not, however, normalised in any way by the existence of nodes at that distance apart – the connected node pairs may not be present at a given distance because no such nodes exist to be connected. To normalise to the number of node pairs a given distance apart would be to generate a probability that nodes that far apart are connected and this is the basis of the Waxman Topology model.

4.3.3.2.1 Applicability of the Waxman Topology Model

Waxman proposed [WAX88] a simple geography based random topology generator which he used for experiments in multipoint routing, and it has become a common random topology model

for various networks [NAL05], including the IP layer. The model specifies that the probability of two nodes a given distance apart (on a geography based Euclidean plane) are connected is:

$$P(u,v) = \beta\, e^{-d(u,v)/(L\alpha)};$$

Equation 19 The Waxman topology connectivity equation

Where α and β are in the range (0,1] and are both parameters of the generated topology. L is the largest Euclidean distance between any two nodes and $d(u,v)$ is the distance between nodes u and v, $P(u,v)$ is then the probability that those nodes should be directly linked. By re-arranging the equation into the form $y = mx + c$ it can be shown that if x is specified as d/L then:

$$\begin{aligned} y &\equiv \ln(P(u,v)); \\ m &\equiv -\alpha^{-1}; \\ x &\equiv d(u,v)/L; \\ c &\equiv \ln(\beta); \end{aligned}$$

Equation 20 The Waxman connectivity equation re-arranged to the straight line equation

where $\ln(.)$ denotes the natural logarithm.

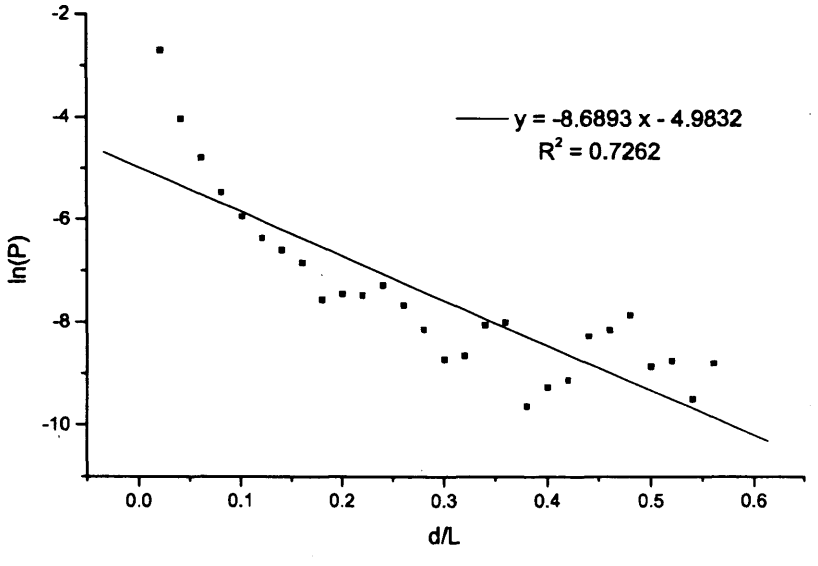


Figure 51 Estimating the parameters for the Waxman topology model: the probability that two nodes are linked (P) vs the normalised distance between them (d/L, which is in bins of 0.02).

Then, by plotting the probability that two nodes in the SDH network are linked, against the distance between them normalised to the maximum distance between any two nodes we can examine the conformity to the Waxman topology equation, as in Figure 51.

The best linear fit to the points suggest that m is -8.6893 and c is -4.9832. The equivalent Waxman parameters are, therefore: $\alpha = 0.1150$ and $\beta = 6.85*10^{-3}$ which satisfies the Waxman parameter boundary conditions.

In the next chapter these values will be used to test the Waxman model's suitability in modelling an SDH network.

It should be noted, however, that the Waxman model for connectivity is far from ideal as the fit's correlation coefficient is only 0.7262. Waxman has a tendency to underestimate the probability of linking nearby nodes.

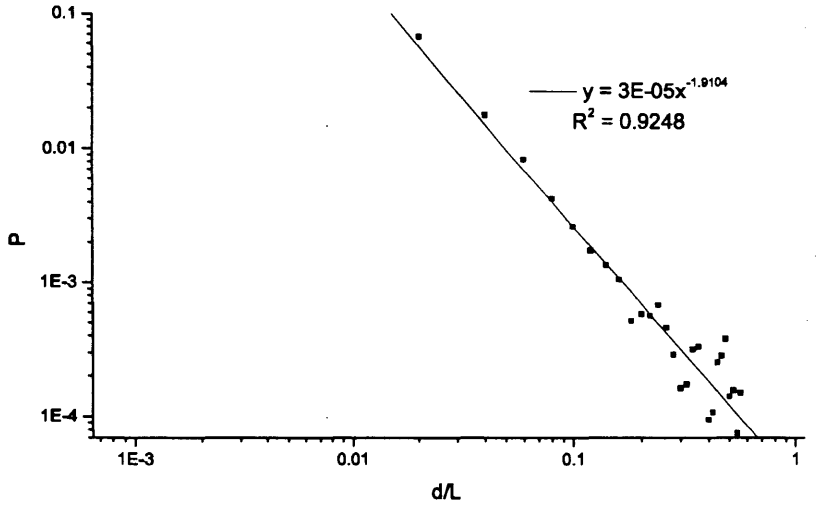


Figure 52 The probability that two nodes are linked against the normalised distance between them (d/L is in bins of 0.02).

For a better model of geographic connectivity we can see a similar plot of probability of connection against normalised distance in Figure 52 but following Equation 21.

$$P(u,v) = \alpha\,(d/L)^{-\beta}\ ;$$

Equation 21 The connection probability of SDH nodes given the Euclidean distance between them

In this case coefficient values of $\beta = 1.9104$ and $\alpha = 3.0*10^{-5}$. The correlation coefficient is better than the Waxman model at 0.9248. It should be noted that Figure 51 and Figure 52 are not directly equivalent (subject to normalising the distance to the largest distance) to Figure 50

since the latter is the distribution of distances of all links while the former plots the probability that a link exists over a given distance, therefore, each bin is the ratio of links that exist over that distance to the total number of possible links over that distance.

The probability of two nodes being linked in the SDH topology is proportional to the normalised distances between them to the power of a constant.

It should be noted that the exponent of the probability of connectivity over a given distance (1.91) is similar to the exponent of distance that is used in the two-body gravity equation which has been used as a model for traffic forecasting [VAL07] when the demand pattern is unknown:

$$F = G\frac{m_1 m_2}{d^2}$$

Equation 22 The two body gravitational equation

Where, in the gravitation model, G is the gravitational constant, m_1 and m_2 are the masses of the two bodies and d is the distance between the bodies, and the result, F, is the force experienced due to gravity. In traffic forecasting a similar model is used where m_1 and m_2 and the number of customers at a node, G is a parameter, d is the distance between the two nodes and F is the forecasted traffic between those nodes.

4.3.3.3 The spatial distribution of topology

In this section we will examine how pervasive the topological traits are at different scales of geography – more specifically by changing the measurement accuracy we will see how coarse the measurement granularity can get before losing the degree distribution power law. This is equivalent to examining the connectivity of exchanges, neighbourhoods, councils, towns, cities, counties and so on.

In Figure 53 we can see the SDH circuit topology with the SDH nodes placed at exactly the locations in the original data set. Figure 54 is the same topology although nodes that are within 4000 distance units of each other are collapsed into a single effective node. If there are links between any nodes in each collapsed set to any node in another collapsed set then the nodes representing the collapsed sets are connected in the new topology.

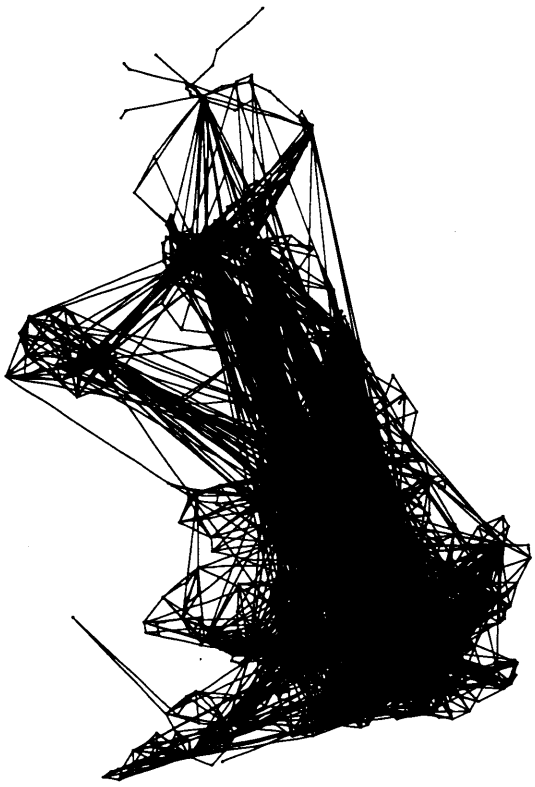


Figure 53 The BT national SDH circuit topology at the finest available geographic granularity

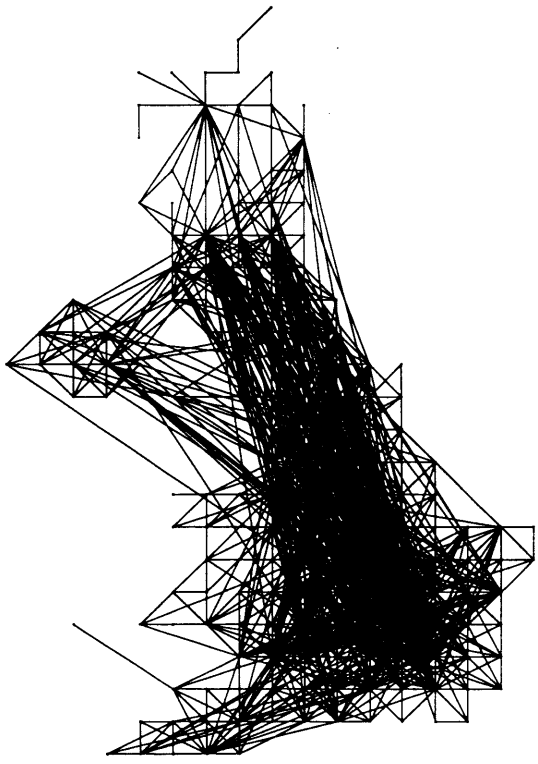


Figure 54 The BT national SDH circuit topology with sites merged to the nearest 4000 distance units (about 40km)

To parameterise this analysis we create a grid where each square is of side M distance units and the nodes are collapsed to the nearest grid junction. To prevent rounding artefacts in the analysis the grid is shifted horizontally and vertically in steps which are one tenth of the grid size and error bars are generated from this.

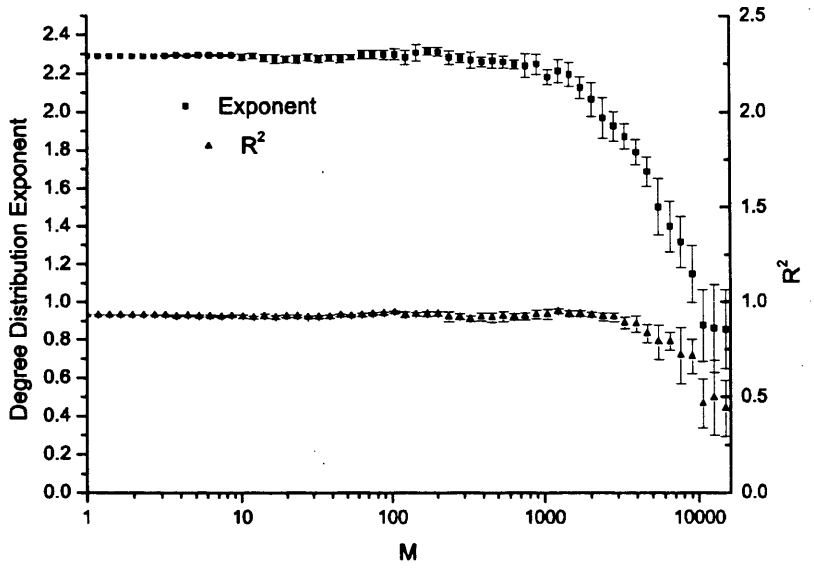


Figure 55 The power-law exponent and correlation coefficient for different grid granularities, M, which is in distance units.

In Figure 55 we can see the exponent and correlation coefficient for a power law fit to the degree distribution of the various topologies formed by adjusting the grid granularity to M distance units. It can be seen that R^2 does not drop below 0.9 until M reaches about 4000 (about 40km).

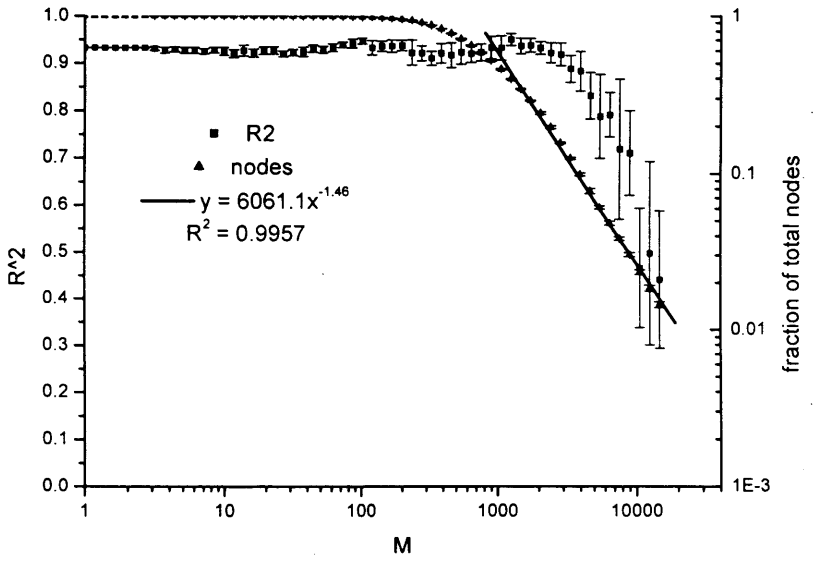


Figure 56 The correlation coefficient of the power-law fit to the degree distribution for different granularity grid sizes, M, in distance units, as well as the number of nodes in the collapsed topology.

From Figure 56 we can see that by the time the correlation coefficient reaches 0.9 at a grid size of 4000; the collapsed topology now only consists of 10% of the number of nodes of the original topology. This is visualised in the topology in Figure 54.

The SDH site degree power-law extends through different geographic scales, as far as the connectivity of SDH node clusters between cities.

An interesting artefact of this analysis is that the counting of total nodes in the topology at different measurement accuracies is equivalent to the box-counting method [RAS90] for finding fractal dimension [PEI92]. The fact that the trace forms a straight line between an M value of 1000 and 10000 suggests that the nodes are actually positioned as a self-similar fractal and the gradient, 1.46, is the fractal dimension. This information is useful in node placement algorithms when generating geography based scenarios for topology generators.

The Euclidean node distribution is self similar with a fractal dimension of 1.46

4.3.3.4 SDH Bandwidth

In this section we will examine the distribution of bandwidth throughout the network and attempt to extrapolate the location of capacity. The bandwidth of a link connecting two nodes in this case is considered to be the total of the capacities of the circuits which traverse this link. In reality this could be provisioned by multiple physical links and they can take diverse physical routes and have different protection at the transmission layer but that is all transparent to us and we only see bandwidth between sites.

The first thing to examine is the distribution of actual bandwidth values as per Figure 57 where the frequency distribution of link bandwidths is plotted in bins of 500 Mbit/s. The existence of many lower bandwidth links and a smaller number of much higher bandwidth links is not unexpected and exhibits a large tail. It can be shown though that much of the extreme part of the tail (on the right) is caused by direct loops in the SDH network – i.e. where a node connects to itself (see Table 9).

What should be noted when examining the plot is that it may not be the Tier 1 links that constitute this tail because by definition the Tier 1 links are the cross country links and because there is potentially a locality of reference in demand pattern (probably more so than in the Internet) many of the demands end and terminate in the same city – the highest capacity links could be large backbones connecting north and south London. This hypothesis will be examined further in this section. As we shall see later in section 4.3.3.7 the cost charged for transporting traffic is a function of a number of things, including the Euclidean distance travelled, hence an implicit locality of reference in demands.

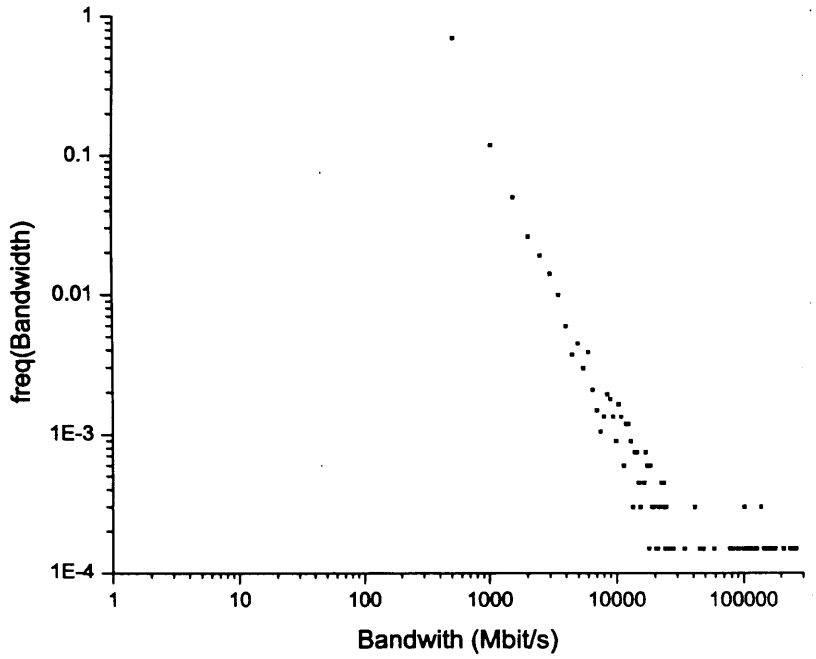


Figure 57 The frequency distribution of the total circuit bandwidths on each link (in bins of 500 Mbit/s)

To examine the highest capacity links more closely we can see in Figure 58 the rank of links in order of decreasing bandwidth.

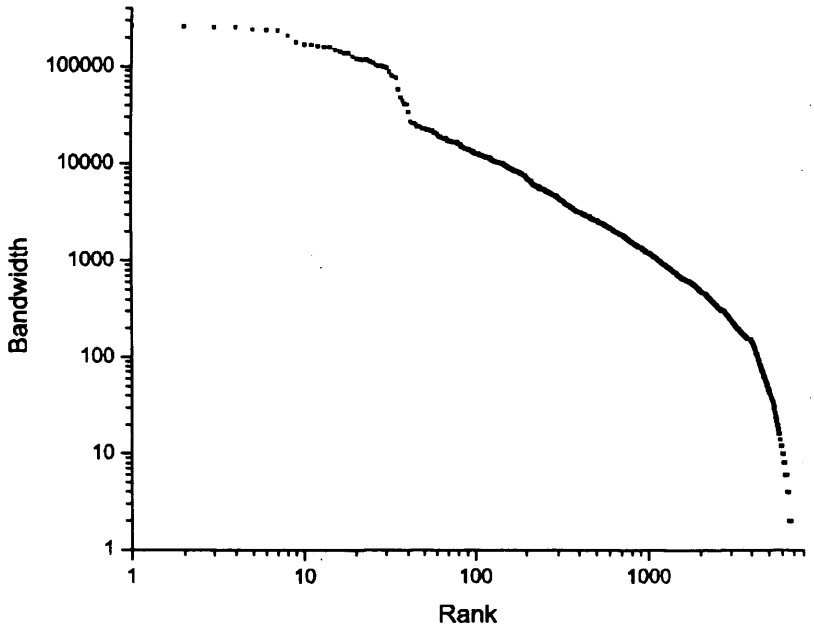


Figure 58 The bandwidth values of all SDH hops (including loops) ranked in increasing order

The step on the left is due to the loops as mentioned earlier. Things to note in this plot are not only the step but also the lower bandwidth of the higher rank links prevent the plot from being approximately power-law compliant.

Next we examine the correlation between metrics which could express the "size" of a node. In Figure 59 we can see a scatter plot of the degree of a node against the average bandwidth of all links connected to it (i.e. sum of bandwidths of all links divided by number of links into a node).

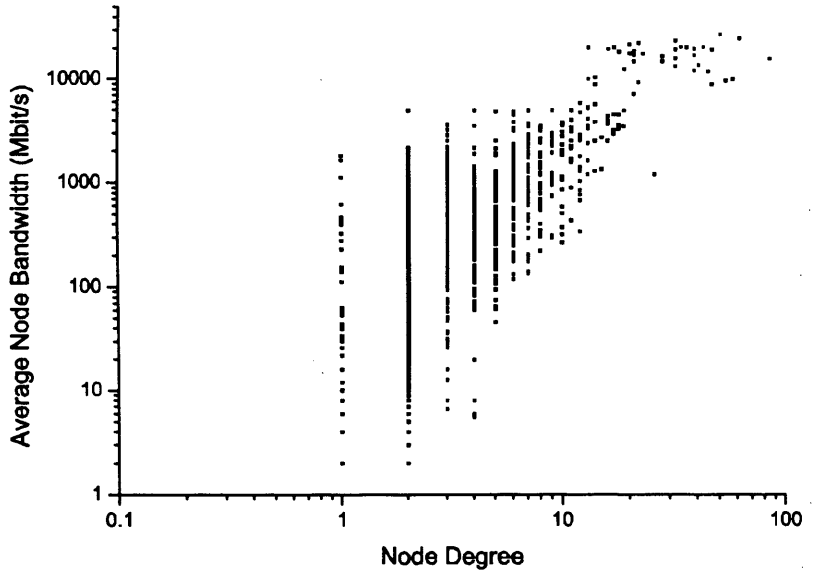


Figure 59 The degree of an SDH node compared to the average bandwidth of its links.

There is little grouping in the lower degree nodes, with average bandwidths of between 2 and 4000 Mbit/s for degrees of 1 to 15, but there is a tendency for the lower bound on bandwidth to increase with degree. What is obvious from the plot is that the high degree nodes all have very high average bandwidth (>4Gbit/s) which tends to suggest these are all some kind of core switching points (bear in mind this is average bandwidth and not total bandwidth, so it is already normalised to degree).

Nodes with a high degree tend to be higher capacity nodes
(those with a higher average bandwidth per link).

To investigate further where the bandwidth is located we see in Figure 60 a plot similar to the assortativity earlier but instead of count we plot the *average* bandwidth between nodes of the given degrees. The degree values in the plot are grouped in bins of 20, so the peak at 100,80 is the average bandwidth of all links who have one endpoint with a degree in the interval [80,100) and the other at [60,80).

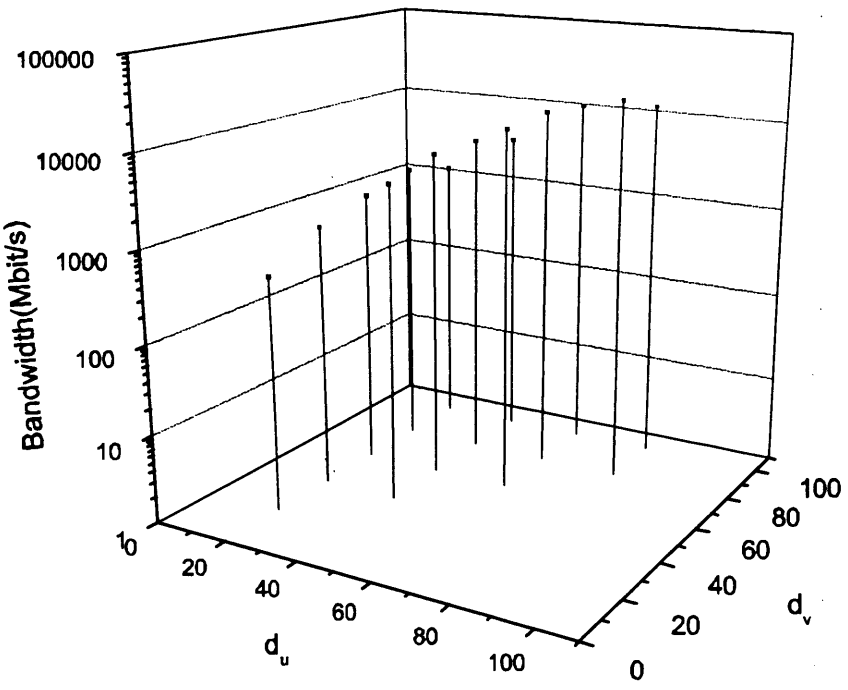


Figure 60 A plot of the average bandwidths of links given the degree values of their end nodes (degree values in bins of 20).

While detail is lost in the plot because of the averaging it is still possible to see that there is a general trend to have higher bandwidth links connecting nodes with higher degrees. The inverse is also true with lower capacity links connecting the nodes with lower degrees.

Links connecting better connected nodes (nodes with a higher degree) tend to have a higher bandwidth.

One hypothesis offered by this thesis is that networks attempt to minimise resource usage, and the cost of transmission is a function of at least distance and the capacity. To examine whether this can be seen in the SDH data we see in Figure 61 a scatter plot of all links with their bandwidth and the straight line distance between the nodes they connect.

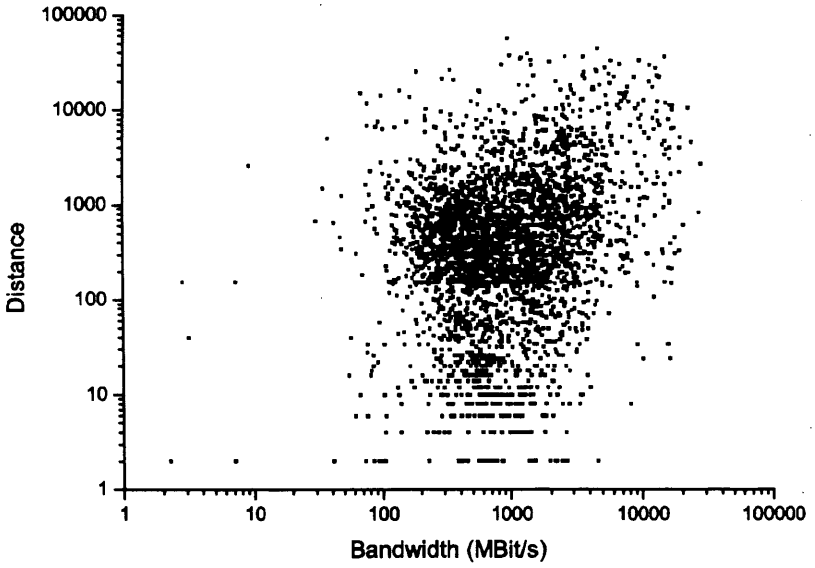


Figure 61 A scatter plot of the total bandwidth between every connected node pair against the geographic distance between them.

Even though it is known that both distance and bandwidth are variables in the cost function of the transmission layer there is little or no correlation between distance and bandwidth according to Figure 61. This could be caused because distance is minimised elsewhere and the utility of the link would be significantly less if it was under-provisioned.

Potentially the higher capacity links should also be shorter because they are the links connecting the large parts of a city, but there is no obvious evidence to that effect.

There is little correlation between the distance traversed by an SDH link and the bandwidth of that link.

4.3.3.5 Network Hierarchy

There have been a number of references in this chapter to the core nodes and the tiers – in this section we will examine whether it is possible to deduce any "core-ness" from the data set.

To start we find the average distance in hops that a node appears from the start or end of a circuit, and plot the frequency distribution of this average distance to edge in Figure 62. The average distance along the circuit is considered rather than the shortest path to any edge node because the circuit would adhere to any grooming to hierarchy that may exist. If shortest paths were

considered then 99.5% of nodes were also ingress or egress (therefore, at a distance of 0 hops to the edge), then 0.4% (about 16) were one hop away and 0.1% were two hops away.
The frequency distribution of average distances to edge nodes is plotted in Figure 62 where it can be seen that the vast majority of nodes are very close to the edge (less than 1 hop on average).

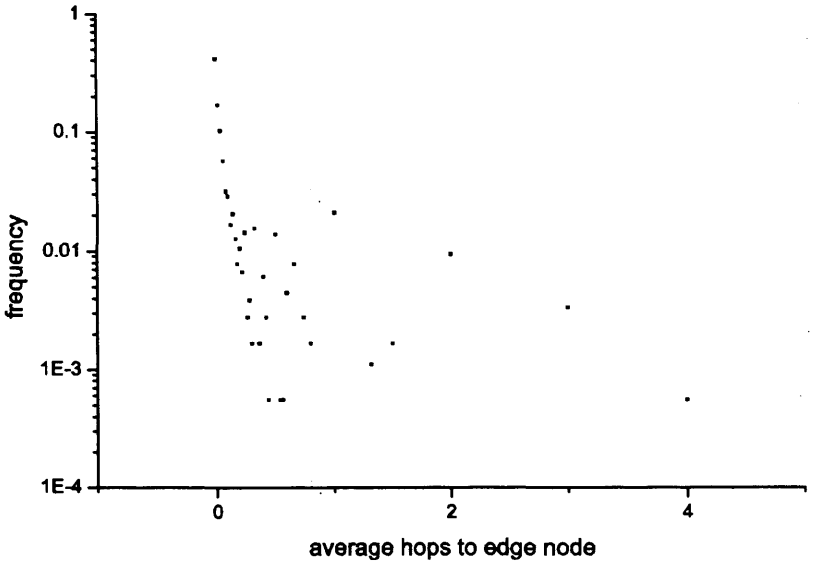


Figure 62 The average distance to the beginning or end of a circuit, in hops, that a node appears at for all circuits.

Now, we consider the location of bandwidth within a circuit. In Figure 63 we can see the average distance into a circuit (from nearer end) that a link appears against the bandwidth of the link (i.e. the sum of all the circuits traversing the link). The plot suggests that the further away from an edge node a link is, the lower the bandwidth. This may be the result of the locality of reference that will be seen in 4.3.3.7 (and specifically Figure 75) – the demands usually attempt to terminate at nearby nodes, which means a lower hop count and higher total bandwidth.
Figure 63 also looks as if there may be two triangular clusters overlaid on top of each other, potentially due to some element of the hierarchy, although those should manifest itself as triangles standing vertically (upper tiers are further away from the edge), rather than stacked horizontally (having a higher bandwidth).

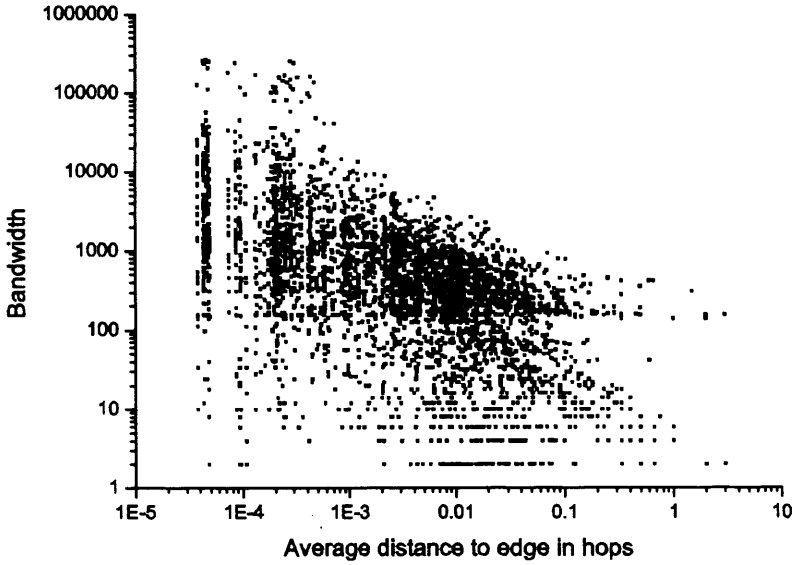


Figure 63 The average distance in hops that a link appears in a circuit against the total bandwidth of that link, averaged from all circuits.

4.3.3.6 Network Routing Efficiency

It is postulated in this thesis that one of the possible reasons for the emergent power-law traits is that they are the result of developing large hubs to minimise network diameter, and in general to minimise resource usage. To see to what extent this can be directly seen in the SDH circuits we will examine the efficiency of their routing.

4.3.3.6.1 Circuit Hop Count

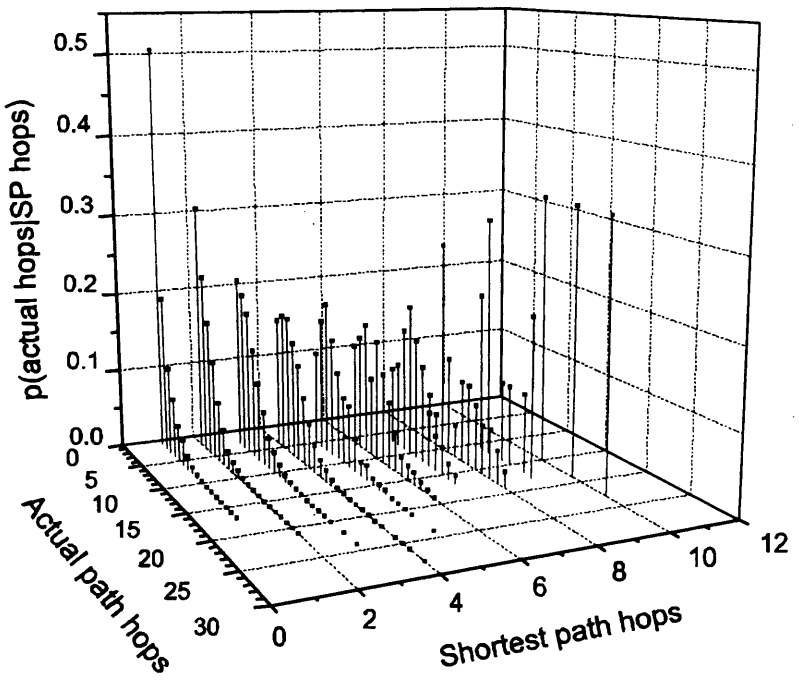


Figure 64 the frequency distribution of the actual number of hops in an SDH circuit vs the number of hops in the shortest path between the end-points (note different scales on hops axes)

In Figure 64 we can see the frequency distribution of the actual path number of hops given that the shortest path hops are a certain value, plotted for every SDH circuit or circuit segment (in the case of there being a customer site in the middle of a path). Circuits that start and end at the same node are removed as the shortest path is then zero hops.

The plot shows that the circuits which have a shortest path of 1 hop are predominantly single hops for the actual path, but with a tail extending up to 18 actual path hops. As the shortest path increases the peak of the most frequent actual hop tends to move away from the shortest path

hops – that is, as the shortest path hops of the circuits increases there is greater inefficiency and the actual hops tend to move away from the shortest path.
The largest disparity between shortest path and actual path is when the shortest path is one hop and the actual path is 18 hops.

The reasons why the shortest path is not followed are numerous, including:

- **Traffic marshalling**: a collection of circuits may be grouped together for reasons of operational or capacity management and this may take the path away from the shortest path.
- **Network design**: Depending on the circuit and its purpose it may be necessary to make the path pass through a different tier (Figure 41) thereby elongating the path.
- **Lack of resources**: the shortest path may not have the required bandwidth available to support this circuit, whereas the capacity required to support the demand may be available away from the shortest path – this is made more likely when considering that capacity is scaled with a granularity that is a factor of four (see Figure 33). Circuits may also be routed away from the shortest path to optimally fill capacity elsewhere in the network without requiring a capacity upgrade on links in the shortest path.
- **Diverse routing**: Since the main and standby circuits are considered separately it may be that one of them, probably the main circuit, follows a short route (less number of hops also means better availability (assuming homogeneous MTBF rates in all equipment)), while the standby circuit must follow a physically diverse path which would take it through different nodes and quite possibly over a longer route.

To put the extent of these non-shortest paths into context we can see in Figure 65 that of all circuits 37.2 % had an actual path length equal to the shortest path length.

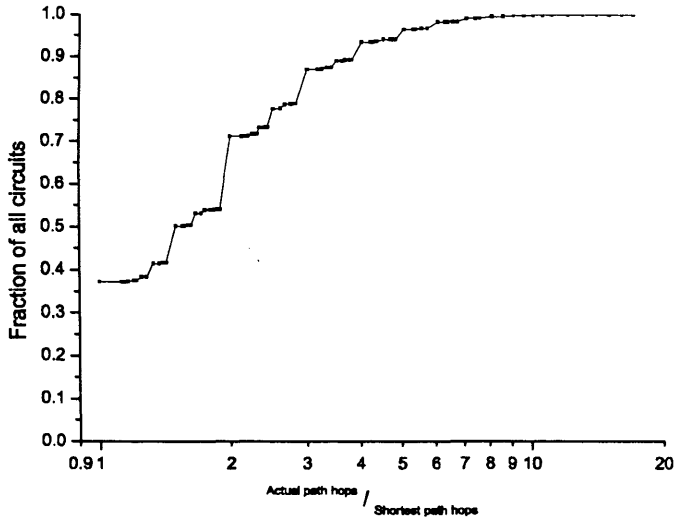


Figure 65 The cumulative fraction of all circuits that have a certain ratio of actual hops to the shortest path hops or less

About 40% of circuits follow the shortest path through the network, of those that do not some can follow paths up to eighteen times longer.

4.3.3.6.2 Euclidean Distance of Circuits and Demands.

Hop count is one metric that would be beneficial to minimise, another is the Euclidean distance travelled. The time to failure, and, therefore, availability of a transmission line is inversely proportional to the physical length, so it would also be beneficial to minimise the Euclidean distance travelled by the circuit. Propagation time is also an attribute which should be minimised in circuits, and therefore the tendency is to optimise for shortest Euclidean routes. While the data set does not contain any physical connectivity details and, therefore, does not contain the transmission line length we show in Figure 66 a plot of the average circuit distance (which is sum of the distances between the nodes comprising the circuit path) given the distance between the end points (the demand distance). The error bars are the highest and lowest circuit distance of any circuit with the given demand distance.

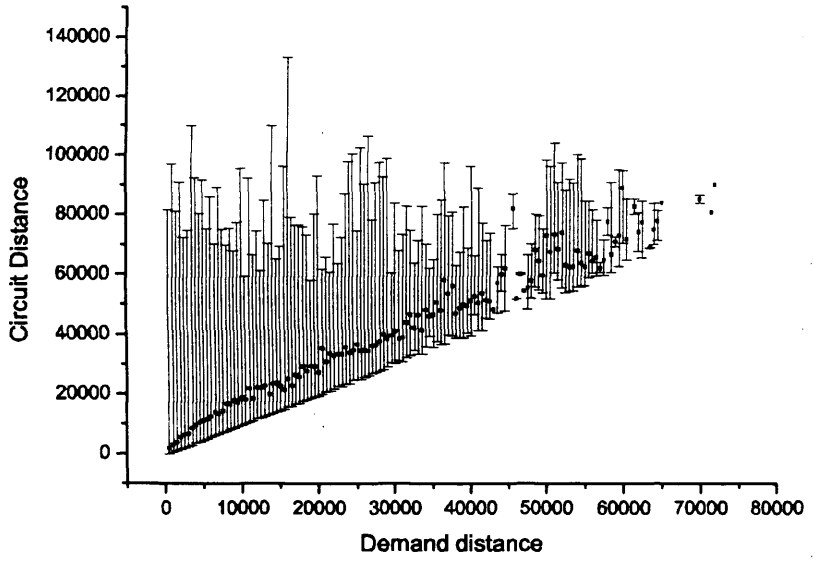


Figure 66 A plot of the average total circuit distance travelled by circuits given the demand distance. Error bars show lowest and highest circuit distances for the given demand distance. (demand distance binned in 500 distance unit bins)

It can be seen that for increasing demand distances the average circuit distance tends to move away from the lowest circuit distance. To see this is a ratio we plot in Figure 67 the cumulative fraction of all circuits that have a given circuit distance to demand distance ratio.

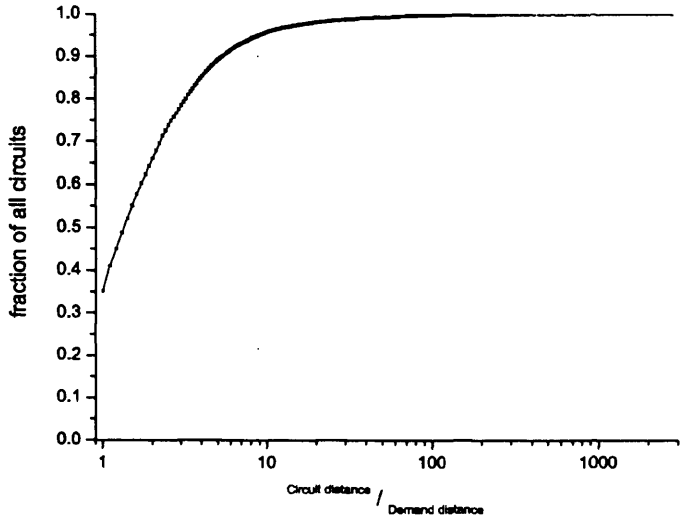


Figure 67 The cumulative fraction of all circuits that have a certain circuit distance to demand distance ratio

It can be seen, therefore, that about 35% of the circuits follow the shortest Euclidean distance possible, although this is probably in large part to the fact that they are probably the single hop circuits. Most circuits are within ten times the direct Euclidean distance, although the largest difference is at 2700 times longer.

About a third of circuits follow the shortest Euclidean path and the rest can take paths up to ten times further.

4.3.3.7 Demand Topology

Another topology that can be extracted from the SDH data set is the demand topology. The demand topology is formed by connecting the ingress and egress nodes of every circuit directly in a topology and ignoring any transit nodes in the circuit. This is the topology that is visible to the above layer. It should be noted that this topology is strictly the demand topology placed on the network rather than the topology of customers, since they are not present (or at least not entirely present) in the data set.

Since this is the layer above the transport layer the processes behind topological development are different. They are potentially less coupled to the distance limitations of the transmission layer, have a different cost function (that imposed by the transport provider) and have many influencing factors that are much more closer coupled to the services and applications as well as the economic and sociological aspects of the end user. This will directly affect the processes behind customer node placement, customer connection affinity (how much two customers would like to connect directly) and connectivity (the successful connection of customers, which is subject to the balance between connection affinity and connection cost).

The location of customers is historically based on geography (rivers, shelter), local resources and social factors, and rarely about the cheap abundance of local bandwidth, however, this does play a significant factor for many businesses who attempt to co-locate with other businesses or points of function like the Internet Exchange Points or the London Stock Exchange.
The connectivity should probably show, for example, an affinity for nearby customers to connect and cluster in towns and cities; however, with the decreasing nature of the cost of long distance communication and the Internet this trend is probably weakening.

The first plot to be considered as with the SDH topology is a ranking of the node degrees in Figure 68.

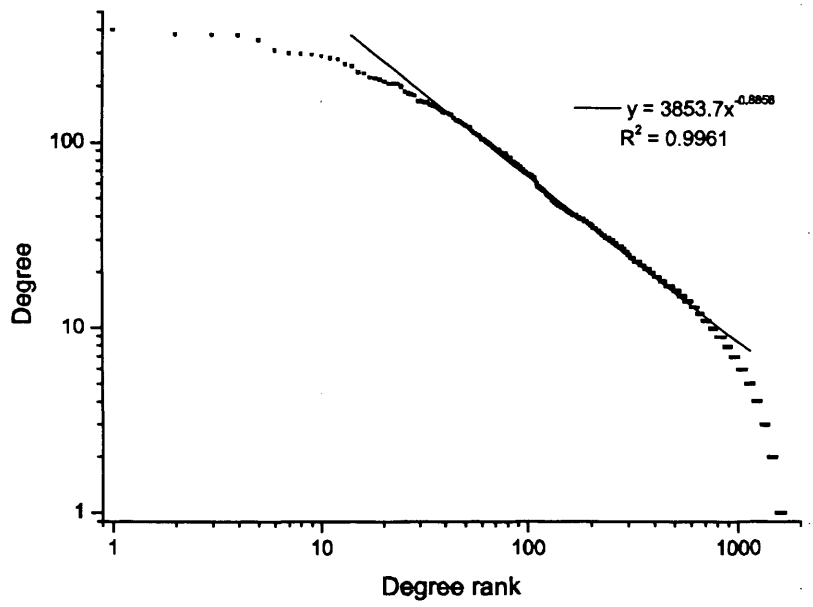


Figure 68 The ranked degree values of the SDH demand topology

The rank plot of the demand topology is not particularly power-law compliant and has a very obvious drop at the high rank values, that is where the degree is low – and this is equivalent to the region on the left part of Figure 69. The exponent of the power-law fit between rank 20 and 1000 is -0.8856.

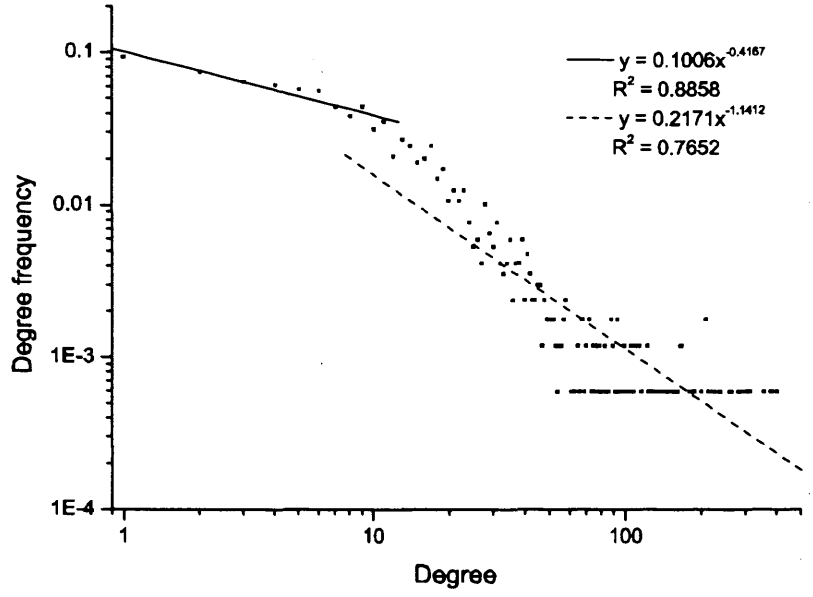


Figure 69 The frequency distribution of demand topology node degrees

In Figure 69 there is a noticeable step in the node degree distribution of the demand topology somewhere between degree 7 and 10. The precise reason for the decreased number of low degree nodes is unknown but it could potentially have something to do with one or more of the following:

- The set-up cost of an exchange which means that for an exchange to be created it requires at least a certain number of demands, or a certain bandwidth of demand, and subject to the range limitations of the access technology it may be cheaper to connect new users/demands to existing nodes that are further away. The reason the left side of the plot does not drop off further may be because those end-points that are there are high capacity business customers (which still appear as BT premises rather than customer premises).
- The change in slope is that this demand topology is the superposition of many different demand types, including voice and data, and their evolution over time. Therefore, the existence of a node may have been dictated a long time ago under PDH planning rules with the limitations of the equipment of that period.
- The demand types could also be the cause for a different reason – the large number of single degree nodes could be the residential telephone exchanges which connect to a switching centre in a hub pattern.
- Yet another possible explanation could be something to do with node type – whether they are terminal multiplexers or ADMs – from Figure 33 we can see that seven low-order paths are multiplexed into a VC3 or TUG-3.

The next plot to examine is the eigenvalue spectrum, which can be seen in Figure 70.

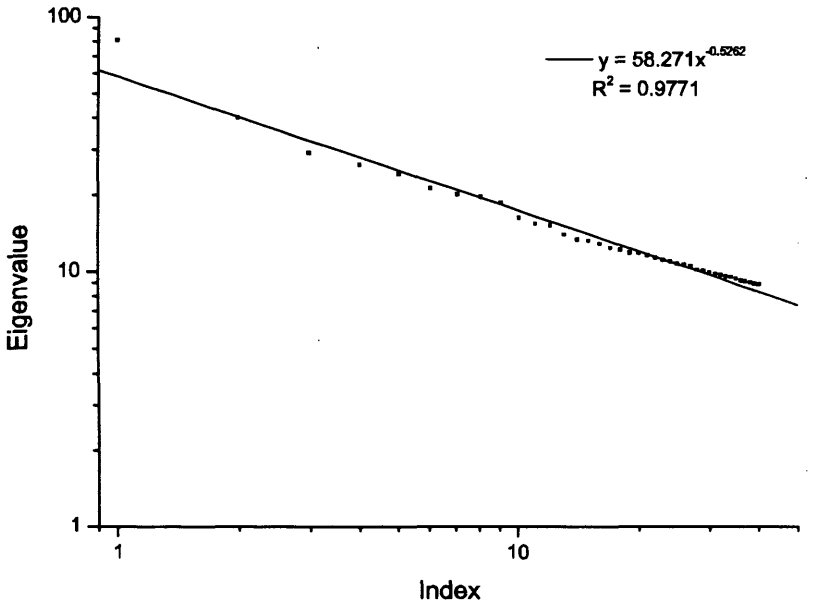


Figure 70 The 40 largest eigenvalues of the demand topology

The plot of the eigenvalues is power-law compliant with a correlation coefficient of 0.9771 and an exponent of -0.5262. The first eigenvalue is potentially of interest since it is significantly higher than the power-law fit.

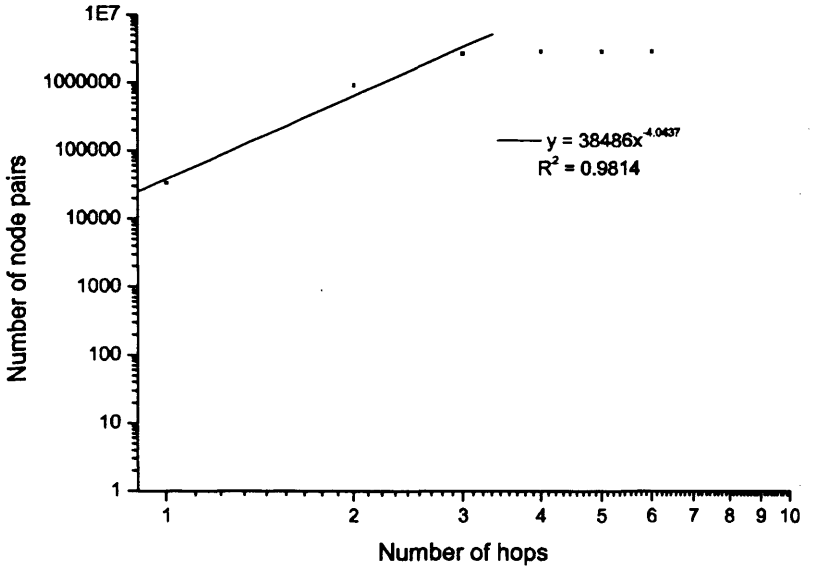


Figure 71 The total number of node pairs a given distance (in hops) apart.

In Figure 71 we can see a cumulative plot of the number of node pairs within a given number of hops. It can be seen that the demand topology actually does not only have a smaller diameter (see also Table 11) but generally has more node pairs within the lower diameters, that is, the exponent is 4.0437, while the exponent for the SDH site topology was 3.186 (Figure 45).

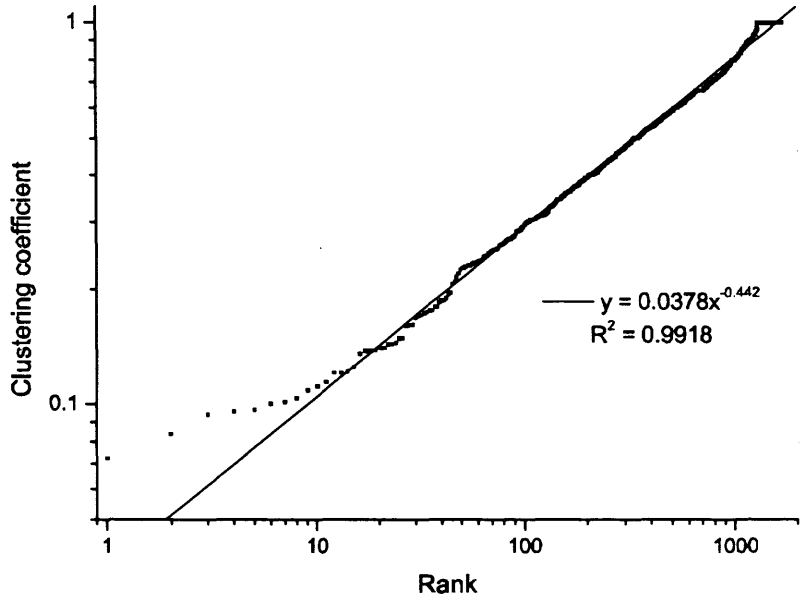


Figure 72 The clustering coefficient of the demand topology

Figure 72 is a plot of the rank of clustering coefficients as per Equation 18. An interesting thing to note in this plot is the very large number of nodes with a clustering coefficient of 1.0, which means that the neighbourhood of nodes around certain nodes are fully meshed. This could be the clustering of demands in towns and cities. In total there are about 400 nodes with such a clustering coefficient – about 20% of all the demand nodes.

Next we look at the frequency distribution of the distance between the endpoints (the Euclidean distances of the links in the demand topology) in Figure 73.

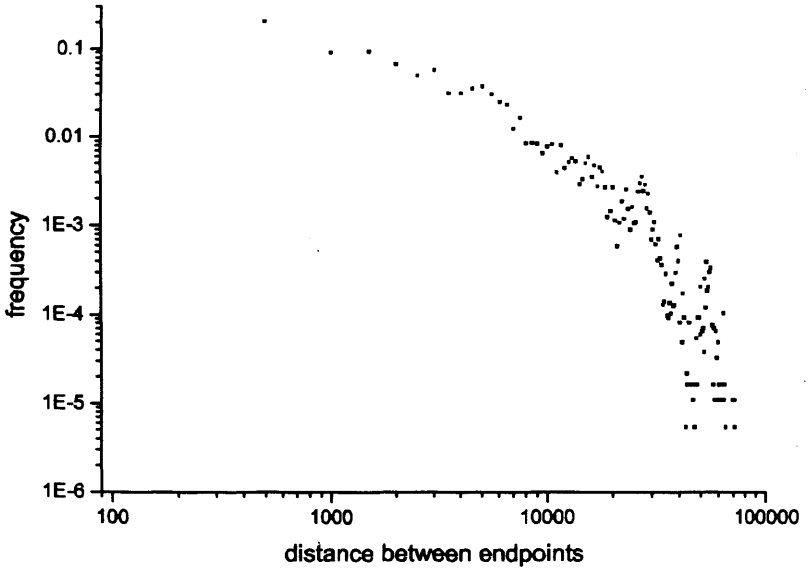


Figure 73 Distribution of geographic distance between demand edge nodes

The drop off is inevitable as the number of possible node pairs such a large distance apart diminishes; as mentioned previously the largest distance between any two nodes (connected by a demand or not) is 116000 distance units, but the maximum distance between any two nodes connected by a demand was 72000 distance units (about 720km).

The normalised count of any two nodes being connected (ie this frequency distribution normalised to the distances between all node pairs) can be seen later in Figure 75.

In Figure 74 we can see the frequency distribution of total bandwidth between edge node pairs (that is, the sum of the bandwidth of demands between common node pairs) – while there are many links (node pairs) with lower values for the total bandwidth between them, there are also a small number of links, or node pairs, with a very high number – these high bandwidth links are most likely the pairings of large population centres. The equivalent bandwidth dirstibution for circuit hops was seen in Figure 57.

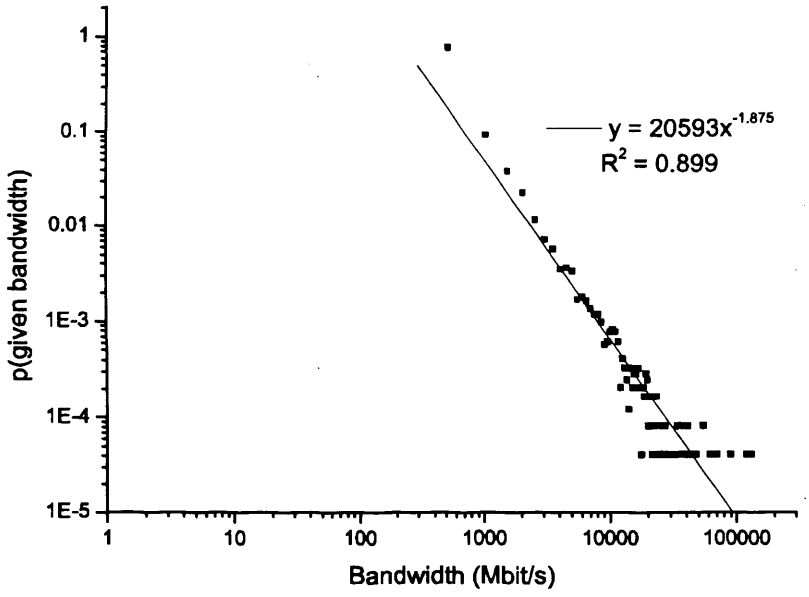


Figure 74 Total bandwidth between edge nodes pairs (in bins of 500Mbit/s)

It is known that as part of the costing of a connection from BT to the customer the capacity and Euclidean distance traversed is taken into consideration. To examine this further we see in Figure 75 a plot of the probability that a circuit (not a demand link) exists between nodes a given distance apart (that is the number of such circuits normalised to the number of node pairs the given distance apart), but split over different capacities. The probability can be above 1.0 because there could more than one circuit between node pairs of the same distance apart.

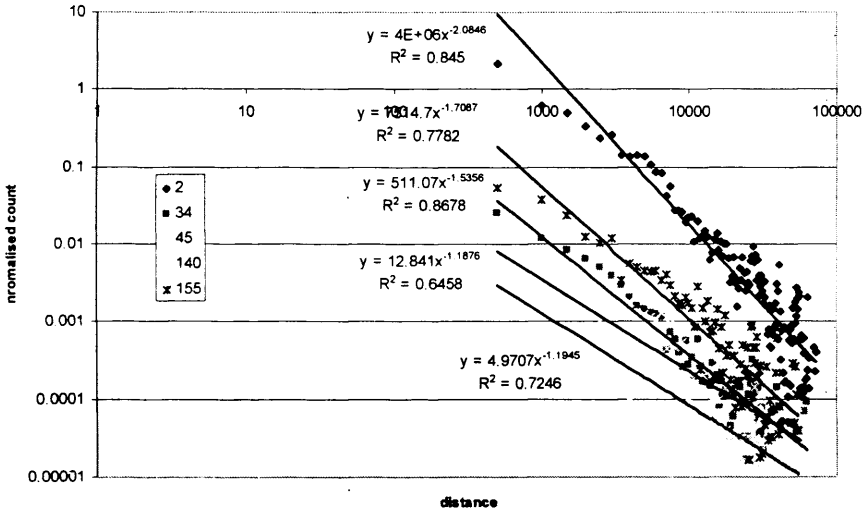


Figure 75 Distribution of geographic distance between circuit edge nodes for various capacities

The circuits of Table 10 are plotted individually based on their capacities (where VC-4 and STM-1 both count as 155 Mbit/s). It can be seen from the plot that the probability of a circuit connecting nodes a larger distance apart is much smaller than connecting nodes closer together; for example a 2Mbit/s circuit is 30 times more likely to connect between nodes 1000 distance units apart than it is nodes 10000 distance units apart. Together with the correlation coefficients, which are not ideal at between 0.6458 and 0.8678 depending on the bandwidth, this would suggest that circuit endpoints follow a similar connectivity distribution as was examined for circuit hops in Section 4.3.3.2.1.

Demands have a strong tendency to terminate at nodes geographically close to the source

4.4 Conclusions

We have seen in this chapter an examination of the SDH transport protocol and that while it does not require any explicit topological features it is common for hierarchies to be formed and rings to be used to provide circuit protection.

The analysis of the BT SDH network has demonstrated a series of power-law traits that were very similar to ones possibly seen in the Internet [FAL99]. The power-law traits are not explicitly planned into the SDH network and are the emergent result of other explicit design processes, potentially the urge to minimise resources. This chapter showed that the resource minimisation was present through the plotting of shortest hop distances against the number of hops in the actual

path, as well as examining the Euclidean distance between ingress and egress nodes and the Euclidean distance that the circuit actually traverses. These resource minimisation properties are expressed via the off-line SDH planning algorithms.

Other power-law traits included the clustering rank which demonstrated some definite structure present in the SDH node topology, as well as very high clustering in a fifth of the demand topology.

The power-law traits cannot be explained through a skewed topology measurement method like has been proposed for the Internet measurements either, since the data set was extracted directly from the SDH configuration database and is an instantaneous snapshot in time.

Another trait that was noticed was the tendency of SDH nodes with a similar degree to connect (high node assortativity), but high degree nodes would not connect so often to low degree nodes (low node disassortativity).

The Euclidean node distribution was shown to be self similar, a fact that will be used during node placement in the later modelling chapters. It has also been shown that the Waxman topology model does not particularly fit the SDH circuit connectivity and a model based on a power is proposed instead. Continuing with geography, it was shown that at least the power-law governing the frequency distribution of node degrees is maintained over a wide range of geographic scales, all the way down to collapsing all nodes in a 40km square into a node and the square's external connectivity was still found to be power-law compliant. For modelling demand topologies it should be noted that we observed a strong tendency to connect to geographically close nodes, which supports the theory of locality of reference.

In terms of node size it was shown that better connected nodes also tend to have a higher average bandwidth, and that high degree nodes tend to have high bandwidth links to other high degree nodes.

These observations will now be used to generate random and structured models in an attempt to produce similar topologies. The observations in this chapter have now laid the foundation to model a multilayer network exhibiting the urge to minimise resource usage in the transport, with a demand topology that is highly clustered and tied to geographic distances. The node placement we now know is self similar and, therefore, connectivity can be more realistic than the random node placement used in other models.

Chapter 5
Single Layer and Multi-Layer Network Topology Modelling

5.1 Introduction

In Chapter 1 it was shown how changes in a layer's cost function resulted in different node placement, and how a change in demand pattern also changed the network configuration. In practical terms the repositioning of nodes is prohibitively expensive and most layer adaptation occurs as link re-configuration. This chapter will build on the inter-layer feedback concept and with target topologies as described by Chapter 3 and Chapter 4 will search for the conditions necessary to achieve similar topologies. The cost functions and connectivity decisions used in the experiments in this chapter are partly gained from the analysis of protocols in Chapter 3 and Chapter 4, and partly from the existing single layer topology generators described in Chapter 2.

The analysis of IP and SDH networks in Chapter 3 and Chapter 4 suggested that instances of both types of telecommunications network layers tend to follow power-laws in various properties and at various measurement scales of the network. The power-laws were found to be very pervasive throughout the network (at different geographic scales in the case of SDH and at both the router and AS level in IP).

This chapter will examine the role that the multi-layer nature of networks can play in the existence of power-laws and present results for simulations of carefully selected scenarios to deduce the effect of feedback between layers.

The chapter initially describes a new topology generator tool, proposed and developed in the course of this research. It differs from existing topology tools in that it considers a demand pattern to be applied to the topology and the load that it exerts, and then performs a number of incremental functions including node addition, link removal and link addition based on this. The use of these three distinct steps allows us to test a range of network evolution scenarios including growth and rewiring. The chapter starts by examining single layer growth, as a baseline for the rewiring experiments (where a link is removed from one location in a topology and replaced elsewhere) later in the chapter, followed by multi-layer growth, then network erosion, which together with link addition are the two components of rewiring. Lastly the chapter examines link rewiring which is the topology evolution that occurs most often in networks due to the relative ease of provisioning compared to the addition of further nodes.

5.2 *The MITIE topology simulation tool*

The various topology models and implementations described in Chapter 2 have their relative merits, however, to investigate fully the feedback mechanisms between layers, in a scalable and versatile way, a new topology generator was created as part of this work, called MITIE. At the time of writing there were no equivalent packages and therefore MITIE was created for this purpose.

The **M**odular **I**nter-layer feedback **T**opology **I**nv**E**stigation tool and simulator (MITIE) is a collection of UNIX shell scripts, gawk scripts and C++ tools created as part of this research work and designed to emulate the incremental processes behind the reconfiguration of network layers. While all topology operations are incremental, they are intended to model the evolution and redesign of the network, attempting to mimic the processes described in Chapter 1, albeit not at such a drastic reconfiguration scale. That is, changes are incremental, with the addition, removal or rewiring of single individual links in every simulator cycle, and a full reconfiguration with the removal, addition and rewiring of multiple links and topological features (such as rings or hubs) is not possible in the current version.

Appendix A provides a full description of MITIE's operation and the options available in topology generation. In its simplest form MITIE performs three tasks optionally and repetitively, in this order:

- Node addition – a node is added to the network of existing nodes and a link is connected between it and another node. The destination node selection algorithm is user configurable and denoted by "**addG:**".
- Link removal – a link is selected to be removed from the topology. The algorithm used to select the link to be removed is user configurable and is denoted by "**remove:**". By default a link will not be selected for removal if its removal will partition the network.
- Link addition – a link is added between two existing nodes in the network – the endpoints are selected by two node selection algorithms (which could be the same algorithm or different algorithms) and denoted by "**addA:**" and "**addB:**".

"**addG:**", "**remove:**", "**addA:**" and "**addB:**" are described more fully in Appendix A, as are the tokens used to denote the selection strategies which are used as shorthand in this chapter. The appendix also contains many details and algorithms used in MITIE and will aid in the understanding of the scenarios in this chapter. A shorthand notation that will be used in this

chapter to denote the node/link selection methods is the GRAB tuple: {addG, remove, addA, addB} (that is the growth, removal and two link-addition endpoints selection methods A and B). It is possible to perform these operations less than once per cycle - so it would be possible to add a node at every cycle (and the link associated with it), but perform the link addition phase only every third cycle. Alternatively, it is possible to skip the node addition entirely and initialise with a topology from an input file and perform link removal, as well as link addition every cycle, thereby effectively performing rewiring on the network. Further usage scenarios are described in Appendix A and it is recommended to read that before continuing with this chapter. MITIE has many more options than those used in this chapter, including the ability to dynamically re-dimension link weights for routing at every cycle.

Some of the scenarios that will be considered in this chapter include:

{AL,0,0,0} – which grows the network and connects the new node to an existing node selected to prefer nodes with a higher average LPL (Load Per Link), i.e. average LPL of a node is the sum of loads transiting, originating and terminating at a node, divided by the degree of the node. This scenario does not remove or add any further links.

{0,arpl,0,0} – remove links with the absolute lowest RPL (Routes Per Link) until no more links can be removed without partitioning the network. There is no node addition or link addition.

{0,ARDP,rD,p} – remove the link with the absolute (do not use preferential random selection, just choose the highest absolute value) highest RDP (RPL-Distance Product; the RPL of the link multiplied by the Euclidean length of it) and then add a link between the node with the higher degree of the two that were just disconnected and a node randomly selected with a probability that is power function of the Euclidean distance between it and the high degree node. There is no node addition.

The multi-layer aspect of these scenarios is better understood if we assume that these are models for an SDH network carrying traffic which uses a shortest path algorithm, say IP with RIP routing (or OSPF with unity link weights, or BGP with no administratively set route selection policies, just the standard AS PATH length) between co-located nodes, and that each point-to-point SDH circuit is a point-to-point IP link. Every cycle one or more of the three node/link operations is performed and this creates/removes links, affecting the flow of demand. Based on the new

topology demands are routed in a different way, affecting link and node loads, on which the decision in the next cycle is made, and so on.
Alternatively, MITIE has more advanced options to allow specific demand topologies, as well as dynamic demand patterns, although this is not examined in this thesis.

A version of MITIE and the concept of separating node addition, link removal and link addition was described in Spencer et al [SPE02] and examined two network development scenarios in an attempt to model two different network types (MANs and LANs). Although the GRAB nomenclature was not used in the paper {D,0,D,D} (new nodes are added preferring high degree nodes and then a second link is added between two nodes both selected by preferring high degree nodes) was used to examine the development of AS topologies and {e,0,D,e} (new nodes are connected using the Waxman probability function (see Equation 12) of distance from the new node and a second link is added from highly connected nodes to a node that is selected based on Waxman probability function from the link source) for LAN or WAN development.

5.3 Experimental method

In the following sections a series of network evolution scenarios are proposed which will be performed as repeated invocations of MITIE and analysed to allow a comparison with the topologies of Chapter 3 and Chapter 4. Some of the growth and erosion experiments may not be by themselves plausible network development scenarios but have been performed as a decomposition of later rewiring scenarios to see if they alone could be responsible for emerging topology traits.
Scenarios are selected based on what we know and have deduced from the analysis of network layers – for example the need to minimise cost, while maximising reach, or the use of a power function in a geography based model, rather than the Waxman probability.
The feedback scenarios are mostly based on the load of links and how this affects network development - it should be noted that, to minimise the number of parameters and aspects under investigation the scenarios assume a demand pattern that is a full mesh of single-traffic-unit demands, and in this case RPL (routes-per-link) and LPL (load-per-link) can be used interchangeably. This also results in a much decreased simulation execution time and lower simulation memory requirements. The routing algorithm used by MITIE is a lowest cost routing between all node pairs based on either Warshall-Floyd [COR01] algorithm or a depth-first search. Unless otherwise specified the cost of a link is unity, therefore, the routing is actually all-pairs-shortest-path.

5.3.1 *On the use of line-fitting, least-squares and R^2 in network metrification*

To compare the generated topologies to the topologies in Chapter 3 and Chapter 4 we require a metric to quantify the likeness. We have seen the use of a power-law line fit to node degree distribution, node degree rank and other topology properties in previous chapters and while neither the IP nor the SDH topologies conform perfectly to these, the fit is considered to be close enough for our modelling target to be a good fit with a power-law. Our measure of representativity is, therefore, the correlation coefficient (least squares method), R^2, of the power-law line fit to the node degree frequency distribution, the node degree rank and the node clustering coefficient rank. Values for R^2 range from greater than 0.0 to 1.0, where 0.0 is no correlation to the fit, and 1.0 is perfect correlation to the fit and all points lie on the fit line. The exponent of the power-law, while significant in actual experiments on the topology, such as failure tolerance [ALB00a], is of secondary importance since there are so many factors that could influence it that this work only seeks to investigate under what conditions the power-law itself occurs.

5.3.2 *Initial network conditions*

Many of the experiments in this chapter will require input configurations, such as initial topologies for the link erosion, or node placement configuration for geography based node or link selection algorithms. These input configurations fall into the following categories:

- Initial seed topology for network growth - In the case of the network growth some experiments will require an initial topology of connected nodes (e.g the BA model, {D,0,0,0}) and the initial topology used here is a chain of five nodes in all cases of network size growth. This was selected because it is small and heterogeneous (the two nodes at the end of the chain have a degree of 1 while the others have a degree of 2).
- Initial topology for network erosion – since the only action is link removal then a highly connected initial topology would be beneficial to maximise the number of choices of links to remove. The use of a full mesh would make an interesting initial topology,, however, it is impractical due to the computational requirements to compute shortest paths through the network. The benefit of not starting with a full mesh is that the network evolution will start at a different point in topology space, as described in section 5.3.3. The networks used in these experiments are:
    - **ER model**: a 500 node network with 5 links per nodes (i.e. initially 2500 links) which is created by MITIE with {0,0,r,r} initialised with a 500 node disconnected network – so random node pairs are connected from an initial 500

node set of disconnected nodes until 2500 links have been added. Of the generated networks, only those which were found not to have any partitions were used as initial topologies in experiments (which was a very easy objective considering the largest cluster property of ER graphs (see section 2.3.2.1)), otherwise the full mesh of demands cannot be fully routed.

- **BA model**: from an initial seed network of the chain mentioned above, nodes were added with preferential attachment (see section 2.3.2.2), with five links being added from each new node to existing nodes, until the network reached the required size of 500 nodes.

There were ten instances of both the ER and BA networks, created by MITIE, and they had the following properties (note that CI is Confidence Interval):

Topology name	Mean minimum degree (left/right 95% CI)	Mean average node degree (left/right 95% CI)	Mean maximum node degree (left/right 95% CI)	Degree rank power-law fit R2 (left/right 95% CI)	Degree distribution power-law fit R2 (left/right 95% CI)	Clustering coefficient power-law fit R2 (left/right 95% CI)
ER_500N_5LPN	2.00 (1.13/2.87)	5.00 (5.00/5.00)	38.75 (34.19/43.30)	0.18 (0.02/0.34)	0.79 (0.77/0.84)	0.81 (0.79/0.83)
BA_500N_5LPN	5.00 (5.00/5.00)	4.99 (4.99/4.99)	75.00 (69.29/80.71)	0.83 (0.80/0.86)	0.99 (0.98/0.99)	0.95 (0.94/0.96)

- Initial topology for network rewiring – since the link to node ratio is to remain constant (one link is removed and one added per cycle) then there should be enough links to provide a good selection of removal candidates, as well as a small enough number that the effect of a single rewiring is not lost in the overall topology (i.e. there are such a high number of links that most routes take single hop routes to their destination, and a rewiring would affect only a very small part of the network). Two links per node per chosen as this is also the average degree of the SDH site topology in Chapter 4. The topologies, created by MITIE, had the following properties (note that CI is Confidence Interval):

Topology name	Mean minimum degree (left/right 95% CI)	Mean average node degree (left/right 95% CI)	Mean maximum node degree (left/right 95% CI)	Degree rank power-law fit R2 (left/right 95% CI)	Degree distribution power-law fit R2 (left/right 95% CI)	Clustering coefficient power-law fit R2 (left/right 95% CI)
ER_500N_2LPN	1.00 (1.00/1.00)	2.00 (2.00/2.00)	43.25 (37.39/49.11)	0.76 (0.74/0.80)	0.83 (0.81/0.84)	0.89 (0.87/0.90)
BA_500N_2LPN	1.75 (1.22/2.28)	1.99 (1.99/1.99)	61.75 (53.756/69.75)	0.74 (0.53/0.95)	0.97 (0.96/0.97)	0.96 (0.96/0.96)

- Network node placement – some of the node or link selection strategies are based on some measure of Euclidean distance and it is, therefore, necessary to specify positions on

a Euclidean plane of nodes. This is done in one of two ways – either by uniformly randomly selecting co-ordinates on a fixed size grid, or by selecting a random subset of the node positions from the BT network analysed in Chapter 4. The BT node geography distribution was shown to be a self-similar fractal pattern with a heavy tailed inter-node distance distribution.

For all of the experiments to investigate every scenario in this chapter, and for all initial conditions (except for the chain seed topology), ten random instances of each initial topology were created and the experiments were performed repetitively with the different instances, and the results averaged.

5.3.3 On Deterministic and Stochastic Node and Link Selection

In this chapter node and link selection schemes are at times absolute, deterministic decisions, such as "arpl" – select the link with the lowest number of routes traversing it (in case of tiebreaker the node number is used so the decision is also deterministic), and at times partly stochastic, such as "rpl" – prefer links with the lower numbers of routes traversing them, using roulette wheel selection to make the biased decision (see Figure 112). The reasons behind this are two fold:

- The face-value of the selection scheme is of interest, so, in the case of "arpl", (select the lowest RPL valued link) if route-per-link is considered as some kind of a link capacity then it may be necessary to find the actual lowest valued link since it is the least used link in the network. Alternatively, the BA model uses degree preference, rather than selecting the absolute highest degree value node because if it were to use the absolute value then no other nodes would have a chance to be promoted to higher degrees – the actual value of the "D" node selection is that it is partly stochastic, but biased.
- If the selection is deterministic then it can be traced in either direction through time, and the trajectory through configuration space can be captured. This is actually the concept of phase space in Dynamical Systems theory and the evolution of the network can then be considered as basins of attraction (the configurations from which trajectories approach the attractor) and attractors (the configuration at which the network settles). This allows for the analysis of dynamic behaviour without being knocked into another basin or attractor by the noise of a stochastic process.

The transfer function from one point in time to the next (that is the link addition/removal and/or node addition process) is what causes the progression along the trajectory, by choosing scenarios with different selection processes the basin and attractor space bifurcates and re-configures. Taking the BA model as an example of bifurcation – in the standard model there are a very high number of attraction basins – the topologies of each final network, and the final destination of the trajectory is actually based on the transfer function and the choices made by the random number generator. If the decision process was deterministic, and would only select the node with the highest degree, then there is only one configuration that will be reached - a full hub configuration. In terms of implementation, however, the random number generator is actually completely deterministic and only generates a fixed sequence of numbers; which sequence is decided by the seed number of the generator. For this reason each experiment is repeated multiple times with different random number generator seeds.

Another aspect to note is that the view of the final topology is incomplete since we are taking an analytical measure of the topology (the power-law fit correlation) and, therefore, at this granularity of measurement all the attractors can look similar – if they are not then we can see this through large error bars.

For the purposes of this chapter the selection of stochastic or deterministic is made based on the specific scenario under test. There is also a comparison of the use of stochastic and deterministic selection in section 5.6.2.2.

5.4 Network Growth Modelling

This is where a network is created by adding nodes or links or both. Network growth based on single-layer random connectivity has already been surveyed in Chapter 2, but experiments with single-layer models are also covered in this section since they are components of later rewiring models, and also serve as a baseline for modelling efficacy comparison.

5.4.1 Single Layer Network and Connectivity Growth

Many of the results in this single layer network growth section are not novel - {D,0,0,0} is the BA model, {r,0,0,0} is a random growth model, closely related to the ER model, {0,0,r,r} is the ER model, {0,0,ge,0} is the Waxman model and {e,0,0,0} is a variation of the Waxman model (although not an identical implementation since the Waxman model does not contain an element of growth – only the distance probability function as used by Waxman is common). The use of a

power-function of Euclidean distance ({p,0,0,0} and {0,0,gp,0} models) is novel though and whose creation was suggested by the link distance distribution in section 4.3.3.2.1 and the fact that the Waxman model did not fit the BT network Euclidean connectivity distribution.

Since single layer network growth is so heavily researched it is not examined too closely here but Appendix B has a large set of results of various scenarios for comparison. The scenarios are based on growth and link addition based on node degree preference, geography and uniformly random node selection, all selection processes that are used later in this chapter as part of the multi-layer modelling. The results are sorted by the decreasing average of the R^2 values for degree frequency distribution, degree rank, and clustering coefficient rank. Examining the highest ranked scenarios as reproduced in Table 12 there are a number of features to note.

Degree preferential attachment for example performs best (expts. 1,2,3,4,5,6,10,11), with the experiments adding more than one link performing better than the tree form (5,6,10,11). When geography is introduced performance is also very good but only for cases where incremental growth is used (compare expt 9 and 98 for power function geography and 21 and 109 for the Waxman function geography – both for BT node placement and 2000 node networks). Node placement does not seem to play a large role (compare experiments 9 and 27 for {p,0,0,0} with incremental growth and 2000 nodes – the average R^2 drops from 0.915 for the BT node placement distribution to 0.905 for a uniformly random distribution of nodes).
It should be noted that the use of incremental growth is implicit in networks that increase in size and use addG:D since the probability of linking to a node with no links is 0.
A mixture of geography and degree performed very well also, with {e,0,D,e} (experiment 7) and {p,0,D,p} (experiment 8) ranking first and second in terms of node degree frequency distribution power-law correlation, significantly better than experiments based purely on degree.

The table in Appendix B has many more experiments, including geography and degree based models with large numbers of links attempting to model the 12 average degree of the BT SDH demand topology in Chapter 4.

Exp #	GRAB tuple cycles:stepG ,stepR,stepA	IG	GEO	Network size/ nodes	Ave. Node Deg.	DFD R^2	DFD R^2 rank	DR R^2	DR R^2 rank	CCR R^2	CCR R^2 rank
1	{D,0,D,D} 10000:1,1,5			4000	1.500	0.848	64	0.955	1	0.969	1
2	{D,0,D,D} 1000:1,1,2			1000	1.499	0.835	68	0.951	4	0.948	4
3	{D,0,D,D} 4000:1,1,2			4000	1.500	0.827	72	0.952	2	0.951	2
4	{D,0,D,D} 2000:1,1,2			2000	1.500	0.826	74	0.952	3	0.950	3
5	{D,0,0,0}			10000	1.000	0.879	38	0.923	19	0.925	5
6	{D,0,0,0}			4000	1.000	0.877	42	0.923	20	0.924	6
7	{e,0,D,e} 2000:1,1,1	•	BT	2000	2.000	0.928	1	0.887	39	0.910	12
8	{p,0,D,p} 2000:1,1,1	•	BT	2000	2.000	0.927	2	0.886	41	0.906	17
9	{p,0,0,0}	•	BT	2000	1.000	0.925	4	0.878	46	0.916	9
10	{D,0,0,0}			2000	1.000	0.871	49	0.923	23	0.923	7
11	{D,0,0,0}			1000	0.999	0.873	47	0.922	25	0.921	8
21	{e,0,0,0}	•	BT	2000	1.000	0.907	16	0.870	51	0.909	13=
27	{p,0,0,0}	•	UR	2000	1.000	0.900	23	0.865	55	0.905	20
97	{d,0,0,0}			2000	1.000	0.792	82	0.777	96	0.827	74=
98	{p,0,0,0}		BT	2000	1.000	0.789	83	0.776	97	0.827	74=
109	{e,0,0,0}		BT	2000	1.000	0.754	98	0.759	105	0.809	91=

Table 12 The single layer modelling scenarios that produce networks with the best power-law fits

The headings are abbreviated thus: Exp # - experiment number, IG – incremental growth, GEO is the geography type, either BT or Uniform Random, DFD – Degree Frequency Distribution, DR – Degree Rank, CCR – Clustering Coefficient Rank.

5.4.2 Multi-layer Network Growth

With a better understand of the behaviour of single layer models and node selection processes this section begins to consider multi-layer modelling scenarios, initially network growth. Multi-layer network growth is where networks grow in size and new nodes are connected to an existing node which is selected based on criteria which is derived from an adjacent network layer, in this case the node's traffic load.

MITIE supports two node selection mechanisms based on load, TL, the total load on the node, and AL, the average load on the adjacent links, which is the total load divided by the degree of the node. Since we are assuming a full demand mesh TR (total routes) and AR (average routes) will be used in place of TL and AL. It should be noted that the TR value of a node is also similar

to the node's betweenness centrality, that is, how central the node is in the topology. Betweeness centrality is however typically calculated as a tally of the total number of shortest paths on which a node could lie, and not based on the single shortest path on which the node lies as is the case here. We will, therefore, examine {TR,0,0,0} and {AR,0,0,0} (no link removal or further addition, and, therefore, the link-to-node ratio is 1.0), but also a variant where a second link is added from the newly added node, with the same node selection scheme, that is {TR,0,g,TR} and {AR,0,g,AR} (we also use the –step:1,1,1 argument to MITIE which causes a secondary link to be added once per new node and link added, so the link-to-node ratio is 2.0). The use of the second link addition means that our graph is not a tree as would be the case with {AR,0,0,0} and {TR,0,0,0} and, therefore, the addition of new nodes and their two links is capable of creating shorter routes through the network, and being a source of multi-layer feedback based on re-routing. In Figure 76 we can see how the R^2 value of the power-law fit to node degree distribution changes as the network grows in size.

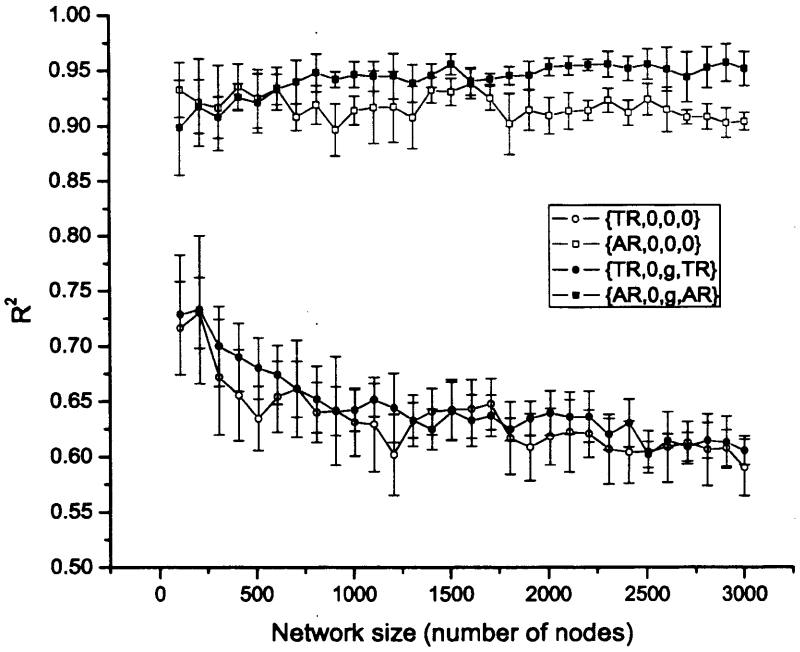


Figure 76 The correlation coefficient, R^2, of the power-law fit of node degree frequency distribution, as load based network growth progresses with {TR,0,0,0}, {AR,0,0,0}, {TR,0,g,TR}, {AR,0,g,AR}.

It can be seen that the network growth schemes based on TR do not perform well, with low R^2 values, with a downward trend, while both AR scenarios perform very well with correlation coefficients greater than 0.9 for most network sizes and showing no sign of falling below this

value as the network grows. The DFD correlation coefficient of {AR,0,g,AR} is, in fact, better at a network size of 2000 nodes than the BA model's degree distribution coefficient of 0.871 (see Table 12). What is interesting to consider in the plot also is the possible sources of the power-law nature: in the BA model, {D,0,0,0}, it is explained [BAR99] that the source of the trait is a "rich get richer effect", however, the total routes (TR) node selection weight is not dissimilar in that the more links a node has connected to it, the greater the routes, which will cause it to be selected more often in future – however, this increase in the number of routes is so high that the attraction to new links too strong – and causes an escalation in degree which is above the power-law fit. This can be seen in Figure 77 as a small number of nodes having very high degrees (the tail on the right) while leaving the rest with lower degrees. The {AR,0,0,0} in comparison shows significantly higher correlation with a power-law fit in Figure 78, even though the metric attracting new nodes' links is a function of the *inverse* of degree (the weight associated with a node in the endpoint selection for AR is the number of routes traversing a node divided by the degree) – so actually has as a component which is the opposite of the BA preferential selection. Growth based on the reciprocal of degree {d,0,0,0} was shown in Table 12 not to be power-law compliant.

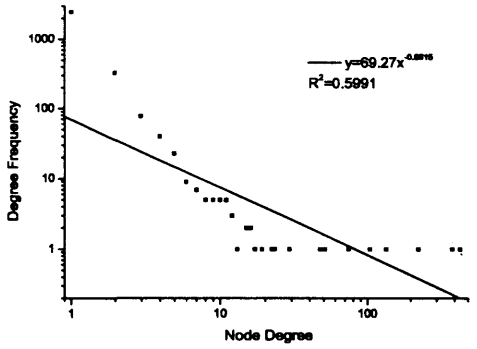


Figure 77 The degree frequency distribution of a 3000 node topology generated with {TR,0,0,0}

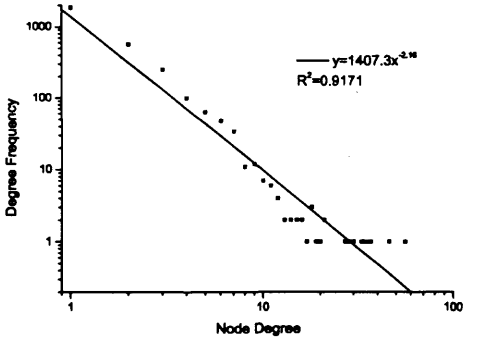


Figure 78 The degree frequency distribution of a 3000 node topology generated with {AR,0,0,0}

Since the correlation of AR based growth is so good the progress of the power-law correlation coefficients for other metrics of the topology are now considered. In Figure 79 the progression in the degree rank power-law correlation coefficient for the four scenarios can be seen.

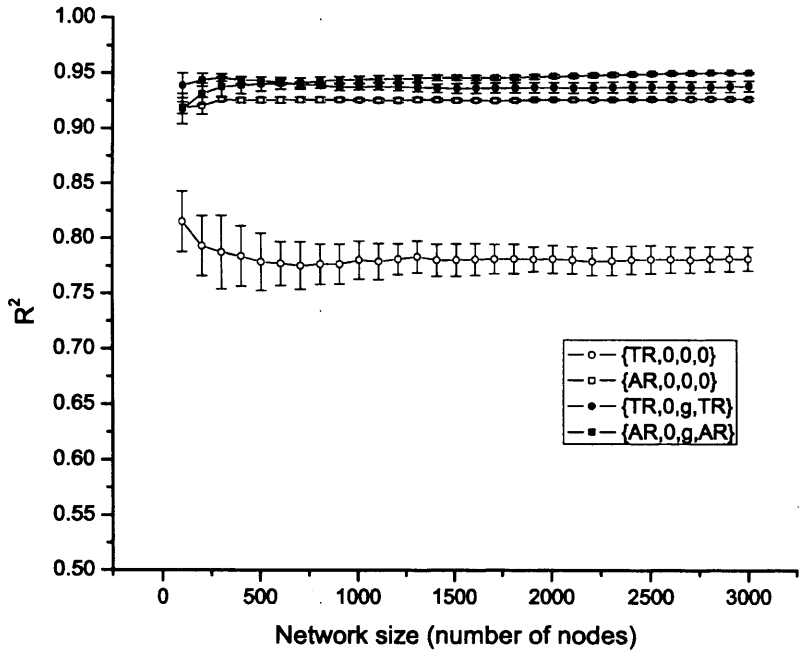


Figure 79 The correlation coefficient, R^2, of the power-law fit to degree rank, as load based network growth progresses with {TR,0,0,0}, {AR,0,0,0}, {TR,0,g,TR}, {AR,0,g,AR}.

It can be seen that in terms of correlation to degree rank power-fit {TR,0,g,TR} performs much better, and on a par with the AR based scenarios, however, {TR,0,0,0} remains with a low correlation, remaining below 0.8 for large network sizes. {TR,0,g,TR}, {AR,0,0,0} and {AR,0,g,AR} all performed better at 2000 nodes than {D,0,0,0} did, which had a R^2 value of 0.923 (see Table 12).

Figure 80 shows the power-law correlation coefficient to the clustering coefficient rank as defined in Figure 46.

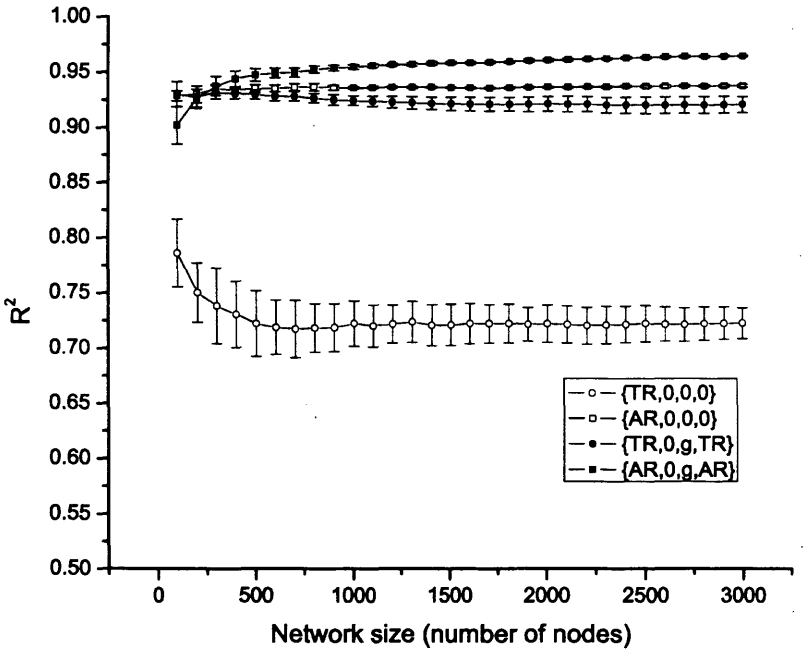


Figure 80 The R^2 correlation coefficient of the power-law fit to clustering coefficient rank, as a function of the network size as the network is grown according to the load based connection algorithms.

With clustering coefficient, like with degree rank, the {TR,0,0,0} performs worst, with {TR,0,g,TR} again being on a par with the AR scenarios, but it can also see that {AR,0,g,AR} performs noticeably better still, with correlation coefficients approaching 0.975. In comparison with {D,0,0,0} at 2000 nodes which had an R^2 of 0.923, both {AR,0,0,0} and {AR,0,g,AR} performed better, but {TR,0,g,TR} was approximately equal, and even showed a gentle downward trend.

It has now been shown that growing networks and connecting new nodes to existing nodes based purely on a load or centrality metric (TR) does not tend to create power-law compliant graphs. From Table 12 it can be seen that node selection based purely on degree produces topologies that are power-law compliant (see the BA model {D,0,0,0}), which shares the rich-get-richer effect with {TR,0,0,0}. However, using node selection based on both load and degree, AR, can yield higher power-law correlation than even the BA model in producing power-law compliant graphs in at least the degree distribution, degree rank and clustering coefficient rank topology metrics.

What is also significant to note is that {AR,0,g,AR} model always performs better than the {AR,0,0,0} model – and it is in this former model that a slightly different type of feedback is occurring. In the {AR,0,0,0} case only one link is added for every new node, and no further links are added at each cycle; the topology is, therefore, a tree, and as such the feedback comes from the increase in load at the all the nodes. In the {AR,0,g,AR} case, however, two links are added from the new node and, therefore, the topology is not a tree but contains alternative paths. These alternative paths then affect the shortest paths of the already present demands by offering alternative routes; so there is also feedback through re-routing of demands, as well as the simple feedback of increased load.

5.5 Network Erosion Modelling

After network growth, the next form of network development type to be considered is that of network erosion, where either nodes or links, or both, are removed from the topology. In this section we concentrate on erosion in terms of link removal only as it is the first part of link rewiring, considered later.

In practical terms link removal could also occur as a development process in telecommunications networks when a disruptive technology emerges and the cost of scaling the capacity of links becomes lower than the cost of scaling the number of ports. This could happen if transport capacity scales to values that can easily support demand (either through physical capacity scaling or through the use of technologies such as VCat (see section 4.2.1.4)), and the installation cost of a physical ports is high, but the cost to scale the ports in terms of speed is low (as would be the case if additional ducts would have to be dug and installed).

For each experiment the initial topology is 500 nodes in size and either generated by the ER or BA models (see section 5.3.2) and has an average of 5 links per node. This initial topology then has links removed until there is only on average one link per node. For most of the experiments bridge links (links whose removal would cause the partition of the network) will be avoided, unless explicitly specified, in the removal process to maintain all-pairs routing integrity.

For the following network erosion experiments topology statistics were extracted every 100 simulation cycles, therefore, for every point on a graph, 100 links have been removed since the last point.

5.5.1 *Single Layer Network Erosion*

In the simplest case the link selected to be removed is selected at random with a uniform probability. For the initial topologies with BA and ER one link at a time is removed and after every 100 cycles (one link removed per cycle) the correlation coefficient of a power-law fit to the node degree frequency distribution is plotted. This can be seen in Figure 81 with two variations – one where links can be removed with no restriction and one where the removal of bridge links is prohibited to prevent graph partitioning. In the case where the network is partitioned the link removal process would lead to situations where nodes are eventually completely disconnected, in which case they are removed from the graph and the power-law fit is performed to the remaining topology. Link erosion continues until the number of links remaining equals the number of nodes less one, and in the case of the partitioning-allowed variant this nearly always will result in multiple isolated graphs, and in the partitioning-not-allowed case this results in a tree topology.

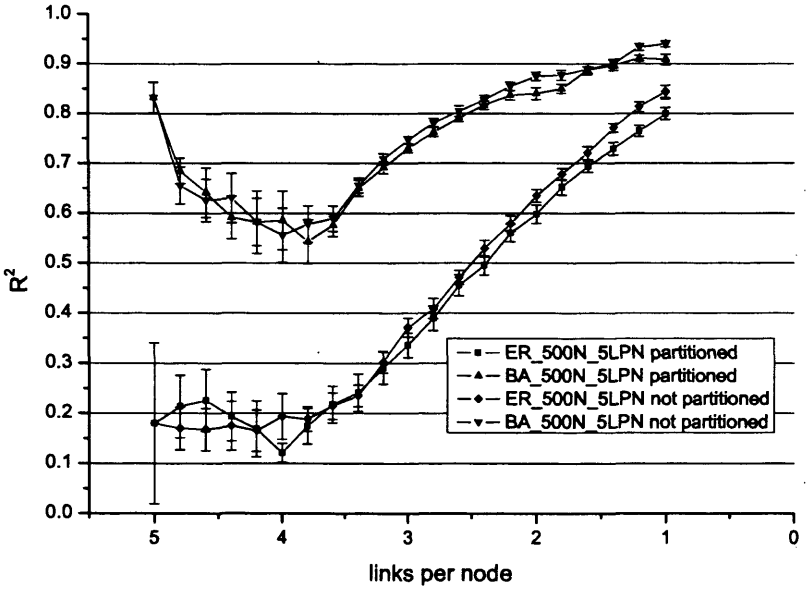


Figure 81 The R^2 correlation coefficient of power-law fit to node degree frequency distribution, as a function of the links-to-node ratio as {0,r,0,0} link erosion progresses. Values plotted for ER and BA inital graphs, with graph partitioning permitted and not permitted

The effect of {0,r,0,0} link removal can be seen in Figure 81. An interesting feature to note is that whatever the initial topology, and whether partitioning is allowed or not, the graph eventually tends to a node degree distribution power-law. A possible explanation for this is that we are performing a form of reverse preferential attachment (see section 2.3.2.2). For a link to be

removed it must first exist, and nodes which have a high degree are more prone to losing links, a "rich get poorer" scenario. The probability of decreasing in degree is, therefore, proportional to the current degree. As the degree decreases, so does the probability of decreasing further. The emergent result is that the degree distribution is approximately power-law compliant (R^2 after erosion is greater than 0.8 for all initial networks). Another interesting feature to note is that while the topology eventually tends toward power-law compliance it initially drops away from the BA power-law compliance R^2 value of around 0.85 towards a much less compliant network (R^2 is around 0.65 at its lowest point), so this erosion scheme cannot maintain an initial power-law.

Since even random link erosion leads to power-law traits this erosion scenario is used as a base line in the subsequent sections to allow comparison between the evolution of further link removal schemes – see section 5.5.3 for a specific comparison. It can also be seen that there is little difference if partitioning is allowed or disallowed, and, therefore, the rest of this chapter continues with partitioning not allowed – this makes good sense also in practical terms of delivered service, since it extends network reach.

The next link removal method introduces a notion of link cost, and tries to decrease network cost by assuming link cost is a function of Euclidean distance, and remove the absolutely longest links, {0,AE,0,0} – there is no roulette wheel selection and preference, just the longest link is removed at every cycle. To examine the effect of node placement cases where nodes are distributed uniformly randomly as well as being extracted from the BT network are considered individually.

The evolution of the network as links are removed can be seen in Figure 82. Like the {0,r,0,0} earlier the graphs eventually tend to power-law topologies, but it is also significant that the networks with nodes placed like the BT network tend to have higher correlation coefficients.

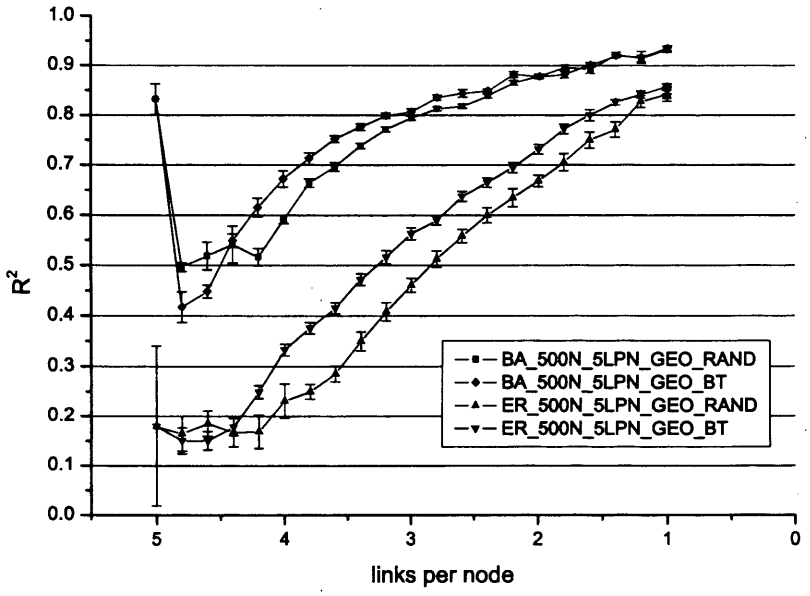


Figure 82 The R^2 correlation coefficient plotted as a function of the links-to-node ratio as {0,AE,0,0} link erosion progresses. Values plotted for ER and BA initial graphs connected with random node placements, and node placements extracted from the SDH network.

It should be noted that the allocation of location to topology in this and the other experiments is random, that is, there is no deliberate linking of degree, centrality or any other aspect of a node to its geographic location. This is because there is no obvious heuristic that could map topology to geography. In the uniformly random cases (BA_500N_5LPN_GEO_RAND and ER_500N_5LPN_GEO_RAND) locations were assigned uniformly at random, while in the BT geography cases (BA_500N_5LPN_GEO_BT and ER_500N_5LPN_GEO_BT) each node in the initial topology was assigned a location from the list of BT nodes, selected uniformly at random. An interesting feature to note here is that for an initial BA topology there is a very large drop in correlation coefficient very early on, while for the ER model there is no such drop – this could be caused by the mismatching of topology characteristics to geographic characteristics. Although the {0,r,0,0} erosion also had an early drop for the BA topology case it was not as large.

5.5.2 Multi-layer Network Erosion

In this section the effect of network load is examined as links are removed causing demands to be rerouted, which influences the future link removal landscape. The simplest form of multi-layer erosion is the removal of links based on the number of routes that traverse them, or rpl (routes-

per-link). This is a valid scenario in network evolution as it is simply the removal of the least used links to minimise cost (assuming the cost of scaling existing links to higher capacities is not prohibitive and that there is no architectural reason not to erode the network, such as the use of backup paths and diverse routing (see section 4.2.1.3.1)).

In this experiment {0,arpl,0,0} was examined which, at every cycle, selected the link with the lowest number of routes traversing it and removed it as long as it is not a bridge link.

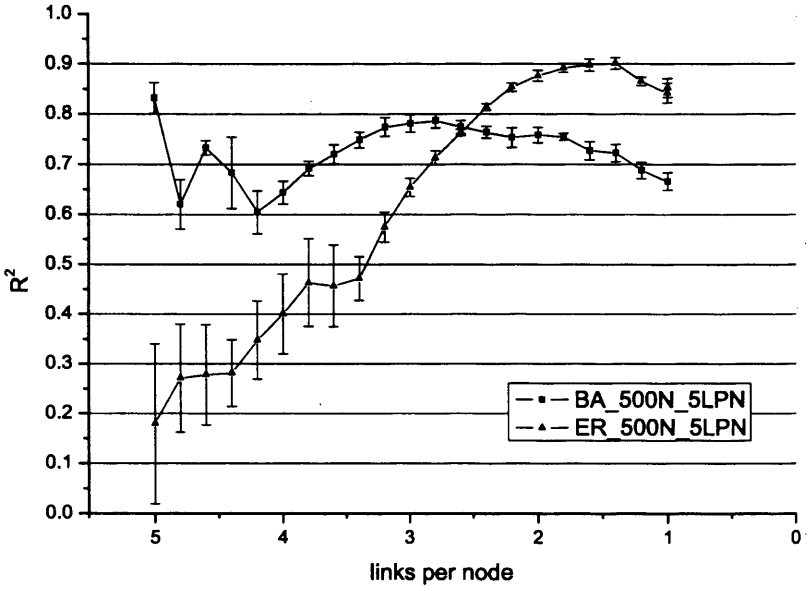


Figure 83 The R^2 correlation coefficient of power-law fit to node degree frequency distribution, as a function of the links-to-node ratio as {0,arpl,0,0} link erosion progresses. Values plotted for ER and BA initial graphs.

The progress of erosion can be seen in Figure 83, where the ER initial topology shows great improvement in correlation, and eventually even passes the BA topology peaking with an R^2 of 0.9, while the BA initial topology on the other hand shows an initial drop, then some improvement, and then a gradual decay in correlation. Whereas the later improvement is partly most likely an artefact of the erosion process, its improvement normalised to the {0,r,0,0} erosion can be seen in Figure 86.

A second link cost metric that could be considered is the RDP or Routes-per-link-Distance-Product (the number of routes traversing a link multiplied by the Euclidean length of that link) – approximately akin to the bandwidth-distance product in section 1.3. The experiment removed the

most expensive (highest RDP) links and the load should then be rerouted over the remaining links, and this feedback changes the cost landscape for the next removal cycle, which continues until a tree topology is reached.

In Figure 84 the evolution of {0,ARDP,0,0} can be seen where the link with the highest RDP is removed at every cycle. The traffic is then rerouted to the remaining links, affecting their RPL and, therefore, their RDP.

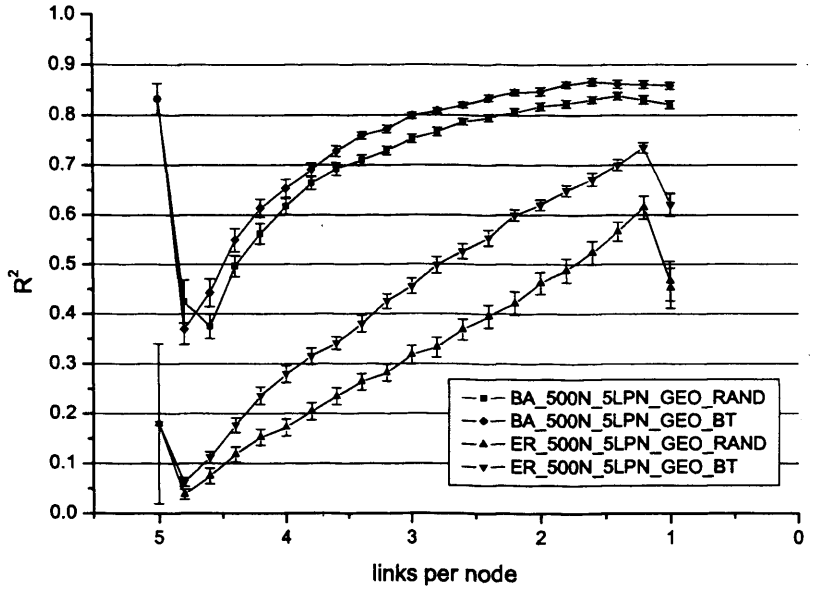


Figure 84 The R^2 correlation coefficient of power-law fit of node degree distribution, as a function of the links-to-node ratio as {0,ARDP,0,0} link erosion progresses. Values plotted for ER and BA initial graphs.

The same sudden initial drop in correlation is present in the BA traces, similar to the BA trace for {0,AE,0,0}, but there is now also a sudden drop for the ER traces as well, which was not until now present in any of the other erosion experiments. The drop right at the end of the erosion of the ER initial topologies is also a notable feature.

5.5.3 Comparison of single and multi-layer erosion

When considering the {0,r,0,0}, {0,AE,0,0}, {0,arpl,0,0} and {0,ARDP,0,0} erosion scenarios it becomes clear from the plots that where the topology is initially a BA graph no link removal technique maintains the initial power-law from the BA topology, whether with or without feedback. It can see that they all drop significantly during the first 100 links removed (the transition from 5.0 links-per-node to 4.8 links-per-node), although most of them rebound to

higher correlation coefficient values later. The decrease in the number of available links to remove is also, however, an issue, as we saw in the partitioning allowed and partitioning disallowed {0,r,0,0} plots which showed the same rebound. The rebound in correlation for {0,r,0,0} has already been theorised in section 5.5.1 and this section will compare the various erosion scenarios to the uniformly random removal case to try to remove artefacts caused by the general loss of links. The improvement can be found as the normalised difference to the mean value of the correlation coefficient for {0,r,0,0} at the given links-per-node that is, for, say, the removal strategy arpl, the % difference at the given lpn to {0,r,0,0} is:

$$\%difference(arpl,r) = 100.\frac{\langle R^2(arpl,lpn)\rangle - \langle R^2(r,lpn)\rangle}{\langle R^2(r,lpn)\rangle};$$

Figure 85 show the percentage differences in mean correlation coefficient of {0,AE,0,0}, {0,arpl,0,0} and {0,ARDP,0,0} to {0,r,0,0} for an initially BA topology.

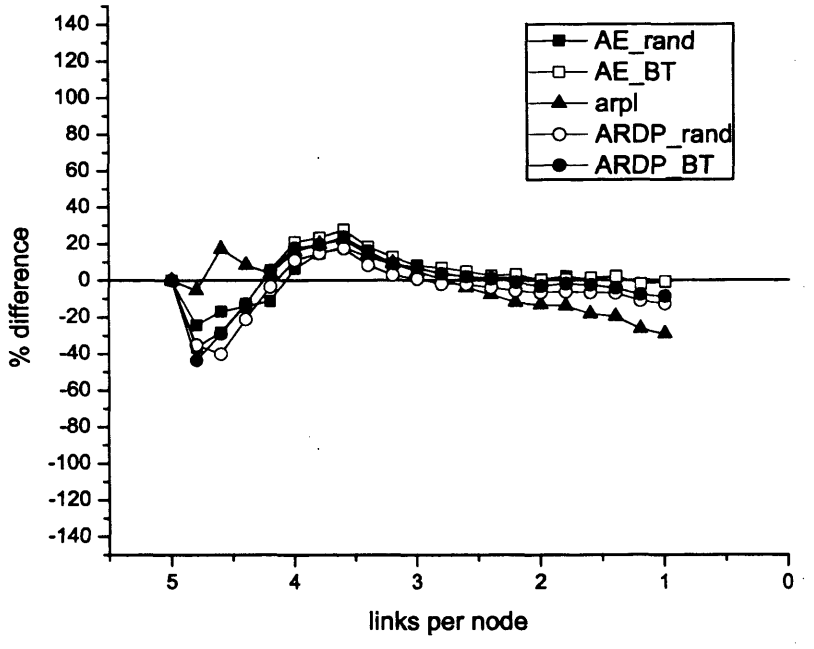


Figure 85 The percentage improvement in the correlation coefficient of power-law fit to degree frequency distribution for various erosion scenarios compared to the {0,r,0,0} erosion, for an initially BA topology

From the results it can see that all of the link erosion methods perform initially worse than random, for links-per-node values of between 5.0 and 4.8, especially those based on geography, and then they all rebound after links-per-node value of 4.0, and with a decreasing choice start to achieve lower correlation coefficients than the random removal. The arpl case, however, performs better for the first half of the erosion process, and then performs the worst of the scenarios considered.

This is in stark contrast to the case when the initial topology is very far from a power-law topology, as can be seen in Figure 86 for the scenarios where the initial topology is an ER graph.

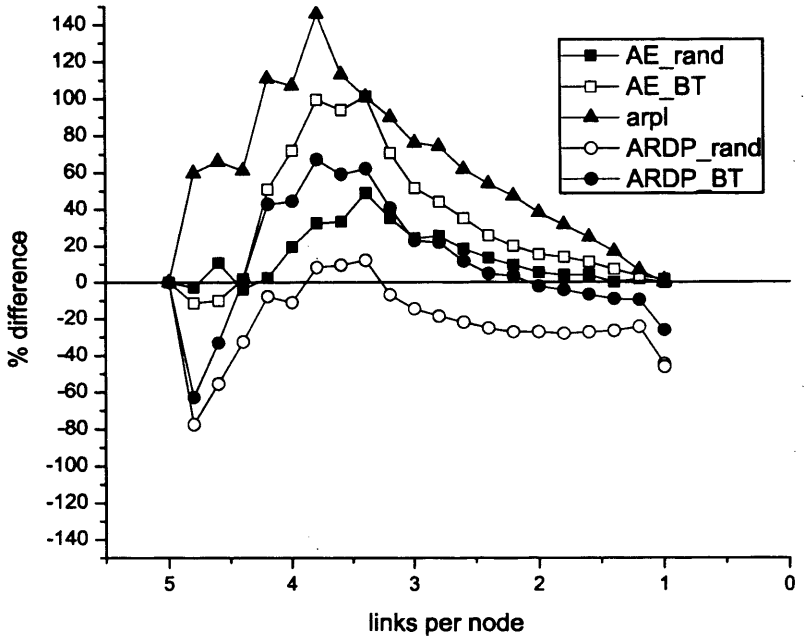


Figure 86 The percentage improvement in the correlation coefficient of power-law fit to degree frequency distribution for various erosion scenarios compared to the {0,r,0,0} erosion, for an initially ER topology

In this case link erosion based strongly on feedback, {0,arpl,0,0} shows the greatest improvement through the entire erosion process and never has a correlation coefficient lower than the {0,r,0,0} case. Another well performing link erosion scheme is {0,AE,0,0} based on the BT geographical distribution – showing significantly higher coefficients throughout most of the erosion. The other erosion scheme based on BT's geographic node distribution {0,ARDP,0,0} also shows a larger

increase in correlation than random (between 4.0 links per node and 2.0), while the schemes {0,AE,0,0} and {0,ARDP,0,0} but based on random node placement show the lowest improvement. This would suggest that there may an intrinsic push towards power-law networks from link load feedback and the presence of a self-similar network geography.

5.6 Network Rewiring Modelling

Bringing link erosion and link growth together this section examines the effect on topology of rewiring based on heterogeneous link and node selection schemes, that is a link may be removed based on, say, geography, because long links are expensive, and reconnected between two nodes based on the load on those nodes, because additional capacity may be required between them. Various removal and addition schemes are simulated to examine the overall trend in the evolution of the network topology. The combinations of selection schemes are based on the earlier examination of network growth and erosion.

5.6.1 Single layer rewiring

As a base line for comparison the effect of random rewiring is first examined. The progress of a {0,r,r,r} rewiring can be seen in Figure 87 (that is, select a link uniformly at random and remove it, and select uniformly at random two nodes and connect them) and it can be seen that for ER and BA random topologies the trend is for the R^2 to drop and the network tends away from a power-law in terms of degree frequency distribution. Also in the figure {0,r,D,D} is plotted (remove link uniformly at random and select two nodes preferring high degree nodes and connect them).

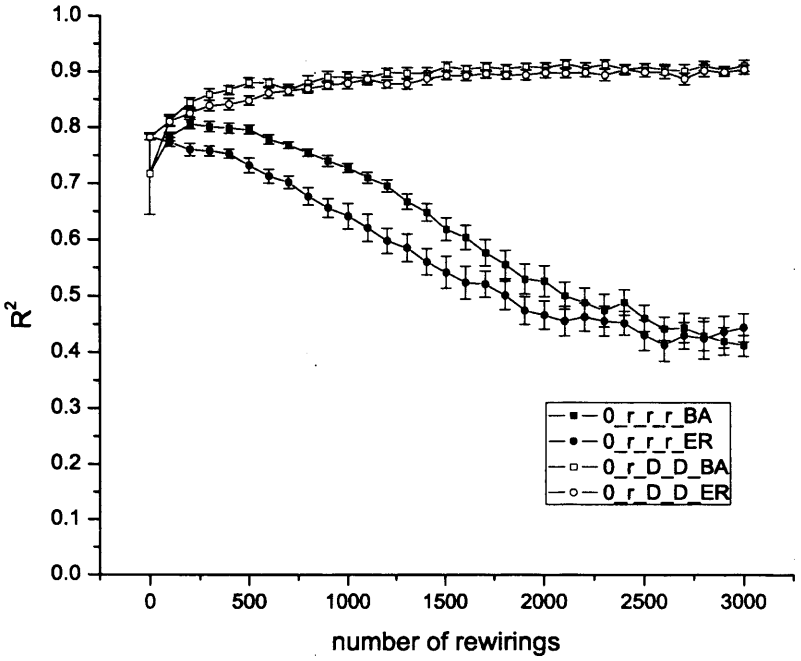


Figure 87 {0,r,r,r} and {0,r,D,D} rewirings on initial ER and BA graphs.

The {0,r,D,D} shows an interesting trend in that it tends toward a good approximation to a power-law fit with an R^2 value of around 0.9. The values plotted are the average of rewirings on a set of initial topologies and with different random seeds, but in Figure 88 we can see the degree distribution plot of a single such instance to confirm that the topology is, in fact, close to a power-law. For comparison, if the ER and BA initial topologies of 500 nodes are taken and grown only in terms of number of links by an additional 3000 links, with the endpoints both selected preferring high degree, but *without* the link removal (i.e. {0,0,D,D}) then the network evolves like in Figure 89 and there is a significant downward trend in R^2 value for degree frequency distribution power-law fit. This suggests that link addition based on degree alone is not the cause of the power-law topology, but the combination of random removal and degree based link addition is.

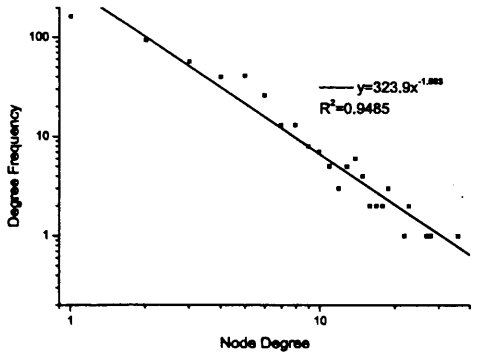


Figure 88 The degree frequency distribution plot for a single instance of a BA topology rewired 3000 times by {0,r,D,D}

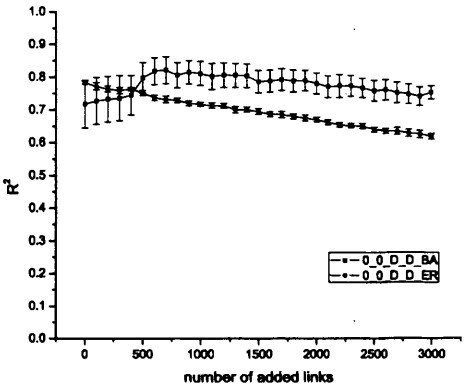


Figure 89 The progression of correlation coefficient to power-law fit of degree frequency distribution as 3000 links are added to 500 node, 1000 link, ER and BA seed topologies with {0,0,D,D}

The hypothesis behind the conformance of {0,r,D,D} but not {0,0,D,D} is that {0,0,D,D} will continue to build a topology until all nodes are connected, but before then it will promote high degree nodes to even higher degree nodes until they are connected to all nodes. Throughout this process it is unlikely that low degree nodes are promoted and the left side of the degree frequency distribution remains, while the right side is stretched further to the right.

The inclusion of a removal stage in {0,r,D,D} which selects a random link maintains the link to node ratio of 2 and has a demoting effect on degree – higher degree nodes are more likely to be affected by the demotion as they have more links, and this maintains equilibrium in terms of degree distribution.

The next scenario to consider is the effect of geography and node placement on rewiring scenarios. In Figure 90 the progress of rewiring can be seen as the absolute longest link (that is the link with the largest Euclidean distance between its two endpoints) is removed at the start of every cycle. A link is then added either by selecting endpoints uniformly at random {0,AE,r,r} or by selecting high degree nodes {0,AE,D,D}. As with {0,r,D,D} the power-law fit correlation coefficient improves for {0,AE,D,D} and settles at a similar value of approximately 0.9.

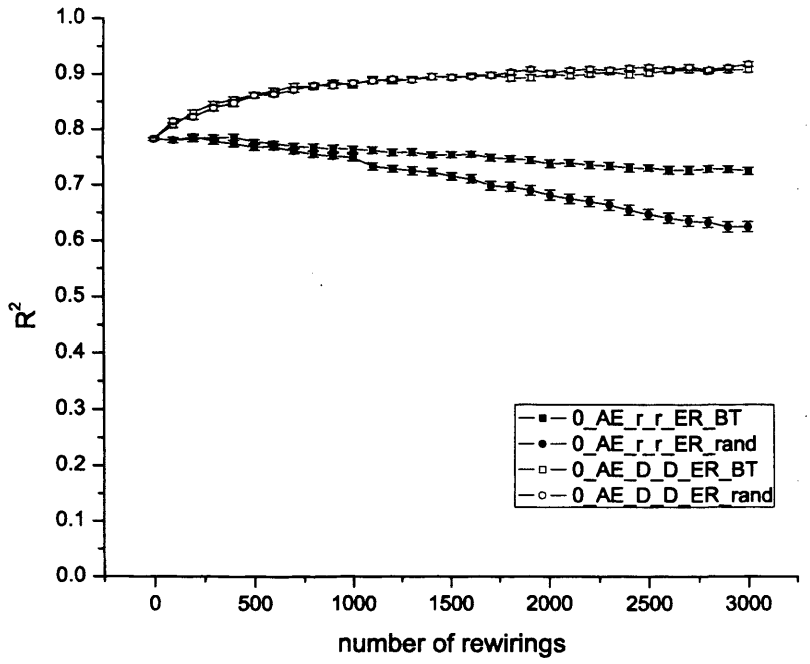


Figure 90 {0,AE,r,r} and {0,AE,D,D} rewirings on intial ER and BA graphs to examine the potential effect of geography on rewiring.

There are two theories why this happens: firstly, the cause could be the same as that hypothesised for {0,r,D,D} since the link removal scheme (AE) is not based on degree and, therefore, more likely to affect the high degree nodes by the nature that they have many links; secondly there may be a similar phenomenon as that postulated for the FKP model (see section 2.3.2.10) where there is a competition between shrinking the network and expanding it – this is, however, unlikely since in {0,AE,D,D} the shrinkage is based on geography and the growth is not based on anything related to centrality (as used in the FKP) since there is no correlation between geography and topological node properties.

5.6.2 Multi-Layer: Load based rewiring

The next series of scenarios examine the multi-layer aspect of networks through the use of traffic load.

5.6.2.1 Random link removal and load based reconnection

In Figure 91 {0,r,AR,AR} and {0,r,TR,TR} are plotted as before for initially ER and BA topologies. Links are removed uniformly at random and added between nodes selected based on either total load, TR, (which under the assumption of a full mesh of single-unit demands is equivalent to a betweeness centrality measure) and the average load, AR, on a node (which because of the division by the degree is a function of the reciprocal of node degree).

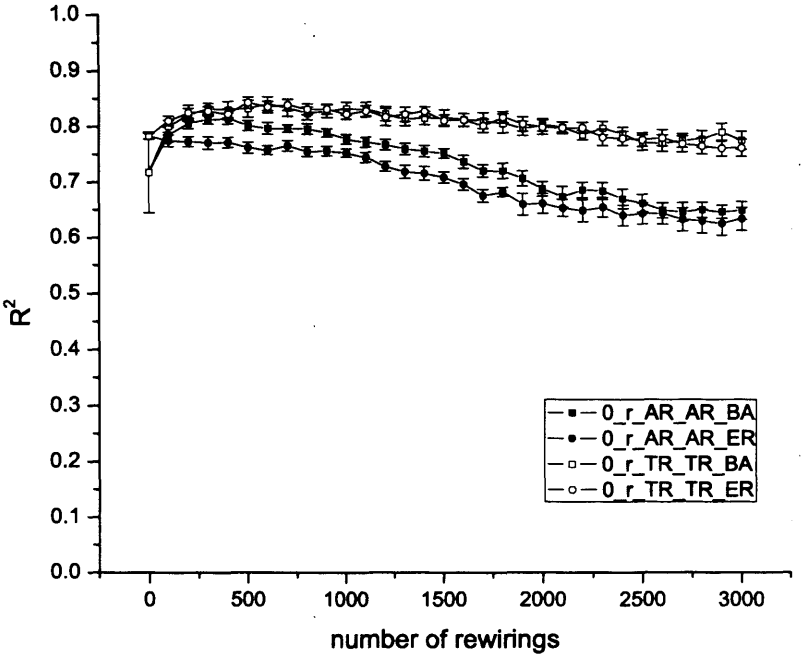


Figure 91 {0,r,AR,AR} and {0,r,TR,TR} rewirings - both multi-layer scenarios as they consider the traffic load traversing nodes

As rewiring progresses the TR scheme tends to, by definition, increase connectivity in the topological core (where betweenness centrality is highest) creating a highly connected core but leaving the edge nodes with less connectivity. The increase in connectivity is, however, offset by the increased probability that the random link selection will affect a high degree node – as mentioned earlier TR node selection has a tendency to super-promote certain nodes in terms of degree. This can be seen more closely in Figure 92 which is the degree frequency distribution plot of the single instance (which happened to be initialised with a BA topology) of {0,r,TR,TR} which had the lowest R^2 by the 3000 rewiring cycle. The plot has a small number of nodes with

higher than power-law compliant degree – the plot has a correlation coefficient of 0.666, but without the three largest nodes this becomes 0.861.

The {0,r,AR,AR} scenario in Figure 91 also has a decreasing correlation coefficient – the degree frequency distribution of the single instance with the lowest correlation coefficient can be seen in Figure 93 and the lower incidence of the highest and lowest degrees could be explained by the way the weight of a nodes in the AR scheme is calculated – it is the total number of shortest path routes on links connected to a node, as TR, but divided by the degree (number of links), and as such, low degree nodes (especially degree-of-1 nodes) have a high weight because they have a low denominator and a high chance of being promoted to a higher degree, and the high degree nodes have a lower weight also because of the high denominator and they are so highly connected that any increase in degree does not result in a significant increase in number of routes traversing the node.

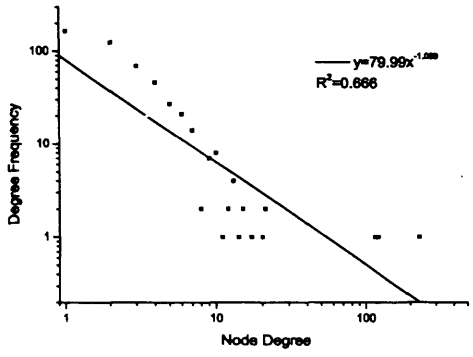


Figure 92 The node degree frequency distribution after 3000 cycles of {0,r,TR,TR} rewiring

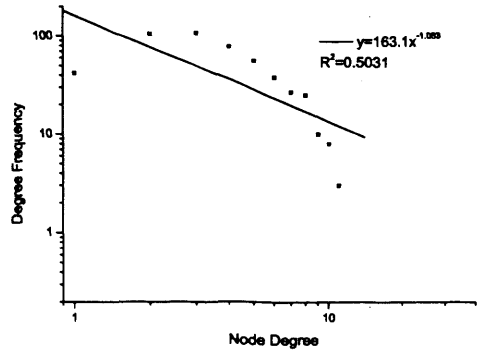


Figure 93 The node degree frequency distribution after 3000 cycles of {0,r,AR,AR} rewiring

5.6.2.2 Load versus Load rewiring

To further involve load in the rewiring the link removal method is now changed to one that considers load, specifically ARPL, that is, select the single highest RPL (routes per link) link in the topology. The concept behind using ARPL as the removal scheme and then load based endpoint selection schemes (AR and TR) is that this is an attempt to rewire the most loaded (and expensive) links in an attempt to make traffic load balance across the rest of the network. The reconnection is between nodes with high load (either measured directly (TR) or as an average load in the vicinity of the node (i.e. average link load), AR). As stated in Appendix A the link that

is removed cannot be immediately replaced in the link addition phase of the cycle. Also note that AR and TR do not forecast traffic load and, therefore, do not select endpoints based on how much traffic will flow along the new link, but based on how loaded the nodes are, and that some of the traffic may choose the new link for transit, and the new link may also attract new traffic to the endpoints as it now creates a new shortest path for other traffic flows.

The progress of correlation coefficient of the power-law fit to degree frequency distribution can be seen in Figure 94 where {0,ARPL,AR,AR} drops on early rewiring and remains at values below 0.65 for either initial topology type. {0,ARPL,TR,TR} however, after an initial drop rises and then continues to rise albeit at a much slower rate, and by 3000 rewirings reaches a correlation coefficient of just above 0.8.

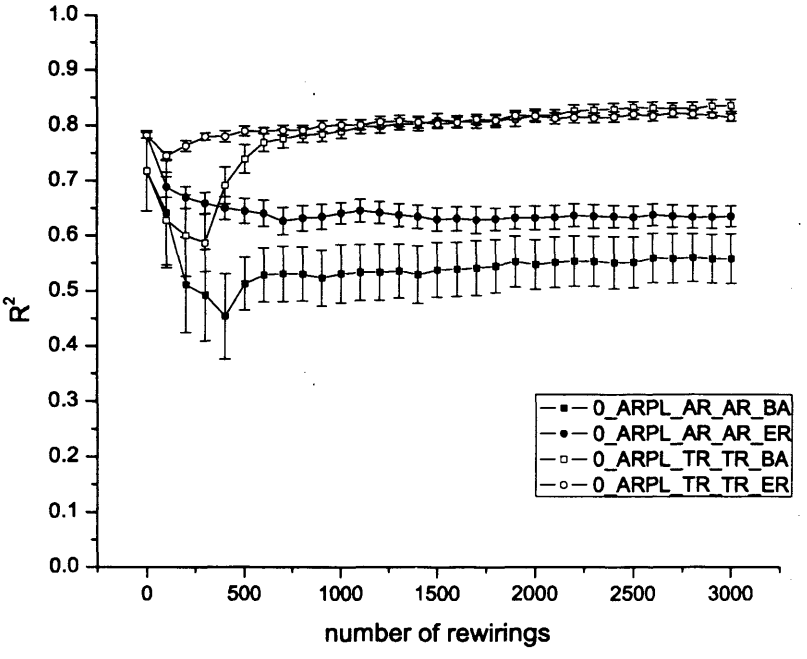


Figure 94 {0,ARPL,AR,AR} and {0,ARPL,TR,TR} rewiring scenarios for ER and BA initial topology types

While the {0,ARPL,TR,TR} shows a positive trend in R^2 it is still some way from other scenarios which achieve values of approximately 0.9. In Figure 95 there is a plot of the instance of {0,ARPL,TR,TR} what has the highest R^2 value of all experiment instances. It can be seen that for much of the distribution the points are well placed for a power law and it is mainly the one degree nodes that are bringing down the correlation coefficient. Without the one degree nodes the correlation coefficient moves from 0.8753 to 0.9405. If the target for a good fit were considered

to be a value greater than 0.9 then it can be shown through a manual target search that the one degree node frequency would have to reach 72 for this to happen. In Figure 96 there is a plot of the number of nodes with a degree of 1 as the rewiring progresses. The plot shows that the general trend for both initial topology types is an increase in frequency of one degree nodes – it can be extrapolated from the graph that it would require approximately 18000 rewiring cycles before the degree-of-one frequency would reach 72.

While this method of extrapolation and assuming continued development of degree-of-1 nodes may be a little far reaching it should also be noted that the reason behind the low incidence of degree-of-1 nodes in the first place is because of the initial topology creation methods. The BA initial topologies are created as {D,0,g,D} and step:1,1,1 and using incremental growth so at every cycle the growth node has two links connected to it – and, therefore, no node has a degree of less than 2. In the ER case the initial topologies are created with {r,0,r,r}, but using incremental growth to a topology size of 500 and –step:1,1,1,1,500,1 and –cycles:1000 – that is grow the network by connected the new node to a random existing node (thereby ensuring full connectivity) and after cycle 500 add another 500 links, one per cycle, with endpoints selected uniformly at random – on which case degree-of-1 nodes are possible but unlikely.

It should also be noted that the biggest increase in R^2 in Figure 94 for the BA initial topology undergoing {0,ARPL,TR,TR} rewiring coincides, at around cycle 500, with the significant boost in the number of degree-of-1 nodes in Figure 96.

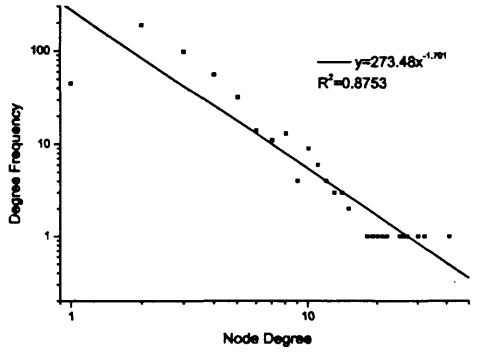


Figure 95 The degree frequency distribution of a single instance after 3000 cycles of {0,ARPL,TR,TR} rewiring starting with a BA topology instance that has the highest R^2 after the 3000 cycles of rewiring

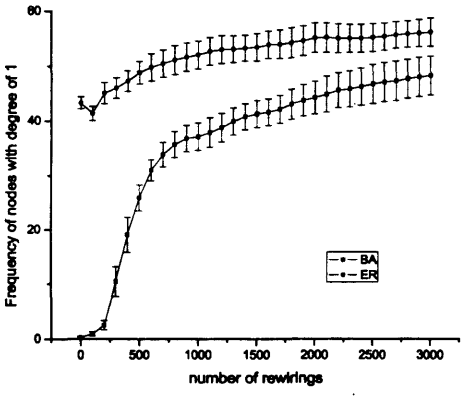


Figure 96 The number of nodes with a degree of 1 as the {0,ARPL,TR,TR} rewiring progresses

In the next scenario, rather than rewiring the most loaded link, the least loaded link is rewired, using the "arpl" link removal selection scheme. As before AR and TR are used to select the endpoints for the location of the rewired link. The progression can be seen in Figure 97. To examine the effect of using a deterministic link selection scheme, i.e. arpl, the results of a similar scenario, but using rpl (i.e. select link preferring links with higher routes-per-link values) can be seen in Figure 98. Such scenarios are an attempt to optimise link usage – where link load is low and the link is not particularly necessary (remember that a bridge link can never be removed) then it should be removed and placed between nodes with high load. The figures show that for either initial topology, after an initial improvement, the networks gradually tend away from power-law compliant topologies, but in the case {0,arpl,TR,TR} this is significantly faster, until a state is reached around cycle 1750 where R^2 is around 0.6, that causes the network to rebound to the earlier peak level and then decline again. The rpl cases do not do this, potentially because it has not reached that state yet.

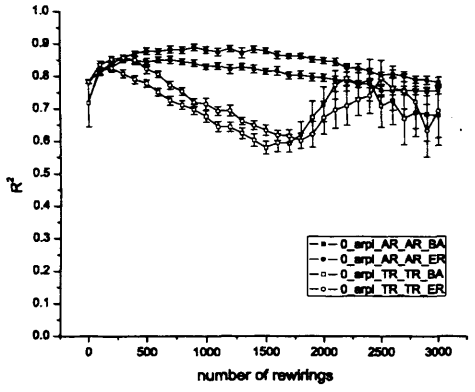


Figure 97 {0,arpl,AR,AR} and {0,arpl,TR,TR} rewiring scenarios for ER and BA initial topology types

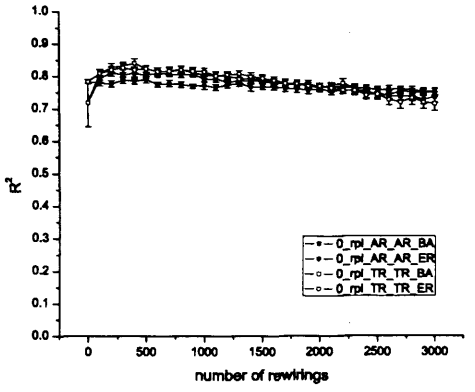


Figure 98 {0,rpl,AR,AR} and {0,rpl,TR,TR} rewiring scenarios for ER and BA initial topology types

Continuing the examination of deterministic decision schemes Figure 99 and Figure 100 demonstrate the effect of using absolute maximum selection for the link addition process through the use of MAXAR and MAXTR. The networks have a tendency to quickly settle with the {0,arpl,MAXAR,MAXAR} reaching a plateau within 300 rewirings, whereas {0,rpl,MAXAR,MAXAR} took until cycle 2000, although settled at a topology with a very different R^2. A rebounding phenomena similar to {0,arpl,TR,TR} can also be seen in {0,rpl,MAXAR,MAXAR}, but since the schemes have no selection processes in common this

would suggest that the phenomena is caused by a more basic event such as the removal of a bottleneck link. Correlation coefficient decay is even more extreme in the MAXTR scenarios with the R^2 dropping to levels around 0.15 which are the lowest ones seen in any of the scenarios in this section.

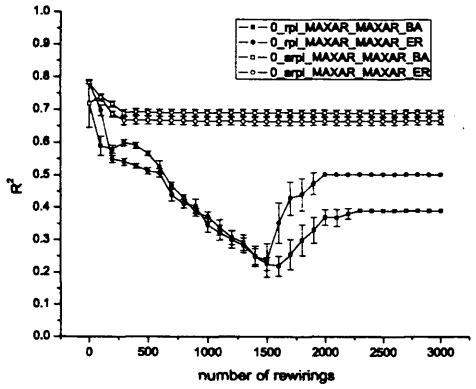


Figure 99 {0,rpl,MAXAR,MAXAR} and {0,arpl,MAXAR,MAXAR} rewiring scenarios for ER and BA initial topology types

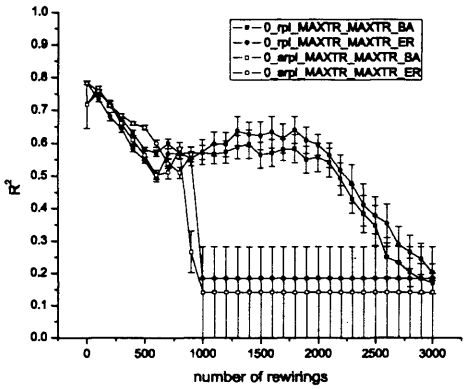


Figure 100 {0,rpl,MAXTR,MAXTR} and {0,arpl,MAXTR,MAXTR} rewiring scenarios for ER and BA initial topology types

One expected phenomenon that can be seen in {0,arpl,MAXTR,MAXTR} and the earlier {0,arpl,MAXAR,MAXAR} is that the use of deterministic selection schemes quickly leads to an attractor in phase space, as discussed in section 5.3.3, and stays there (since there is no stochastic element that may knock the trajectory into another basin).

5.6.3 Multi-Layer: Geography and Load based rewiring

This section introduces the concept of geography to the topology rewiring schemes, as well as heterogeneity, by considering geography and load, but initially in separate processes.

5.6.3.1 Geography versus Load rewiring

Figure 101 starts by showing the development of {0,AE,AR,AR} under the ER and BA initial topology types and with uniformly random and random-subset-of-BT node placement configurations. The general trend in R^2 is to drop, which is what also happened in {0,r,AR,AR}, the parity of response is similar to that of {0,r,D,D} and {0,AE,D,D}, albeit the latter is a positive trend. In Figure 102, however, where the progression of {0,AE,TR,TR} can be seen the use of TR in link addition has a slightly different effect with the general trend being an increase in R^2, albeit a very slow one.

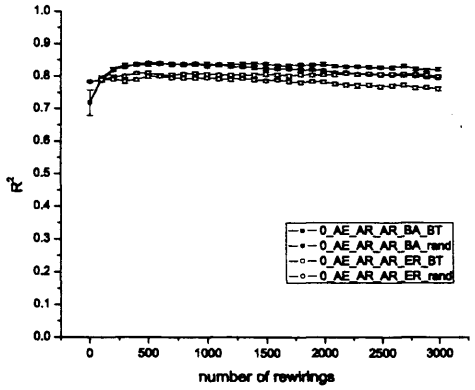


Figure 101 The correlation coefficient of a power-law fit to the degree frequency distribution of {0,AE,AR,AR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

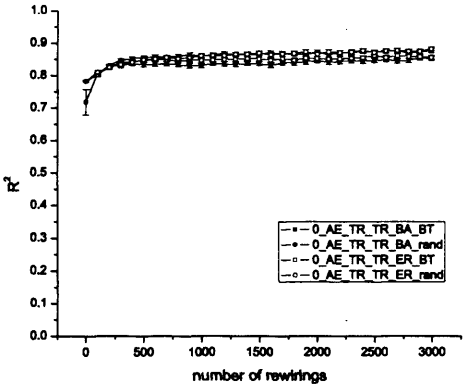


Figure 102 The correlation coefficient of a power-law fit to the degree frequency distribution of {0,AE,TR,TR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

To examine more closely the effect of {0,AE,TR,TR}, we can see the degree frequency distribution of two instance of the rewiring, with the lowest R^2 instance of 0.7815 in Figure 103 (where the initial topology was a BA graph) and the highest R^2 instance of 0.9446 in Figure 104 (where the initial topology was an ER graph).

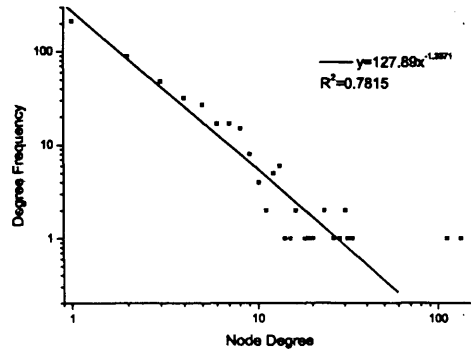


Figure 103 The degree frequency distribution for the instance of {0,AE,TR,TR} rewiring with the lowest R^2 correlation coefficient

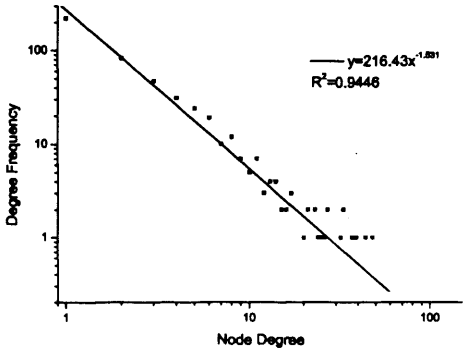


Figure 104 The degree frequency distribution for the instance of {0,AE,TR,TR} rewiring with the highest R^2 correlation coefficient

In this case, unlike the earlier {0,ARPL,TR,TR} case, the deviation from power-law conformance in the lowest R^2 instance is caused by only two nodes which both have very high degrees and skew the graph, and not a lack of degree-of-1 nodes. If these high degree points were removed from the fit in Figure 103 then the R^2 would change from 0.7815 to 0.8960. Figure 104 is an example of these nodes being demoted as part of the rewiring process.

It should also be noted that since the AE removal scheme is a deterministic process it is most often dominated be a single long link and the other links are at a low risk from being removed, and this could be the cause of the slowness of the improvement in {0,AE,TR,TR}.

Figure 104 does, however, show that {0,AE,TR,TR} is very capable of evolving topologies toward power-law compliant topologies. The reason behind this is most likely that a power-law topology is the equilibrium state between two opposing forces, one governed by the need to decrease link distances, causing demands to be rewired, and the other a response to that rewiring and the creation of distance-agnostic links in the network, most of which will inevitably be longer than the shortest links. To strengthen this hypothesis it can be seen in Figure 91 that if the influence of geography is removed and the link removal is purely random then the {0,r,TR,TR} rewiring has a downward trend in terms of degree frequency distribution power-law correlation coefficient. Further, the existence of AE rewiring alone, without TR, in {0,AE,r,r}, as was seen in Figure 90, had no suggestion of a power-law trend, while pure network growth by TR, {TR,0,0,0} and {TR,0,g,TR}, in Figure 76 also had no tendency toward power-law compliance. Therefore, of the contributing processes, when taken individually, none had a tendency toward power-law compliance, but together are capable of very good power-law conformance as in Figure 104. For comparison to the single layer case, {0,r,D,D} and {0,AE,D,D} both tended to power law, but when considering the "rich get richer" phenomenon of "D" and "TR" node selection there is little in common between the two in actual application because of the extreme nature of TR's "get richer" effect (Figure 77), and {TR,0,0,0} and {0,r,TR,TR} showed no significant power-law tendency – the tendency for power-law compliance in {0,AE,TR,TR} would suggest that the response to the high strength (relative to "D") of the TR "get richer" effect is actually stronger and balances out – and this would suggest the response is somehow an adaptation to the different rewiring scheme.

To continue the examination of the power-law compliance the degree rank and clustering coefficient rank power-law fit correlation coefficient development as {0,AE,TR,TR} rewiring progresses can be seen in Figure 105 and Figure 106 respectively.

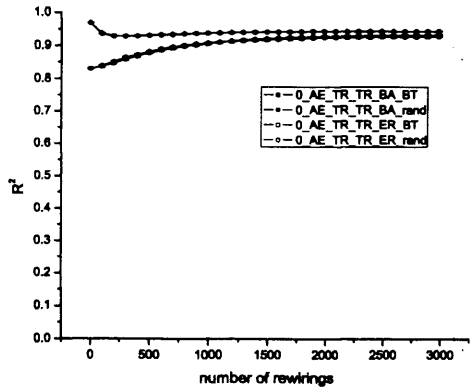


Figure 105 The correlation coefficient of a power-law fit to the degree rank of {0,AE,TR,TR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

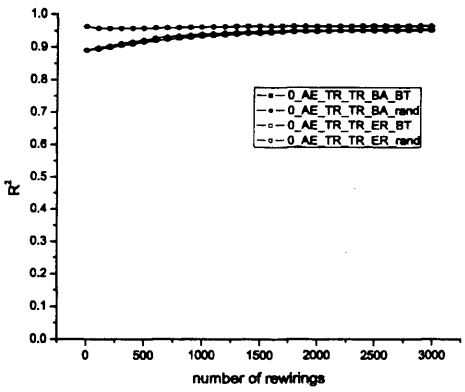


Figure 106 The correlation coefficient of a power-law fit to the clustering coefficient rank of {0,AE,TR,TR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

Both in terms of clustering coefficient rank and degree rank the {0,AE,TR,TR} scenario showed a good fit to a power-law, independent of initial topology and underlying geographical node placement, although after 3000 cycles the values of correlation coefficient of clustering rank and degree rank are lower than those of the initial BA topologies.

5.6.3.2 Load-Geography hybrid versus Load rewiring

Another approach to incorporating geography and load is to join them in a single metric, rather than have them as opposing strategies as in {0,AE,TR,TR}. While AE was based on the fact that network cost optimisation is attempted based on removing the links with the longest Euclidean distance, it does not take into consideration any concept of link capacity. To this end the idea of bandwidth-distance product is used in the form of RPL-distance product, that is, the "capacity", which will be considered to be the number of routes on a link, multiplied by the Euclidean length of the link. The weight of selecting a link for removal now becomes the RPL-distance product of a given link divided by the sum of RPL-distance values for all links, as in Equation 23.

$$P(i) = \frac{RPL_i.d_i}{\sum_j^L RPL_j.d_j};$$

Equation 23 The probability that a link i of L links is removed in the RDP removal scheme

The progression of degree frequency distribution during {0,ARDP,AR,AR} and {0,ARDP,TR,TR} rewiring can be seen in Figure 107 and Figure 108 respectively. The ARDP link selection process is simply the deterministic selection of the link with the highest RDP value, rather than a stochastic preferential selection method.

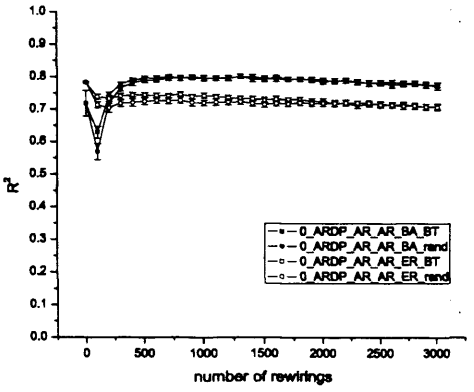


Figure 107 The correlation coefficient of a power-law fit to the degree frequency distribution of {0,ARDP,AR,AR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

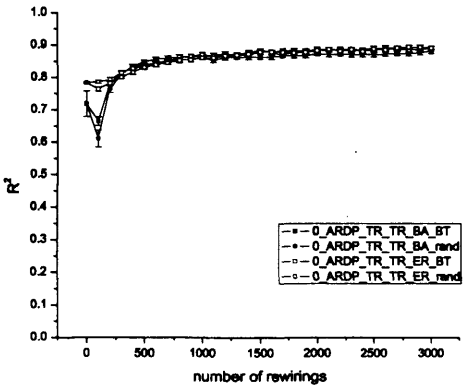


Figure 108 The correlation coefficient of a power-law fit to the degree frequency distribution of {0,ARDP,TR,TR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

As with the AE link selection scenarios the scenarios with AR node selection tend to move away from power-law compliance, and the TR based scenarios tend to better power-law compliance. The early drop in correlation coefficient that was also present in the ARDP erosion in Figure 84 appears here as well, as is mostly likely caused by some kind of mismatch in initial topology to node placement allocation. The progress of degree rank and clustering coefficient rank for {0,ARDP,TR,TR} can be seen in Figure 109 and Figure 110 respectively.

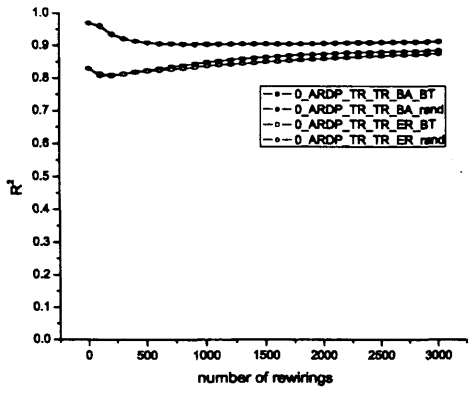


Figure 109 The correlation coefficient of a power-law fit to the degree rank of {0,ARDP,TR,TR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distribution

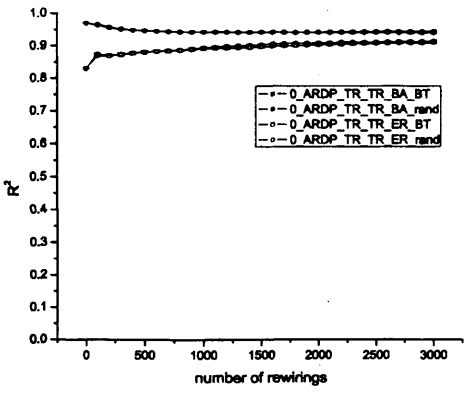


Figure 110 The correlation coefficient of a power-law fit to the clustering coefficient rank of {0,ARDP,TR,TR} rewiring scenarios for ER and BA initial topology types and uniformly random and BT node distributions

As before, good conformance is seen throughout, which may suggest that the exact form of link cost is not so important, just as geographic node distribution is not important. Rewiring with RDP based removal does still perform differently from the AE based rewiring, since the early drop appears, but the final result is nevertheless similar.

5.7 Comparison of Single and Multi-Layer modelling methods

This chapter has thus far presented results for a number of scenarios, some single-layer, being random, based on geography, or relying only on the local properties of a node, like degree, and

some multi-layer, being based on link or node load, and optionally a function of geography as well, in the case of RDP.

Network growth based on preferential attachment in the form of degree has shown, power-law compliance (which was previously known since this is the BA model), but preferential attachment in terms of load or betweenness centrality, TR, did not achieve power-law compliance and it was shown that this was due to the rich-getting-richer effect, but getting too rich too quickly. The result that {AR,0,0,0} resulted in power-law compliance is significant and unexpected, because it is a function of the inverse of degree and, therefore, not caused by a richer-get-richer effect, at least not based on the single-layer degree metric, but could only be caused by the traffic load that was being rerouted – or put another way the multi-layer feedback. The case of multi-layer feedback is strengthened by the fact that {AR,0,g,AR} performed even better than {AR,0,0,0}, and the only difference was the existence of alternative paths and the ability to re-route traffic.

Network erosion was shown to have the inherent property of tending to a power-law network, even if links are selected purely at random. This was theorised to be a form of "rich-get richer" but in reverse. The comparison of single and multi-layer erosion scenarios in Figure 85 and Figure 86 has shown that the use of multi-layer scenarios has produced relatively more power-law compliance than erosion based on random removal. Erosion based on link load showed the most improvement, although when incorporated with geography, the results became less favourable (ARDP entries in Figure 86).

The rewiring scenarios showed that based on degree the single-layer degree-based rewiring can be effective in producing power-law compliance. More significantly, however, in the multi-layer case a number of scenarios based on TR, which in the network growth case were ineffective, suddenly became effective. The postulation was that the elements of feedback in rewiring have adapted to the high strength of TR node selection attraction. This was for link removal based purely on geography, but the use of a hybrid of load and geography, ARDP, also created similar power-law compliance.

5.8 Conclusion

This chapter has shown by example and counter-example, by decomposition and close analysis of specific cases that in certain cases network development based on metrics derived from adjacent layers can achieve power-law compliance in degree frequency distribution, degree rank and

clustering coefficient rank. The removal of specific parts of conforming multi-layer scenarios has resulted in the loss of conformance (e.g. {AR,0,0,0} vs. {d,0,0,0}), and the addition of multi-layer selection elements to otherwise non-conforming scenarios (e.g. {0,AE,r,r} vs. {0,AE,TR,TR}), has resulted in conformance, suggesting that multi-layer effects are seen in the simulated network development. Although single-layer scenarios were also shown to be capable of power-law compliance they were deduced to have a different source of the property. Achieving power-law compliance in terms of degree is unsurprising if node selection is itself based on degree, but this chapter has shown that conformance is also possible based on unrelated metrics, like the link load (either AR or TR depending on the network development type).

The discussion of deterministic versus stochastic link/node selection has shown that deterministic processes can have an overriding effect on network development, and as such the use of AE may be better replaced by a process which selects preferring the longer distance links, although AE has not shown any tendency to lock the progress of the network into any specific attractor.

Chapter 6
Conclusions

This thesis has examined network modelling from the perspective of multiple interacting layers, rather than the typical approach of examining layers in isolation. Chapter 1 introduced the concept of network layers and layer interaction, and proposed that there is a feedback system between adjacent layers, whose influences are expressed as demand and delivered service. Any layer has finite resources to support a demand, so the demand is then serviced to a given quality level (packet loss ratio, available bandwidth etc..), and the layer providing service will then either re-configure itself or the layer expressing the demand will change the characteristics of the demand. Chapter 1 showed in a simple analytical example how changes in demand influence the optimal network configuration and with a change in the cost function which is to be minimised it was shown that the optimal configuration changes again. The physical (geographic) relocation of nodes is, however, expensive to implement and most changes in network configuration are performed in terms of link reconfiguration, and this was the focus of the rest of the thesis.

Chapter 2 gave a review of current topology modelling techniques, for comparison with multi-layer modelling in Chapter 5, and to better understand the network layers Chapter 3 and Chapter 4 examined two layers to decompose their protocols and methods to extract features which influence their design, cost functions and possible dynamic behaviour, as well as performing a analysis of the topology to obtain a target for modelling.

In terms of examining IP topologies the conclusions that were drawn (partly from the quoted Faloutsos' results) is that IP networks at both the router and AS level tended towards topologies from which certain distributions of metrics had a good correlation with a power-law fit. The chapter also examined methods of topology measurement, as well as offered initial results from such a measurement. The chapter then explored the efficacy of the use of projections into arbitrary topologies and showed that the repeated traceroute method of measurement can skew the measured topology toward power-law traits, although this occurred mainly when there were only a few projections and it was soon possible to differentiate between ER and BA underlying graphs with only a few more projections. The examination of the protocols associated with IP and IP itself suggested that there is often a tendency to cluster and collect routers into areas (e.g. OSPF areas) or domains (e.g. AS domains) with the more commonly used routing protocols. A common intra-domain routing protocol, OSPF, in some default configurations, would push traffic onto high capacity, potentially backbone, links through its use of inverse capacity link weights.

Shortest path routing in IP (minimum hops in RIP, unity link weights usage in OSPF, AS PATH length based route selection in BGP) suggested that IP protocols tend to attempt to minimise resource use, and justified the use of all-pairs-shortest-path routing in MITIE. MITIE also supports inverse capacity link weight assignment as per OSPF. For added flexibility MITIE can evaluate the link capacity used in the link weight inversion calculation through a number of methods – either by assuming the link is 100% matched to demand and setting the link capacity as the actual load on the link, or by taking this load and rounding it up to the next higher capacity to give a certain granularity – in IP this granularity could scale as powers of 10, as per Ethernet (Table 8), or if carried by a packet-over-SDH/SONET system the capacity scales as a factor of 4 (Figure 33). This would allow for the investigation of the existence of Braesse's paradox in telecommunications networks (although it was also shown previously to exist in experiments performed with qBGPSim, a simulator derived from MITIE in [GRI07]). The power-law compliance discovered by Faloutsos' et al in this chapter serves, together with results from the next chapter, as a good target for modelling, and Chapter 5 used these power-laws as a measure of respresentativeness of models.

The SDH network analysis in Chapter 4 was a novel investigation into an SDH network and examined, not only topology, like the Chapter 3 did for IP and the Internet, but also network resource configuration in terms of node placement and capacity allocation, as well as examining the topology formed by the demand matrix (and hence the available topology for routing of the layer above). A number of traits discovered were introduced into MITIE, including node selection methods based on assortativity, Euclidean length and cost function. The Euclidean length distribution of links between SDH sites was found not to follow a commonly used geography based model, the Waxman model, with shorter links being more common than forecast by the Waxman model, instead a new power-based geography model was proposed and implemented in MITIE. An examination of the design target for the SDH network planner has shown it to be highly structured containing various topological components, most notably rings, and hierarchies. It was also found to have very careful planning to maintain guaranteed service levels, including route diversity for protection. Because of the high cost of operating transport networks and the long design times the chapter also examined the possible influence of legacy networks, which in a national network such as BT's would have a lasting effect on current and future planning. The result of the SDH topology analysis, which was unexpected in so far as the resulting traits are not explicitly designed into the network, is that the BT SDH network also follows the same power-laws found by Faloutsos in the Internet, and many more besides. The chapter also showed that the degree distribution power-law is very deeply ingrained in the network, as it is present at all

geographic scales from that of individual SDH sites, all the way up to the connectivity of cities. The node placement of the SDH sites was shown to follow a heavy tailed distribution.
Since the SDH network is so strictly planned and configured and there is no dynamic re-configuration or re-routing (other than pre-planned protection switching) the source of the power-law traits must be a more implicit force expressed somehow through the network design algorithms, such as the need to minimise resources, to maximise reach, to optimise cost, to adapt to adjacent layers, or a combination of these. The traits could not be caused by the topology measurement method as was a potential issue in Internet topology measurement, as the data set was a direct extraction from the network configuration database.

The source of these traits, and whether they are similar for IP and SDH networks can be debated, and commonalities and differences drawn with other power-law networks [ALB99] [JEO00] [BAR02] [LIL01] [ALB00b] [ABE99] [BAR99] but based on the examination of the IP and SDH technologies and example networks for IP and SDH at least, the best current hypothesis is that the traits exist as part of an optimisation of cost and utility. This was seen explicitly in IP as the use of shortest path routing; although the cost function was not straight forward since financial cost (related to physical link length, link availability, and so on) did not always equate to topological cost because of the complexities of link weight allocation, inter-domain routing policies and so on – and the utility function was also complex with prefix reachability and packet-level quality of service being components of it, and whose utility differs for different applications and user types.
In the case of SDH, being adjacent so close to the physical layer, the cost of physical connectivity is perhaps more of an influence since the transmission is shared between less services (transport network types) and also simpler to price (since physical distance is a more easily measured metric). The comparison of circuit length versus shortest topological distance and physical circuit length versus shortest physical path lengths showed that while distance is optimised, the other requirements such as architecture and available bandwidth meant that many, but not most of the circuits followed the shortest paths. The utility function of SDH, however, is also very complex with the imposition of many constraints (diverse routing, explicit structures and hierarchy) to deliver required connectivity properties such as availability and capacity scalability.

In other terms, the conditions under which this optimisation occurs are dependant on the demand requirements and the available resources, that is, the adjacent layers. MITIE, as previously described, contains algorithms based on the protocol observations from Chapter 3 and Chapter 4 and was created to simulate a simplified form of topology reconfiguration to examine the effect

of adjacent layers. The simplification of topology reconfiguration used in MITIE is that three operations occur optionally in a cycle – node addition, link removal, and link addition. These three operations are the components of most network evolution scenarios, including growth (either in terms of nodes, or in terms of links), erosion (that is, contraction in terms of links), and redesign (or link rewiring). Only network size contraction is not possible (the removal of nodes). The decision processes behind each of these operations, such as which node a link is to connect to or which link is to be removed or rewired can vary, some being biased toward topological traits, such as node degree, some being biased toward more physical traits such as physical line length. Although Chapter 3 and Chapter 4 also examined more complex architectural structures, such as the OSPF areas, or protection architectures of SDH and the utility function of those (as the availability rate in 4.2.1.3.2), these were not implemented in MITIE due to their complexity.

MITIE was used in Chapter 5 to experiment with a range of network development scenarios to examine the inter-layer feedback conditions under which power-law compliant networks develop. The scenarios, not only the individual processes, were partly selected from the understanding of IP and SDH in the earlier chapters, but also in a way to investigate the conditions required for feedback.

6.1 Inter-layer Feedback Conditions and Power-Laws

The network development scenarios examined in Chapter 5 fell into three categories: network growth, network erosion and network rewiring. The modelling target, as suggest by Chapter 3 and Chapter 4 was a topology which had a degree frequency distribution that followed a power-law, and then if compliance was good, the other metrics were examined for compliance.

Of the network growth scenarios that used an element of inter-layer feedback, a node selection scheme based purely on node load, which in the case of the full mesh of demands used here is equivalent to betweeness centrality (TR, or total routes per node), did not tend to produce power-law compliance, as Fabricant et al [FAB02] also discovered when attempting to grow networks based on various measures of centrality. Node selection based on average link load around a node (AR), however, did produce topologies which conformed very well to power-law fits - the measure of average load was the total number of routes through a node divided by the degree, and as such were a function of the reciprocal of degree. Another growth model based on degree, the BA model (a single-layer model), which uses degree directly also produced power-law compliance but we are now considering the inverse of degree, so the "rich-get-richer" explanation was no longer sufficient. Using node selection based on the reciprocal of degree did not produce power-law compliance and as such the conclusion was that the inclusion of load (and, therefore,

feedback) created the power-law compliance. Another significant thing to note of the use of AR as the node selection scheme was that when only one link is added from the new node there is good compliance to a power-law, however, when there are two nodes added from the new node to existing nodes selected based on average load then the compliance was even higher, and in the latter case not only was increased load changing the affinity of connection, but since there were two links added the topology was no longer a tree and newly added nodes created alternative paths and the routing table and, therefore, load distribution changed according to the new topology.

The erosion scenarios also produced significant results. By uniformly randomly removing links from either the ER or BA topologies the topology began to tend toward power-law compliance. This had nothing to do with inter-layer feedback and it was postulated that the trait was caused by the effect of "richer-get-poorer-faster" or more precisely "the-richer-have-more-to-lose", just like "rich-get-richer" is the cause of the BA power-laws, this was happening in reverse – the high degree nodes had a higher chance of being selected as the end-point of a link chosen at random. To examine the effect of feedback the remaining erosion scenarios had to be normalised to the random erosion scenario to see whether feedback had an increased tendency to power-law. The normalised comparison of scenarios for the erosion of BA initial topologies showed that load based erosion (arpl - the scenario most directly linked to load based feedback) showed the most improvement, but in general there was not much to improve on since the initial topology was already relatively power-law compliant and the gains were not very large. Erosion of initially ER topologies, however, showed significant improvement in power-law compliance as erosion progressed; how much depended on exactly what criteria were used for the link removal. Erosion of ER based purely on load showed the most improvement in compliance, with the compliance being always higher than the random erosion case, however, erosion based purely on geography (AE) also showed significantly higher compliance than random erosion. Again, it would be valid to conclude that some element of feedback is responsible for the tendency to power-law in the load based case. The fact that the erosion based on geography scenario also showed a tendency towards power-law is an unrelated phenomenon, since there was no feedback, and it could be postulated that it occurred because the removal of the longest links tended to create hubs reaching only into the local vicinity – this view is supported by the fact that if the node placement were uniformly random there was less of a tendency towards power laws, and when the nodes were placed in a heavy tailed distribution (that is, they were naturally clustered) the tendency toward power-law compliance was stronger.

Erosion scenarios which were based on a more realistic metric of link cost, the bandwidth-distance product (as used in Chapter 1), were however less likely to tend toward power-law compliance with either no tendency stronger than the random removal scenario, for the uniformly random node placement case, or a weak tendency, in the heavy-tailed node placement case. The bandwidth-distance erosion scenarios both (that is, with uniformly random and with heavy tailed node placement) showed weaker tendencies toward power laws than either geography only, or load only based erosions, and did not even reach approximate power-law compliance once the tree topology was reached.

The growth and erosion scenarios, that is, node growth and link removal, are actually closely related to the two constituent processes of rewiring – link removal and link growth. Significant results from the rewiring include the discovery that the process of rewiring on a single layer basis alone can make the network tend to a power-law, but only under certain conditions, such as uniformly random removal with replacement link end-point selection based on degree. A scenario with no removal and just repeated link addition based on degree alone tended away from power-law compliance and the inclusion of the removal stage caused the development of the topology to tend towards a power-law – this was caused by the fact that random removal affected high degree nodes more and tended toward an equilibrium where the overall topology was power-law compliant.

In the multi-layer scenarios where there was uniformly random removal, and addition based on total or average load neither scenario would tend to a power-law, while as already mentioned the single layer random removal with addition based on degree did – there was obviously something missing from the load based scenarios that precluded the tendency to power-law. The addition of removal based on highest load, when paired with addition based on total-load, did however have a tendency to power-law so the conclusion is that some kind of counter-balance or related and opposite force was required for power-laws to emerge. By reversing the direction of this counter-force and removing based on lowest load the power-law went away again. The inference, therefore, was that the forces had to be similar and opposite. The removal based on uniformly random selection and then add based on degree also supported this since the opposite of "add based on degree" could be considered to be uniformly random removal, since it would inherently affect the high degree nodes more.

The removal stage was then changed to be geography based and the power-law compliance for scenarios with addition based on total load remained – this time the opposing forces were based

on geography – the longest links were being removed, but the addition based on total load would most likely attempt to connect node pairs which were further apart. This postulate is however questioned by the scenario where the longest link was removed, to be replaced by a uniformly random link – this would also add a link between nodes which could be further apart, however no power-law compliance was achieved.
The last set of rewiring scenarios were those where removal was done on bandwidth-distance product – power-law compliance was highest where links were removed which had the "highest cost" and replaced between nodes with high total load – which would draw some load away from the other high cost links, reducing their cost. The opposing forces in this case are the decrease in cost due by removal of most expensive links, and the increase in cost by adding links between areas of high load (since the demand mesh remains constant).

Examining the modelling work as a whole it should be noted that node degree, which has often been suggested to be some kind of measure of the importance, or "bigness", of a node is not always a drop-in replacement for the actual load on the node, and will achieve similar power-law compliance results only if the counter-acting processes are similarly matched as those against the load based processes. Another thing to notice is that the measures of load, either average or total load, are also not equivalent to each other, and which one produces power-law compliance and which does not is often also dependant on which counter-acting process is present – this time however it is more of a question of efficacy in counter-action. The last overall conclusion is that node location distribution is not irrelevant in the scenario examined here – while often uniformly randomly and heavy-tailed distributions acted similarly there were a number of cases where there were significant differences in their performance (e.g. distance-based erosion).

6.2 Future work

The work in this thesis has only touched on the broad research topic that is multi-layer feedback modelling which still has many unanswered questions, and this thesis offered a wide area sweep of the many cases and details present in multi-layer modelling. There are, however, still many aspects that would be of great value to research further.

6.2.1 Types of Feedback

The use of load-based selection, changing the underlying network, and examining the effect of re-routing is only one type of multi-layer feedback. MITIE is also capable of dynamically allocating

and re-allocating link weights based on a number of algorithms, including inverse capacity, as is used by OSPF (see section 3.3.3.1.2), inverse capacity with a given granularity (granularities such as in Table 8), weights based on Euclidean distance (potentially a good link weight for transport or transmission networks), and a bandwidth-distance product that includes link capacity granularity. The use of non-unity link weights and the use of these algorithms is another form of multi-layer feedback because now properties of the transport network are directly affecting the routing of the traffic/load layer, which in turn dictates the new weights.

6.2.2 Heterogeneous demands

Another interesting aspect to consider is the use of non-unity-bandwidth non-full-mesh demands, as the current full mesh of single-unit demands is effectively centrality betweenness and should instead be an arbitrary, uneven demand pattern. The re-routing of high bandwidth demands should have a greater effect on the network than the re-routing of low bandwidth demands.

6.2.3 Hybrid node and link selection processes

Fabricant et al [FAB02] examined what proportion of geography and centrality were necessary to achieve power-law compliance for their growth models – it would be trivial to implement a form of weighted hybrid selection in MITIE, where –addA: could be specified more than once in the command line together with a series of weights and the values used in the roulette wheel for selection are simply the weighted sum of the weights generated by each individual –addA: policy. Such a hybrid node selection policy would then allow us to investigate, for example, growth based on the convolution of the distance to a candidate node and the total load of that node.

6.2.4 Multi-layer, not just dual layer

MITIE could be enhanced to daisy-chain multiple instances which communicate with each other and each handles a single layer, communicating any changes to the adjacent layers and the real effect of many stacked layers could be examined. Each instance would use very different link and node selection processes, with geography and physical link cost being the most important selection factor in the lowest layer, and the available bandwidth being the most important utility in the highest layer. Work on percolation in hierarchical system and self-organisation by Shnirman [SHN98] demonstrated a model that would suggest that heterogeneity in the decision

processes at each layer could potentially lead to very large avalanches across the layers, but most will be small – so a truly multi-layer investigation should also examine heterogeneity.

6.2.5 Realistic network architectures and structures

The examination of the protocols in this thesis showed that explicit structure appears in SDH and to a lesser extent in IP, as protection rings, hierarchies, OSPF areas, IXPs, AS domains. It would be beneficial to have a version of MITIE that actually uses known network planning algorithms to create topologies, based on the demand, such as IGEN by Quotin [IGEN] – although in this case the algorithms are part of the chain of layers and the other layers adapt differently, as described in section 6.2.4. The issue of co-location should also be addressed – this thesis assumed direct co-location between one layer and the next, however this is not the case usually and the effect should be studied.

Appendix A
The Modular Inter-Layer Feedback Topology Investigation Tool and Simulator

The **M**odular **I**nter-layer feedback **T**opology **InvE**stigation tool and simulator (MITIE) is a collection of shell scripts, gawk scripts and C++ tools created by myself which are designed to emulate the processes behind network topology evolution, with a number of features that distinguish it from existing topology tools:

- Rather than support a single type of connectivity growth, based on say, node degree, a large selection of connectivity strategies are available, including ones which are based on geography, topology, shortest paths and network load. Different strategies can be used at different processes of the evolution, for example the link growth stage could connect high degree nodes to another node chosen based on Euclidean distance. The wide range of connectivity strategies allows us to investigate different network topology types, such as ducts, which are much more closely linked to geography than, say, IP which may be modelled with some type of preferential attachment based on node degree.
- MITIE is designed to route traffic and calculate load on links, emulating a network layer above the underlying topology layer – as such it is capable of making connectivity and re-connectivity decisions based on link load, which forms the closed-loop feedback element of the simulation.
- When routing demands link weights can be recalculated as a function of various attributes such as distance to the network edge, link utilisation or hierarchical positioning, further affecting the feedback process and allowing us to investigate the effect of node centrality, hierarchy and other aspects that affect demand routing.

Aside from the main application, there are a number of other applications and gawk (GNU AWK, a text processing language) and Unix scripts to facilitate the manipulation and analysis of input files (geography, seed topology etc) and parse the output.

MITIE design

MITIE was implemented in C++ for its speed and ability to manipulate large amounts of memory quickly and efficiently. The use of the C++ Standard Template Library (STL) also

simplifies the implementation. MITIE was designed to be executed non-interactively, configured through the command line, and send all output to the standard output in a format that was designed to be parsed by further programs so that the required analysis could be extracted or further performed. It was designed to be executed in large batches in parallel on multiple computers, and have the output piped to individual files, and for the raw output in the files to be further analysed after simulation. This approach, together with some sort of distributed execution environment (see the 'jobs' directory of the source tar-ball for a simple job creation and distribution implementation) allows for the experimentation on various parameters with many different values and random number generator seeds to generate many output metrics and their related statistics.

MITIE's basic function is to read in an initial topology and then perform, cyclically, three operations:

- **A growth stage:** a node is optionally added to the network and a link is added from that node to a node in the existing network. The node selected for the second endpoint is chosen based on a given strategy, denoted '`addG`'.
- **A link removal stage:** a link is optionally removed from the network. The link selection strategy is denoted '`remove`'.
- **A link addition stage:** a link is optionally added to the network. The endpoints of the link are selected based on two strategies, denoted '`addA`' and '`addB`', which can be different or the same.

These three stages are repeated in the order listed until either a maximum number of cycles are reached, or the network is fully meshed, or the network is a tree and any further link removal would partition the network.

The arrangement into three separate stages like this can make MITIE very flexible, allowing us to investigate:

Topology generation: By having only a growth stage we could create an Albert-Barabasi (see section 2.3.2.2) network by selecting `addG` to choose nodes according to node degree, and omit the link removal and addition stages. It would be possible to investigate the effect of the choice of seed network easily because this is a parameter in the simulation.

Topology erosion: By having a well developed initial topology (that is one which is more than just a small seed topology) loaded we could remove links based on a strategy like choosing links with the lowest load, to minimise network cost (minimise number of links).

Connectivity growth: By loading a seed topology it would be possible to investigate the evolution of connectivity growth by using the link addition stage and no network growth or link removal stages.

Topology rewiring: By having no growth stage and specifying a link removal and link addition stage it is possible to investigate topology rewiring. Rewiring has many options in MITIE – it would be possible for example to specify a power-law like demand matrix and perform rewiring based on link load – thereby investigating the effect of demand and whether it is mirrored in the topology of the underlying network.

There is also an option to perform each stage only every *n*th (where there is a different *n* for each stage) cycle so more fine grained control over network evolution is possible.
Further more specific example scenarios are given later in this Appendix.

MITIE Connectivity Strategies

The strategies available for each stage are varied and not all strategies are available for every stage. The link and node selection algorithms are expressed at the command line by tokens; their validity and meaning are listed below:

For **addG** the following tokens are valid:

0, N, r, d, D, k, K, tr, TR, ar, AR, tl, TL, al, AL, e, p

For **addA** the following tokens are valid:

0, r, d, D, k, K, g, G, tr, TR, ar, AR, tl, TL, al, AL, ge, gp, GE, GP

For **addB** the following tokens are valid:

r, d, D, k, K, tr, TR, ar, AR, tl, TL, al, AL, e, p

where the tokens signify:

0	do not perform this stage
N	do not add any links (just the new node)
r	Select a node uniformly at random
d	preferentially choose lower degree nodes
D	preferentially choose higher degree nodes

k	the lower degree node of the two nodes that were most recently disconnected.
K	the higher degree node of the two nodes that were most recently disconnected
g	select the node that was added in the most recent growth phase
G	the existing node that was connected to in the most recent growth phase
tr	preferentially choose nodes with lower total RPL
TR	preferentially choose nodes with higher total RPL
ar	preferentially choose nodes with lower average RPL (total RPL/degree)
AR	preferentially choose nodes with higher average RPL (total RPL/degree)
tl	preferentially choose nodes with lower total load per links
TL	preferentially choose nodes with higher total load per links
al	preferentially choose nodes with lower average load per links (total load per links/degree)
AL	preferentially choose nodes with higher average load per links (total load per links/degree)
MAXAR	select the node with the highest average RPL (total RPL/degree), on tiebreak choose one at random from tiebreak set.
MAXTR	select the node with the highest total RPL, on tiebreak choose one at random from tiebreak set.
e	The probability of destination node based on exponential normalised Euclidean distance from source node (Waxman)
p	The probability of destination node based on power normalised Euclidean distance from source node
ge	The probability of two nodes being connected is based on Waxman geo - favouring smaller distances. There is no need to specify addB.
gp	The probability of two nodes being connected is based on power geo - favouring smaller distances. There is no need to specify addB.
GE	The probability of two nodes being connected is based on Waxman geo - favouring larger distances. There is no need to specify addB.
GP	The probability of two nodes being connected is based on power geo - favouring larger distances. There is no need to specify addB.

Where RPL (Routes Per Link) is the number of routes traversing a given link given all-pairs lowest cost (lowest weight) routing, therefore, **TR** preferentially selects nodes which have the highest sum of number of routes traversing the links that connect to it, i.e. the total number of routes terminating or transiting this node. **AR** (average RPL) is this value (total number of routes) divided by the degree of the node.

LPL (Load Per Link) is similar to RPL but it is the sum of the capacities of all demands traversing a given link – where the capacity is specified by a traffic matrix specified in the command line. Demands are routed along the lowest cost path, like the RPL, but do not necessarily exist between every node pair.

For "**remove**" the following tokens are valid:

0, rpl, RPL, arpl, ARPL, lpl, LPL, lrpl, ALPL, w, W, r, e, E, p, P

where the tokens signify:

0	do not remove any links.
RPL	prefer links with the highest number of traversing routes.
rpl	prefer links with the lowest number of traversing routes.
ARPL	the link with the highest number of traversing routes.
arpl	the link with the lowest number of traversing routes.
LPL	prefer links with the highest total capacity of demands traversing it.
lpl	prefer links with the lowest total capacity of demands traversing it.
ALPL	the link with the highest total capacity of demands traversing routes.
alpl	the link with the lowest total capacity of demands traversing routes.
W	the absolute highest link weight.
w	the absolute lowest link weight.
r	A link chosen uniformly randomly
e	select a link according to the waxman (exponential) probability of normalised euclidean distance.
E	select a link according to the inverted waxman (exponential) probability of normalised Euclidean distance.
p	select a link according to the power probability of normalised euclidean distance.
P	select a link according to the inverted power probability of normalised euclidean distance.

A selection of other "**addA**", "**addB**", "**addG**" and "**remove**" strategies are also implemented, for example:

phdd	prefer high degree difference links
ldhd	prefer load degree low difference links
ldld	prefer low degree low degree difference links

These are for experimenting with network assortativity however these *are not fully tested* and not examined further. For further details please examine the MITIE source code.

MITIE Implementation

The operation of MITIE can be seen in Figure 111 where the order of execution of tasks can be seen. The flow diagram follows the cycle of processes described earlier but also adds some steps before the link/node addition/removal process which prepares link weights for the routing algorithm if necessary, then performs the routing (APSP, i.e. All-Pairs Shortest Path). Once routed the routes and the demand matrix are used to calculate the number of Routes Per Link (RPL), the number of Routes Per Node (RPN), and optionally the Load Per Link (LPL) and Load Per Node (LPN), which is based on the input demand matrix.

The APSP algorithm used can be selected at the command line – by default a depth first search (DFS) algorithm is used to find the lowest weight paths, which performs best in sparser networks, but there is also an implementation of a variation on the Floyd–Warshall algorithm which

performs better in highly connected networks. In most cases the DFS algorithm performs very well and is the recommended algorithm to use, with the Floyd–Warshall only for special cases.

For selection processes (`addG`, `addA` or `addB`) that are not absolute decisions, (i.e. the lowest load link (`-remove:alpl`) or the node with the lower degree of the two endpoints from the previous link removal (`-addA:k`) are absolute decisions) i.e. those that state "prefer" in their description, the implementation of the selection procedure is based on the roulette wheel selection algorithm [COR01]. So the probability of a link or node being selected from a set is directly proportional to the weight assigned to it.

Roulette wheel selection works by generating a set of options from which a choice is to be made and assigning weights to them – the elements in this set are either nodes, or potential links (node pairs). These weights are normalised to sum to 1.0 and a uniformly random number between 0 and 1 is generated. By subtracting the weights (in an arbitrary order) from this number until it drops below zero a set member is chosen as the location where the roulette ball "fell". By having larger weights the chances that the "ball" "falls" on a given set member increases. This is the method by which certain nodes or links are favoured in the selection process.

This process can be seen in Figure 112.

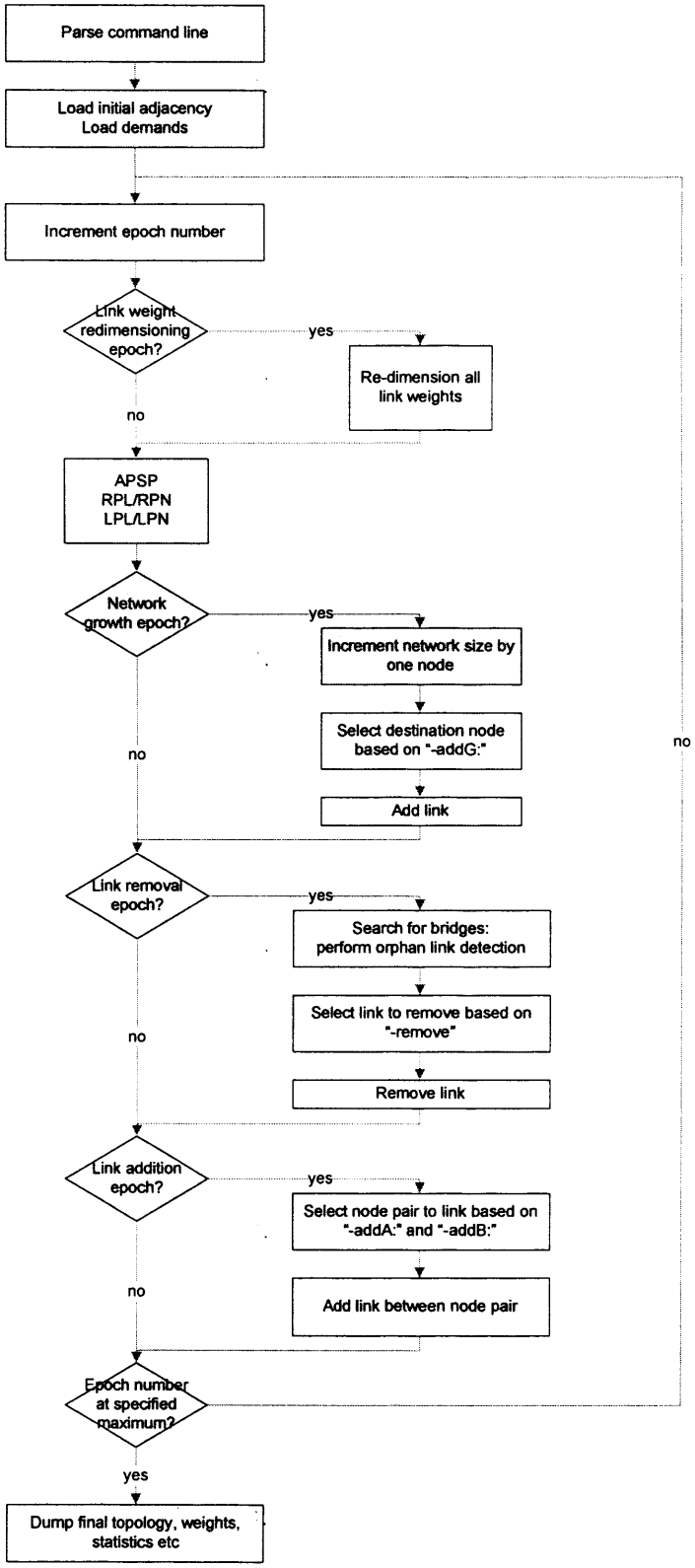


Figure 111 The main execution path for MITIE

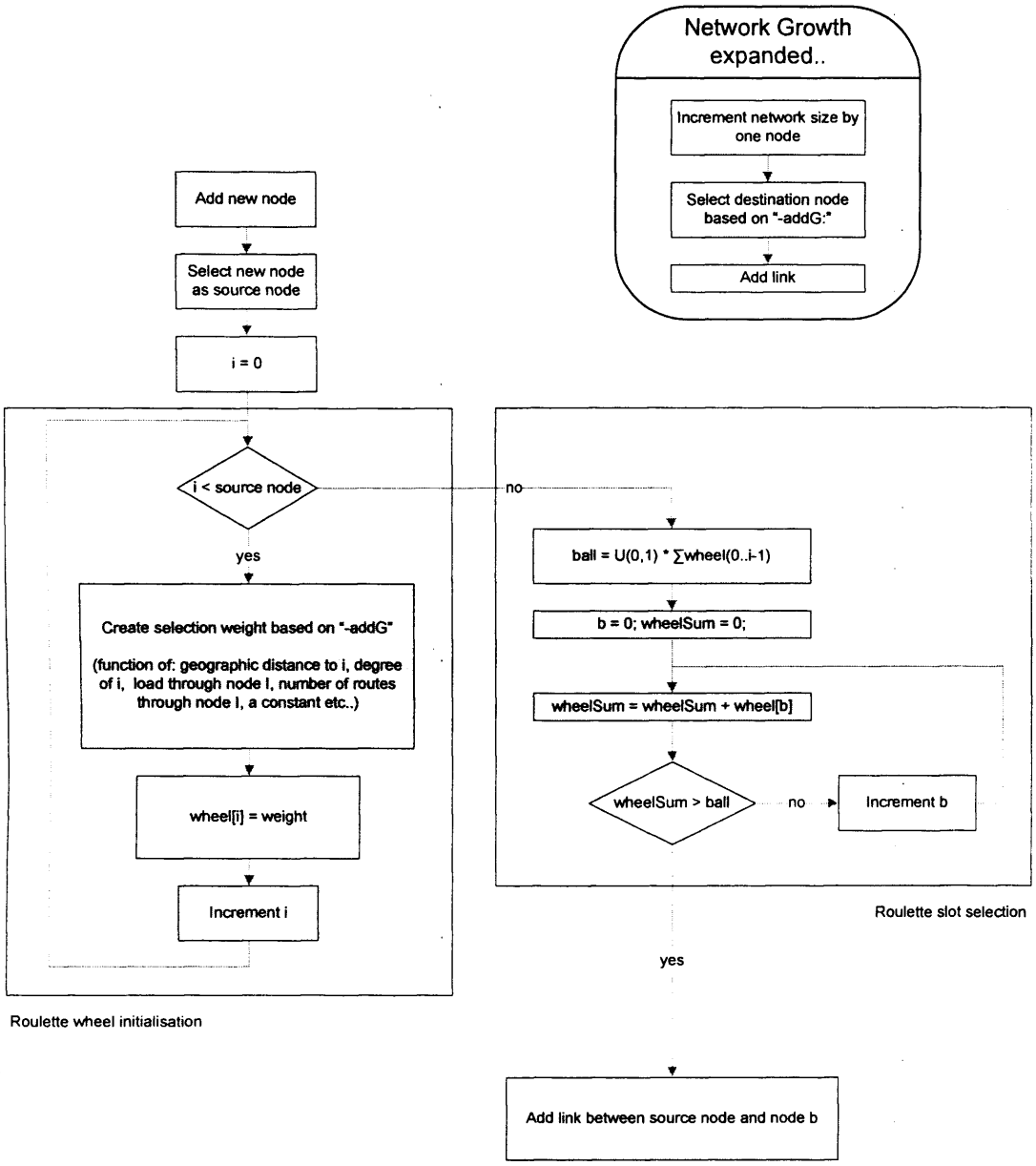


Figure 112 Roulette wheel node or link selection in the growth process of MITIE

It is possible that during the rewiring or link removal that the network is partitioned and this would prevent the routing of all nodes to all nodes (APSP, for RPL). To prevent this it is necessary to test whether links that are to be removed are bridge links (bridge links are links which when removed would partition the network) or not. Rather than removing a link,

performing APSP until failure or a timeout condition, which is generally a very costly operation (in terms of execution time) we attempt to mark edge trees as being bridge links. The algorithm is decomposed in Figure 113.

This cycles through all the nodes checking whether there is only one link that is not marked as a bridge. If there is only one then it is also marked as a bridge and the search continues. This process is repeated until no new links are marked as bridges. This is equivalent to a breadth first search for cycles and can substantially reduce the set of links that are in the set of possible links to remove, thereby preventing many retry attempts if a link is later discovered to be a bridge. This algorithm was implemented because in the case of networks with many hub components (which is very likely in power-law networks) the probability of selecting a bridge is very high.

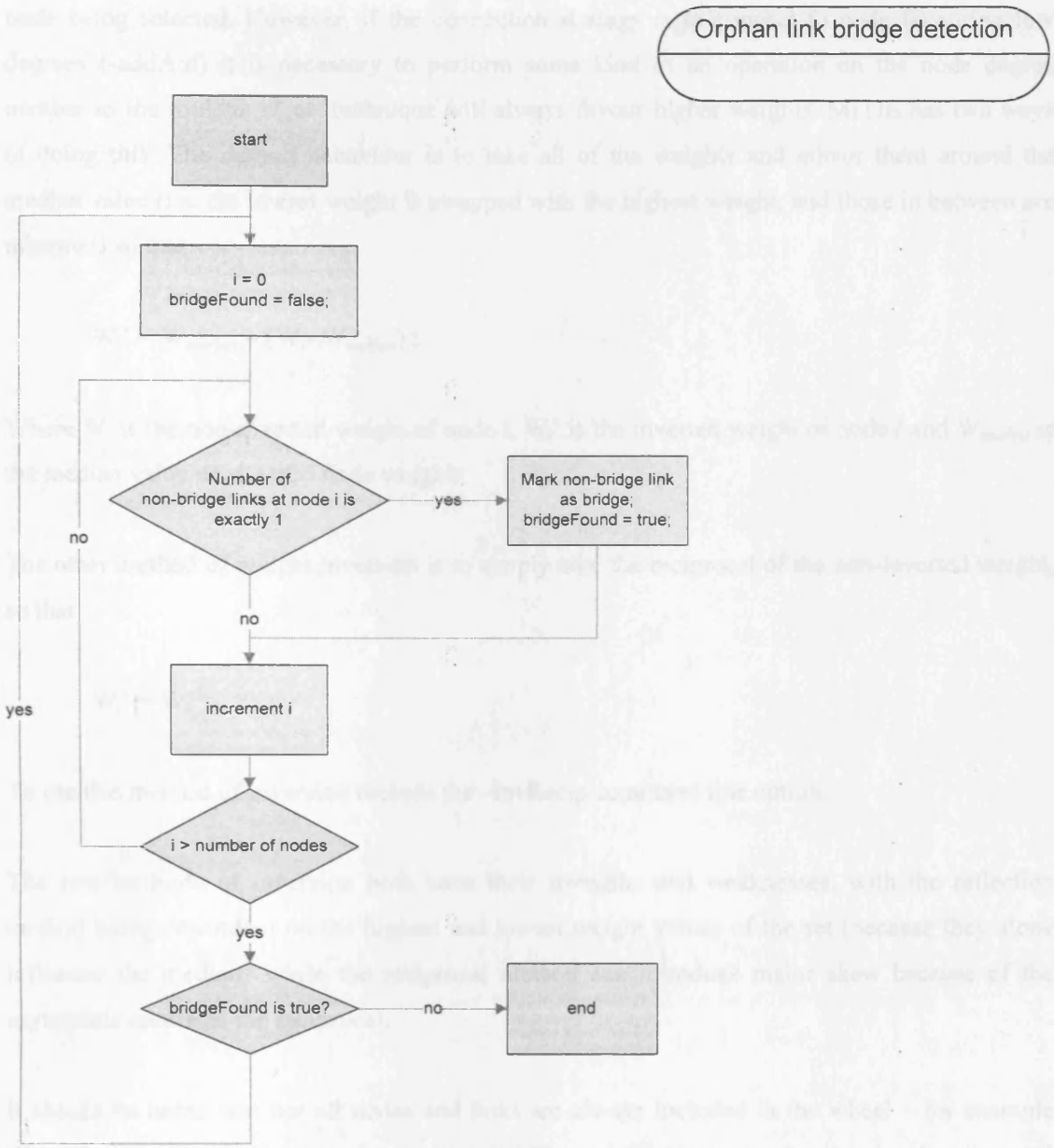


Figure 113 MITIE orphan link bridge detection algorithm

Link and node selection inversion

When a set of weights have been put into the roulette wheel for link or node selection it is sometimes necessary to invert their weight. So, in the case of –addA:D the degree of each node is directly used as the weight in the wheel – the larger the weight the higher the probability of that

node being selected. However, if the connection strategy is to connect to node favouring low degrees (-addA:d) it is necessary to perform some kind of an operation on the node degree number as the roulette wheel technique will always favour higher weights. MITIE has two ways of doing this. The default behaviour is to take all of the weights and mirror them around the median value (i.e. the lowest weight is swapped with the highest weight, and those in between are mirrored) so that

$$W_i' = W_{median} - (W_i - W_{median}) ;$$

Where W_i is the non-inverted weight of node i, W_i' is the inverted weight of node i and W_{median} is the median value of all valid node weights.

The other method of weight inversion is to simply take the reciprocal of the non-inverted weight, so that

$$W_i' = W_i^{-1};$$

To use this method of inversion include the –invRecip command line option.

The two methods of inversion both have their strengths and weaknesses, with the reflection method being dependant on the highest and lowest weight values of the set (because they alone influence the median) while the reciprocal method can introduce major skew because of the asymptotic nature of the reciprocal.

It should be noted that not all nodes and links are always included in the wheel – for example when choosing the destination of a new link (addB or addG), the nodes that are already connected to the source node cannot be added to the roulette wheel. Similarly if a link was found to be a bridge then it cannot be included in the "remove" roulette wheel.

MITIE Command Line Arguments

Since MITIE is designed to be non-interactive and completely controlled from the command line and since it has a large set of features the command line parameters are extensive. Below is a full list of command line options.

--help

Print the main command line options

--helpGRAB

Print options related to the addG, remove, addA and addB

-f:<input weight matrix file>

Load this link weight matrix to initialise the network

-verbose

Enable logorrhoea mode. (specify more than once for more verbosity)

-seed <integer>

To specify the random number generator seed.

-cycles <integer>

To specify the number of remove/add cycles to perform (default is to continue indefinitely until tree or full mesh)

-size <integer>

To explicitly specify the number of nodes in the initial network. Useful when requiring a disconnected initial network for, for example, connecting based on geography.

-addG:{0,N,r,d,D,tr,TR,ar,AR,tl,TL,al,AL,e,p} <int N>

Specifies the strategy used when selecting which node is to be linked to the new node that is being added. For an explanation of the tokens see the section above. The trailing integer specifies the size to which the network is to grow. If 0 is specified then no integer is required. The default is "0" (do not add any links).

-addA:{0,r,d,D,k,K,g,G,tr,TR,ar,AR,tl,TL,al,AL,ge,gp,GE,GP}

Specifies the strategy used when selecting a node for the first endpoint of the new link that is being added. Default is "0" (do not add any new links)

-addB:{r,d,D,k,K,tr,TR,ar,AR,tl,TL,al,AL,e,p}

Specifies the strategy used when selected a node for the second endpoint of the new link that is being added. Default is "D" (favour nodes with a higher degree)

-remove:{0,rpl,RPL,arpl,ARPL,lpl,LPL,lrpl,ALPL,w,W,r,e,E,p,P}

Specifies the strategy used when selecting a link to remove from the topology. The default is "rpl" (prefer links with a low routes-per-link value).

-dontUseDFS_APSP

use Floyd–Warshall like APSP (may be faster with highly connected graphs)

-invRecip

By default inversion of roulette wheel values is done by reflection and translation. This option will take the reciprocal of the values instead.

-doTiming

Measure the wall time taken to execute certain operations, such as APSP and bridge detection.

-step:<int stepG>,<int stepR>,<int stepA>

Specifies how many cycles will pass between successive growth, removal and addition processes, ie growth will occur every stepG cycles, removal every stepR cycle and link addition every stepA'th cycle.

-step:<int offG>,<int stepG>,<int offR>,<int stepR>,<int offA>,<int stepA>

Specifies how many cycles will pass between successive growth, removal and addition processes, but also specifies an offset, or delay, such that growth will only begin after the offG'th cycle. This option is useful when wanting to start rewiring (removal with addition) only after the network has grown to a certain size.

-incGrowth

Enables incremental growth - connections in links created by -addG can only be to existing nodes (nodes with a lower index).

-loadGeoPosn:<filename>

Specifies the file from which the x,y euclidean co-ordinates of each node are to be read. The file must contain the same number of lines as there are maximum number of nodes and each line must have two integers, the x y pair.

-loadGeoDist:<filename>

Specifies the file from which the x,y euclidean *distances* between each node pair are to be read. The file must contain a white space or comma seperated N by N matrix of integers, where N is the maximum number of nodes.

-loadDemands:<filename>

Specifies a file containing an N by N matrix of integers which specifies the traffic matrix between all node pairs. N is the maximum number of nodes. If this is not specified then a full mesh unit matrix is assumed.

-Alpha_B <double>

Set the alpha parameter of euclidean based probability calculations for -addB:e,p

If not specified the default value is $5*10^{-5}$.

-Alpha_G <double>

Set the alpha parameter of euclidean based probability calculations for -addG:e,p

If not specified the default value is $5*10^{-5}$.

-Alpha_R <double>

Set the alpha parameter of euclidean based probability calculations for -remove:e,E,p,P

If not specified the default value is $5*10^{-5}$.

-Beta_B <double>

Set the beta parameter of euclidean based probability calculations for -addB:e,p

If not specified the default value is 1.9104.

-Beta_G <double>

Set the beta parameter of euclidean based probability calculations for -addG:e,p

If not specified the default value is 1.9104.

-Beta_R <double>

Set the beta parameter of euclidean based probability calculations for -remove:e,E,p,P

If not specified the default value is 1.9104.

Output data control options

-dumpDistances

The minimal distances (number of hops) between every node pair will be dumped to stdout at every cycle.

-dumpFullAdj

Dumps the full adjacency matrix to stdout at every cycle. (marked as w_<cycle#>_)

-dumpFinalWeights

Dump the entire weights matrix at the end of the simulation (marked as _wtF_)

-dontDumpSparseAdj

Do not dump adjacencies list at every cycle (in the format of "a_N_ <node> <degree D> <node1> <node2>..<nodeD>")

-dumpDegreeDist

output the degree frequency distibution every cycle in the format "dd_<cycle#>_ <degree> <count>".

-dumpStep <integer>

Specifies how many cycles will pass between successive data dumps (adjacency list etc..) use with –dumpFinalWeights –dumpDistances etc..

-dumpRouteDeltas <integer>

Every given number of cycles find and print the number of routes that have changed since the last time the number of routes changes was printed.

-finalDump:dawDl#d#a#w#D#l

Specify which data is to be printed at the end of the cycles. Where the options specify:

do not print the following output

d distances matrix

a sparse adjacency matrix

w link weights matrix

D degree distribution

l link load

Link dimensioning options

-newLinkWt:{min,ave,max,<int>,{dte,hte}:(<int>,<int>,<int>..)}

Any new links will take a link weight decided by:

min The minimum value of any links attached to either of the new link's endpoints

ave The mean average of all current links attached to both of the new link's endpoints

max The maximum value of any links attached to either of the new link's endpoints

<int> A fixed value

dte:(<int>,<int>,<int>..) The link weight is decided according to the link weight distance to the network edge. This is done by first finding, for each node the lowest link weight distances to an edge node (see -netEdgeDegree) and arranging these distances in increasing order. This list is then partitioned into the same number of sections as there are integers listed in this argument. Those distances in the first segment are then associated with the first integer and so on for all integers listed. For the new link the minimum link weight distance to the edge is found and the weight associated with that segment is assigned to the new link.

hte:(<int>,<int>,<int>..) Equivalent to dte: but the minimum hops to the edge are used rather than the link weight distance

-dimensionWt

Dimension link weights on initialisation according to "-dimensioning".

-redimensionWtFreq:<int>,<int>

First integer specifies the number of cycles from start before full redimensioning, second integer is how many cycles between subsequent full redimensions.

-dimensioning:{<int>,{dte,
hte,rpn,rpl}:(<int>,<int>,<int>..),ED,{Linv,Rinv,RDPinv}:<int>}

(default: -dimensioning:1)

When a complete redimensioning is performed (see –dimensionWt) this specifies how the link weights are found:

<integer>: new link takes the weight the fixed value specified

dte:(..): distance to edge

hte:(..): hops to edge

rpn:(..): links are assigned weights based on the number of routes passing through the link's endpoint nodes.

rpl:(..): links are assigned weights based on the number of routes per link currently traversing the link

Linv:<int>: ceil link load to <int>, then set weight to max(load(link_*)/load(link_i)

Rinv:<int>: ceil rpl to <int>, then set weight to max(rpl(link_*))/rpl(link_i)

RDPinv:<int>: ceil rpl to <int>, then set weight to max(rpl(link_*)*d_link_*)/(rpl(link_i)*d_link_i)

RDP:<int>: ceil rpl to <int>, then set weight to a function of RDP

ED: normalised euclidean distance

-netEdgeDegree:<int>[%]

If no "%" is specified then this specifies the maximum degree a node could have to be considered an edge node (default 1). If a "%" is specified then the integer specifies what proportion of nodes, sorted in increasing order of degree is labelled as an edge node. Edge nodes are required when using "hte" or "dte" as options in the –newLinkWt or –dimensionWt switches.

Example usage

Below are a list of example execution scenarios and their command lines.
The "\" character at the end of a line means that the command continues on the next line, and is not part of the actual command.

Albert-Barabasi topology generator:

```
./mitie -f:ring5.adj -addG:D 1000 -remove:0 -addA:0 -incGrowth
```

Loads a seed topology from `ring5.adj` (assume this contains a topology of five nodes connected in a ring with five links) and at every cycle adds one node, connecting it to nodes favouring nodes with a higher node degree until there are 1000 nodes in total. Since only one link is added as part of the growth stage the final link to node ratio will be 1.0. Although `-incGrowth` is specified the result would be the same without it as making all nodes valid nodes for selection in the roulette wheel would still mean they do not get chosen because their degree is 0 as they are initially disconnected.

Albert-Barabasi topology generator with arbitrary average degree:

```
./mitie -f:ring5.adj -addG:D 1000 -remove:0 -addA:g -addB:D \
     -incGrowth -cycles 1000 -step:1,1,2
```

Loads a seed topology from `ring5.adj` (assume this contains a topology of five nodes connected in a ring with five links) and at every cycle adds one node, connecting it to nodes favouring nodes with a higher node degree until there are 1000 nodes in total. At every second cycle (the 2 in the `-step` switch) we also add an additional link between the newly added node (as specified by `-addA:g`) and a node that is randomly selected, favouring higher degree nodes (`-addB:D`). Since one node and a link is added per cycle, and an additional link is added at every other cycle the link to node ratio is 1.5. In this case `-cycles 1000` must also be specified to limit the number of cycles to perform – the program will not automatically terminate after performing the growth to 1000 nodes because we are also adding links with `-addA`.

Waxman topology generator:

```
./mitie -f:null1000node.adj -loadGeoPosn:random_1_1000.geo \
      -addG:0 -remove:0 -addA:ge -cycles 2000
```

Since the growth stage will always use the new node as one of the endpoints it would not be correct to use it to generate a Waxman topology (although there are Euclidean geography based connectivity strategies for `-addG`) since **both** endpoints must be selected based only on the distance between them. Instead MITIE is initialised with `null1000nodes.adj` which is a file containing an adjacency matrix of 1000 nodes with no links, the geographic positions of the nodes are loaded from `random_1_1000.geo`, and MITIE completes 2000 cycles of link addition using the '`ge`' strategy – that is, it selects node pairs based on an exponential function of the distances between them. The generated network is, therefore, a Waxman topology with a link to node ratio of 2.0 (1000 nodes, 2000 cycles, each of which add a single link). Since we are adding links to a disconnected network and not actually growing the network in node size it is entirely possible that after the 2000 cycles not all nodes have been selected as endpoints and the network is, therefore, disconnected.

Growth and then rewiring

```
./mitie -f:ring5.adj -addG:D 1000 -remove:arpl -addA:D -addB:D \
     -step:0,1,995,1,0,1 -cycles 3000
```

The special feature demonstrated here is the cycle offset used in the step switch. In this configuration the topology will be grown partly as an Albert-Barabasi in terms of node growth, but at every step a second link will be added between nodes selected by favouring the higher degree nodes (note this is not strict BA model behaviour as that would have the growth node as one of the endpoints – see earlier example). `addG` states that the network growth is to continue up to 1000 nodes, hence will end at the 995 cycle (since five nodes were imported from the seed topology). At cycle 995 the '`remove`' process will begin (see third integer in '`step`' argument) to execute every cycle, while the addition process continues at every cycle, until the maximum number of cycles, 3000, is reached. The result is that the network is first grown based on node degrees and then rewired based on removing the one link with the lowest routes per link, and then connecting a new link between the highest degree nodes (effectively rich club growth).

What is not possible currently is to add links at a given rate during the growth process (to be able to reach an arbitrary link-to-node ratio) and then change the link addition rate to match the removal rate (so that the net outcome would be rewiring and a constant link-to-node ratio). To perform this task it is recommended to perform the network growth and rewiring as two separate invocations of MITIE, such that:

```
./mitie -f:ring5.adj -addG:D 1000 -addA:D -addB:D -remove:0 \
     -step:1,1,3 -cycles 995 -seed 42 | grep _aF_ \
     | gawk -f adjL2adj.awk > top_D_1000_0_D_D_seed_42.adj
./mitie -f: top_D_1000_0_D_D_seed_42.adj -addG:0 -remove:arpl \
     -addA:D -addB:D -step:0,1,0,1,0,1 -cycles 2000
```

Here the topology at the end of the growth stage is extracted (it is prefixed with "`_aF_`" in the output), converted from an adjacency list (the format of the "`_aF_`" lines) into a square adjacency matrix and stored in `top_D_1000_0_D_D_seed_42.adj`. This file is then the used to initialise the second invocation which performs rewiring for 2000 cycles.

For the second invocation it may also be interesting to add the `-dumpRouteDeltas <integer>` argument to be able to plot over time the number of routes that have been affected by the rewiring. Note that the integer specified with `-dumpRouteDeltas` says that the number of routes changed after every said number of cycles is to be dumped, and that this is not equivalent to the number of routes changed in one cycle multiplied by the number of cycles between consecutive tests.

Appendix B
Single Layer Modelling results

In this appendix we present a list of various experiments performed with MITIE. The results serve as a baseline for comparison to multi-layer results in Chapter 5 and are discussed further there.

The abbreviations used in the table headers are:

Expt. #	Experiment number – simply an index into the experiments so they can be more quickly referenced from within Chapter 5.
GRAB tuple	This is the node growth, link removal and link addition schemes used in the experiment, followed by the number of cycles performed for the experiment and the step at which growth, removal and addition were performed.
IG	Specifies whether incremental growth was used (•) or not.
GEO	The geography used when performing an experiment based on geographic distribution – this is either ER, a plane with nodes placed uniformly at random, and BT which are random subsets of the BT's SDH network node placements.
Target network size	This is the total number of nodes in the network
Actual network size	This is the actual number of connected nodes in the network
Ave. degree (..)	The average degree of nodes in the network, with the average degree minimum and maximum in parenthesis
DFD R^2	The correlation coefficient of the power-law fit to the node degree frequency distribution.
DFD R^2 rank	The position this experiment would occupy if the DFD R^2 were ranked in decreasing order
DR R^2	The correlation coefficient of the power-law fit to the node degree ranking.
DR R^2 rank	The position this experiment would occupy if the DR R^2 were ranked in decreasing order
CCR R^2	The correlation coefficient of the power-law fit to node clustering coefficient ranking
CCR R^2 rank	The position this experiment would occupy if the CCR R^2 were ranked in decreasing order
Ave R^2	This is the mean average of the DFD R^2, DR R^2 and CCR R^2.
Ave R^2 rank	This is the rank of the list of Ave R^2 values, sorted in decreasing order. The table is ordered by this rank.

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
1	{D,0,D,D} 10000:1,1,5			4000	4000.0	1.500 (1.000-207.200)	0.848 (0.842,0.854)	64	0.955 (0.955,0.955)	1	0.969 (0.969,0.969)	1	0.924	1
2	{D,0,D,D} 1000:1,1,2			1000	1000.0	1.499 (1.000-108.500)	0.835 (0.829,0.840)	68	0.951 (0.950,0.952)	4	0.948 (0.947,0.949)	4	0.911	2
3	{D,0,D,D} 4000:1,1,2			4000	4000.0	1.500 (1.000-256.650)	0.827 (0.823,0.832)	72	0.952 (0.952,0.952)	2	0.951 (0.951,0.951)	2	0.910	3
4	{D,0,D,D} 2000:1,1,2			2000	2000.0	1.500 (1.000-166.825)	0.826 (0.821,0.832)	74	0.952 (0.952,0.952)	3	0.950 (0.950,0.951)	3	0.909	4
5	{D,0,0,0}			10000	10000.0	1.000 (1.000-169.075)	0.879 (0.872,0.886)	38	0.923 (0.923,0.924)	19	0.925 (0.925,0.926)	5	0.909	5
6	{D,0,0,0}			4000	4000.0	1.000 (1.000-106.725)	0.877 (0.870,0.884)	42	0.923 (0.923,0.924)	20	0.924 (0.924,0.925)	6	0.908	6
7	{e,0,D,e} 2000:1,1,1	•	BT	2000	2000.0	2.000 (1.000-51.650)	0.928 (0.925,0.930)	1	0.887 (0.886,0.888)	39	0.910 (0.909,0.910)	12	0.908	7
8	{p,0,D,p} 2000:1,1,1	•	BT	2000	2000.0	2.000 (1.000-52.325)	0.927 (0.924,0.930)	2	0.886 (0.885,0.887)	41	0.906 (0.905,0.907)	17	0.906	8
9	{p,0,0,0}	•	BT	2000	2000.0	1.000 (1.000-14.525)	0.925 (0.922,0.928)	4	0.878 (0.877,0.879)	46	0.916 (0.915,0.917)	9	0.906	9
10	{D,0,0,0}			2000	2000.0	1.000 (1.000-74.875)	0.871 (0.863,0.879)	49	0.923 (0.922,0.923)	23	0.923 (0.922,0.924)	7	0.906	10
11	{D,0,0,0}			1000	1000.0	0.999 (1.000-52.975)	0.873 (0.864,0.881)	47	0.922 (0.921,0.923)	25	0.921 (0.920,0.922)	8	0.905	11
12	{e,0,D,e} 2000:1,1,1	•	UR	2000	2000.0	2.000 (1.000-53.244)	0.926 (0.924,0.928)	3	0.885 (0.885,0.886)	42	0.904 (0.904,0.905)	22	0.905	12
13	{p,0,0,0}	•	BT	1000	1000.0	0.999 (1.000-12.375)	0.921 (0.916,0.925)	7	0.877 (0.876,0.879)	48	0.916 (0.914,0.917)	10	0.905	13
14	{e,0,D,e} 1000:1,1,1	•	BT	1000	1000.0	1.999 (1.000-41.050)	0.921 (0.917,0.924)	8	0.886 (0.885,0.887)	40	0.899 (0.898,0.901)	25	0.902	14
15	{p,0,D,p} 1000:1,1,1	•	BT	1000	1000.0	1.999 (1.000-41.575)	0.917 (0.913,0.921)	10	0.884 (0.883,0.886)	44	0.901 (0.900,0.902)	24	0.901	15
16	{p,0,D,p} 2000:1,1,1	•	UR	2000	2000.0	2.000 (1.000-50.256)	0.922 (0.920,0.923)	6	0.878 (0.878,0.879)	45	0.902 (0.902,0.903)	23	0.901	16
17	{e,0,D,e} 1000:1,1,1	•	UR	1000	1000.0	1.999 (1.000-43.119)	0.919 (0.917,0.921)	9	0.885 (0.884,0.886)	43	0.894 (0.893,0.895)	27	0.899	17
18	{e,0,0,0}	•	BT	1000	1000.0	0.999 (1.000-11.175)	0.912 (0.907,0.917)	13	0.871 (0.870,0.873)	49	0.910 (0.909,0.912)	11	0.898	18
19	{p,0,D,p} 1000:1,1,1	•	UR	1000	1000.0	1.999 (1.000-40.969)	0.917 (0.915,0.919)	11	0.878 (0.877,0.878)	47	0.896 (0.895,0.897)	26	0.897	19
20	{e,0,0,0}	•	UR	1000	1000.0	0.999	0.910	14	0.870	50	0.909	13=	0.896	20

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
						(1.000-10.919)	(0.907,0.912)		(0.869,0.871)		(0.909,0.910)			
21	{e,0,0,0}	•	BT	2000	2000.0	1.000 (1.000-12.250)	0.907 (0.902,0.911)	16	0.870 (0.869,0.871)	51	0.909 (0.908,0.911)	13=	0.895	21
22	{e,0,0,0}	•	UR	2000	2000.0	1.000 (1.000-12.144)	0.907 (0.905,0.909)	15=	0.868 (0.868,0.869)	52	0.909 (0.908,0.909)	14	0.895	22
23	{p,0,0,0}	•	UR	1000	1000.0	0.999 (1.000-10.537)	0.904 (0.901,0.906)	18	0.866 (0.865,0.866)	54=	0.906 (0.905,0.906)	18	0.892	23
24	{r,0,0,0}	•		2000	2000.0	1.000 (1.000-11.800)	0.902 (0.897,0.906)	20	0.866 (0.865,0.867)	54=	0.906 (0.905,0.907)	16	0.891	24
25	{r,0,0,0}	•		1000	1000.0	0.999 (1.000-10.775)	0.897 (0.890,0.904)	26	0.868 (0.866,0.869)	53	0.907 (0.905,0.909)	15	0.891	25
26	{e,0,g,e} 22000:11,1,1	•	UR	2000	2000.0	11.992 (12.000-167.000)	0.904 (0.903,0.906)	17	0.921 (0.921,0.922)	27=	0.845 (0.844,0.846)	57	0.890	26
27	{p,0,0,0}	•	UR	2000	2000.0	1.000 (1.000-11.681)	0.900 (0.897,0.902)	23	0.865 (0.864,0.865)	55	0.905 (0.905,0.906)	20	0.890	27
28	{r,0,0,0}	•		4000	4000.0	1.000 (1.000-12.925)	0.899 (0.895,0.903)	25	0.864 (0.864,0.865)	56	0.905 (0.905,0.906)	19	0.889	28
29	{e,0,g,e} 22000:11,1,1	•	BT	2000	2000.0	11.992 (12.000-167.000)	0.907 (0.903,0.910)	15=	0.924 (0.924,0.925)	17	0.836 (0.834,0.837)	65	0.889	29
30	{e,0,g,e} 90000:45,1,4	•	UR	2000	2000.0	12.242 (12.000-170.000)	0.899 (0.898,0.900)	24	0.921 (0.920,0.921)	30	0.843 (0.842,0.844)	60	0.887	30
31	{r,0,0,0}	•		10000	10000.0	1.000 (1.000-14.200)	0.893 (0.887,0.898)	31	0.863 (0.862,0.863)	57	0.904 (0.904,0.905)	21	0.887	31
32	{e,0,g,e} 90000:45,1,4	•	BT	2000	2000.0	12.242 (12.000-170.000)	0.901 (0.898,0.904)	21	0.923 (0.922,0.923)	22	0.831 (0.830,0.833)	71	0.885	32
33	{e,0,g,e} 22000:11,1,1		BT	2000	2000.0	11.992 (12.000-168.000)	0.886 (0.883,0.888)	35	0.927 (0.927,0.928)	13	0.837 (0.835,0.838)	64	0.883	33
34	{e,0,g,e} 22000:11,1,1		UR	2000	2000.0	11.992 (12.000-167.656)	0.878 (0.877,0.880)	40	0.925 (0.925,0.926)	16	0.843 (0.842,0.844)	59	0.882	34
35	{e,0,g,e} 90000:45,1,4		BT	2000	2000.0	12.242 (12.000-170.775)	0.875 (0.872,0.878)	44	0.926 (0.925,0.927)	14	0.833 (0.832,0.835)	66	0.878	35
36	{e,0,g,e} 90000:45,1,4		UR	2000	2000.0	12.242 (12.000-170.787)	0.868 (0.866,0.869)	53	0.924 (0.924,0.924)	18	0.842 (0.841,0.843)	61	0.878	36

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
37	{e,0,g,e} 11000:11,1,1	•	UR	1000	1000.0	11.985 (12.000-167.000)	0.875 (0.873,0.878)	43	0.919 (0.919,0.920)	32	0.838 (0.836,0.839)	63	0.878	37
38	{e,0,D,e} 2000:1,1,1		BT	2000	2000.0	2.000 (1.000-46.875)	0.893 (0.890,0.896)	30	0.849 (0.847,0.850)	61	0.885 (0.884,0.886)	30=	0.876	38
39	{e,0,g,e} 45000:45,1,4	•	UR	1000	1000.0	12.235 (12.000-170.000)	0.866 (0.864,0.868)	55	0.918 (0.918,0.919)	33	0.838 (0.836,0.840)	62	0.874	39
40	{e,0,D,e} 2000:1,1,1		UR	2000	2000.0	2.000 (1.000-48.881)	0.890 (0.889,0.892)	33	0.848 (0.848,0.849)	62	0.882 (0.881,0.883)	32	0.874	40
41	{p,0,D,p} 2000:1,1,1		BT	2000	2000.0	2.000 (1.000-45.725)	0.896 (0.893,0.900)	27	0.848 (0.847,0.849)	64	0.875 (0.873,0.876)	39	0.873	41
42	{e,0,g,e} 11000:11,1,1	•	BT	1000	1000.0	11.985 (12.000-167.000)	0.879 (0.875,0.883)	39	0.922 (0.921,0.923)	26	0.807 (0.804,0.809)	93	0.869	42
43	{p,0,D,p} 2000:1,1,1		UR	2000	2000.0	2.000 (1.000-48.419)	0.884 (0.883,0.886)	36	0.842 (0.842,0.843)	67=	0.877 (0.877,0.878)	35	0.868	43
44	{e,0,D,e} 1000:1,1,1		BT	1000	1000.0	1.999 (1.000-36.775)	0.877 (0.873,0.881)	41	0.849 (0.847,0.850)	60	0.875 (0.874,0.877)	37	0.867	44=
45	{p,0,g,p} 22000:11,1,1	•	UR	2000	2000.0	11.992 (12.000-167.000)	0.900 (0.899,0.901)	22	0.917 (0.917,0.918)	35	0.784 (0.782,0.785)	113	0.867	44=
46	{p,0,D,p} 1000:1,1,1		BT	1000	1000.0	1.999 (1.000-35.975)	0.880 (0.877,0.884)	37	0.847 (0.844,0.849)	65	0.868 (0.866,0.871)	47	0.865	45
47	{e,0,g,e} 11000:11,1,1		UR	1000	1000.0	11.985 (12.000-168.350)	0.838 (0.836,0.840)	66	0.923 (0.922,0.923)	21=	0.833 (0.831,0.835)	68	0.865	46
48	{e,0,D,e} 1000:1,1,1		UR	1000	1000.0	1.999 (1.000-38.056)	0.874 (0.872,0.876)	46	0.848 (0.847,0.849)	63	0.870 (0.869,0.871)	44	0.864	47
49	{p,0,g,p} 90000:45,1,4	•	UR	2000	2000.0	12.242 (12.000-170.000)	0.894 (0.893,0.896)	29	0.917 (0.916,0.917)	36	0.782 (0.780,0.783)	118	0.864	48
50	{e,0,g,e} 45000:45,1,4	•	BT	1000	1000.0	12.235 (12.000-170.000)	0.867 (0.864,0.871)	54	0.921 (0.920,0.922)	29	0.803 (0.800,0.806)	97	0.864	49
51	{p,0,g,p} 22000:11,1,1		UR	2000	2000.0	11.992 (12.000-167.925)	0.872 (0.871,0.874)	48	0.923 (0.922,0.923)	24	0.789 (0.787,0.790)	110	0.861	50
52	{e,0,g,e} 45000:45,1,4		UR	1000	1000.0	12.235 (12.000-171.000)	0.827 (0.825,0.829)	73	0.921 (0.921,0.922)	28	0.833 (0.832,0.835)	67	0.861	51

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
53	{e,0,g,e} 11000:11,1,1		BT	1000	1000.0	11.985 (12.000-167.825)	0.849 (0.844,0.853)	63	0.926 (0.925,0.926)	15	0.806 (0.803,0.809)	95	0.860	52
54	{p,0,D,p} 1000:1,1,1		UR	1000	1000.0	1.999 (1.000-37.500)	0.865 (0.863,0.867)	56	0.842 (0.841,0.842)	68	0.870 (0.869,0.871)	46	0.859	53
55	{p,0,g,p} 11000:11,1,1	•	UR	1000	1000.0	11.985 (12.000-167.000)	0.868 (0.866,0.871)	52	0.915 (0.914,0.915)	37	0.789 (0.787,0.791)	109	0.857	54
56	{p,0,g,p} 90000:45,1,4		UR	2000	2000.0	12.242 (12.000-170.963)	0.861 (0.860,0.863)	57	0.921 (0.921,0.922)	27=	0.786 (0.784,0.787)	112	0.856	55
57	{p,0,D,p} 4000:1,1,1	•	BT	2000	2000.0	2.999 (1.000-66.125)	0.874 (0.871,0.877)	45=	0.826 (0.825,0.828)	82	0.866 (0.864,0.867)	48	0.855	56
58	{e,0,D,e} 4000:1,1,1	•	BT	2000	2000.0	2.999 (1.000-64.125)	0.870 (0.867,0.874)	50	0.833 (0.832,0.834)	74	0.861 (0.860,0.862)	52	0.855	57
59	{p,0,g,p} 45000:45,1,4	•	UR	1000	1000.0	12.235 (12.000-170.000)	0.861 (0.859,0.863)	58	0.914 (0.913,0.914)	38	0.789 (0.787,0.790)	111	0.854	58
60	{e,0,D,e} 4000:1,1,1	•	UR	2000	2000.0	2.999 (1.000-67.081)	0.870 (0.868,0.872)	51	0.832 (0.832,0.833)	75	0.855 (0.854,0.856)	54	0.852	59
61	{e,0,g,e} 45000:45,1,4		BT	1000	1000.0	12.235 (12.000-170.625)	0.833 (0.829,0.838)	69	0.923 (0.922,0.924)	21=	0.800 (0.798,0.803)	103	0.852	60
62	{p,0,g,p} 22000:11,1,1	•	BT	2000	2000.0	11.992 (12.000-167.000)	0.923 (0.920,0.926)	5	0.940 (0.940,0.941)	10	0.691 (0.689,0.693)	139	0.851	61
63	{p,0,g,p} 11000:11,1,1	•	BT	1000	1000.0	11.985 (12.000-167.000)	0.892 (0.887,0.896)	32	0.941 (0.940,0.942)	9	0.714 (0.710,0.717)	135	0.849	62
64	{p,0,g,p} 22000:11,1,1		BT	2000	2000.0	11.992 (12.000-168.525)	0.903 (0.900,0.906)	19	0.945 (0.944,0.945)	7	0.698 (0.696,0.701)	137	0.849	63
65	{p,0,g,p} 11000:11,1,1		BT	1000	1000.0	11.985 (12.000-167.700)	0.874 (0.870,0.878)	45=	0.947 (0.946,0.947)	5	0.725 (0.722,0.728)	133	0.849	64
66	{p,0,D,p} 4000:1,1,1	•	UR	2000	2000.0	2.999 (1.000-63.100)	0.860 (0.858,0.861)	59	0.827 (0.826,0.827)	79	0.859 (0.858,0.860)	53	0.849	65
67	{p,0,g,p} 90000:45,1,4	•	BT	2000	2000.0	12.242 (12.000-170.000)	0.916 (0.913,0.918)	12	0.940 (0.939,0.940)	11	0.690 (0.688,0.692)	140	0.848	66
68	{p,0,g,p} 11000:11,1,1		UR	1000	1000.0	11.985 (12.000-	0.830 (0.828,0.832)	71	0.919 (0.919,0.920)	31	0.794 (0.792,0.795)	107	0.848	67

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
						168.244)								
69	{p,0,D,p} 2000:1,1,1	•	BT	1000	1000.0	2.999 (1.000-52.700)	0.849 (0.845,0.853)	62	0.827 (0.825,0.829)	78	0.864 (0.862,0.866)	50	0.847	68
70	{p,0,g,p} 45000:45,1,4	•	BT	1000	1000.0	12.235 (12.000-170.000)	0.888 (0.884,0.892)	34	0.940 (0.939,0.940)	12	0.710 (0.707,0.714)	136	0.846	69
71	{p,0,g,p} 14000:7,1,11	•	BT	2000	2000.0	1.635 (1.000-22.075)	0.830 (0.827,0.833)	70	0.828 (0.827,0.829)	77	0.879 (0.878,0.880)	34	0.846	70
72	{p,0,g,p} 90000:45,1,4		BT	2000	2000.0	12.242 (12.000-171.575)	0.896 (0.893,0.898)	28	0.944 (0.943,0.944)	8	0.694 (0.692,0.697)	138	0.845	71
73	{e,0,D,e} 2000:1,1,1	•	BT	1000	1000.0	2.999 (1.000-52.725)	0.854 (0.850,0.857)	61	0.831 (0.829,0.833)	76	0.848 (0.846,0.850)	56	0.844	72
74	{p,0,g,p} 45000:45,1,4		UR	1000	1000.0	12.235 (12.000-171.031)	0.817 (0.814,0.820)	76	0.918 (0.918,0.918)	34	0.793 (0.791,0.795)	108	0.843	73
75	{e,0,D,e} 2000:1,1,1	•	UR	1000	1000.0	2.999 (1.000-55.062)	0.848 (0.846,0.850)	65	0.833 (0.832,0.834)	73	0.844 (0.843,0.845)	58	0.842	74
76	{p,0,g,p} 45000:45,1,4		BT	1000	1000.0	12.235 (12.000-170.525)	0.858 (0.855,0.862)	60	0.945 (0.944,0.946)	6	0.719 (0.716,0.723)	134	0.841	75
77	{p,0,D,p} 2000:1,1,1	•	UR	1000	1000.0	2.999 (1.000-52.131)	0.835 (0.833,0.837)	67	0.827 (0.826,0.828)	80	0.854 (0.853,0.855)	55	0.839	76
78	{p,0,g,p} 42000:21,1,25	•	BT	2000	2000.0	1.839 (1.000-24.675)	0.776 (0.771,0.780)	90	0.852 (0.851,0.854)	59	0.886 (0.884,0.887)	29	0.838	77
79	{p,0,g,p} 7000:7,1,11	•	BT	1000	1000.0	1.635 (1.000-18.650)	0.807 (0.802,0.812)	79	0.826 (0.825,0.828)	81	0.876 (0.875,0.878)	36	0.837	78
80	{e,0,g,e} 14000:7,1,11	•	BT	2000	2000.0	1.635 (1.000-17.600)	0.802 (0.797,0.807)	80	0.818 (0.817,0.819)	84	0.875 (0.874,0.876)	38	0.832	79
81	{e,0,g,e} 14000:7,1,11	•	UR	2000	2000.0	1.635 (1.000-17.756)	0.798 (0.796,0.801)	81	0.816 (0.815,0.816)	86	0.872 (0.871,0.873)	42	0.829	80
82	{e,0,g,e} 7000:7,1,11	•	BT	1000	1000.0	1.635 (1.000-15.925)	0.783 (0.777,0.790)	87	0.818 (0.816,0.820)	83	0.872 (0.870,0.874)	41	0.825	81
83	{e,0,g,e} 7000:7,1,11	•	UR	1000	1000.0	1.635 (1.000-16.156)	0.784 (0.781,0.787)	86	0.817 (0.816,0.817)	85	0.870 (0.869,0.871)	45	0.824	82
84	{p,0,g,p} 21000:21,1,25	•	BT	1000	1000.0	1.839 (1.000-20.450)	0.728 (0.721,0.734)	110	0.855 (0.854,0.857)	58	0.882 (0.880,0.884)	33=	0.822	83
85	{p,0,g,p} 14000:7,1,11	•	UR	2000	2000.0	1.635 (1.000-16.988)	0.788 (0.786,0.791)	84	0.811 (0.810,0.812)	88	0.865 (0.864,0.866)	49	0.821	84
86	{e,0,g,e} 42000:21,1,25	•	UR	2000	2000.0	1.839 (1.000-19.556)	0.725 (0.722,0.728)	111	0.841 (0.841,0.842)	70	0.885 (0.885,0.886)	30=	0.817	85

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
87	{e,0,g,e} 42000:21,1,25	•	BT	2000	2000.0	1.839 (1.000-19.175)	0.722 (0.717,0.728)	112	0.841 (0.840,0.843)	69	0.886 (0.885,0.888)	28	0.817	86
88	{p,0,D,p} 4000:1,1,1		BT	2000	2000.0	2.999 (1.000-57.175)	0.822 (0.817,0.826)	75	0.792 (0.790,0.794)	95	0.832 (0.830,0.835)	69=	0.815	87
89	{p,0,g,p} 7000:7,1,11	•	UR	1000	1000.0	1.635 (1.000-15.481)	0.770 (0.767,0.773)	92	0.811 (0.810,0.812)	87	0.863 (0.862,0.864)	51	0.815	88
90	{e,0,D,e} 4000:1,1,1		BT	2000	2000.0	2.999 (1.000-57.950)	0.810 (0.807,0.814)	78	0.797 (0.796,0.799)	92	0.832 (0.830,0.834)	70	0.813	89
91	{e,0,D,e} 4000:1,1,1		UR	2000	2000.0	2.999 (1.000-61.650)	0.812 (0.810,0.813)	77	0.798 (0.797,0.799)	90	0.829 (0.828,0.830)	72=	0.813	90
92	{p,0,g,p} 42000:21,1,25	•	UR	2000	2000.0	1.839 (1.000-18.863)	0.714 (0.711,0.717)	113	0.837 (0.836,0.837)	72	0.874 (0.874,0.875)	40	0.808	91
93	{p,0,D,p} 4000:1,1,1		UR	2000	2000.0	2.999 (1.000-60.025)	0.791 (0.789,0.793)	82	0.796 (0.795,0.797)	93	0.832 (0.831,0.833)	69=	0.806	92
94	{e,0,g,e} 21000:21,1,25	•	UR	1000	1000.0	1.839 (1.000-17.850)	0.693 (0.689,0.698)	114	0.842 (0.841,0.843)	67=	0.882 (0.881,0.883)	33=	0.806	93
95	{e,0,g,e} 21000:21,1,25	•	BT	1000	1000.0	1.839 (1.000-17.050)	0.681 (0.673,0.689)	115	0.844 (0.842,0.846)	66	0.885 (0.883,0.886)	31	0.803	94
96	{p,0,D,p} 2000:1,1,1		BT	1000	1000.0	2.999 (1.000-45.800)	0.787 (0.781,0.793)	85	0.791 (0.788,0.794)	96	0.829 (0.826,0.831)	72=	0.802	95
97	{e,0,D,e} 2000:1,1,1		BT	1000	1000.0	2.999 (1.000-46.425)	0.777 (0.773,0.781)	89	0.798 (0.795,0.801)	91	0.820 (0.817,0.824)	78	0.798	96
98	{p,0,0,0}		BT	2000	2000.0	1.000 (1.000-7.600)	0.789 (0.782,0.797)	83	0.776 (0.774,0.778)	97	0.827 (0.824,0.829)	74	0.797	97
99	{p,0,g,p} 21000:21,1,25	•	UR	1000	1000.0	1.839 (1.000-17.050)	0.680 (0.676,0.684)	116	0.839 (0.838,0.839)	71	0.872 (0.871,0.873)	43	0.797	98
100	{e,0,D,e} 2000:1,1,1		UR	1000	1000.0	2.999 (1.000-48.369)	0.773 (0.771,0.775)	91	0.799 (0.798,0.801)	89	0.817 (0.815,0.818)	82	0.796	99
101	{p,0,0,0}		BT	1000	1000.0	0.999 (1.000-7.025)	0.783 (0.774,0.792)	88	0.772 (0.769,0.774)	99	0.822 (0.819,0.825)	77	0.792	100
102	{p,0,D,p} 2000:1,1,1		UR	1000	1000.0	2.999 (1.000-47.406)	0.748 (0.746,0.750)	101=	0.796 (0.794,0.797)	94	0.825 (0.824,0.827)	76	0.790	101
103	{0,0,r,e} 1000:1,1,1		BT	1000	864.5	1.000 (0.000-7.875)	0.749 (0.741,0.757)	100=	0.763 (0.760,0.766)	102=	0.819 (0.815,0.822)	80=	0.777	102=
104	{0,0,r,e} 1000:1,1,1	•	BT	1000	864.5	1.000 (0.000-7.875)	0.749 (0.741,0.757)	100=	0.763 (0.760,0.766)	102=	0.819 (0.815,0.822)	80=	0.777	102=
105	{0,0,r,e} 1000:1,1,1		UR	1000	864.0	1.000 (0.000-8.125)	0.747 (0.743,0.752)	102=	0.764 (0.762,0.765)	101=	0.819 (0.817,0.820)	79=	0.777	103=
106	{0,0,r,e} 1000:1,1,1	•	UR	1000	864.0	1.000 (0.000-8.125)	0.747 (0.743,0.752)	102=	0.764 (0.762,0.765)	101=	0.819 (0.817,0.820)	79=	0.777	103=
107	{0,0,r,p}	•	BT	1000	861.1	1.000	0.750	99=	0.760	103=	0.817	81=	0.776	104=

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
	1000:1,1,1					(0.000-8.100)	(0.742,0.758)		(0.757,0.763)		(0.814,0.820)			
108	{0,0,r,p} 1000:1,1,1		BT	1000	861.1	1.000 (0.000-8.100)	0.750 (0.742,0.758)	99=	0.760 (0.757,0.763)	103=	0.817 (0.814,0.820)	81=	0.776	104=
109	{e,0,0,0}		BT	2000	2000.0	1.000 (1.000-6.750)	0.754 (0.748,0.761)	98	0.759 (0.758,0.761)	105	0.809 (0.808,0.811)	91=	0.774	105
110	{e,0,0,0}		UR	1000	1000.0	0.999 (1.000-6.606)	0.757 (0.753,0.762)	95	0.758 (0.756,0.759)	109	0.808 (0.806,0.809)	92	0.774	106
111	{r,0,0,0}			1000	1000.0	0.999 (1.000-6.650)	0.759 (0.751,0.766)	93=	0.756 (0.753,0.758)	113	0.806 (0.803,0.808)	96	0.773	107
112	{e,0,0,0}		UR	2000	2000.0	1.000 (1.000-6.850)	0.757 (0.753,0.761)	96	0.756 (0.755,0.757)	112=	0.806 (0.805,0.807)	94	0.773	108
113	{0,0,r,p} 2000:1,1,1	•	BT	2000	1719.8	1.000 (0.000-8.550)	0.745 (0.737,0.752)	104=	0.759 (0.757,0.761)	106=	0.816 (0.814,0.818)	83=	0.773	109=
114	{0,0,r,p} 2000:1,1,1		BT	2000	1719.8	1.000 (0.000-8.550)	0.745 (0.737,0.752)	104=	0.759 (0.757,0.761)	106=	0.816 (0.814,0.818)	83=	0.773	109=
115	{e,0,0,0}		BT	1000	1000.0	0.999 (1.000-6.500)	0.748 (0.739,0.757)	101=	0.760 (0.757,0.762)	104	0.809 (0.807,0.812)	91=	0.772	110
116	{0,0,r,e} 2000:1,1,1		BT	2000	1722.1	1.000 (0.000-8.400)	0.741 (0.733,0.748)	106=	0.758 (0.757,0.760)	107=	0.814 (0.812,0.816)	84=	0.771	111=
117	{0,0,r,e} 2000:1,1,1	•	BT	2000	1722.1	1.000 (0.000-8.400)	0.741 (0.733,0.748)	106=	0.758 (0.757,0.760)	107=	0.814 (0.812,0.816)	84=	0.771	111=
118	{r,0,0,0}			2000	2000.0	1.000 (1.000-6.925)	0.759 (0.752,0.765)	93=	0.752 (0.750,0.754)	116	0.803 (0.801,0.804)	98	0.771	112
119	{0,0,r,p} 1000:1,1,1	•	UR	1000	869.7	1.000 (0.000-7.862)	0.741 (0.736,0.745)	105=	0.758 (0.756,0.759)	108=	0.814 (0.812,0.815)	85=	0.771	113=
120	{0,0,r,p} 1000:1,1,1		UR	1000	869.7	1.000 (0.000-7.862)	0.741 (0.736,0.745)	105=	0.758 (0.756,0.759)	108=	0.814 (0.812,0.815)	85=	0.771	113=
121	{r,0,0,0}			4000	4000.0	1.000 (1.000-7.350)	0.758 (0.751,0.764)	94	0.751 (0.750,0.752)	117	0.802 (0.801,0.803)	100=	0.770	114
122	{0,0,r,e} 2000:1,1,1		UR	2000	1725.6	1.000 (0.000-8.281)	0.738 (0.735,0.742)	107=	0.756 (0.756,0.757)	111=	0.812 (0.811,0.813)	87=	0.769	115=
123	{0,0,r,e} 2000:1,1,1	•	UR	2000	1725.6	1.000 (0.000-8.281)	0.738 (0.735,0.742)	107=	0.756 (0.756,0.757)	111=	0.812 (0.811,0.813)	87=	0.769	115=
124	{r,0,0,0}			10000	10000.0	1.000 (1.000-7.550)	0.756 (0.751,0.761)	97	0.748 (0.747,0.749)	122	0.799 (0.798,0.800)	104	0.768	116
125	{0,0,r,p} 2000:1,1,1	•	UR	2000	1738.2	1.000 (0.000-8.238)	0.738 (0.734,0.742)	109=	0.753 (0.752,0.754)	115=	0.809 (0.808,0.810)	91=	0.767	117=
126	{0,0,r,p} 2000:1,1,1		UR	2000	1738.2	1.000 (0.000-8.238)	0.738 (0.734,0.742)	109=	0.753 (0.752,0.754)	115=	0.809 (0.808,0.810)	91=	0.767	117=
127	{p,0,0,0}		UR	2000	2000.0	1.000 (1.000-6.744)	0.747 (0.743,0.751)	103	0.749 (0.748,0.750)	121	0.801 (0.800,0.802)	102	0.766	118

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
128	{p,0,0,0}		UR	1000	1000.0	0.999 (1.000-6.375)	0.738 (0.733,0.743)	108	0.750 (0.749,0.752)	118	0.802 (0.801,0.804)	99	0.764	119
129	{p,0,g,p} 14000:7,1,11		BT	2000	2000.0	1.635 (1.000-13.700)	0.663 (0.653,0.673)	117	0.730 (0.727,0.732)	123	0.796 (0.793,0.798)	105	0.730	120
130	{p,0,g,p} 42000:21,1,25		BT	2000	2000.0	1.839 (1.000-15.675)	0.576 (0.567,0.586)	122	0.773 (0.771,0.775)	98	0.826 (0.824,0.828)	75	0.725	121
131	{p,0,g,p} 7000:7,1,11		BT	1000	1000.0	1.635 (1.000-12.125)	0.609 (0.596,0.622)	118	0.725 (0.723,0.728)	124	0.795 (0.792,0.797)	106	0.710	122
132	{p,0,g,p} 21000:21,1,25		BT	1000	1000.0	1.839 (1.000-14.075)	0.515 (0.503,0.527)	130	0.771 (0.768,0.774)	100	0.827 (0.825,0.829)	73	0.704	123
133	{e,0,g,e} 14000:7,1,11		BT	2000	2000.0	1.635 (1.000-11.600)	0.599 (0.589,0.609)	119	0.716 (0.714,0.719)	126	0.783 (0.781,0.786)	114	0.700	124
134	{e,0,g,e} 14000:7,1,11		UR	2000	2000.0	1.635 (1.000-11.750)	0.597 (0.592,0.602)	120	0.715 (0.714,0.716)	128	0.782 (0.781,0.783)	117	0.698	125
135	{e,0,g,e} 42000:21,1,25		BT	2000	2000.0	1.839 (1.000-13.200)	0.510 (0.497,0.523)	132	0.756 (0.754,0.758)	112=	0.813 (0.811,0.814)	86	0.693	126
136	{e,0,g,e} 42000:21,1,25		UR	2000	2000.0	1.839 (1.000-13.231)	0.504 (0.497,0.511)	133	0.754 (0.753,0.755)	114	0.811 (0.810,0.811)	90	0.690	127
137	{e,0,g,e} 7000:7,1,11		UR	1000	1000.0	1.635 (1.000-10.938)	0.565 (0.558,0.572)	123	0.717 (0.715,0.718)	125	0.783 (0.782,0.785)	115	0.688	128
138	{p,0,g,p} 14000:7,1,11		UR	2000	2000.0	1.635 (1.000-11.387)	0.580 (0.574,0.586)	121	0.708 (0.707,0.709)	129	0.775 (0.774,0.776)	119	0.688	129
139	{e,0,g,e} 7000:7,1,11		BT	1000	1000.0	1.635 (1.000-10.650)	0.551 (0.536,0.567)	125	0.716 (0.713,0.719)	127	0.783 (0.780,0.786)	116	0.683	130
140	{p,0,g,p} 42000:21,1,25		UR	2000	2000.0	1.839 (1.000-12.838)	0.481 (0.474,0.488)	137	0.749 (0.748,0.750)	120	0.802 (0.801,0.803)	101	0.677	131
141	{e,0,g,e}		UR	1000	1000.0	1.839 (1.000-12.450)	0.449 (0.442,0.457)	141	0.757 (0.755,0.758)	110	0.811 (0.810,0.813)	89	0.672	132=
142	{p,0,g,p} 7000:7,1,11		UR	1000	1000.0	1.635 (1.000-10.500)	0.535 (0.527,0.542)	126	0.708 (0.706,0.709)	130	0.775 (0.773,0.776)	120	0.672	132=
143	{e,0,g,e}		BT	1000	1000.0	1.839 (1.000-12.350)	0.443 (0.426,0.460)	142	0.756 (0.753,0.759)	112=	0.812 (0.809,0.815)	88	0.670	133
144	{p,0,g,p} 21000:21,1,25		UR	1000	1000.0	1.839 (1.000-12.056)	0.429 (0.421,0.437)	144	0.749 (0.747,0.751)	119	0.802 (0.801,0.804)	100=	0.660	134
145	{0,0,r,p} 3680:1,1,1		BT	2000	1945.0	1.840 (0.000-12.325)	0.553 (0.542,0.564)	124	0.675 (0.673,0.677)	131	0.751 (0.748,0.753)	122	0.659	135
146	{0,0,r,e} 3680:1,1,1		BT	2000	1946.0	1.840 (0.000-11.950)	0.533 (0.521,0.544)	127	0.672 (0.669,0.674)	132	0.739 (0.737,0.742)	124	0.648	136
147	{0,0,r,e} 3680:1,1,1		UR	2000	1948.6	1.840 (0.000-11.831)	0.533 (0.528,0.538)	128	0.669 (0.668,0.670)	133	0.737 (0.736,0.738)	125	0.646	137
148	{0,0,r,p}	•	BT	2000	1960.4	2.000	0.522	129=	0.666	134=	0.744	123=	0.644	138=

Expt. #	GRAB tuple Cycles:stepG, stepR,stepA	IG	GEO	Target network size, nodes	Actual network size, nodes	Ave. degree (ave min D-Ave max D)	DFD R^2 (left and right 95% CI)	DFD R^2 rank	DR R^2 (left and right 95% CI)	DR R^2 rank	CCR R^2 (left and right 95% CI)	CCR R^2 rank	Ave R^2	Ave R^2 rank
	4000:1,1,1					(0.000-13.200)	(0.510,0.534)		(0.664,0.669)		(0.742,0.747)			
149	{0,0,r,p} 4000:1,1,1		BT	2000	1960.4	2.000 (0.000-13.200)	0.522 (0.510,0.534)	129=	0.666 (0.664,0.669)	134=	0.744 (0.742,0.747)	123=	0.644	138=
150	{0,0,r,p} 3680:1,1,1		UR	2000	1955.4	1.840 (0.000-11.575)	0.513 (0.507,0.518)	131	0.663 (0.662,0.664)	137	0.736 (0.734,0.737)	126	0.637	139
151	{0,0,r,p} 2000:1,1,1	•	BT	1000	979.7	2.000 (0.000-12.200)	0.482 (0.470,0.493)	136=	0.665 (0.662,0.667)	135=	0.752 (0.749,0.754)	121=	0.633	140=
152	{0,0,r,p} 2000:1,1,1		BT	1000	979.7	2.000 (0.000-12.200)	0.482 (0.470,0.493)	136=	0.665 (0.662,0.667)	135=	0.752 (0.749,0.754)	121=	0.633	140=
153	{0,0,r,e} 4000:1,1,1		BT	2000	1960.5	2.000 (0.000-12.625)	0.494 (0.483,0.505)	134=	0.662 (0.660,0.665)	138=	0.730 (0.728,0.732)	129=	0.629	141=
154	{0,0,r,e} 4000:1,1,1	•	BT	2000	1960.5	2.000 (0.000-12.625)	0.494 (0.483,0.505)	134=	0.662 (0.660,0.665)	138=	0.730 (0.728,0.732)	129=	0.629	141=
155	{0,0,r,e} 4000:1,1,1		UR	2000	1962.5	2.000 (0.000-12.469)	0.490 (0.485,0.496)	135=	0.660 (0.659,0.661)	140=	0.727 (0.726,0.728)	131=	0.626	142=
156	{0,0,r,e} 4000:1,1,1	•	UR	2000	1962.5	2.000 (0.000-12.469)	0.490 (0.485,0.496)	135=	0.660 (0.659,0.661)	140=	0.727 (0.726,0.728)	131=	0.626	142=
157	{0,0,r,e} 2000:1,1,1		UR	1000	981.2	2.000 (0.000-11.988)	0.464 (0.458,0.471)	139=	0.664 (0.662,0.666)	136=	0.730 (0.729,0.732)	128=	0.620	143=
158	{0,0,r,e} 2000:1,1,1	•	UR	1000	981.2	2.000 (0.000-11.988)	0.464 (0.458,0.471)	139=	0.664 (0.662,0.666)	136=	0.730 (0.729,0.732)	128=	0.620	143=
159	{0,0,r,e} 2000:1,1,1		BT	1000	980.6	2.000 (0.000-11.900)	0.464 (0.450,0.477)	140=	0.662 (0.660,0.665)	139=	0.729 (0.726,0.732)	130=	0.618	144=
160	{0,0,r,e} 2000:1,1,1	•	BT	1000	980.6	2.000 (0.000-11.900)	0.464 (0.450,0.477)	140=	0.662 (0.660,0.665)	139=	0.729 (0.726,0.732)	130=	0.618	144=
161	{0,0,r,p} 4000:1,1,1	•	UR	2000	1968.2	2.000 (0.000-12.175)	0.469 (0.463,0.475)	138=	0.654 (0.653,0.655)	142=	0.727 (0.726,0.728)	132=	0.617	145=
162	{0,0,r,p} 4000:1,1,1		UR	2000	1968.2	2.000 (0.000-12.175)	0.469 (0.463,0.475)	138=	0.654 (0.653,0.655)	142=	0.727 (0.726,0.728)	132=	0.617	145=
163	{0,0,r,p} 2000:1,1,1	•	UR	1000	984.4	2.000 (0.000-11.656)	0.439 (0.432,0.447)	143=	0.657 (0.656,0.659)	141=	0.732 (0.731,0.734)	127=	0.610	146=
164	{0,0,r,p} 2000:1,1,1		UR	1000	984.4	2.000 (0.000-11.656)	0.439 (0.432,0.447)	143=	0.657 (0.656,0.659)	141=	0.732 (0.731,0.734)	127=	0.610	146=

Bibliography

[8023i] The IEEE 802.3 Working Group "10BASE-T 10 Mbit/s over twisted pair" 1990

[8023u] The IEEE 802.3 Working Group "100BASE-TX, 100BASE-T4, 100BASE-FX Fast Ethernet at 100 Mbit/s" 1995

[8023ab] The IEEE 802.3 Working Group "1000BASE-T Gbit/s Ethernet over twisted pair at 1 Gbit/s" 1999

[8023an] The IEEE 802.3 Working Group "10GBASE-T 10 Gbit/s Ethernet over unshielded twisted pair" 2006

[8023ba] IEEE P802.3ba 40Gb/s and 100Gb/s Ethernet Task Force http://ieee802.org/3/ba/index.html

[ABE99] J. Abello, P.M. Pardalos, and M.G.C. Resende, "On maximum cliques in very large graphs", in "External memory algorithms," J. Abello and J. Vitter (eds.), DIMACS Series on Discrete Mathematics and Theoretical Computer Science, vol. 50, pp. 119-130, American Mathematical Society, 1999.

[ADA01] L. Adamic, R. Lukose, A. Puniyani, B. Huberman, "Search in power-law networks", Physical Review E, VOLUME 64, 046135.

[AIE00] W. Aiello, F. Chung and L. Lu: "A random graph model for massive graphs", in proc. of the 32rd Annual ACM Symposium on Theory of Computing, pp. 171–180, 2000.

[ALB00a] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks", NATURE, vol. 406, pp. 378-381, July 2000.

[ALB00b] R. Albert and A.-L. Barabási, "Topology of evolving networks: Local events and universality", Physical Review Letters 85, 5234 (2000).

[ALB99] R. Albert, H. Jeong, and A.-L. Barabasi. "Diameter of the world-wide web." Nature, 401:130--131, September 1999

[ANA02] Anagnostakis K, Greenwald M and Ryger R: 'On the sensitivity of network simulation to topology', Proceedings of the 10th IEEE/ACM Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS) (2002).

[BAR02] A.-L. Barabasi, H. Jeong, R. Ravasz, Z. Neda, T. Vicsek, A. Schubert, "Evolution of the social network of scientific collaborations" PHYSICA A 311 590 (2002).

[BAR99] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks", SCIENCE, vol. 286, pp. 509-512, Oct. 1999.

[BAR99b] Barabási, A-L.; Albert, R.; and Jeong, H. "Mean-Field Theory for Scale-Free Random Networks." Physica A 272, 173-187, 1999.

[BOL84] Béla Bollobás, "The Evolution of Random Graphs", Transactions of the American Mathematical Society, Vol. 286, No. 1 (Nov., 1984), pp. 257-274

[BRA05] Braess D., Nagurney A. and Wakolbinger T., "On a Paradox of Traffic Planning", Transportation Science Vol. 39, No. 4, November 2005, pp. 446–450

[BU02] T. Bu and D. Towsley. "On distinguishing between internet power law topology generators." In INFOCOM, 2002.

[CAH98] R. S. Cahn. Wide Area Network Design: Concepts and Tools for Optimisation. Morgan Kaufmann, 1998.

[CAR99] J. M. Carlson and J. Doyle. "Highly optimized tolerance: a mechanism for power laws in designed systems." Physics Review E, 60(2):1412–1427, 1999.

[CHE02] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, W. Willinger, "The Origin of Power Laws in Internet Topologies Revisited", Proceeding of IEEE Infocom 2002.

[CHU01] F. Chung and L. Lu, "The diameter of sparse random graphs", Adv. in Appl. Math. 26 (2001), 257-279. MR 1826308 (2002c:05138)

[COR01] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Cliff Stein, "Introduction to Algorithms (Second Edition)" , MIT Press 2001 ISBN 978-0262032933

[DAC00] Carlos Da Costa, "Planif 2000, a tool for planning SDH networks". ConfTele 2001 23-24 April 2001, Portugal

[DIMES] The DIMES project http://www.netdimes.org/ - A distributed Internet Mapping project

[DOA93] Matthew Doar and Ian Leslie, "How bad is naïve multicast routing?", In Proceedings of IEEE Infocom 1993 pp 82-89.

[EDW98] Edwars, J.A., "Protection Interworking for SDH networks", IEE Telecommunications 1998, pp202. 29 March – 1 April 1998, no 451.

[ERD60] Erdos, P.; Rényi, A. (1960). "The Evolution of Random Graphs". Publ. Math. Inst. Hung. Acad. Sci., 5: 17–61.

[FAB02] Alex Fabrikant, Elias Koutsoupias, Christos H. Papadimitriou , "Heuristically Optimized Trade-Offs: A New Paradigm for Power Laws in the Internet" LNCS Vol. 2380, Proceedings of the 29th International Colloquium on Automata, Languages and Programming pp 110 - 122, 2002

[FAL99] M. Faloutsos, P. Faloutsos, and C. Faloutsos. "On Power-Law Relationships of the Internet Topology." Proceedings of ACM SIGCOMM '99, pages 251–262, Aug. 1999.

[FOR03] Fortz B., Soriano P., Wynants C. "A tabu search algorithm for self-healing ring network design", European Journal of Operational Research 151 (2003) 280-295.

[G.7041] ITU-T G.7041, "G.7041 : Generic framing procedure (GFP)" 2005

[G.7402] ITU-T G.7042, "Link capacity adjustment scheme (LCAS) for virtual concatenated signals" 2006

[G.707] ITU-T G.707, "Network node interface for the synchronous digital hierarchy (SDH)" 1988-2007

[G.708] ITU-T G.708, "Sub STM-0 network node interface for the synchronous digital hierarchy (SDH)" 1999

[G.807] ITU-T recommendation G.807/Y.1302 "Requirements for the Automatic Switched Transport Network (ASTN)"

[G.841] ITU-T G.841, "Types and Characteristics of SDH Network Protection Architectures," 1995.

[G8080] ITU-T recommendation G.8080/Y.1304 "Architecture for the Automatic Switched Optical Network (ASON)"

[GHITLE] "GHITLE: Generator of Hierarchical Internet Topologies using Levels" Cedric de Launois, Department of Computer Science and Engineering (INGI), Université Catholique de Louvain, Belgium http://ghitle.info.ucl.ac.be/

[GOV00] Govindan R., Tangmunarunkit H., "Heuristics for Internet Map Discovery", Proc IEEE Infocom 2000, Tel Aviv, Israel

[GRI07] D. Griffin, J. Spencer, J. Griem, M. Boucadair, P. Morrand, M. Howarth, N. Wang, G. Pavlou, H. Asgari, P. Georgatsos, Inter-domain Routing through QoS Class Planes, IEEE Communications, special issue on Quality of Service Routing Algorithms for Heterogeneous Networks, Vol. 45, No. 2, pp. 88-95, IEEE, February 2007.

[GRO04] W. D. Grover. Mesh-Based Survivable Networks. Prentice Hall PTR, 2004.

[GTITM] E.W.Zegura, K. Calvert. "Modeling Topology of Large Internetworks " http://www.cc.gatech.edu/projects/gtitm/

[IGEN] "IGen: Topology generation through network design heuristics", Bruno Quoitin, CSE Department, Université Catholique de Louvain, Belgium. http://www.info.ucl.ac.be/~bqu/igen/

[INET20] Cheng Jin, Qian Chen, Sugih Jamin, "Inet: Internet Topology Generator" University of Michigan Technical Report CSE-TR-433-00 2000.

[INET30] Jared Winick, Sugih Jamin, "Inet-3.0: Internet Topology Generator" University of Michigan Technical Report CSE-TR-456-02 2002.

[INET] "Inet: 3.0: Autonomous System (AS) level Internet topology generator" Jared Winick et al. University of Michigan, http://topology.eecs.umich.edu/inet/

[JEO00] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.-L. Barabasi, "The large-scale organization of metabolic networks", Nature 407, 651 - 654 (05 Oct 2000).

[LAK02] A. Lakhina, J. Byers, M. Crovella, I. Matta, "On the Geographic Location of Internet Resources.", IEEE Journal on Selected Areas in Communications Vol. 21. No. 6 August 2002

[LAK03] A. Lakhina, J. Byers, M. Crovella, P. Xie, "Sampling Biases in IP Topology Measurements", IEEE Infocom 2003

[LIL01] F. Liljeros, C. Edling, L. Nunes Amaral, H. Eugene Stanley, Y. Aberg "The web of human sexual contacts", Nature 411, 907 - 908 (21 Jun 2001).

[MAG01] D. Magoni and J.–J. Pansiot: "Analysis of the Autonomous System network topology", ACM SIGCOMM Computer Communication Review, vol. 31, no. 3, pp. 26–37, July 2001.

[MCK97] Lee W. McKnight and Joseph P. Bailey, "Internet Economics", May 1997. MIT Press ISBN 0-262-13336-9

[MED01] A. Medina, A. Lakhina, I. Matta, and J. Byers, "BRITE: An Approach to Universal Topology Generation," in Proceedings of MASCOTS '01, August 2001.

[MIT94] D.J. Mitzel and S. Shenker. "Asymptotic Resource Consumption in Multicast Reservation Styles". Proc. of ACM SIGCOMM '94, 1994.

[MUR01] Murray, M. claffy, k., "Measuring the Immeasurable: Global Internet Measurement Infrastructure", Passive and Active Network Measurement Workshop (PAM) 2001

[NAL05] M. Naldi, "Connectivity of Waxman Topology Models", Computer Communications 29 (2005) 24-31

[NEW03] M. Newman, "Mixing patterns in networks", Phys. Rev. E 67, 026126 (2003)

[NIE06] Nieto-Hipolito, J.-I. et al. "AStop: a new topology generator at the Autonomous Systems level", IEEE Electronics, Robotics and Automotive Mechanics Conference, 2006. Volume 2, Issue , Sept. 2006 Page(s):355 – 360

[NLANR] NLANR, National laboratory for applied network research routing data 1999. http://moat.nlanr.net/Routing/rawdata/

[PAL00] Christopher Palmer and J. Gregory Steffan. "Generating network topologies that obey power-law." Proceeding of GLOBECOM, 2000.

[PAN98] J. Pansiot and D. Grad ,"On routes and multicast trees in the Internet", ACM Computer Communication Review, vol. 28, no. 1, pp. 41--50, Jan. 1998

[PAR01] K. Park and H. Lee. "On the effectiveness of route-based packet filtering for distributed dos attack prevention in power-law internets." In In. Proc. of ACM SIGCOMM, 2001.

[PEI92] Peitgen, H. O., Jürgens, H. and Saupe, D.: "Length, area and dimension: Measuring complexity and scaling properties", Chapter 4 in: 'Chaos and Fractals: New Frontiers of Science', New York, Springer-Verlag (1992).

[QBGP] M. Boucadair (Ed.), "QoS-Enhanced Border Gateway Protocol", draft-boucadair-qos-bgp-spec-01.txt, Work in Progress, July 2005.

[QUO07] Bruno Quoitin, "Towards More Representative Internet Topologies", Technical Report 7th September 2007, http://www.info.ucl.ac.be/~bqu/igen/

[RAD01] Rados I, Turalija, Sunaric T, "Availability Model of Bidirectional Switched Ring" International Conference on Transparent Optical Networks 2001 p312.

[RAS90] Rasband, S. N. "Fractal Dimension.". Ch. 4 in "Chaotic Dynamics of Nonlinear Systems" New York: Wiley, 1990

[RFC0768] J. Postel, "User Datagram Protocol", IETF RFC 768 http://www.ietf.org/rfc/rfc0768.txt

[RFC0791] J. Postel ed, "The Internet Protocol", IETF RFC 791 http://www.ietf.org/rfc/rfc0791.txt

[RFC0792] J. Postel, " Internet Control Message Protocol", IETF RFC 792 http://www.ietf.org/rfc/rfc0792.txt

[RFC0793] J. Postel, "Transmission Control Protocol", IETF RFC 793 http://www.ietf.org/rfc/rfc0793.txt

[RFC0827] E. Rosen, "Exterior Gateway Protocol", IETF RFC 827, http://www.ietf.org/rfc/rfc0827.txt

[RFC1157] J. Case, M. Fedor, M. Schoffstall, J. Davin, "A Simple Network Management Protocol", RFC 1157, http://www.ietf.org/rfc/rfc1157.txt

[RFC1195] R. Callon, "Use of OSI IS-IS for Routing in TCP/IP and Dual Environments", IETF RFC 1195 http://www.ietf.org/rfc/rfc1195.txt

[RFC1322] D. Estrin, Y. Rekhter, S. Hotz, "A Unified Approach to Inter-Domain Routing", IETF RFC 1322, http://www.ietf.org/rfc/rfc1322.txt

[RFC1519] V. Fuller, T. Li, J. Yu, K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", IETF RFC 1519, http://www.ietf.org/rfc/rfc1519.txt

[RFC1771] Y. Rekhter, T. Li, "A Border Gateway Protocol 4 (BGP-4)", IETF RFC 1771, http://www.ietf.org/rfc/rfc1771.txt

[RFC1918] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, E. Lear, "Address Allocation for Private Internets", IETF RFC 1918 http://www.ietf.org/rfc/rfc1918.txt

[RFC2460] S. Deering, R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", IETF RFC 2460 http://www.ietf.org/rfc/rfc2460.txt

[RFC2616] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", IETF RFC 2616, http://www.ietf.org/rfc/rfc2616.txt

[RFC3945] E. Mannie, Ed. "Generalized Multi-Protocol Label Switching (GMPLS) Architecture" IETF RFC http://www.ietf.org/rfc/rfc3945.txt

[RFC4271] Y. Rekhter, Ed.m T. Li, Ed., S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", IETF RFC 4271 http://www.ietf.org/rfc/rfc4271.txt

[ROCKE] "Rocketfuel: An ISP Topology Mapping Engine" http://www.cs.washington.edu/research/networking/rocketfuel/

[ROUTV] "University of Oregon Route Views Project" http://www.routeviews.org/

[SHN98] M.G. Shnirman and E.M. Blanter "Self-Organized Criticality in a Mixed Hierarchical System". Physical Review Letters 14 December 1998, Vol. 81, Number 24. pp.5445-5448

[SIG03] Siganos, G. Faloutsos, M. Faloutsos, P. Faloutsos, C. , "Power laws and the AS-level Internet topology", IEEE/ACM Transactions on Networking, Volume: 11, Issue: 4 pp 514-524 Aug. 2003

[SKITT] http://www.caida.org/tools/measurement/skitter/

[SOR99] Patrick Soriano and Christelle Wynants and Rene Seguin et al. "Design and Dimensioning of Survivable SDH/SONET Networks", Chapter 9 in "Telecommunication Network Planning", Kluwer Academic Publishers 1999

[SPE02] Spencer J. and Sacks L., "Modelling IP Network Topologies by Emulating Network Development Processes", Proceedings of IEEE Softcom 2002, Split, Croatia, October 2002.

[SPE99] Spencer J., Antonopoulos A., Sacks L., O'Reilly J.J., "Resource Allocation in WDM OMS-SPRing Architectures with Arbitrary Demand Patterns"; European Conference on Networks and Optical Communications, NOC '99, June 1999.

[TAN02] Tangmunarunkit H., Govindan R., Jamin S., Shenker S., Willinger W., "Network Topology Generators: Degree-Based vs. Structural" in Proceedings of ACM INFOCOM 2002.

[VAL07] Andy Valdar, "Understanding Telecommunication Networks" IET Telecommunications series, 2007, ISBN 0863413625

[VAZ02] A. Vazquez, R. Pastor-Satorras, and A. Vespignani, "Large-scale topological and dynamical properties of Internet," Phys. Rev. E, vol. 65, no. 066130, 2002.

[WAS94] Wasem, O.J., Tsong-Ho W., Cardwell, R.H, "Survivable SONET Networks – Design Methodology", IEEE Journal on Selected Areas in Communications Vol 12. No.1. 12:205—212 January 1994.

[WAT98] D. J. Watts and Steven Strogatz (June 1998). "Collective dynamics of 'small-world' networks". Nature 393: 440–442.

[WAX88] B. Waxman. Routing of Multipoint Connections. IEEE J. Select. Areas Commun., SAC-6(9):1617–1622, December 1988.

[YOO02] S. Yook, H. Jeong, and A.-L. Barabasi, "Modeling the Internet's large-scale topology", PNAS 99 13382-13386 (2002).

[ZEG96] Ellen W. Zegura, Ken Calvert and S. Bhattacharjee. "How to Model an Internetwork." Proceedings of IEEE Infocom '96, San Francisco, CA.

[ZEG97] E. Zegura, K. Calvert and M. Donahoo. "A Quantitative Comparison of Graph-based Models for Internet Topology." IEEE/ACM Transactions on Networking, Volume 5, No. 6, December 1997.

[ZHO03] S. Zhou and R. J. Mondragon, "Towards modelling the Internet topology - the Interactive Growth model" in Proc. of 18 Intl. Teletraffic Congress (ITC18), Berlin, German: Elsevier, Sept. 2003, pp. 121–130.

[ZHO04] S. Zhou and R. J. Mondrag´on, "The rich-club phenomenon in the Internet topology," IEEE Commun. Lett., vol. 8, no. 3, pp. 180–182, March 2004.

[ZHO05] S. Zhou and R. J. Mondragon, "The positive-feedback preference model of the AS-level Internet topology" Proceedings of IEEE International Conference 2005. pp 163-167.