Plextek Limited, London Road, Great Chesterford, Essex, CB10 1NY, UK
Telephone: +44 (0)1799 533200 Fax: +44 (0)1799 533201
Website: http://www.plextek.co.uk Email: post@plextek.co.uk

Electronics Design & Consultancy

Final Report - Study of Efficient Mobile Mesh Networks

Ofcom Contract Ref. C31400/001

16 Jan 2006

Document Name AYR005

Version 02

Plextek Ltd

Spectrum Trading Associates

UCL Consultants Ltd

Distribution:

Ahmad Atefi Ofcom

Registered Address
London Road
Great Chesterford
Essex, CB10 1NY, UK

Company Registration
No. 2305889
VAT Registration
No. GB 532 2188 67

Executive Summary

This report deals with mobile meshes at frequencies below 3.5GHz. This implies quite different constraints to either fixed meshes, or meshes using higher frequencies. The reader who wishes to compare fixed and mobile meshes is referred specifically to Appendix G.

After an introduction to the fundamentals of mesh communications technology, this report begins to look at four hypotheses of interest to Ofcom in relation to mobile meshes. These first three hypotheses examine:

- Do meshes self generate capacity as new nodes join?
- Are meshes more spectrally efficient?
- Do directional antennas confer *significant* benefits for *handhelds* below 3.5GHz?

'No' is the answer in all cases because these hypotheses, whilst having a theoretical basis, can be shown to rely on inappropriate assumptions when applied in the real world. They are first disproved by PHY and MAC level considerations alone. The later addition of protocol and application constraints further reinforces the rejection of each hypothesis and gives insight into the degree of mobility which may be handled before a mesh fails.

The fourth hypothesis asks

- May meshes improve spectrum utilisation?

'Yes' is the answer in any general case where short line-of-sight links are involved, and this is expected to include many practical meshes. Utilisation embodies the wider concept of best use of spectrum, rather than simple spectral efficiency.

Interference and coexistence issues are considered next in the report, although physical layer issues are found not to be mesh specific. At the higher layers however, there are potential benefits from the reconfigurability aspect of mesh networking. Unfortunately, this is not always so attractive when the effect of regular such reconfiguration is evaluated on the transport protocol, which can stall and lose efficiency. The report returns to further pivotal effects of user mobility in a later section, when considering how meshes might see widespread adoption.

Notwithstanding the above, however, there remain properties of meshes which make them uniquely attractive, including their quick set up and tear down capability and especially their ability to provide coverage in cluttered environments, such as an urban environment. A chain of mesh nodes may 'hop' around corners in the urban environment in a way which cellular systems cannot. It is by exploiting this aspect that meshes are expected to find wider application.

In considering application scenarios, no new services were identified, within the remit of this report, which mesh alone could support. Rather it seemed more likely that mesh would contribute by delivering present services in a new way. Five suitable areas were identified where mesh adoption is thought to be most likely. These are based on a mesh network's valuable attributes of coverage and/or reduced reliance on infrastructure:

- cellular multi-hopping
- WLAN hotspot extension
- community networking
- home and office indoor networking
- zero- or low-infrastructure environments

A good example is home/indoor networking. This is presently available via current technology, but would fit a mesh especially well, since existing infrastructure is limited, there is a relatively small, closed user group and many traffic flows could be peer to peer.

At this point, it was pertinent to return to protocol level issues in more depth, since one key question was pressing:

- What degree of mobility could a mesh support?

It became abundantly clear that there is a stark difference between mesh and cellular operation when considering user mobility and network performance. In a mesh, the network performance (e.g. availability/outage, service level) depends on the users themselves. It depends both on their mobility and the traffic carried. This makes it strictly impossible for an operator to offer a service level guarantee in a pure mesh, unless the dependence on users is mitigated somehow. Given that the problem occurs due to the uncertainty in node behaviour (i.e. the users), the solution is to inject some certainty in the form of fixed nodes. How many to add depends on the projected user mobility and traffic levels. In a practical situation, it seems likely that a precautionary design margin will be required, since accurate prediction of user behaviour is likely to be limited.

In adding infrastructure to gain the ability to guarantee service levels, the opportunity should be taken to ensure that it is added in such a way that the additional nodes also ensure scalability (in the same sense as cellular, via the incremental base station / access point concept), plus the infrastructure should be sufficient to 'seed' the network from day zero, when few real users may be available to provide mesh connectivity.

There are several other considerations when planning to implement a mesh which do not fall neatly under any heading so far considered: An example is the power consumption of what may well be a handheld device. Because mesh nodes are required to relay traffic which is not their own, especially when closest to the access point, more demands are made on battery performance. A situation may occur where a user who is not generating his own traffic may nonetheless experience high battery drain due to relaying the traffic of others. A further problem comes via what that user's response may then be: A selfish user who turns off his node when not in use will compromise the mesh for all other users. Clearly some provision needs to be made to encourage good group behaviour from individual users.

Avenues to encourage innovation and thus shorten time-scales are examined from technology and regulatory viewpoints. But, interestingly, there are also human factor aspects, due to the unfamiliar way in which meshes operate, such as the selfish user effects and user expectations of service level variation. It is potentially such 'softer', user-focussed issues which could have hard-to-predict effects on the overall time-scales for mesh adoption. In terms of regulation, the move away from the command and control method of spectrum management is expected to encourage innovation, but some unexpected regulatory obstacles may need clearing, since much legislation predates mesh networking concepts. Mesh networking, other than for personal area networks, is not seen as a driver for licence-exempt spectrum. In terms of technology it is noted that physical layer radio technology is much less lacking than software and networking technology. Examples are cross layer issues for improving protocol efficiency for real time services, traffic modelling for meshes and cognitive radio approaches for network co-existence.

The report conclusions and recommendations for future work are drawn together in sections 8 and 9. Section 1 outlines the report for the benefit of the reader.

Finally, two technical papers were accepted at international conferences, based on the material in this report. These are reproduced in Appendix F. The longer, 10 page IEEE DySpan paper is an attempt to bring together many of the important technical aspects contained in the full report.

Plextek

Revision History

Version	Date	Description	Author	Approved by
01	12 Oct 05	Release to Ofcom	SGM (ed.)	SGM
02	16 Jan 06	Appendix on comparison with fixed report added. Editorial corrections.	Steve Methley	Steve Methley

Version 02	Approved by	SGM

Contents

EXECUTIVE SUMMARY **2**

REVISION HISTORY **4**

1 REPORT ORGANISATION **8**

1.1 WORK ITEMS 8
1.2 REPORT CONTEXT 8
1.3 REPORT BACKGROUND 8
1.4 GUIDANCE FOR THE READER 9

2 INTRODUCTION **10**

2.1 A VISION OF FUTURE MOBILE COMMUNICATIONS 10
2.2 WHAT IS AN AD HOC MESH NETWORK? 11
2.2.1 Forms of mesh 11
2.2.2 Planned versus ad hoc 12
2.2.3 Characteristics of an ad hoc pure mesh network 12
2.2.4 Characteristics of a hybrid mesh network for access: The Access Mesh 13
2.2.5 Example applications for mesh 13
2.2.6 Standards activities and major commercial deployments 13
2.2.7 Other enablers for widespread mesh adoption 14
2.2.8 Mesh attractions and myths 14

3 FUNDAMENTALS OF MESH TECHNOLOGY **15**

3.1 OVERVIEW 15
3.2 PRACTICAL FUNDAMENTALS 15
3.2.1 Physical versus logical meshes 15
3.2.2 Intra- and extra-mesh traffic flows 15
3.2.3 Every node is a relay 18
3.2.4 Every node is a router 19
3.2.5 Network coverage 19
3.2.6 Per user throughput and latency 19
3.2.7 PHY and MAC level issues 19
3.2.8 Ad hoc routing protocols overview 19
3.2.9 Transport protocols overview 20
3.2.10 Non-technical aspects 20

4 CAPACITY CONSTRAINTS (OFCOM WORK ITEM 2) **21**

4.1 INTRODUCTION TO CAPACITY AND SCALABILITY 21
4.2 HYPOTHESIS TESTING ... THAT CUSTOMERS SELF-GENERATE CAPACITY 21
4.2.1 Introduction to scaling and capacity principles 22
4.2.2 Introduction to several myths concerning self generation of capacity 23
4.2.3 Analysis of the pure mesh 25
4.2.4 Analysis of the myth of self-generation of capacity 32
4.2.5 Hybrid mesh network 34
4.2.6 Access mesh network for extra-mesh traffic 36
4.2.7 Conclusions 38
4.3 HYPOTHESIS TESTING ... THAT MOBILE MESHES ARE MORE SPECTRALLY EFFICIENT 40
4.3.1 Introduction 40
4.3.2 Comparative efficiency of pure mesh and cellular 42
4.3.3 Comparative efficiency of access mesh and cellular 44
4.3.4 Conclusions on spectral efficiency 49
4.4 HYPOTHESIS TESTING ... THAT DIRECTIONAL ANTENNAS CONFER SIGNIFICANT BENEFITS FOR MOBILE MESH NETWORKS BELOW 5GHZ 50
4.4.1 Antenna directionality to increase capacity 50

4.4.2 *Directional antenna application to hybrid networks* ... 55
4.4.3 *Conclusions* ... 56
4.5 HYPOTHESIS TESTING ... THAT MESHES COULD IMPROVE SPECTRUM UTILISATION ... 57
4.6 SUMMARY OF CAPACITY CONSTRAINTS ... 58

5 SUSCEPTIBILITY AND CO-EXISTENCE ISSUES (OFCOM WORK ITEM 5) ... 59

5.1 INTRODUCTION TO SUSCEPTIBILITY ... 59
5.2 SUSCEPTIBILITY TO INTERFERENCE - PHY AND MAC ... 60
5.2.1 *Multiple access control* ... 62
5.2.2 *Medium access control (MAC)* ... 63
5.2.3 *Conclusions* ... 64
5.3 SUSCEPTIBILITY TO INTERFERENCE - TRANSPORT AND ROUTING ... 65
5.3.1 *Transport* ... 65
5.3.2 *Routing* ... 67
5.4 OTHER CO-EXISTENCE APPROACHES ... 69
5.4.1 *Knowledge based approaches* ... 69
5.5 SUMMARY OF SUSCEPTIBILITY ISSUES ... 69

6 THE POTENTIAL CASE FOR MOBILE MESH ADOPTION (OFCOM WORK ITEM 4) ... 71

6.1 REMAINING KEY QUESTIONS FOR MESH ADOPTION ... 71
6.1.1 *Review* ... 71
6.1.2 *Would a mesh enable new services?* ... 71
6.1.3 *What degree of mobility could a mesh support?* ... 72
6.1.4 *Could a mesh guarantee a quality of service?* ... 82
6.1.5 *Conclusions* ... 84
6.2 COVERAGE ... 84
6.2.1 *Rural environment with open propagation* ... 87
6.2.2 *Urban environment with obstacles* ... 87
6.2.3 *Extension to a mixed environment* ... 88
6.3 LIKELY ADOPTION SCENARIOS ... 89
6.3.1 *Cellular multi-hop* ... 89
6.3.2 *WLAN hotspot extension* ... 95
6.3.3 *Community networking* ... 96
6.3.4 *Home and office indoor networking* ... 96
6.3.5 *Zero- or low-infrastructure environments* ... 96
6.4 WIDER PRACTICAL ISSUES FOR DEPLOYMENT ... 96
6.4.1 *Initial mesh roll-out* ... 97
6.4.2 *Battery life* ... 97
6.4.3 *Network management and billing* ... 97
6.4.4 *Security and trust* ... 98
6.5 SUMMARY OF CONSIDERATIONS FOR MOBILE MESH DEPLOYMENT ... 98

7 HOW TO ENCOURAGE INNOVATION (OFCOM WORK ITEM 8) ... 99

7.1 INTRODUCTION ... 99
7.2 INNOVATION GOALS - WHAT, WHERE AND HOW? ... 100
7.3 ENABLERS FOR INNOVATION ... 100
7.3.1 *Technology enablers* ... 100
7.3.2 *Human factor aspects* ... 102
7.3.3 *Investment and pricing* ... 103
7.3.4 *Technology understanding - operator exploitation* ... 104
7.4 REMOVING BARRIERS TO INNOVATION ... 104
7.4.1 *Regulatory* ... 104
7.4.2 *Standards* ... 109
7.5 SUMMARY AND TIME-SCALES ... 109

8 CONCLUSIONS ... 111

9 SUGGESTIONS FOR FUTURE WORK ... 113

APPENDIX A DEFINITIONS 114

APPENDIX B GLOSSARY AND ABBREVIATIONS 115

APPENDIX C REFERENCES 117

APPENDIX D MOBILITY MODELS 121

D.1 THE RANDOM WAYPOINT MOBILITY MODEL 121
D.2 THE RANDOM DIRECTION MOBILITY MODEL 122
D.3 THE CITY SECTION OR MANHATTAN MOBILITY MODEL 122

APPENDIX E MODULATION AND MULTIPLE ACCESS 124

E.1 CARRIER FREQUENCY 124
E.2 SIGNAL BANDWIDTH 126
E.3 TIME STRUCTURE 128
E.4 MODULATION 129
E.5 TRANSMIT POWER 129
E.6 GEOMETRY 130

APPENDIX F INTERNATIONAL CONFERENCE PUBLICATIONS 131

F.1 IEE 3G AND BEYOND (SHORT PAPER) 132
F.2 IEEE DYSPAN (LONG PAPER) 138

APPENDIX G COMPARISON GUIDELINES FOR FIXED AND MOBILE MESHES 148

G.1 GUIDELINES FOR COMPARISON OF RESULTS 148
G.2 SUMMARY 149

1 Report Organisation

1.1 Work items

Ofcom work items to be covered are

- WI 2 Capacity Constraints (section 4)
- WI 5 Interference and Coexistence (section 5)
- WI 4 Key Benefits, Problems and Time scales (section 6)
- WI 8 Encouraging Innovation (section 7)

Note that original work item numbering has been kept, but that some works items were deleted at Ofcom's request before the start of work on this report.

1.2 Report context

Several exceptions to what this report is to cover are as follows:

- not closed, standalone user groups (e.g. emergency, military)
- not sensor networks
- not fixed networks
- not above 3.5 GHz

The report examines the qualities of mesh networks and identifies the key solution time scales for the challenges involved with making meshes with an external access capability. The report also considers how meshes could live alongside existing radio systems, in terms of interference and coexistence strategies. As well as identifying technical challenges, there is also issue of encouraging the community to actively bring meshes to fruition, in other words to be innovative.

As requested this report does attempt to keep mathematics to a minimum. Where maths is shown, it remains practical to follow the flow of the report without needing to fully understand the maths.

Where issues have no conclusion possible within the time frame of this report, the authors have been encouraged to include and identify these as areas for future work.

1.3 Report background

There are two sub projects under Efficient Mesh Networks in the SES program, one consortium is biased towards fixed networks and the Plextek consortium is biased towards mobile networks. The fixed/mobile distinction affects findings strongly. Both include planned and ad hoc approaches and have the same objectives:

- Looking at the capacity constraints of mesh networks. Examining the hypothesis that for a mobile mesh the more consumers use a service, the more capacity the network has: effectively, every mobile user, laptop user, etc., becomes a base station
- Investigating whether mesh systems would change the way spectrum is managed, e.g. would the wider use of mesh systems imply that there should be more licence-exempt spectrum?
- Examination of the key problems in the delivery of fixed and mobile mesh systems and predictions as to when these problems might be resolved and mesh systems might become

widespread.

The project start-up meeting with Ofcom further clarified the scope: Ofcom take as a known fact that the spectral efficiency of networks for wide area, fixed access is high, and so do not see a need to analyse this in detail. Neither do Ofcom wish to give much attention to low-data-rate telemetry applications, since this is such a low utilisation of spectrum that it is of little interest to the macro-economics of spectrum usage. Ofcom requested that the Plextek consortium focus its study on mobile mesh networks – exploring the limits on service types, throughput, quality of service, etc., and with focus on 0.5-3.5 GHz.

In summary, the study should answer the following questions, with particular reference to mobile/un-tethered applications:

- Are meshes more spectrally efficient than alternatives ?
- Can meshes enable the use of (higher) frequency bands, and/or support services-types that alternatives can not ?
- Are meshes practical, and what are the enabling technologies ?
- What needs to be done to promote mesh systems?
- ‘Mobility’ can mean vehicular mobility and/or quasi-static nomadic use. The study should endeavour to find what degree of “mobility” could be supported by mesh networks.

Ofcom stated that the second study of mesh networks, let to a different consortium, was not duplication the scope of this study by the Plextek consortium, but was instead focusing on modest-throughput fixed meshes and evaluation of the current bus-transport information system installed by MeshNetworks in Portsmouth, UK.

1.4 Guidance for the reader

Each report chapter is designed to be relatively stand-alone. The main chapters and their intended audiences are summarised as follows:

	Chapter	**Audience**
0	Executive summary	All readers
2	Introduction and context	All readers
3	Fundamentals of mobile meshes	Most readers
4	Key capacity hypotheses - testing	Technical readers
5	Susceptibility issues	Technical readers
6	The case for mesh adoption	Most readers
7	How to encourage innovation	Most readers
8	Conclusions	All readers
9	Suggestions for future work	Technical readers

Table 1 Guidance for the reader

Plextek

2 Introduction

2.1 A vision of future mobile communications

It is generally accepted that the wider vision of mobile communications describes the future as an integration of all mobile and wireless nodes (e.g. cellular, WLAN, PAN etc) with an IP core. This is a revolutionary integration of the cellular approach with the WLAN approach as depicted in Figure 1, where proprietary interfaces and protocols are largely removed.

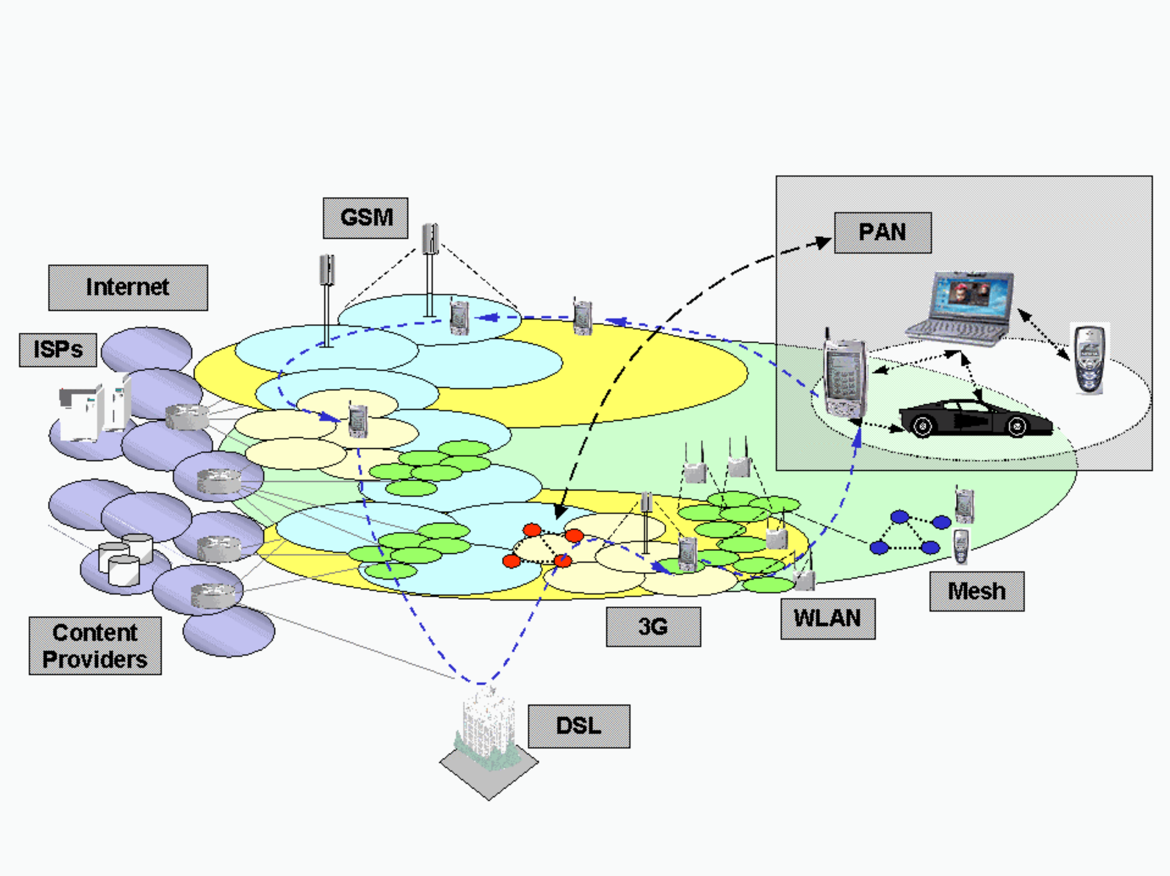


Figure 1 Future mobile integrated vision using IP core

In the literature, this is sometimes referred to as '4G', or 'B3G', meaning beyond 3G. Such terms are usually proffered by the existing cellular focus groups, but WLAN parties have a similar vision. As shown in Figure 1, this report finds that meshes may be expected to provide another access route into the core, alongside WLAN, 3G etc.

Note that the core supports access at very many different data rates from the different devices. This demands a granularity of service which is well supported by TCP/IP services. The core of the network is used to connect the users to the Internet via the Internet Service Providers (ISPs). Content providers may be closely aligned with ISPs in some business models. The types of network linked into the core are:

- PAN - personal area network. This may for example use Bluetooth and may even be mesh-like in its structure. Speed is currently under 1Mb/s (although 100Mb/s is predicted) and range is short, e.g. 10 metres. No infrastructure is needed, except for an access point.

- GSM/3G - 2nd and 3rd generation cellular mobile. Data speeds will initially be under 2Mb/s with the average around several hundred kb/s. Range is high, on the scale of several km. It relies on deployed infrastructure (base stations).
- DSL - digital subscriber line. This will typically be a point to point connection running over fixed copper pairs carrying Ethernet or ATM frames. Speed is currently around 2Mb/s.
- WLAN - wireless local area network. This provides potentially the fastest access towards the core; up to 54Mb/s is available, with 11Mb/s being widespread. Its range is around 100 metres. Some WLANs enable a choice of whether to use infrastructure or not, but the majority use an access point infrastructure.
- Mesh - presently only very sparsely deployed into the 'early adopter' market. Their performance capability is to be investigated within this report. Potentially high speed, good coverage and no infrastructure needed, except for an access point.

  (Later, in section 2.2, the distinction will be drawn between a pure mesh with no external connectivity and an access mesh, with external connectivity. Whilst pure meshes will continue to serve their niche markets such as the military or emergency services teams, by their very nature they have no obvious position within the future mobile system vision, since they provide no connectivity to the core).

One key point from Figure 1 is that there are two ways at present to obtain relatively high speed wireless access to the core - via cellular and via WLAN. They have different performances in terms of quality, speed and range, but also critically in terms of cost. Application performance is also different; cellular is biased for real time voice and WLAN for non-real time data. But the idea to have a dual mode handset for attachment via WLAN at close range and via cellular at longer range is very attractive to the user. However, such 'fixed-mobile convergence' presents a real challenge to present operators who may well have to change their business model and adapt their technology. This is re-visited in section 7.

Mesh is relatively new to the vision and may enable a radically fresh approach to high speed wireless access for the future. Hence this report characterises the complex set of aspects relating to a mesh network, to help show what solutions meshes may offer to the overall future mobile vision.

2.2 What is an ad hoc mesh network?

The meanings of the terms mesh and ad hoc are first clarified.

2.2.1 Forms of mesh

A distinction is made between three basic mesh type architectures, since this frequently affects the discussion within this report. The types are mentioned here and described in section 3.2.2.

- Pure Mesh - all traffic is intra-mesh, i.e. the mesh is isolated. All traffic is relayed.
- Hybrid Mesh - a pure mesh with a hierarchy, in order to improve efficiency via backbone routes.
- Access Mesh - a (hybrid) mesh where considerable traffic is extra-mesh, e.g. traffic accesses other networks and the Internet.

Traffic flows and hence appropriate architecture depend on whether the content to be accessed is inside or outside the mesh. In other words the type of mesh required in a given situation is driven by the user and application needs for content.

Most of the early published research was funded by the military to look at pure meshes.

2.2.2 Planned versus ad hoc

A second distinction is concerned with the design rationale of the network. A planned network such as cellular has a predicted maximum level of users and prescribed cells in which they may operate. The benefit of this is that the signal quality is predictable and it flows from this that guarantees of the quality of service delivery can be made. The downside is that infrastructure is needed; in other words the operator must first make a provision for anywhere the user requires a service.

An unplanned network allows ad hoc connections. The benefit is that no infrastructure is needed and the users themselves may extend the area of coverage. The downside is that without planning, there is no control over interference from other users. Hence signal quality is beyond the control of any one party, therefore guarantees over the quality of the delivered service cannot be made.

2.2.3 Characteristics of an ad hoc pure mesh network

Table 2 summarises the characteristics of a pure, ad hoc mesh network.

Ad hoc	Unplanned. Therefore the coverage and interference environment is uncontrolled, which is the exact opposite of the cellular case. This directly raises quality of service issues.
No separate infrastructure	All functionality is provided and carried out within the mesh. This includes power control, routing, billing, management, security etc. There is no centralised equivalent to the base station or AAA centre (for example) of cellular networks. (Not true for an access mesh).
Mobility	Nodes are free to move and even disappear. The network interconnections may be thus very dynamic.
Wireless	In order to support mobility. Wireless could be radio or optical. This report concentrates on radio. Radio links are lower quality than wired links, packet loss in radio is 'normal', whereas on wired connections, loss is equated to congestion. Transport protocols (as developed for wired networks) may thus have the 'wrong' reaction when used on radio networks.
Relay	All nodes may be required to relay information for other nodes. This will lessen the bandwidth available to each node user.
Routing	All nodes will be required to participate in a routing protocol. This may be either proactively by maintaining up to date tables or reactively by creating routes on demand (which may also be cached in tables). Routing creates an overhead, which will depend on the protocol, the traffic and the mobility of the nodes.
Multi-hop	A corollary of relay and routing, multi-hop is an enabler of coverage, especially in the cluttered environment.
In-homogeneity	Not all nodes need be equal, beyond the subset of capabilities needed for basic mesh operation: Some nodes may have additional network connections (external connections in the Access Mesh case).

Table 2 Ad Hoc Mesh Network Characteristics

2.2.4 Characteristics of a hybrid mesh network for access: The Access Mesh

A hybrid mesh is similar to Table 2, except some nodes will additionally have external network connections, enabling them to be access points to external networks such as other meshes, infrastructure or the Internet. Factors influencing the provision of such access points are investigated in this report.

This type of mesh is found to have such high relevance throughout this report, that the term ‘Access Mesh’ will henceforth be used in reference.

2.2.5 Example applications for mesh

Mesh application examples include:

- Disaster scene or military team communications
- Sensor networks
- Hotspot extension of urban public wireless access
- Rural Community Networks
- Lecture hall or convention hall networks
- Integrating PAN devices (e.g. PDAs phones) into the WAN
- Home networks

All but the first two are access meshes rather than pure meshes, since they require external network access. This is important since much published literature addresses only the isolated pure mesh case. The more likely adoption scenarios are discussed in section 6.3.

2.2.6 Standards activities and major commercial deployments

Standards, be they either formal or *de facto*, are required to ensure the inter-working needed to create a large homogenous market and thus enable scale economies in production. Standards activities and notable proprietary efforts are listed below. They will be referred to later in the report.

IEEE 802 activities	802 covers PHY and MAC only. Mesh is covered in 802.11 and 802.16. Both versions have limited mesh performance. 802.16 has very recently started a ‘mesh extension of base station coverage’ working group.
IETF activities	Transport and Routing protocols e.g. AODV are being developed here.
3GPP activities	Little known activity in mesh mode
ETSI activities	Little known activity in mesh mode
Mesh Networks	Now owned by Motorola. Several installations.
Mesh Dynamics	Start up, using a 3 radio mesh for best capacity claims. Several installations.
Spanworks	Software to enable application working over meshes.
Locust networks	Freeware mesh network stack. Very popular in the technical early

	adopter market.

Table 3 Brief summary of proprietary deployments and standards activities

2.2.7 Other enablers for widespread mesh adoption

Apart from the foregoing technical considerations, there are also economic, social and political factors which will affect any widespread adoption of mesh networks, Table 4. Many of these factors are visited in section 7, Innovation.

Economic	Finding the money making model
Social	Attitudes to radio emissions Selfish user issues
Political and regulatory	Service level expectations Operator buy-in Legacy legislation

Table 4 Economic, Social and Political considerations

2.2.8 Mesh attractions and myths

Perhaps the largest attraction of meshes is that they can be entirely unplanned in pure form. This is useful to the military and to disaster recovery teams who neither need infrastructure access for content nor want to rely on its presence for operation. It is far less clear what these benefits could lend to the roll-out of a mass market mesh network. On the other hand to a service provider or regulator, the lure of a network which promises no planning phase must be high and thus merit investigation.

Such investigation forms the basis for section 4 of this report.

It has often been said, as if a truism, that meshes increase capacity. The reasoning is usually along the lines of *each new user brings additional capacity to the mesh,* or *each new user effectively becomes a base station.* This report critically examines such statements and separates the reality from the 'something for nothing' type of mythology. The difference between network capacity and the user throughput actually available is outlined. The difference between a node with independent, external infrastructure access such as a base station and a normal mobile user node is emphasised.

Despite some unfortunate myths, meshes do have some strongly attractive features, notably in the area of coverage, where they offer complimentary performance to that of cellular systems, see section 6.2. It is for this reason that meshes should find application in some scenarios, as part of a larger picture of mobile access technology.

3 Fundamentals of mesh technology

3.1 Overview

Note that a glossary of terms is provided in Appendix B.

The discussion within this report takes a simple view of the segmentation of the 'stack' or levels of functionality in a mesh node. The four levels considered are: PHY, MAC, protocol and application. The PHY is the PHYsical radio part. The MAC is the Medium Access Controller, to control use of the channel. The protocol involves the decision on routing paths and the method of transporting the data as it becomes available. The application is whatever the node user wishes to use. An example stack would be an 802.11 radio (PHY) and 802.11 ad hoc MAC carrying TCP/IP to allow the transport of email application data.

Beyond the communications 'stack', there are two more important areas to consider:

- network management and billing
- security

These have particular problems associated with the ad hoc nature of meshes.

3.2 Practical Fundamentals

For the purposes of this study meshes have been defined as those systems that utilise node to node links to form at least part of the network, c.f. section 2.2.

Meshes using node-node links are different from what is typically deployed at present. They are more than the sum of the set of links which form them, in the sense that the introduction of meshes brings with it new issues and options. This section reviews the fundamental features of meshes indicating the options available and how they might impact on capacity and performance. More detailed analysis of various features occurs later in the document, see sections 4 and 6.

3.2.1 Physical versus logical meshes

A mesh may be made physically or logically. Physical meshes are those made by physical level constraints, for example by directional antennas or perhaps by constraint of the signal path by terrain or medium. The wired internet is clearly a perfect physical mesh in that transmitting on one link does not interfere with any others. On the other hand a logical mesh is configured above the physical layer. There is not necessarily any physical constraint to a station's neighbours imposed by the system. Omni-directional antennas in an open field could be connected as a logical mesh, although their interference footprint would clearly be quite different from that of a physical mesh and the full benefits of physical meshing should not be expected. (The unrealised benefit would be that the omni-directional antennas could equally be re-configured as e.g. point to multi-point, in the logical sense, if that were ever required.)

3.2.2 Intra- and extra-mesh traffic flows

Candidate network architectures for mesh networks stem from considering the traffic flows which need to be supported and whether the drivers are for spectral efficiency and/or area coverage. Traffic flows depend on whether the content to be accessed is inside or outside the mesh network.

Thus two fundamental traffic flows exists: defined here as “extra-mesh” and “intra-mesh”.

Most of the commercially-significant applications to date e.g. telephony, messaging, internet access would present *extra-mesh* traffic flows as all require access to external infrastructure. Hence within the area covered by a mesh there would be Access Points providing a route to the infrastructure resources.

For some applications however there is no need to access external (or central) resources e.g. team communication. For such applications traffic flow can be contained within the mesh with users communicating with each other either directly if range permits or by hopping via peers. In current markets such intra-mesh traffic tends to be limited to closed user groups such as:

- File-sharing: on-campus, intra-company, local community, etc.
- Home networking
- Military, emergency and disaster relief activities which cannot depend on infrastructure.

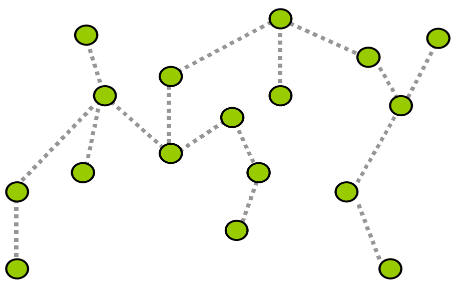
Clearly, the vast majority of traffic which is of commercial significance to network / service providers today would be classed as extra-mesh traffic, with an interface to the public telephone network and the internet. Thus, if there is to be substantial uptake in mesh networks driven by intra mesh traffic there is a need for a “killer application” in this space – such as, possibly, music sharing or video sharing – and also means for service providers to gain revenue from such applications.

It should be noted that even in the case of wholly intra-mesh traffic, there may be associated network management and/or billing traffic which must flow to a network management centre and so becomes extra-mesh. Thus, in practice one is very likely to require a hybrid architecture supporting both intra- and extra-mesh traffic types to varying degrees.

Intra-mesh and extra-mesh traffic is next examined in more detail.

3.2.2.1 Intra-Mesh traffic architectures

In this case the sources and sinks of all traffic are within the mesh network - i.e. there is no requirement for connection to an external network, such as for connection to the internet, public telephone network, command and control centre, etc. For such intra-mesh traffic the mesh may comprise entirely of subscriber nodes, as illustrated in Figure 2:



Routing through Mobile Nodes
(other non-participating nodes not shown)

Figure 2: Subscriber nodes forming route connections

In this case, traffic concentrations will occur only as the result of user-concentrations: for example around business centres, communities, retail centres, etc.

Notwithstanding the applications issues, the integrity and coverage of the network can be enhanced by the addition of fixed nodes added to assist with local traffic flows, as illustrated in Figure 3

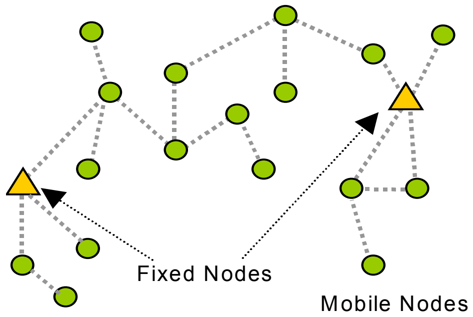


Figure 3: Adding fixed relay nodes within a network

These fixed nodes might be added to:

- enhance connectivity or coverage when user-nodes are sparsely distributed. This may be the case during early roll-out of the service when there is inadequate customer-base to provide sufficient connectivity of the mesh or geographical coverage (in this context they are often referred to as 'Seed Nodes' [Radiant 2005].
- ensure a minimum degree of coverage and connectivity, independent of customer density. This may be required, for example, to address the lack of subscriber nodes which arises when users commute in and out of city / residential / recreational areas.
- enhance throughput capability in regions of high customer-density.
- enhance coverage by aiding routing around obstacles, such as in the urban environment.

To support traffic flows over long path lengths, which would otherwise require a large number of short hops (with consequent reduction in spectral efficiency and / or increase in end-to-end delay – see discussion later in this report), one could deploy a hierarchy of fixed relay nodes having long communication range between them (either wired or wireless). This forms a "backbone network" in a fixed-mobile Hybrid mesh architecture as illustrated in Figure 4:

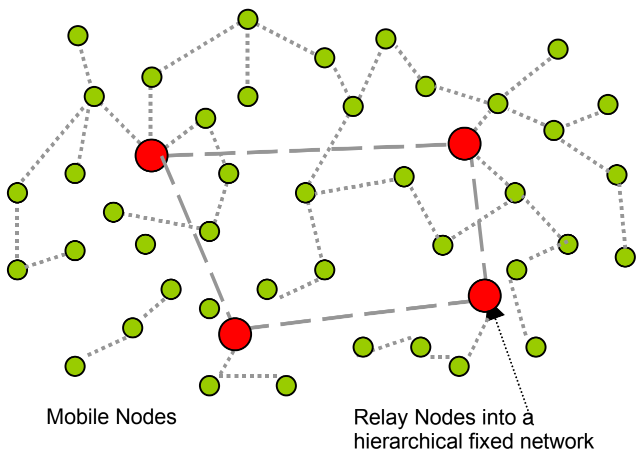


Figure 4: Hierarchical mesh with backbone network

The decision on when to route via the backbone network would be taken by the routing protocol (in accordance with load distribution, delay targets, etc.), and so this architecture employs hierarchical routing.

Note that the traffic flow is no longer evenly distributed throughout the mobile nodes, but is concentrated around nodes in the vicinity of the Relay Nodes. This exacerbates the effects of any local bottlenecks, but perhaps more significantly it has major implications on the required quantity, performance, power consumption, etc of nodes in proximity to the Relay Nodes. Thus the quality-of-services becomes more dependent on the availability and behaviour of subscriber nodes in the vicinity of Relay Nodes.

3.2.2.2 Extra-Mesh traffic flow

In this case traffic enters / leaves the Mesh via one or more Access Points connected to a public or private access network, as illustrated in Figure 5:

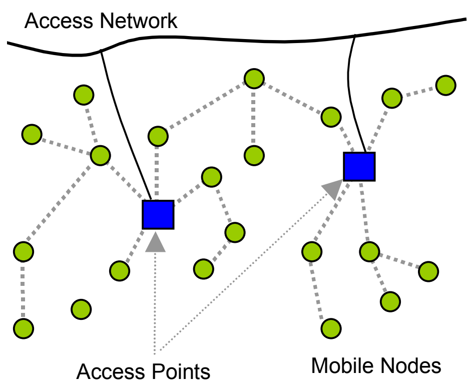


Figure 5 Extra mesh traffic flow via access point: The "Access Mesh"

As per the hierarchical mesh network, the traffic flow is no longer evenly distributed throughout the mobile nodes, but is concentrated around nodes in the vicinity of the Access Points. Again this has major implications on achieving satisfactory quality-of-service when it is subject to the availability and behaviour of subscriber nodes in the vicinity of Access Points.

3.2.3 Every node is a relay

To be able to form a mesh it is necessary for subscriber units to act as relays. This has several consequences.

Firstly, by acting as a relay a subscriber unit takes on a workload over and above that required to support its user's requirements. The general consensus from both academic and commercial published work is that subscriber units will need to be able to relay traffic of a volume a few times above that of their own service, i.e. they must handle not only their user generated traffic but potentially that of several other users as well. This implies that mesh using systems will require more capable subscriber units than equivalent PMP systems.

The second consequence of having subscriber units relay traffic is that they must somehow be co-ordinated so that:

a) those pairs of units that are supposed to be communicating do in fact do so, and

b) that communicating pairs do not adversely affect each other.

There are several multiplexing schemes that could be used and, over and above that choice, meshes could be designed to operate with either an overall controller (centralised or distributed) or by essentially a random access mechanism with local arbitration.

Finally, the consequence of having system performance depending on subscriber units is an increased difficulty in maintenance and upgrade. For example with the introduction of EDGE onto GSM it was possible to upgrade base stations, and then allow subscribers to sign up to the new services. With a mesh system users are dependent on the installed base of subscribers and so new services cannot be provided unless all or at least a substantial fraction of existing users are persuaded to change their units.

3.2.4 Every node is a router

In an ad hoc mesh, each and every node needs to know what route traffic needs to follow in order to reach the destination. This is directly analogous to routing in the Internet, although only some Internet nodes need be routers. Mobility adds volatility to the routing problem. Ad hoc routing protocols are introduced in section 3.2.8.

3.2.5 Network coverage

A key feature of meshes is the potential they have for achieving high levels of coverage by routing around problem areas.

A downside to using meshes for coverage extension is that the offered quality of service inherits a dependence on node numbers and movements. This adds an extra level of variability which means that the provisioning level may need to be increased to offer a given level of service. This is covered in detail in section 6.

3.2.6 Per user throughput and latency

Multi-hop meshes typically suffer from greater latency than PMP systems, and the more hops the greater the latency. It is possible to mitigate this by, for example, transmitting shorter packets of traffic more regularly but such actions to reduce latency are likely to require compromises elsewhere such as spectral efficiency. Balancing latency against other metrics is also discussed in 4.2.2.

For mesh systems which offer resilience (see section 5.3) by use of multiple routes there is an issue of varying latency when the latency changes on switching from one route to another. Some applications do not tolerate varying latency.

3.2.7 PHY and MAC level issues

In any network the choice of carrier modulation scheme is critical to overall performance, and the choice of narrow band or broadband techniques has major implications on co-existence.

Additionally, multiple access techniques are required to access the physical medium. In a PMP network there is centralised co-ordination of many of the PHY and MAC processes. However in a distributed network such as a mesh there is significantly less opportunity for co-ordination and so many of the traditional attributes of multiple access schemes are not apparent and/or are less easy to implement and manage.

3.2.8 Ad hoc routing protocols overview

In an ad-hoc network environment, each node participates within the network not only as a possible source and sink of traffic, but also as a relay and a router, enabling the forwarding of traffic between nodes in the network.

Ad-hoc routing protocols are used in environments where there is not necessarily a well-controlled infrastructure network but where the administrative regime is fairly uniform, i.e. there is one common routing policy. Whilst the first of these is a good match for mesh networks, the second may not necessarily be so. Nevertheless, looking at ad-hoc networking protocols is a good starting point for considering the network layer protocol requirements for mesh environments, especially mobile mesh environments. Examining the structure and operation of ad-hoc networks can offer insights into ways in which networking could be supported within the mesh environments.

3.2.9 Transport protocols overview

Residing between the application and the network, the transport protocol is an essential part of the information transfer process. It allows multiplexing of several applications in the same network and, if desired, it implements congestion control and/or packet level reliability. Today's applications are designed to work with the same transport protocols as the Internet of 20-30 years ago but today's Internet is a more varied infra-structure not only in the types of technologies it uses for transmission but also in the nature of the applications running over this infra-structure.

Mesh networks will potentially be made of several wireless links where noise will cause loss rates significantly larger than in a wired scenario. They may also have more hops which can potentially introduce more delay in the total transfer. This will inevitably have an impact on elastic and inelastic applications which should be taken into account when choosing and designing a mesh network, see section 6.

3.2.10 Non-technical aspects

As well as the various performance issues of meshes there are also some significant 'non-technical' aspects which will affect the development of the technology.

One of the most interesting issues is the co-operation aspect required of users: For the mesh to function it requires nodes to allow the use of their equipment to support other users. This is especially significant in the case of mobile applications where battery capacity is at a premium: It is conceivable that a user might find his battery exhausted by supporting other users before making a call on his own account. This could give rise to a strategy amongst users of not switching on their equipment until wanting to make a call. However this would reduce the density of active units compromising the performance of the mesh. It also would throw an increased load onto the altruistic users speeding up the exhaustion of their batteries or their conversion to selfish operation. Approaches have been published on techniques for "encouraging" proper behaviour, but come with an overhead, see section 6.4.2.

Security will also be an important concern. Meshes will inherit all the security issues of radio systems in general and add some of their own. In particular data will be transiting third party equipment not belonging to either the user or the operator. It will therefore be more vulnerable to capture, delay and manipulation. Security concerns will also centre around user authentication. In a strictly ad hoc network there is no parallel to the present central authentication function of cellular networks which use AAA (access, authentication and authorisation) servers such as those using RADIUS (Remote Authentication Dial-In User Service).

4 Capacity Constraints (Ofcom Work Item 2)

4.1 Introduction to Capacity and Scalability

The generally proclaimed benefits of multiple hop mesh networks include:

- providing connectivity in adverse coverage conditions
- conserving transmit energy resources[1]
- reducing interference
- increasing link throughput by multiple hopping

A key problem in assessing the literature is that different assumptions are made in different published papers: a direct comparison is thus at risk of being inconsistent. Noting this caveat, this report examines several hypotheses which are of interest to Ofcom. These hypotheses cover:

- capacity self generation
- spectral efficiency
- omni-vs.-directional antenna benefits
- spectrum utilisation

The results of testing these hypotheses are tabulated in the Summary of Capacity Constraints, section 4.6.

The capacity and scaling characteristics are covered initially from only the PHY and MAC level, since this is sufficient to prove or disprove the hypotheses. The additional influences of protocols and applications are addressed in chapter 6 and do not alter the results here.

4.2 Hypothesis Testing ... that customers self-generate capacity

There are huge attractions to having 'self generation of capacity' in a radio network. Notably, that the network is self-sustaining and that it could avoid the so-called 'tragedy of the commons'[2]. Such a tragedy relates to the days when common land was used for the grazing of livestock with free access for all. The danger is that free access to a finite resource can result in that resource being fully consumed or compromised further such that it loses its usefulness to all. What then, if each user were somehow to add grazing capacity as they joined the common?

This section examines the hypothesis that in a mesh network the subscriber base self-generates capacity. It is concluded that:

<u>For a pure mesh subscribers cannot self-generate capacity at a rate sufficient to maintain a target level of per-user throughput regardless of network size and population. The only viable ways scalability can be achieved are by providing additional capacity either in the form of a secondary backbone (fixed) mesh network – so forming a "Hybrid Mesh", or an</u>

[1] Most often claimed for sensor networks, outside the scope of this report

2 Note that not everyone believes in the tragedy of the commons. Interviews with UK network operators suggest that they believe the commons policy worked well historically and they expect it to work well with radio spectrum with one key proviso – that there will always be access to premium spectrum for those who will pay for it, if and when it is needed.

<u>access network – so forming an "Access Mesh". In these two configurations scalability is possible and has similar characteristics to that of a cellular network</u>.

The analysis in this chapter is first conducted on the basis of employing omni-directional antennas: the case for directional antennas is covered in section 4.4.

The reader is reminded of Appendix B which contains a glossary and list of abbreviations.

4.2.1 Introduction to scaling and capacity principles

The issues of capacity and throughput are considered in the light of two intrinsically different types of service: intra-mesh and extra-mesh services. The work shows that there are very differing results for each. For intra-mesh traffic, although there is a concept that as more mesh nodes enter the network the underlying "installed" capacity increases, the in-efficiency with which this can be shared amongst the subscribers means that there is a net reduction in the available per-node throughput as the node population increases (unless one can limit such parameters as node density or end-to-end traffic path length).

But the performance for extra-mesh traffic is very different. The network's performance for extra-mesh traffic is very akin to that of a traditional cellular network, in which the traffic capacity of each cell is set by the base station and that is shared amongst the population served by that cell – capacity is scaled up by sectoring or adding more base stations.

Capacity growth in a standard cellular network is illustrated in Figure 6.

In classical single-cell PMP network the capacity per unit area is defined by the capacity of the base station covering that area. Let this be B bps. Then the mean traffic throughput per subscriber in the cell is proportional to k_1B/N, where N is the number of subscribers in the cell and k_1 is a function of the system overheads for multiple access, channel coding, etc.

Capacity can be increased by deploying more base stations, and by sectoring cells, such that the mean capacity per subscriber is of the order of $k_1.k_2.M.B/N$, where M is the number of cells/sectors and k_2 is a function of the interference between cells/sectors.

Thus in terms of scalability a cellular network's offered per-user throughput decreases nominally as $1/N$ until additional base stations are added

Geographic coverage is extended by deploying base stations over a wider area.

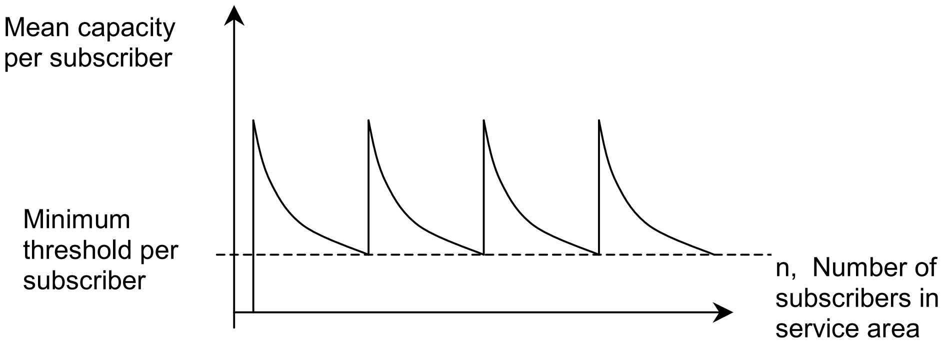


Figure 6: Capacity scaling in a Cellular Network

It will be shown that the above illustration would be expected to be very similar in shape (but not necessarily scale) for a mesh network supporting only **extra-mesh traffic,** in which Access Points replace the base station. The key differences for the extra-mesh mesh network will be:

- the rate of decay of available per-node throughput within each area served by the Access Points (shown as $1/n$ in the above illustration for cellular) may be steeper than $1/n$ because of the relaying process going on in the mesh,
- coverage and service availability in the mesh network will be dependant upon the size and distribution of the population of nodes,
- coverage area around Access Points is extended via multi-hopping through nodes so that there are benefits from spatial diversity and improved link budget over short hops.

For a mesh network supporting only **intra-mesh traffic** the growth characteristics are expected to be very different. As illustrated in Figure 7: usable per-node throughput will rise initially as the first few nodes come on board and the network achieves some degree of connectivity, but thereafter it is likely to be a diminishing function of increasing node population, unless specific measures can be taken to avoid this happening.

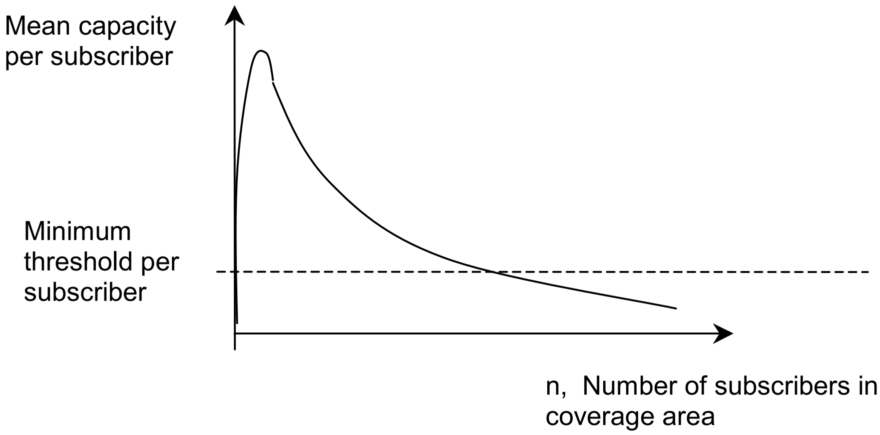


Figure 7: Capacity Scaling in a Mesh Network

In a mesh network, extending geographic coverage over a wider area is dependent solely on geographic expansion of the customer base, whilst coverage (i.e. service availability) within the area of the mesh is dependent on the density and spatial distribution of nodes. Clearly, there are several key questions central to this concept:

- What is the scale factor on the vertical axis of these coverage characteristics – does mesh offer more per-user capacity than cellular (for a given spectrum allocation)?
- Are meshes scalable, i.e. under what circumstances, if any, can per-user capacity be held above a target service threshold as population grows?

4.2.2 Introduction to several myths concerning self generation of capacity

Although this report concludes that meshes do not scale adequately without some form of additional infrastructure, we highlight some concepts which appear, initially, to contradict this.

Four published approaches are briefly reviewed below and, whilst each presents a coherent argument based on its stated assumptions, it will be shown that those assumptions do not always translate well to practical applications. The four approaches examined are:

Approach	Assumption Challenged
Grossglauser and Tse [2001]	Unbounded delay
Shepard [1996]	Unbounded spectrum
Negi and Rajeswaren [2004]	Unbounded spectrum
Gupta and Kumar [2000]	Strict localisation of traffic

4.2.2.1 Mobile couriers

This approach was taken as an input to an economics paper [Jones et al 2003] which points to the benefits if a 'tragedy of the commons' could be avoided. Jones et al initially set a scene for an economic assessment by showing how the combination of two well known papers on ad hoc networking could point to self generation of capacity. The two well known papers whose findings are combined are:

- Gupta and Kumar's [2000] identification that if traffic flows remain localised despite growth of the network then per-user throughput can stay constant as users are added.
- Grossglauser and Tse's [2001] assertion that mobility can be harnessed to achieve constant throughput as users are added.

In Grossglauser's approach the long-range movement of nodes is used as the means of delivering traffic: traffic is transferred from source to a "courier" and then from courier to recipient; thus all traffic is transferred via just two radio-hops. Clearly, the drawback is that there can be no guarantees on delivery or latency, and so the technique has very limited application. This method is further discussed in section 4.2.4. The localisation of traffic assumption is covered in 4.2.4.3.

4.2.2.2 Constraint of localisation of traffic

Gupta and Kumar [2000] identify that if traffic is confined to localised groups in a relatively large mesh, then scaling will be achieved. Clearly this is very application-specific and it is arguable that this is no longer a large mesh, but a collection of smaller meshes. Further discussion appears in 4.2.4.3.

4.2.2.3 Shepard [1996]

Shepard [1996] has a relatively 'out-of-the-box' approach in suggesting a mesh in which collisions are not the be-all and end-all for the MAC. He sees multiple transmissions as a signal-to-noise issue, rather than a requirement to back off and try again. He does this by using spread spectrum transmission, hence multiple transmissions simply raise the noise floor, as in any CDMA system. He develops a complete theory to enable meshes to scale to millions of nodes with an allegedly respectable throughput. The problem is that it is extremely spectrally inefficient, due to the large processing gain required. Other problems include high computational effort and the distributed MAC complexity, which is an open issue. He also concludes that small hop meshes are best, so that interference may be managed locally by his distributed algorithm.

Problems here are the invention of the distributed algorithm itself, but more importantly that an increase in spectrum must be traded for the necessary increase in signal-to-noise ratio. In any case the throughput of a large mesh is still only in the several kb/s range in Shepard's own worked example with a realistic spectral bandwidth.

Shepard's specific scheme is not further discussed in this report, but the idea of having a MAC which is more sophisticated than being simply collision based (CSMA/CA), like 802.11, does clearly merit further attention. It is noted in section 5.2 that such a MAC approach will result in better efficiency of access to the available spectrum, albeit with an attendant need for increased network knowledge by the user nodes or by bespoke management nodes.

4.2.2.4 Negi and Rajeswaren [2004]

A broadly similar approach with similar problems is that of using "infinite" spectral bandwidth (bandwidth substantially greater than required to support the raw transmission rate used on links), for example in the ultra wide bandwidth (UWB) sense. This approach is discussed in section 4.2.4.2

4.2.3 Analysis of the pure mesh

The vast majority of published research on meshes stems from interest in ad hoc networks and so considers traffic flows that are contained entirely within the network, i.e. "intra-mesh" traffic.

Mathematical representation is difficult and complex because of the large number of parameters that must be specified (e.g. radio range per link as a function of bandwidth per link, transmit power control procedures, radio environment (propagation law, clutter losses), signal-to-interference margins vs. carrier modulation type and data rate, multiple access scheme (TDMA, CDMA, FDMA), traffic models, user mobility models, traffic routing strategy, etc. As such there can be no single representation: instead one must constrain parameters in order to construct a tractable problem. Common limitations adopted in the literature are one or more of the following:

- Static users
- Fixed link bandwidth
- Fixed link range
- Uniform random, or regularised, distribution of nodes
- Uniform random distribution of traffic

In consequence, each network analysis generally serves to indicate just one or two trends in system behaviour.

The alternative to mathematical modelling is system simulation. Again there is a considerable amount of activity in this field but the vast majority of it employs some form of IEEEE 802.11 air interface and MAC layer – taken from the ad hoc networking arena. It is generally agreed that the 802.11 series and associated MAC and routing algorithms are sub-optimal for efficient mesh networking. Again it is necessary to reduce the number of variables by constraining parameters as listed above. Thus, again, such simulations merely serve to indicate performance for a specific configuration. For the above reasons, it is essential to treat any theoretical and practical work with caution and to fully understand the assumptions that have been made, before drawing any conclusions.

Gupta and Kumar [2000] are the authors of much of the pivotal theoretical analysis of mesh networks which excited early work and is taken as the reference work by several subsequent researchers. But Gupta's model is highly idealised and leads to optimistic conclusions which the author himself warns against.

Note:

Gupta (and others) talks of "Arbitrary Networks" and "Random Networks", and two protocol types: "Protocol Model" and "Physical Model".

The "Arbitrary Network" is defined as one in which node locations, traffic sources and destinations, and traffic demands are arbitrarily arranged to achieved best performance (by some metric). The "Random Network" is one in which these parameters each have a statistical distribution. Therefore the "Random Network" is more relevant to the analysis of a real network.

The "Protocol Model" assumes that the protocol is able to prevent neighbouring nodes transmitting at the same time on the same channel within a specified range. The "Physical Model" sets a minimum interference ratio (SIR) for successful reception. Therefore, the Physical Model is considered the more relevant to practical implementations.

Key features of Gupta's theoretical model are:

- Uniform, random distribution of Nodes in a unified propagation environment
- All nodes transmit at the same power level.
- All nodes have the same relay throughput rate (bps).
- A target signal to interference (SIR) ratio is set.
- All nodes are static.
- Log-normal fading is not included.
- Nodes employ omni-directional antennas.
- All overheads for channel coding, routing and management traffic are ignored.

Gupta's work is specifically concerned with scalability: he does not address absolute capacity. His key conclusions are:

1. For a network of n identical randomly located nodes, each capable of relaying usable data at W bps over a fixed range, the upper bound for the throughput capacity of the total network is of order[3] of $c_1 W\sqrt{n}$ bps. Thus the network's capacity increases in proportional to the square route of the node population.
2. This capacity is shared amongst the nodes such that the upper bound for the *average* throughput $\lambda(n)$ obtainable by each node for a randomly chosen destination is of order of $c_2 W/\sqrt{(n.\log n)}$ bps for the Random Network with Physical Model. Thus the per-user throughput decreases with increasing node population.
3. The parameters c_1 and c_2 are functions of the signal-to-noise threshold, β, required of the carrier modulation scheme and the rate of decay of RF signal power (the propagation law), γ, such that for a high signal-to-noise threshold the capacity limits are reduced, whilst for a high propagation law the capacity limits are increased – as one would expect.

The above results are frequently quoted by researchers, but one must remember that these are idealised theoretical upper bounds on performance. Gupta himself states that:

> *"The results in this paper allow for a perfect scheduling algorithm which knows the locations of all nodes and all traffic demands, and which co-ordinates wireless transmissions temporally and spatially to avoid collisions which would otherwise*

[3] The phrase "of order of" in this context means that to a first-order approximation the value tends towards this value as n approaches infinity. Parameters C_2 , C_3 etc are functions of the signal-to-noise ratio β for successful reception and the propagation law, γ

> *result in lost packets. If such perfect node location information is not available, or traffic demands are not known, or if nodes move, then the capacity can only be reduced".*

Gupta's results are highly optimistic. Other researchers have attempted to analyse more practicable scenarios and find a more pessimistic performance. An example is theoretical work by Arpacioglu and Zygmunt [2004] in which he concludes that the average per-user throughput λ(n) obtainable for a randomly chosen destination has a faster rate of decay approximately proportional to $1/n$, rather than $1/\sqrt{n}$ or $1/\sqrt{(n.\log n)}$ predicted by Gupta. In essence the primary reason for the more pessimistic performance is that in Gupta's model the path loss is modelled as d^{γ} , but d is allowed to reduced towards zero as the density of nodes is increased, thus there is an anomalous *decrease* in path loss for r<1. Arpacioglu and Zygmunt eliminate this by setting path loss as $(1+d)^{\gamma}$. This leads to the result that the network's capacity does not increase monotonically with n, but rather that there is an upper limit to the number of simultaneous transmissions that can be supported in a given area – and that is independent of n. Intuitively this is correct when one sets practical limits on propagation attenuation law, and required signal-to-interference margin.

But regardless of these disagreements over the order of proportionality with n, this family of models all agree that average per-user throughput diminishes towards zero as the number of nodes increases – <u>thus the mesh network does not scale indefinitely</u>.

It is interesting to consider what parameters, if any, might be changed to avoid this demise. For this discussion we consider a simplified view of a network model: using Arpacioglu and Zygmunt's [2004] model the dependencies on system parameters can be logically and simplistically stated as:

average throughput λ(n) is proportional to functions of (γ, W, G/β, $1/L$, $1/r$, A, and $1/n$)

where γ = propagation attenuation law

W = channel transmission rate

G = channel processing gain

B = required signal to noise ratio

L = mean end-to-end path length

r = mean per-hop link length

A = area covered by network

n = number of nodes

This implies that unless one or more of the parameters grow with n then per-user throughput will be asymptotic to zero:

- W cannot grow arbitrarily large because of thermal noise constraints and limits on transmission power.
- G/β depends on the properties of the communication system and increasing it generally makes it necessary to decrease W.
- Reducing hop length r (e.g. by constraining transmit power) increases spatial re-use but at the expense of increased hop-count and hence increased relay traffic. However it turns out (Gupta and Kumar [2000], Arpacioglu and Zygmunt [2004] and others) that the preference is to reduce r to increase spatial re-use. But there is of course a limit here in that if r is too small then the network can become disconnected, i.e. minimum r is related to the inverse of node density (A/n).
- In random traffic flow models with uniform node density the mean end-to-end communication

path length, *L*, is assumed to grow with coverage area *A* (*L* proportional to $\sqrt{A}$). This reduces capacity because of increased hop count. Thus, if one could conceive of services with more localised traffic (e.g. amongst localised communities) then *A/L* will increase more rapidly with increasing *A*. This will help to improve scalability.

- The remaining parameter that might scale with n is the area *A*. Arpacioglu and Zygmunt ([2004] Corollary 3) suggests that three factors are required to achieve a non-zero throughput with increasing n: (i) the attenuation law γ needs to be greater than 3, (ii) the hop count *H* needs to be independent of n, (iii) area, *A*, needs to increase with n (i.e. the node *density* needs to be nearly constant or reducing with increasing *A*).
  However, (iii) requires that as the subscriber base increases those subscribers spread themselves out more thinly. It is not easy to see on what basis this might happen in any practical deployment.
- The propagation attenuation law γ strongly influences the above conclusions. A higher attenuation factor γ will permit higher throughput capacity (Gupta and Kumar [2000], Arpacioglu and Zygmunt [2004]).

From the above list of options, one can see that there appears to be very little prospect of avoiding the asymptotic reduction in per-user throughput with increasing subscriber base.

4.2.3.1 The primary underlying causes of limited capacity

It is useful to identify the primary reasons why per-user throughput decreases with increasing population in a mesh supporting intra-mesh traffic.

When considering *average* rates of traffic flow, this decreasing throughput is *not* due to blockage caused by the limited relay throughput of nodes. Instead it is due to spectrum sharing within the mesh and the need to relay traffic through multiple hops.

However, when considering *actual* traffic flows the throughput may be constrained by the limited transport capability of nodes.

Consider the activity around a single node. The use of an element of time/bandwidth resource to communicate across a hop will impose an interference boundary around the transmitting node within which that resource cannot be re-used. Other nodes wanting to communicate within this interference zone must use other elements of time/bandwidth resource. If the transmission rate of nodes in this bandwidth, *B,* is *W* bps then the maximum total throughput through this interference zone is of order *W.* Other traffic paths can pass through this zone, but the total is limited to *W* bps. If node density is such that there are m other nodes in this zone then the zone's throughput can be shared - providing a mean of *W/m* to each. If this is zone is of area*, a*, then one can consider having a maximum throughput of order *W/a* , (bps per unit area) in this zone. Clearly, then, it is advantageous to keep this area, *a*, as small as possible. This confirms the conclusion of other researchers that short hop lengths and high propagation attenuation factor are conducive to high throughput capacity of the network *(but subject to other limitations such as routing overheads, route volatility, susceptance to mobility etc., see section 6.1).*

Thus it can be seen that – subject to the influences of node density and the peak rates of *actual* traffic flows– throughput is not limited by blocking at individual nodes but by the throughput capacity of the interference zones.

The fact that node relay-throughput is not a limiting factor is illustrated in a system simulation by Hekmat and Mieghem [2004]. This paper addresses mesh throughput from the standpoint of signal-to-interference levels within a network, and from this attempts to determine values for hop-count, capacity and throughput per node. (Its weakness is that it uses a regular lattice of nodes rather than random distribution, but it does attempt to model some practical values for data rate and bandwidth – based on 802.11b (CDMA)).

Several graphical results usefully show some key trends. Of relevance to this present discussion is Figure 8 which illustrates a "saturation point" at which the achievable per-node throughput (as limited by interference from other nodes) equals the required per-node throughput to support a given offered traffic level. This achievable capacity is approximately half that of the "stand-alone" throughput capacity of a node, W (W = 2Mbps in Hekmat and Mieghem's simulation), i.e. network saturation is caused by mutual interference, not by node saturation.

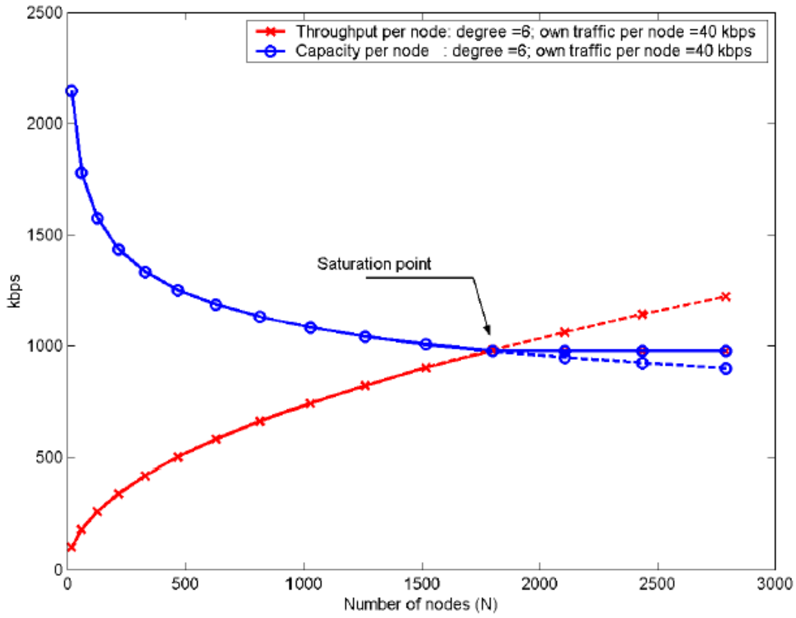


Figure 8: Comparing Available and required Relay Throughput Per-Node (802.11 system with 22MHz channel width, before CDMA de-spreading, and 2Mbps relay rate)

Returning to the concept of the interference zone around a node: one can extend this concept to a multi-hop route passing through the mesh. Figure 9 illustrates such a route as comprising a sequence of transmit/receive boundaries:

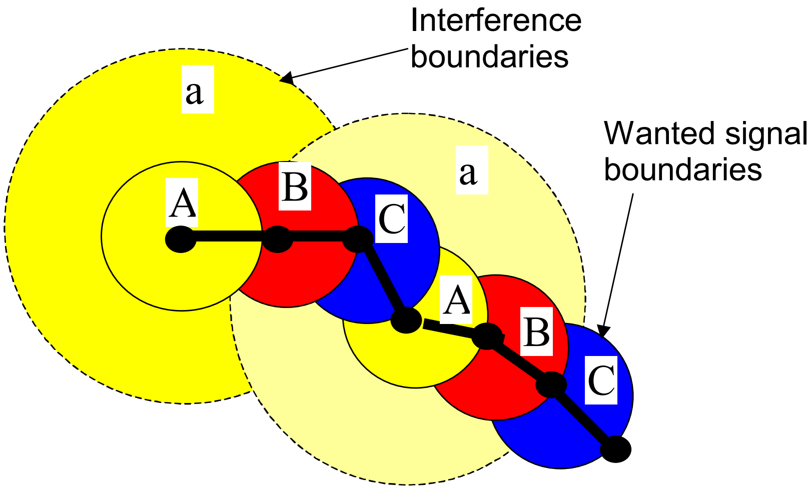


Figure 9: Spectrum resource re-use along a traffic route and associated interference zones around transmitters

In this illustration the colours and letters A, B, C represent usage of different bandwidth/time resource to pass traffic along a route . The smaller circles indicate the omni-directional boundary of the wanted signal on each link and the larger circles indicate the interference zones corresponding to each of these (for clarity only the interference boundaries for resource A are shown).

In this example three resource-elements are needed because the interference zone around a transmitter is assumed to extend to nominally twice the communication distance and so a minimum of three hop lengths separation is required between co-resource users[4] – this represent a near-optimal situation for a "string-of pearls" route in which hops are nominally aligned and are of the same length. In practice the resource utilisation is likely to be substantially higher: Li et al [2001] suggest a theoretical lower limit of four hop lengths but simulations using a modified 802.11 MAC required up to seven.

Thus if the number of different resource elements needed per route is represented as b (b = 3 to 7, suggested above), then for a traffic rate of T bps a resource capacity of bT is used for each route.

One can then deduce that the total number of such traffic routes, m, which could cross within any one interference zone, area a, around an active Node is limited by:

$$\sum_{i=1}^{m} bT_i \leq W ,$$

Thus, for example a system employing Nodes with 20Mbps relay throughput, supporting traffic rates up to 1Mbps, and using an average b = 5 spectrum resource elements per traffic route, could support only four such traffic routes passing through each interference zone around active Nodes

Furthermore, one can see from Figure 9 that a single traffic route lays down a footprint of adjoining interference zones along its path and so extends this problem throughout its length.

The crossover "bottleneck" caused by a specific interference zone could to some extent be reduced

[4] This factor 3 is derived on the following basis. Let required SNR be β , let propagation attenuation factor be γ then the theoretical ratio of interference range vs wanted link length, Δ, given by $10\gamma \log\Delta = \log\beta$. E.g. if β=13dB and γ=4 then the range-ratio is approx. 2. **This neglects log normal fading.**

by diverse routing around it, using a suitable load-balancing routing protocol, but since the interference zone is large (e.g. circa 2-hop radius in this example) routing around it is likely to impose a considerable increase in route length (via hop-count) – with associated degradation of performance. This issue is illustrated in Figure 10 for just three crossing routes.

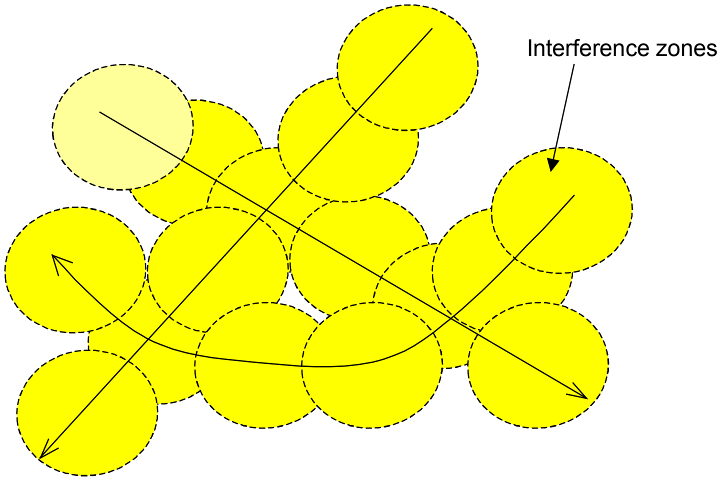


Figure 10: Routing three paths to avoid a three-route occupancy of one interference zone

From this analysis one can also see yet another issue: that as the offered traffic rate, T, approaches a high proportion of the node relay rate, W, such that bT approaches W, then all the spectrum resource is employed on this route. Thus a single traffic route imposes an un-crossable boundary through its length and the mesh becomes *disconnected* as illustrated in Figure 11*:*

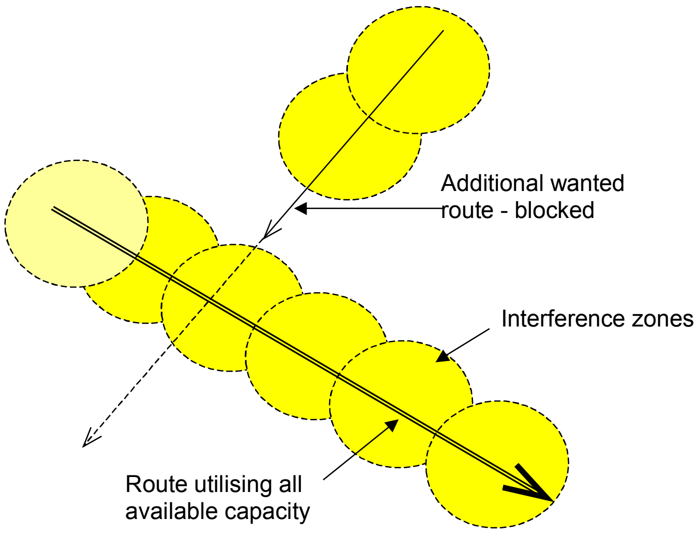


Figure 11: Single high throughput route causing partitioning of the network

All the above artefacts will combine to dictate the upper limit on the maximum traffic rate that can be supported by the mesh. (Remember that the theoretical analysis of Gupta and Kumar [2000] and others determines only the *average* throughput that can be offered to all nodes – its says nothing about the (higher) traffic-rate that could be offered to a subset of the nodes). **This is a topic which would benefit from further investigation.**

4.2.3.2 Practical performance

The above theory establishes some basic trends, but for any practical implementation these upper bounds will not be achieved because of such factors as routing overheads, mobility, and propagation environment.

The following two practical examples were identified from published literature.

- MeshDynamics [2005] offer white papers covering two views of scaling. Their conclusion is a very substantial rate of degradation of per-user throughput proportional to 1/n^2, however they accept that other researchers conclude that it is closer to 1/n. The 1/n^2 result includes the effect of routes not being the idealised 'string of pearls' illustrated in Figure 9. MeshDynamics have a three-radio solution which allegedly solves the scaling issues, i.e. throughput scaling is unity, independent of user numbers. But this is in fact a Hybrid Network comprising a fixed wireless-relay backbone network.
- Gupta et al [2001] have empirically found the scaling of 802.11 (specifically) to be 1/n^1.68. By applying the MeshDynamics discussion, it is no great surprise to find this scales between 1/n and 1/n^2, given the MAC is a collision avoidance MAC, ensuring only one Tx/Rx is active concurrently.

4.2.4 Analysis of the myth of self-generation of capacity

Returning to the apparent mechanisms for overcoming the non-scalability of the pure mesh, identified in Section 4.2.2, the following sub-sections review three of the candidate techniques.

4.2.4.1 Intra-mesh traffic: the 'mobile-courier' scenario

A foundation paper [Grossglauser and Tse 2001] is often cited under the heading of 'mobile mesh'. However, Grossglauser's model specifically uses the mobility of nodes to act as intermediate couriers of data between source and destination. Datagrams are passed from source nodes to near neighbours and delivery occurs when the courier nodes encounter the target recipients. Under this idealised model the per-node throughput remains constant – independent of the number of nodes, i.e. the mesh network is fully scalable in terms of capacity.

However, an obvious consequence of this model is that the end-to-end packet delivery delay is related to the transit time of nodes moving throughout the area covered by the mesh – statistically the *mean* delivery time is of order of *2d/v* where *d* is the diameter of the mesh network and v the mean velocity of nodes within it. In a practical situation, of course, the courier nodes may never encounter the recipient – in which case traffic is lost. The author accepts that this is clearly not acceptable for voice, or other real-time communications, and so directs the concept to non-critical store-and-forward messaging applications. (Note also that the outward and return paths are completely uncorrelated in time and space, and so the concept of two-way communications is only loosely applicable).

Although of limited application in its basic form, the technique might be enhanced to reduce the transport delay and increase the probability of message delivery by, for example, the following improvements:

1. The originating node passes its datagram to <u>all</u> of its in-range neighbours so that there is a number of nodes acting as couriers – so increasing the probability of intercepting the recipient

within a given time (this is developed by Sharma and Mazumdar [2004]).

2. The courier nodes pass their datagrams on to neighbouring couriers at some event trigger such as a specified time interval, or the extent of their movement through the mesh. This spreads the data further throughout the network in a "virus" fashion – again increasing the probability of intercepting the recipient within a specific time interval. This approach starts to bridge the gap between Grossglauser's basic single-courier approach and a fully connected real-time route through the mesh (this is developed by Sharma and Mazumdar [2004]).
3. For (2) above, the courier process might be augmented by nodes retaining a database of all other nodes they have had contact with and so selecting couriers on the basis of those that have had recent contact with the recipient.

Each of the above enhancements to reduced end-to-end delay will increase the traffic activity on the network and so will decrease its capacity somewhat. Thus this approach illustrates fundamental trade-offs between throughput capacity and delay.

4.2.4.2 The potential of using 'infinite' bandwidth

An interesting alternative scenario is suggested Negi and Rajeswaren [2004], in which the channel bandwidth, B (Hz), is very large and increases as n increases. Negi and Rajeswaren suggest that this approach could achieve a per-node throughput that is an *increasing* function of n, i.e. the network is fully scalable[5].

But this is a false conclusion in that the transmission bandwidth, B, must *increase as n increases*. This work does not, therefore, depart from the fact that with *fixed* bandwidth the per-user throughput will be a decreasing function with increasing n.

However, we included the work here because it may give useful clues to an intermediate solution. One aspect of Negi and Rajeswaren's model is that the transmission bandwidth is set wide enough to render transmission from neighbouring nodes below the thermal noise floor. This has potentially interesting implications for very wideband technologies such as DSSS, Fast Frequency hopping, and Ultra-wideband (UWB) transmission **As such it raises topics for future investigation.**

Specific features of Negi and Rajeswaren's model are:

- Link adaptation is used such that the data rate over a link corresponds to the Shannon capacity limit[6]. (Note that this is unattainable in practice and so a scale-factor is required on these results).
- Transmission bandwidth and spreading gains are sufficiently wide to permit all interference to be tolerated even when *all nodes transmit simultaneously.*
- Transmission bandwidth and spreading gains *are increased with increasing node density*.
- Propagation law γ is greater than 2.
- Automatic transmit power control (APC) is used.

4.2.4.3 The benefit of traffic localisation

The theoretical analysis of Gupta and Kumar [2000] and others assume a random association

---

[5] Specifically he concludes that the average throughput $\lambda(n)$ obtainable by each node for a randomly chosen destination has an upper bound of order $c_a.n^{(\gamma-1)/2}$, and a lower bound of order $c_b.\{n^{(\gamma-1)}/(\log n)^{(\gamma+1)}\}^{1/2}$ (where c_a and c_b are constants of proportionality) as n and B tend to infinity.

[6] Shannon Capacity limit = B.log(1+SIR) where SIR is the signal-to-interference ratio, B is the bandwidth.

between source and destination nodes. Thus path lengths range from nearest neighbour (one-hop) to the full diameter of the area covered (many hops), and so as the network size increases geographically and/or in terms of node-density the number of hops per path must increase and this is one of the primary factors which cause the reduction in capacity with increasing number of nodes.

It is clear, then, that if traffic flows were more localised amongst neighbouring nodes, regardless of the geographic size of the network, then the number of hops per path would not increase pro-rata with size and so the network would scale better - i.e. the more localised the traffic flows, the more capacity can be supported and the less this is affected by growth in population.

This suggests identifying applications whose traffic is predominantly amongst close neighbours. However, whilst this might have some prospect for *fixed* mesh applications in which the geographic location of users if fixed and known, it has far lower prospect for mobile applications. Possible scenarios for mobile applications are where there is local traffic flow to/from fixed sources such as retail outlets, transport management/information points, public information points, etc.

For the case where the end-to-end path length does increase as the network is scaled, an alternative technique is to route this longer-range traffic over a hierarchical fixed network – as discussed in Section 4.2.5.

4.2.5 Hybrid mesh network

As stated earlier, one of the key limitations to the efficient flow of intra-mesh traffic through a network is that, as geographic size and/or user-density increase the conflict between using small hops for spectral efficiency through spatial re-use, and longer hops for reduced delay and better route connectivity becomes un-reconcilable.

One means of addressing this conflict is to add a fixed infrastructure in the form of an overlaid network configured as a fixed mesh network. This results in a Hybrid Network, as illustrated in Figure 12.

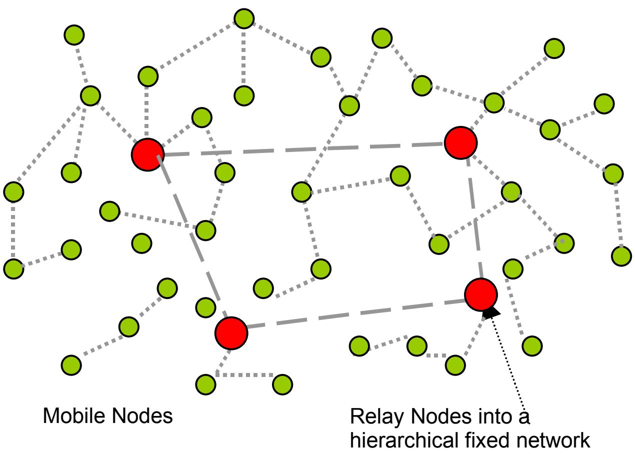


Figure 12: Hybrid Network: Intra-Mesh traffic with Infra-structure support

Interconnect to the fixed mesh network is via Relay Nodes. A Relay Node is a Layer 3 device, similar to a client node except that it is not itself a source or destination of traffic, and possibly has higher throughput. A routing strategy is adopted by which, if the traffic destination node is in the same 'cell' as the associated source node then the traffic is routed via the peer-to-peer hops, but if the traffic destination is in a distant cell then it is routed via the infrastructure.

In such a network the overlaid fixed mesh adds capacity and so the overall capacity and scalability are greatly enhanced.[7]

Some useful foundation theoretical work on this architecture has been presented by Liu et al [2003]. Liu et al conclude the following:

1. If the quantity of Relay Nodes, *m*, increases at a rate less than √*n* then although there is a substantial increase in the capacity of the network the rate of decay of per-user throughput with increasing nodes, n is not improved substantially – i.e. capacity is improved but scalability is not. Specifically:

   The network capacity is nominally proportional to $$\sqrt{\frac{n}{\log\left(n/m^2\right)}}$$ …..(1)

   The capacity gain factor over Gupta's is of the order $$\sqrt{\frac{\log n}{\log\left(n/m^2\right)}}$$ …..(2)

2. If the quantity of Relay Nodes, m, increases at a rate greater than √*n* then there is useful improvement in the scalability as well as capacity. Specifically

   Network capacity increases polynomially with *m* (i.e. proportional to $\sqrt[a]{m}$ where *a* is less than 1)

   The capacity gain factor is of the order $$m\sqrt{\frac{\log n}{n}}$$ ….(3)

3. In the limit, when *m* increases at the same rate as *n*, the throughput per user remains constant – i.e. the network is fully scalable. This characteristic is self-evident from the fact that in this case each Relay Node serves a constant number of nodes, therefore each node retains a constant share of the total transmission bandwidth available. At this point one has the equivalent of a cellular network where the network capacity scales directly with the number of base stations, where:

   Network capacity is proportional to *m* (i.e. proportional to *n* in the limit)

Boppana and Zheng [2005] offer a simulation of such an architecture using 802.11 (WiFi) which, although technology- and scenario-specific, does illustrate the attainable improvement in capacity and per-node throughput with the additional of the hierarchical fixed mesh network. Adjacent pairs of fixed Relay Nodes are interconnected by a wired link, not wireless. (Note that the bandwidth of each wired link in the hierarchical network was set at 2Mbps – the same as the bandwidth of the radio links of the mobile nodes.) The same radio spectrum is used for peer-to-peer links between mobile nodes and for mobile-to-Relay Node links.

For a simulated 1000-node network extending to 36km^2 in a 4th-order propagation environment (and neglecting log-normal fading) the network throughput capacity is increased by 63% by adding 9 fixed Relay Nodes with 12 interconnecting paths, and increased by 190% by adding 25 Relay Nodes with 40 interconnecting paths. These figures are of course specific to this particular model, but they do serve to indicate the potential benefits of a Hybrid network.

---

[7] Note that although this architecture looks similar to the "cellular-plus multi-hopping" or "mesh plus Access Points for extra-mesh traffic" architectuures, we see a significant difference in that the Relay Nodes are routers as per the subscriber nodes and therefore are included in the scope of the routing protocol.

There are substantive differences between the working assumptions in Lui's theory and Boppana's simulation and so direct comparison between the two can be misleading. Furthermore the theoretical analysis is idealised and represents the limiting case for large networks. Nevertheless we present a comparison of the capacity gains in Table 5:

	Capacity gain, Boppana	Capacity gain, Liu
1000 nodes, no Relays	Reference level	Reference level
1000 nodes, m=9 Relays	57%	66%
1000 nodes, m=25 Relays	190%	384%

Table 5: Capacity Improvements in a Hybrid Network

Further benefits noted [Boppana and Zheng 2005] are a reduced percentage of link breakages (caused by user-mobility) and reduced end-to-end delay (through lower hop-count).

There is a potential down-side due to the traffic bunching around the Relay Nodes into the hierarchical network (similar to extra-mesh traffic flowing through Access Points). Clearly, in this respect, the network's performance is conditional on there being an equitable balance between local and long-range traffic.

This hierarchical network could use radio, optical, or wired links between Relay Nodes. Clearly, if the hierarchical network uses radio links then there is additional radio spectrum required for this network and so there is an impact on overall spectral efficiency. Whether or not the additional spectrum required is matched by the achieved gain in capacity - such that overall spectral efficiency is not degraded - will depend on the scope for frequency re-use within the hierarchical network. **This merits further investigation.**

4.2.5.1 Conclusion on the Hybrid Mesh network

It is clear that adding a hierarchical infrastructure to a mesh network will substantially enhance it its performance in terms of capacity, scalability, reliability, and ability to support long-distance traffic flows.

This hybrid architecture is the only solution to supporting general purpose mobile networking for peer-to-peer traffic and to achieving full scalability of mesh network capacity (i.e. constant per-user throughput as customer base increases).

4.2.6 Access mesh network for extra-mesh traffic

All of the above analysis relates to traffic that is routed between peer-to-peer users either directly or via a fixed infrastructure mesh network.

However it has already been pointed out that there are very few user applications that would suit that model – particularly in the commercial and consumer sector. Far more relevant to today's applications for wireless communications are services which require access to a public Access Network, such as for national/international telephony and the Internet. (As discussed elsewhere in this report, there are also the potential requirements for network management and billing traffic that must flow to/from a management centre via a public Access Network)

For such applications traffic flows are centred around Access Points or gateways, as illustrated in Figure 13:

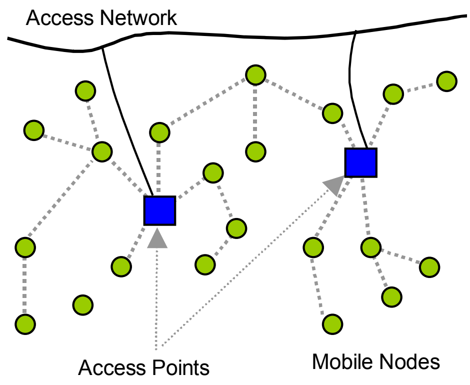


Figure 13: Mesh network with Access Points for Extra-Mesh traffic flow

There is now a concentration of traffic around the Access Points and an associated higher burden of traffic flow through those client nodes that provide the last hop to the Access Points.

This traffic concentration substantially reduces the capacity of the underlying mesh network (compared to the intra-mesh traffic case) and the per-user throughput. This arises because, in the absence of any sectoring of the Access Point coverage, all of the hops into the Access Point are in re-use contention. Hence the maximum combined throughput of all of these hops is not more than *W*.

This capacity is shared amongst the *n* nodes serviced by the Access Point – resulting in a scaling factor proportional to $1/n$. (The network is closely akin to a cellular network in which the capacity is set by the base station and this is shared amongst the cell's occupants).

However, in practice the interference domain of the first tier of links into the Access Point may also encompass most of the second tier links into the nodes and possibly some of the third tier, thus the maximum available network capacity may be limited to $W/2$ or $W/3$.

Thus one can say that the per-user throughput is limited to approximately cW/n, where W is the relay throughput rate of a node and c is a constant of proportionality (as a function of system parameters, re-use interference, multi-hops and routing overheads).

An example simulation of a small network with a single Access Point configuration is given by Jun and Sichitiu [2003] This is presented in the context of a *fixed* mesh, but it serves to illustrated the traffic-concentration issue. A network of 30 nodes using 802.11b at a raw transmission rate of 11Mbps and an aggregate relay throughput of 5.1Mbps (i.e. W = 5.1Mbps) is shown to achieve results which are asymptotic to a network capacity of about 1.5Mbps, i.e. an average throughput per node of about 50kbps. This implies a value for *c* in the region of 0.3.[8]

[8] Note: this is for a static mesh with pre-determined, static, traffic routes: there would clearly be additional overheads (i.e. lower value for *c*) for a mobile, dynamic, network.

4.2.7 Conclusions

Mesh and ad hoc networks comprise user-nodes which relay traffic between themselves in order to achieve end-to-end traffic routes. As such, each new node joining the network adds capacity and connectivity. However, the raw statement that *as the number of nodes increase so does the capacity of the mesh* is disingenuous. It is true that network transport capacity increases with increasing node population, but in sharing this resource amongst the users there is a net decrease in the available throughput per node.

The conclusion presented in this report is that meshes have no especially good properties with respect to scaling. In particular as node density and geographic size increase, the traffic rate available to any particular user decreases.

In fact a contrary viewpoint may be given in the form: *as the number of nodes increases, the integrity of the network increases through improved connectivity but its per-user throughput decreases*.

This report shows that this lack of scalability can only be overcome by either adding additional capacity in the form of a hierarchical network or containing the end-to-end traffic flows to localised regions within the network

All current theory and measurement of ideal, novel and practical meshes conclude that ad hoc mesh networks comprising only peer-to-peer communication links do not scale well with increasing node population unless there are specific limitations on the density of nodes; the propagation environment and the traffic models.

It is principally the need to share the spectrum coupled with the need to relay other traffic which leads to a scaling of between 1/n and 1/n^2 in per-user throughput. To avoid this asymptotic decrease, either communication must be kept within small regions or bandwidth must be added in a planned manner beyond that which the users ‘add’ themselves- leading to a Hybrid Network of fixed and mobile interconnects.

Additionally and beyond these PHY capacity constraints, there exist practical MAC and routing challenges which further push for meshes which have a low hop count – and hence localised traffic flows.

As a final point: whilst the capacity of a cellular system can, within limits, be expanded post-deployment, mesh systems are less flexible. For a mesh, the network capacity is a function of the capability of the nodes. Operators/suppliers must therefore take a view at deployment time as to the long term subscriber density and desired services and provide nodes that can support these. As performance depends on the subscriber's equipment it cannot be easily upgraded as required. It follows therefore that nodes may have to be substantially over-dimensioned from the outset - with associated R&D and product cost implications.

Finally, Table 6 summarises the detailed findings with respect to mesh scaling.

	Finding	Comments
1	Average per-user throughput decreases as the node population increases.	Estimated figures range from $1/\sqrt{n}$ to $1/n^2$, but the underlying trend is that throughput decreases asymptotically with increasing n and so the mesh network cannot scale indefinitely. This is due to mutual interference and to relaying traffic over multiple hops.
2	If traffic flows are localised in the network, then throughput can be independent on the total number of nodes, n.	This requires identification of specific user-applications which do not demand ubiquitous flows across the whole area of the network.
3	Highest frequency reuse occurs when hop-length is minimised.	But short hop-length increases vulnerability to link breakages due to mobility. There is thus a careful balance to be defined between these conflicting characteristics.
4	A higher attenuation propagation environment enables higher network capacity.	This is true to the extent that spatial reuse is increased in a high attenuation environment. However performance is governed by the density of nodes: if this is too low for the prevailing environment then connectivity is poor and routing is unreliable.
5	The only mechanism for substantially improving the scalability of a mesh network and its ability to carry long-haul traffic is to overlay a hierarchical fixed mesh network, resulting in a Hybrid Network.	Clearly this involves deploying a fixed, planned, infrastructure and so is a departure from a stand-alone ad hoc network. The fixed network could employ wire/fibre interconnect, in which case overall spectral efficiency is not impacted. Or it could employ wireless interconnect. High spectral efficiency in the fixed network could be achieved by narrow-beam directional point-to-point links
6	To support Extra-Mesh traffic requires connection to an Access Network.	Again this involves deploying a fixed, planned, infrastructure and so is a departure from a stand-alone ad hoc network.

Table 6 Summary of detailed findings re. capacity scaling

4.3 Hypothesis Testing ... that mobile meshes are more spectrally efficient

This section addresses the topic of network spectral efficiency. It brings practical issues together with theory and concludes that

1. there seems little reason to believe that *practical mobile meshes* will be intrinsically more spectrally efficient than traditional PMP networks
2. the mobility of users in the mesh can lead to a much bigger trade-off between efficiency and availability than current cellular systems, since mesh users are also system routers. (Adding fixed relay-nodes as part of the infrastructure can mitigate this effect).
3. any preference for implementing a mesh should be based on other benefits such as coverage or lack of infrastructure, etc.

The section begins with an introduction to the problem, followed by an outline of the fundamental supporting arguments for meshes offering higher spectral efficiency, and then describes the practicalities that may serve to negate this advantage. This includes reference to other, published work in the area.

4.3.1 Introduction

As stated, the hypothesis *'...that mobile meshes are more spectrally efficient',* is essentially un-testable without consideration of further information. For example there are many possible deployment scenarios varying in range, numbers of subscribers, traffic flows etc. There are also several possible definitions of spectral efficiency; this report takes bps/Hz/km^2. Additional complexity arises from the fact that practical systems comprise not just a topology, but a whole range of components such as protocols, radios, antennas etc. Consideration of just one element may be misleading. For example, fixed mesh systems claim high spectral efficiency but this efficiency is known to arise principally from the use of highly directional antennas, e.g. Radiant Networks [e.g. Radiant 2005], rather than the use of the mesh architecture *per se*.

With so many variables the danger of comparison is that, by judicious choice of scenario and parameters, proponents of any given system may 'prove' the superior efficiency of their topology whether it be mesh or PMP. A similar conclusion was reached by an earlier study [Aegis 2000] which also compared PMP and mesh architectures, although in a different context.

In order to provide a level and fair comparison, this section first addresses the question 'more spectrally efficient' *than what*? Figure 14 helps illustrate a simplified comparison of mesh with cellular, assuming a convenient, if unlikely, scaleable propagation environment.

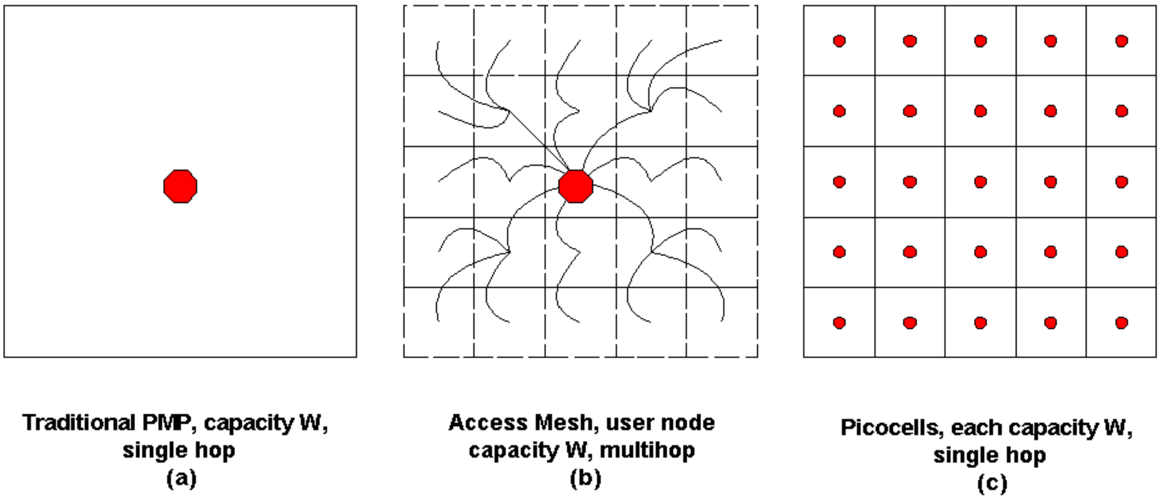


Figure 14 Cellular, picocellular and Access mesh network examples

Figure 14(a) shows a single traditional cell. The cell is drawn as a square for simplicity.

Figure 14(b) shows an access mesh approach with a snapshot of 25 mobile users distributed conveniently, but unrealistically, at the centres corresponding to the picocell sites of Figure 14(c).

Figure 14(c) shows simple square picocells, where 25 remote picocell antennas are provided to serve the whole area. Figure 14(c) is included to show that adding more actual base stations is better than relying on the myth that each added mesh node 'effectively becomes a base station', see sections 2.2.8 and 4.2.2.

The users are free to move over the whole area in all cases.

Although 25 users have been shown for simplicity in the mesh diagram, subsequent discussion assumes a significant number of users i.e. such that the user density is sufficient for the functioning of a mesh system. To simplify the discussion it is also assumed[9] that all active users are allocated the same channel bandwidth. It should be emphasised that this is in terms of Hz, the unit of the underlying resource (spectrum), not bps as used by some other authors.

To simplify the analysis, the premise is that the vast majority of commercially significant applications (e.g. telephony, internet access) would generate extra-mesh traffic if implemented on a mesh. The analysis therefore focuses again on the Access Mesh in which all traffic flows to/from some central gateway, as shown.

If deploying a cellular system, Figure 14(a), the approach would be to locate a base station at some appropriate central position with enough power/sensitivity to meet the required objectives. If a fixed rate service is offered, e.g. a GSM voice call, then the worst case users determine the link budgets, i.e. those furthest from the base station in terms of path loss. The data rate on the channel (or alternately the width of the channel) is dictated by this worst-case link budget. For users nearer the base station the transmit power may be reduced. The spectral efficiency ignoring inter-cell effects can then be simply calculated as user bit rate per channel bandwidth per unit area of cell. Some cellular systems e.g. 3G offer a graded service to users whereby users nearer the base station

[9]One could keep the delivered bit rate constant and vary the required channel bandwidth, the argument is the same. It is noted that there are various possible access schemes such as TDMA, FDMA, CDMA. Each of these have their own benefits/drawbacks in different scenarios. However for the purposes of considering spectral efficiency all of these access schemes can be recast as frequency division; having a common basis aids comparison.

are offered higher bit rates. This grading of service will serve to increase the spectral efficiency.

Now consider the introduction of hopping, i.e. a mesh system, Figure 14(b). In principle the worst that can be done is the same as the cellular system, since it can default to a one hop service to all users. In fact for users near the base station or access point it probably will use a single hop. For further away users, however, it is possible to improve the delivered bit rate by transferring the data in a number of hops, see section 4.3.2. Such gains however depend on user behaviour (distribution, mobility), quite unlike the predictability of a cellular system.

Note that the efficiency of the picocell system, Figure 14(c), may never be approached by the mesh system, since the average path length of the mesh will always be larger and the radio network is not taxed by any backhaul or relay of other user's traffic.[10] Fixed infrastructure is however needed to backhaul traffic (see also discussion of backhaul in section 7.4.1.3).

It is noteworthy that, even when effort is taken to compare like with like, a complication remains due to the fact that meshes in particular have a performance (including efficiency) which depends on instantaneous user-node distribution and mobility. However this is not unlike the adaptive behaviour of 3G systems as mentioned above. Whilst a grading of service will serve to increase the spectral efficiency for some user distributions, it is not a guaranteed effect. The same is true for a mesh; a worst case ought to be used for mesh planning. The difference between worst case and best case in a mesh can be very high, since it is driven by user mobility as already stated. This leads to the efficiency-availability trade-off discussed in section 6.1.3.

The following two sections of this report indicate the relative spectral efficiencies of pure mesh and access mesh networks to PMP equivalents.

4.3.2 Comparative efficiency of pure mesh and cellular

Section 4.2 illustrated that the usable capacity of a pure mesh[11] handling intra-mesh traffic does not scale well with increasing customer-base. (The theoretical upper bound for a static idealised network is a scaling of $\sqrt{n}$ (Gupta and Kumar [2000]). The alternative theoretical analysis of Arpacioglu and Zygmunt [2004] suggests capacity is constant. A practical simulation of Boppana and Zeng [2005]) suggest a small capacity growth-rate of the order of $n^{0.18}$, whilst practical simulations of Gupta et al [2001] and MeshDynamics [2005] and indicate a decreasing capacity proportional to $1/\sqrt{n}$ and $1/n$ respectively.) It must be noted that these analyses ignore overheads due to node mobility.

Note that all of the above theoretical analyses derive only the *mean* throughput available to each node. They do not indicate what can be achieved for applications that demand much higher (fixed) rates for a fraction of the nodes for some percentage of the time, as is a more typical usage. e.g. that a mesh of a 1000 nodes can support an aggregate throughput of 100kbps does not necessarily imply that it can support at any instant of time 100 paths of 1000kbps. To emphasise the point it does not guarantee that at any instant of time it can support any random set of 1000 x 100kbps paths.

However, none of these analyses quantify the absolute capacity/spectral-efficiency (they only address scalability). Thus the important question is: what is the absolute spectral efficiency of a small mesh and how does this compare with a PMP network?

---

[10] Whilst the picocell wins on spectral efficiency grounds, it is often discounted on financial grounds. It should be clear that this is a wholly different basis on which to argue and one which may be already changing with time as RF-on-fibre picocells become available.

[11] One without any infrastructure or fixed relay-done support

But direct comparison is not possible, not only because of the very large range of parameters involved and assumptions about functionality that must be made, but because a PMP network can continually increase its capacity by deploying more infrastructure (basestations). Thus a first-stage comparison must be that of a single basestation PMP cell and a pure mesh serving the same population over the same geographic area.

Unfortunately most of the published work addressing absolute rather than relative amounts of capacity makes use of IEEE802.11 protocols. Nevertheless the Boppana and Zheng [2005] simulation, for example, can be used to provide a first order estimate of absolute capacity of a pure mesh. The simulations assumed a 2Mbps node capability. Consider a hypothetical situation in which an isolated GSM cell provides telephony services between users within its coverage area (i.e. emulating intra-mesh traffic flow). For GSM telephony, of the order of 8kbps in each direction is needed so a 2Mbps capacity a base station can support 125 duplex channels. GSM systems provision base stations on the basis of about 0.06 Erlang/user, so 2Mbps should support of the order of 1600 users (leaving some headroom for peak loads). This figure is close enough for extrapolation from the Boppana and Zheng simulation of 1000 users to be worthwhile. Their simulation achieved a total usable capacity of about 500kbps so that, using the capacity-scaling law deduced from their results (capacity $\propto n^{0.18}$), it would be anticipated that a mesh of 1600 users will be able to support an aggregate throughput of about $500\times(1600/1000)^{0.18}=544$ kbps which is equivalent to only 34 duplex channels.

In other words, even when providing an intra-mesh service the mesh only achieves about 25% of the efficiency that a single-cell PMP system might be expected to achieve. And this excludes the effects of node mobility.

NB: The Boppana and Zheng simulation used and IEEE802.11 MAC[12]. Whilst this is almost certainly unfair, since the IEEE802.11 MAC does not optimise mesh performance, it is difficult to see how optimising MAC performance alone could overcome the apparent large disparity.

But, furthermore, if the customer-base increases either in density within the same area or increased numbers over an extended area, then the mesh network capacity stays near constant or diminishes – according to the scaling law achieved – whilst the PMP network can deploy more basestation capacity. Thus the pure mesh network cannot possibly match the PMP equivalent.

The only way that the pure mesh can compete is by departing from purity by adding infrastructure in the form of an overlaid backhaul network to transport the longer-range intra-mesh traffic – as illustrated by the hybrid mesh of Figure 12, with its hierarchical relaying architecture. This architecture effectively breaks the overall mesh down into a number of smaller meshes and so improves scalability and capacity – as discussed in Section 4.2.5. However, this assumes that additional bandwidth is available either in a wired overlaid mesh or a wireless overlaid (fixed mesh). In the case of using a wired overlaid mesh it is accepted that, if this is formed from fixed high-gain antennas, it can be very spectrally efficient. Thus the combined spectral efficiency of the mobile and fixed mesh can be relatively high.

---

[12] The preponderance of 802.11 in mesh work is noteworthy. In part this may be due to expediency, many researchers will have ready access to simulators and equipment that use 802 protocols and thus may choose to use them as is rather than endeavouring to develop systems optimised for meshes. However it may reflect an interesting difference in mind set between groups. On the one hand groups such as MANET are considering meshes essentially for LANs for which mobility and ease of use are paramount virtues and spectral efficiency may be sacrificed. On the other hand commercial operators with a fixed spectrum allocation will take great pains to achieve the maximum in spectral efficiency so as to maximise potential revenue. Currently most academic work derives from the first group.

4.3.3 Comparative efficiency of access mesh and cellular

4.3.3.1 Basic attributes of multi-hopping

One of the traditionally used scenarios for suggesting that mesh operation into an Access Point might be more spectrally efficient than a PMP cell is the concept that increased throughput can be achieved over a series of short hops rather than one long hop. It will be demonstrated that this is only true for an idealised single-path scenario, and is diminished by the dissimilar antenna gains of Access Points and mobiles, the overhead of relaying traffic from multiple users, the spectrum re-use contention within the mesh, and finally, of course, the routing overheads to combat node mobility.

For the case of hopping between nodes of like type, consider node-to-node links in a mesh. If two hops of roughly equal length replace a single hop, as shown in Figure 15, then:

- only half the time-bandwidth product of spectral resource is available for each hop, and this acts to reduce the delivered data rate by a factor of 2
- but as each hop is half the length of the original link, the link budget is improved. This improvement can be used to improve spectral efficiency either by increasing the transmission rate on each hop or reducing the transmit power. For example, in a third-law propagation environment the link budget is improved by x8 (~9dB); this would permit a four-fold increase in transmission rate by changing from QPSK to QAM64. Alternatively, with spread-spectrum the coding gain could be reduced to realise a similar increase in transmission rate.

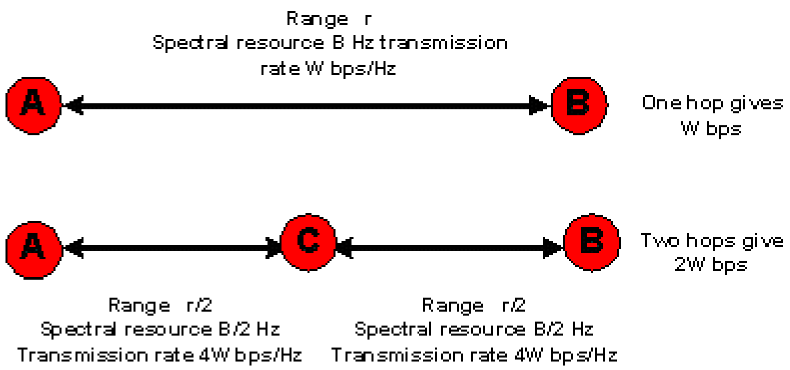


Figure 15: Two-hop vs. one-hop rate improvement between mesh nodes

Overall these two factors imply that twice as much data can be transferred using two shorter hops: i.e. spectral efficiency is doubled. But this only prevails when the path length is exactly halved. If instead there is asymmetry in the two-hop path lengths then the link-budget gain in the longer hop will diminish and so the higher rate becomes unsupportable. This “sweet spot” in the path length split is illustrated in the graph of link budget in Figure 16. Taking the lower of the two lines defines a peak at 0.5 (i.e. the centre), which tails off away from the centre in either direction, hence the term sweet spot. Overall, the two hop chain is only as good as the weakest link, which must always defined as the lower of the two lines.

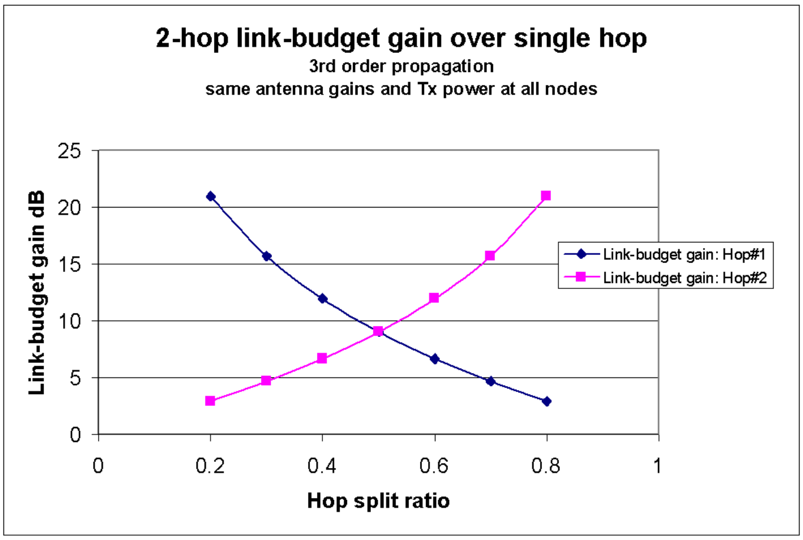


Figure 16: Two-hop link budget gain over single hop

Note that although there are now more mesh transmitters using the same power the interference situation is not necessarily degraded as each transmitter only occupies a fraction of the bandwidth. The amount of "spectral pollution" is reasonably unchanged, although its footprint is different.

But the comparative performance is further eroded for the case of multi-hopping into a mesh Access Point or cellular base station as represented in Figure 17.

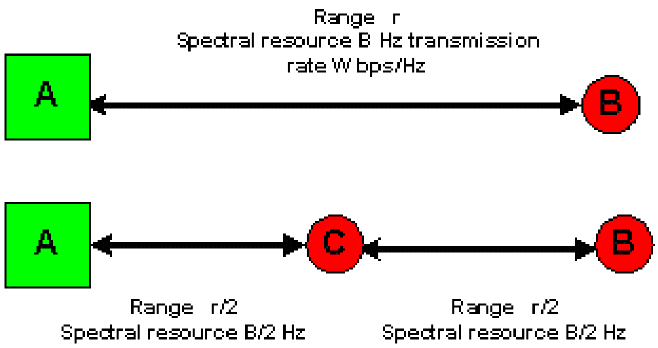


Figure 17: Two-hop vs. one-hop into high gain Access Point

The hop(s) between mobiles lack the higher antenna gain and height of the link into the Access Point (item **A** in Figure 17). Due to this imbalance the “sweet spot” does not occur at the 50:50 path-length split. The graph of Figure 18 illustrates this for the case when the Access Point antenna gain is just 13dB above the mobile nodes’ gain – the “sweet spot” has moved to approximately 75:25 path length ratio and the optimal link budgets on the two hops are only about 4dB above the single-hop case. With this small link-budget gain the transmission rate might be little more than doubled. Thus the best-case throughput rate of the two-hop route is the same as the single-hop.

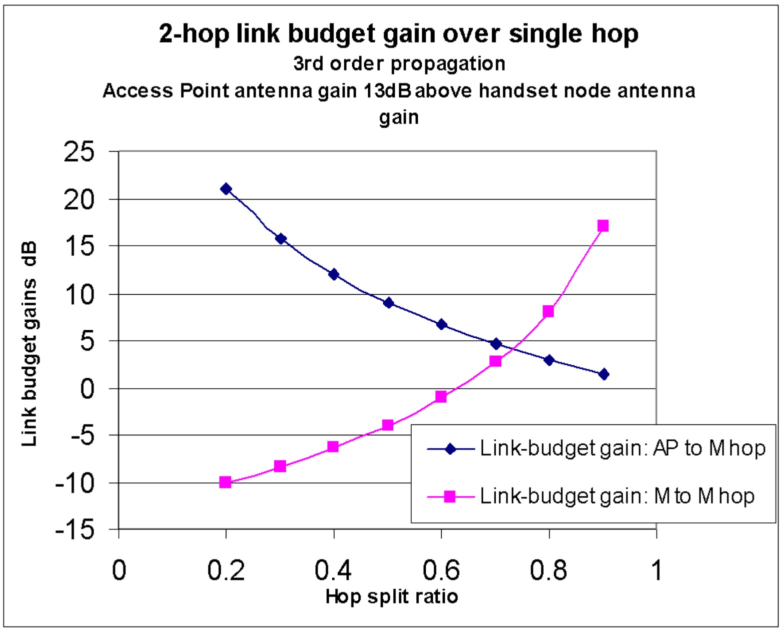


Figure 18: Two-hop vs. one-hop link budgets with high antenna gain

Continuing the analysis to a *three* hop scenario. Each hop would now be allocated 1/3rd of the spectral resource. The "sweet spot" occurs at about 60:20:20 path length split and at this point the link budgets are about 7dB above the single hop. This link budget gain might just support a tripling of transmission rate and so again achieve about the same throughput rate as the single hop case.

On the basis of the above, albeit highly simplified analysis we conclude that multi-hopping may rarely be much more spectrally efficient than single hop.

4.3.3.1.1 Extended route length

An implicit assumption in the above simplified analysis is that the multi-hop path lengths is the same as the single hop length. In practice, of course this will not be the case: nodes will be unevenly distributed and routes may circumvent building and terrain clutter. The detrimental effect of increased route length is illustrated in the graph of Figure 19 which illustrates the reduction in link budget gain at the "sweet spot" of Figure 18 as the route length is increased (results given for 3rd and 4th order propagation laws).

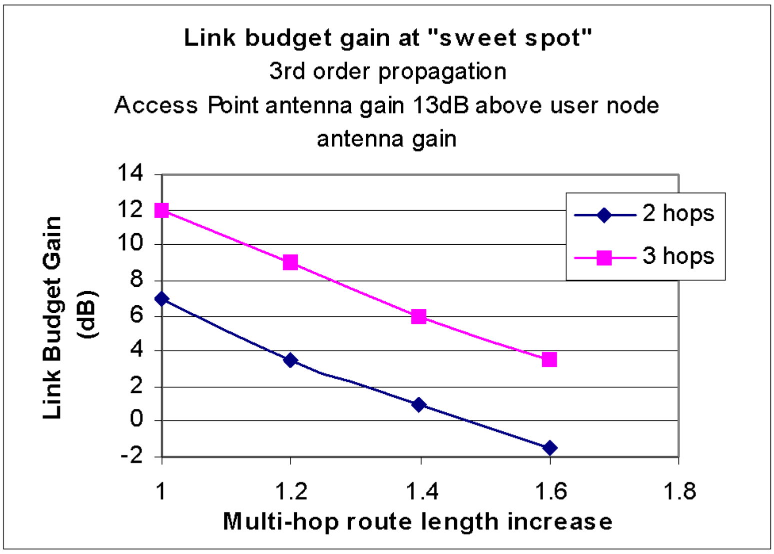


Figure 19 "Sweet Spot" link budget gain vs. extension in total route length

4.3.3.1.2 Propagation law

The gains in link budget through multi-hopping increase for higher propagation law. This is illustrated in Figure 20. for idealised 2- and 3-hop paths.

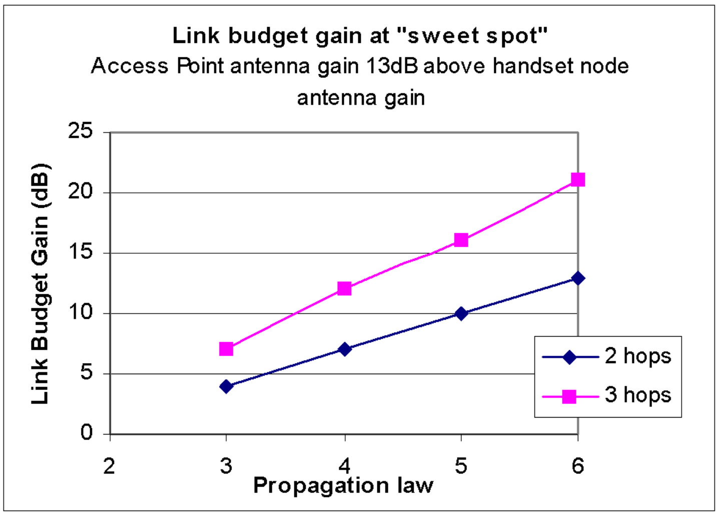


Figure 20 "Sweet Spot" link budget gain vs. propagation law

This is the basis on which multi-hopping can be used to circumvent high clutter losses on a single-hop path – i.e. the benefit of spatial diversity in a mesh network. It is a primary argument in favour of mesh networking and is the basis for coverage improvement in the "Cellular with Multi-hopping" scenario discussed in section 6.3.1.

4.3.3.1.3 Error Propagation

It may appear that the consequences of multiple hops for error rates have been neglected. To sustain a given error rate on a link comprising multiple hops it might be expected that the error rate on each hop must be substantially improved to compensate. For example, if a link requiring a BER of 10^{-3} is implemented using three hops then the required BER of each hop is 3×10^{-4}. The implications in terms of improved link budget depend on the particular system. However in general, and especially if spectral efficiency is important, there will already be a reasonable amount of error control coding deployed. One of the effects of coding is to steepen the BER curve in the region of required error performance. From consideration of typical curves it can be shown that the required link budget improvement to maintain end to end performance is in fact negligible, being some small fraction of a dB, when error control coding is in use, and thus of no real consequence in this discussion.

4.3.3.2 Mesh networking issues

The above discussion has addressed the fundamentals of multi-hopping in isolation to overall network issues. In a multi-user network a number of practicalities will further erode spectral efficiency, these include:

- traffic concentration
- mobility and routing overheads
- efficiency/availability trade-off

4.3.3.2.1 Traffic concentration

This was discussed in section 4.2.6. We suggest that this resource contention around the mesh Access Point will erode capacity more than in the cellular case.

4.3.3.2.2 Mobility and Routing Overheads

The introduction of hopping brings a requirement for dynamic routing. Even leaving aside the need for route discovery; this adds an overhead that will diminish spectral efficiency[13].

Mobility leads to a user dependent routing overhead, and overheads for retransmissions. This vulnerability to mobility is illustrated in section 6.1.3.

The consequential routing overheads may be as high as 100% in a simple packet-based approach (see section 6.1.3.3).

4.3.3.2.3 Efficiency-availability trade-off

In current cellular systems, 100% coverage availability is not guaranteed; instead outages of 2% to 5% are often used in planning. This is done to conserve base station power levels and hence spectrum efficiency. Further increases in availability and capacity are achieved by in-filling cell coverage with picocell sites but this has severe commercial limitations. The resulting compromises set upper limits to cellular coverage, capacity and spectral efficiency.

The multi-hopping inherent within a Mesh can increase coverage availability, e.g. section 6.1.4., but this service level enhancement is dependent on the statistics of user-terminal availability, location and mobility. The issue is that each user node is completely free to move - and may move to a position where coverage is lost. This must also affect any nodes that are downstream of this node. With mesh systems the Operator is not in control of all the network routing elements. The resultant compromise may well include a notably larger probability of mesh network outage, relative to current cellular systems. This uncontrollability of mobile user-node behaviour has lead several researchers to conclude that *"a pure multi-hopping solution has little chance of ensuring*

[13] For measuring efficiency in bps/Hz/km^2 only user bits should be counted, not bits used by the protocol.

stable and guaranteed service to users" [Lungaro and Wallin 2003]. This is further investigated in section 6.1.4.

4.3.3.3 The influence of link-adaptation

One key point about the models of Gupta, Arpacioglu and others is that, in order to simplify the analysis, they assign a specific relay throughput rate (bps) to each node regardless of the link budget on the radio path. Thus they specifically do *not* take account of the potential advantages of link adaptation, whereby throughput rate is set according to path-length/path-loss and signal-to-noise ratio. In this sense it would be more appropriate to assign a specific *bandwidth* to the relay process.

Link adaptation can take the form of one or more of the following techniques when link budget is good and/or when interference levels are low (i.e. the available SIR is high):

- Higher order modulation scheme (e.g. more bits per symbol)
- Reduced coding gain in spread-spectrum transmission

The concept of "good" link budget can included the following:

- Reduced propagation path loss in the prevailing environment.
- Short path length between terminals

Intelligent link adaptation, especially when based on feedback from higher layers[14], offers several potential advantages, but is beyond the scope of this report. It is suggested as an item for future work.

4.3.4 Conclusions on spectral efficiency

There seems little justification for a belief that practical mobile meshes will, in general, provide a significantly more spectrally efficient means of carrying commercially-significant applications of today and the foreseeable future. This is for several key reasons:

- Realistic meshes will have a performance which will depend on user distribution and mobility behaviour. In the worst case they could be no better than current cellular systems, and potentially worse. In the best case, where the environment allows, the efficiency could be very high on occasion, but the danger is that availability could dip to very low values as users move. In other words, this mobility of users within the mesh can lead to a much bigger trade off between efficiency and availability than current cellular systems, since mesh users are also system routers.
- Gains predicted by multiple hopping may be reduced in the real world due to the practical penalties in adaptive bit rate schemes and by the diluting effect of unequal length hops in a path. Mobile to mobile link budgets are also likely to be far less then base station to mobile budgets for practical reasons

By creating small mesh islands through the use of fixed seed and routing nodes, dependence on user behaviour can be lessened. Several current commercial mesh deployments do this, e.g. the approach of mounting mesh routers regularly on street lighting in urban deployments. Of course, this is moving away from the pure mesh towards one with infrastructure.

[14] i.e. with a cross-layer perspective and metrics

4.4 Hypothesis Testing ... that directional antennas confer significant benefits for mobile mesh networks below 5GHz

There are four potential incentives for the use of directional antennas in any radio network, each of which has relevance to mobile mesh networks:

1. Reduced mutual interference and hence greater spatial reuse of frequency. This is the primary mechanism which is in use for increasing the spectral efficiency of fixed Mesh networks [Aegis 2000].
2. Higher antenna gain, enabling reduced Tx power consumption for a given range and transmission rate. This can be significant, especially noting that the gain benefit can be realised at both ends of the link and so the net effect on link budget is equivalent to twice the antenna gain.
3. Higher antenna gain, permitting greater range for a given Tx power input. Using antenna gain to increase range rather than reduce interference in a mesh network can be counterproductive to overall spectral efficiency, as discussed in section 4.4.1, nevertheless it may be advantageous to reduce throughput delay under certain traffic load conditions (e.g. when the network is not operating near its maximum capacity).
4. Higher antenna gain, permitting higher transmission rate for a given Tx power input. Using antenna gain to enable higher transmission rate (via link adaptation) is also a potential benefit as a means of either increasing the throughput of nodes or reducing the on-air time for a given rate – so improving spectral utilisation.

But, when considered in the context of mobile mesh networks, these potential benefits are diluted by the limitations of:

a) Achievable gain and directionality of practical antennas for hand-held products – particularly below frequencies in the region of 5GHz.

b) Overheads for antenna control.

c) Selective "deafness" over some radio paths, resulting in confusion and ambiguities in the MAC layer for route exploration and CSMA/CA schemes.

As a consequence of the above this report concludes that whilst antenna directionality can increase mesh network spectral efficiency, the benefits may not be substantial in a practical deployment of hand-held mobile devices at the carrier frequencies of interest.

4.4.1 Antenna directionality to increase capacity

The following analysis considers directional antennas as a means of reducing mutual interference within the network, rather than extending range. Thus it does not exploit the directional *gain* of the antenna: specifically it assumes that the effective radiated power (ERP) in the main beam and the receiver sensitivity referred to the antenna ("radiated sensitivity") is the same as the omni-directional case. In this sense, the application offers the additional potential benefit of reducing RF power consumption.

A starting point in the analysis is to consider an idealised antenna having negligible side lobe responses. This can be represented by the "flat top" model – where the antenna beam in the azimuth (horizontal) plane is represented as an arc of a circle subtending an angle equal to the 3dB beam width of a polar response. The leads to a simplistic interfering / non-interfering alignment of beams as illustrated in Figure 21:

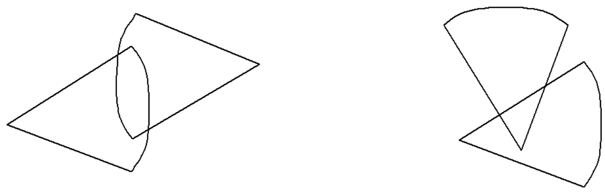


Figure 21: Interference Model for Directional Antennas

For a network of randomly deployed nodes equipped with such antennas, the theoretical upper limit on the improvement of throughput capacity is as large as $4\pi^2/\alpha\beta$. (where α and β are the beam widths of the transmit and receive antennas respectively). This is derived in a useful paper by Yi et al [2003], and can be deduced directly by noting that the probability of a Nodes falling within beams is reduced by a factor $\alpha/2\pi$ and $\beta/2\pi$ compared to the omni-directional case , and thus the reduction in mutual interference is the proportional to the product of these two.

Note: To illustrate the magnitude of improvement implied by this, consider that for mobile/hand-held products in the bands of interest here (below 6 GHz) the minimum achievable antenna beam width is likely to be in the region of 90deg ($\pi/2$). Then the above upper bound on capacity improvement from the idealised antenna is x16.

However, for any practical antenna, and more-so for mobile/hand-held products in the bands of interest here (0.5-3.5 GHz) there will be finite side lobe responses which will seriously erode the above potential gains. As a first step towards analysing side lobe effects one can model the antenna beam as one having as uniform side lobe response outside the main beam, as illustrated in Figure 22:

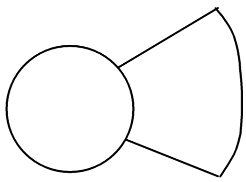


Figure 22: Radiation Pattern for Composite Antenna Model

The key manifestation of this finite side lobe response in the network is to extend the interference boundary around Nodes [Yi et al 2003]. The physical extent of this boundary is governed also by the attenuation factor of the propagation environment. If an antenna has a mean side lobe level which is κdB below the main beam then in a propagation environment with attenuation rate γ (i.e. path loss proportional to (range)$^\gamma$ the differential coverage range, Δ_R, between main beam and side lobe is given by:

$$\kappa = 10.\ \gamma.\log(1/\Delta_R), \qquad \text{i.e.} \quad 10.\ \log(\Delta_R) = -\ \kappa/\gamma$$

It is postulated, from practical work at Plextek and data from the antenna-supply industry, that for mobile/hand-held products operating below approximately 6GHz the side lobe response is unlikely to be more than about 10dB-15dB below the main beam. So, taking a likely figure for side lobe level of κ=13dB: then in a fourth-law propagation environment Δ_R is only 0.5. Thus, the interference boundary for the side lobes is only, nominally, half that within the main beam.

Some pivotal theoretical analysis of a mesh network with directional antennas has been carried out [Yi et al 2003] that builds directly on the theoretical work of Gupta and Kumar [2000]. Yi's analysis indicates that the theoretical maximum capacity gain factor for an idealised random network is of the order of:

$$1/[(\Delta_R)^2 + \{(1-(\Delta_R)^2\}\ \alpha\beta/4\pi^2]$$

Again considering the case of 90deg beam widths with -13dB side lobes this implies a capacity gain in the region of x3.3 (compared to a theoretical gain of x16 for the zero-side lobes case). This illustrates the detrimental effect of finite side lobe levels. Theoretical capacity-gain vs. antenna beam width and side lobe level is illustrated in Figure 23:

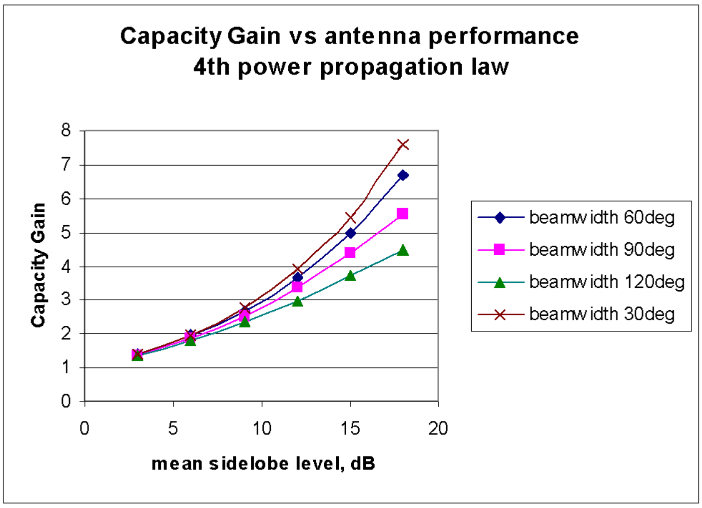


Figure 23: Theoretical Capacity Gain vs. Antenna Performance

The above illustrates that the capacity gain factor is a more sensitive function of side lobe level than it is of beam width. Furthermore, as beam width is reduced the side lobe level dominates performance, thus indicating that there is little benefit in decreasing beam width without equal attention to reducing side lobe levels.

But this capacity-gain performance is also a function of the propagation environment. The range difference, Δ_R, between main beam and side lobes decreases with increasing propagation attenuation law and so the benefit of side lobe attenuation diminishes. This follows from the premise that for a given density of nodes the ratio of the number of nodes residing inside the main beam coverage area to the number residing in the side lobe coverage area diminishes with increases propagation law. This is illustrated in Figure 24 for a beam width of 90deg; using the above formula derived in Yi et al [2003]. From this one can see that the benefit from antenna directionality decreases with increasing propagation attenuation factor. (Note: all of the curves are asymptotic to the theoretical maximum gain of x16 noted above.)

It must be noted that the capacity-gain curves in Figure 24 are normalised to the omni-directional antenna case for each propagation law, thus the curves do not imply that the network has the same capacity, for omni-directional antenna, regardless of propagation law: instead there are scale-factors to be applied to the vertical axis for each curve. In fact the corollary to this is that the high attenuation environments enable greater spatial reuse and hence higher spectral efficiency than a low attenuation environment [Gupta and Kumar 2000, Arpacioglu and Zygmunt 2004]; however, the low attenuation environment will reap more benefit from the use of directional antennas, because of reduced interference over the longer propagation ranges.

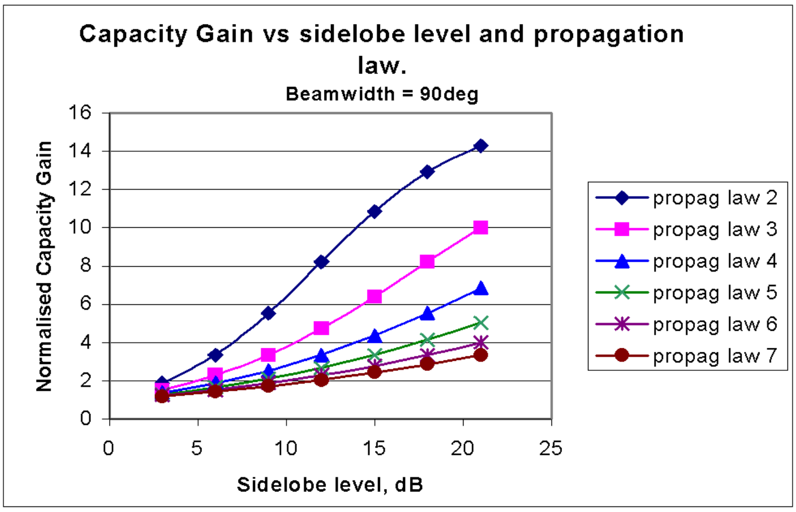


Figure 24: Capacity Gain vs. Propagation Environment

4.4.1.1 Null-steering

The above analysis has been presented in the context of steered-beam or switched-beam directional antennas.

An alternative technique is to steer a null in the direction of the dominant interfering signal. (In fact a preferred form of implementation is to steer the array so as to maximise the received signal-to-interference ratio; using such metrics as received symbol quality, training burst quality, per-packet error rate, etc.).

Null-steering is essentially limited to being a receive-mode function, for the following reasons:

1. In general one cannot assume that the direction of the source of dominant interference that is dictating null steering at a node is necessarily the same direction that will be vulnerable to interference from the node's transmission.
2. Null-steering requires very accurate phase and amplitude alignment between each element of the array [Plextek 2005]. This alignment is made on the basis of signals received at each element in the array and pass through associated receive paths in the radio. There are then problems replicating these phase and amplitude characteristics in the transmit path. Thus it is generally not possible to realise the same antenna pattern during transmit mode – in which case one is obliged to revert to a steered (or switched) beam during the transmit phase.

In consequence of the above, the antenna will, generally, be operated in steered-beam/switched-beam or omni-directional mode during transmit.

To date this report has not found any 3rd party work specifically addressing the benefits of null-steering in a mesh network. However, it may be postulated that the capacity-gain performance will be similar to the case of beam-steering because, again with practical small antennas below 3 - 5GHz, the attenuation within the null will not be particularly high - possibly in the region of 15dB-20dB, and the side lobe attenuation will still be in the region of 10-15dB.

Thus it is unlikely that null-steering will give any improvement in network capacity compared to the switched-beam/steered-beam case.

4.4.1.2 Practical aspects eroding performance

In practice there will be a number of factors that erode the potential capacity increases cited above:

Operational issues:

1. Log-Normal fading (due to clutter in the radio path) introduces a substantial variation on path loss, which may severely distort the figures for the mean level of interference assumed in the above analysis. However, the rate of change of such clutter losses may be relatively slow in relation to the MAC processes in which case, to a first approximation one might consider the dynamic system as quasi-static as far as the MAC processes are concerned. In which case log-normal fading may not degrade the operation of directional antennas too significantly.
2. Rayleigh fading (fast fading due to multipath dynamics will not affect *beam steering* unduly because the wide beam width envisaged for hand-held terminals will encompass a number of multipaths and so the mean signal level and its fast fading characteristics will not be affected significantly. However, Rayleigh fading *will* impact the performance of *null-steering,* since the null must track the bearing of the most significant multipath and therefore must be fast enough and accurate enough to maintain tracking. This is a severe undertaking for mobile terminals in a cluttered urban environment. For this reason null-steering is most applicable to the scenario of very localised terminals whereby the most significant interferer is closely located in line-of-sight conditions – since in this case the bearing of the most significant multipath is relatively constant.
3. MAC Deafness. A problem that is likely to arise when antenna beam width is reduced is "deafness" at the MAC layer. In a CSMA/CA scheme nearby nodes that lie outside the main beams of a prevailing communication link measure a reduced level of prevailing signal (via side lobe radiation) and thus may acquire the channel for reuse – in so doing they run a risk of interfering with the prevailing signal. Thus the system becomes more susceptible to errors in the channel acquisition and spatial re-use process. This is directly associated with the "brittleness" factor cited in section 6.1.3.2, as the extent of spatial reuse is increased – by whatever means.
4. In any practical route-exploration process employing CSMA/CA and/or RTS/CTS signalling there will be a need for Nodes to either transmit and receive omni-directionally or to scan the directional beam at a suitable speed to capture wanted and unwanted signal profiles. This will increase interference levels during such activities and so undermine some of the capacity gain.
5. There will be a finite overhead for antenna control:
   - For beam-steering the required agility and accuracy of beam steering must increase with decreasing beam width. For wide beam width antennas envisaged here, this overhead should be relatively low in most environments because the beam width will be sufficient to encompass the significant multi-path components, and so the steering dynamics can be relatively slow.
   - However, for null-steering the overhead is likely to be high (and per-packet) because the null must align with the most significant multipath component and the orientation of this can change rapidly in a cluttered environment under Rayleigh fast-fading conditions.

   Note: The multipath scattering that exists in urban environments means that the directions of arrival of wanted and unwanted signals are not directly related to the orientation of the sources and so absolute position-based systems (e.g. GPS-based) will be of limited benefit in such environments.

4.4.1.3 Path reciprocity

If one is using FDD transmission then there will only be a loose correlation between the multipath components at the Rx and Tx frequencies, i.e. there may be poor path reciprocity Thus any antenna steering based on the angle of arrival of the wanted received signal will be only loosely correlated with the preferred angle of transmission. This factor is one of several which lead to a preference for a TDD radio interface in ad hoc networks.

4.4.1.4 Hand-held vs. vehicle-mounted antennas

The comments made above regarding the limitations on performance of small hand-held antennas, need not necessarily apply to vehicle mounted antennas – where there is more physical space and improved ground plane, (and more available dc power). Having said that, one is interested in achieving directionality in a 360deg arc in the horizontal plane so, in practice, useful implementations are likely to be confined to roof-mounted antenna arrays.

4.4.1.5 Requirements for Z-plane directionality

As a further caveat to the above discussions, one must recognise that in practical deployments mesh users will be distributed in the vertical plane; not just the horizontal. As a consequence one would ideally want antenna directionality to cover the vertical axis. This is a significant further complication to implementation and one that is unlikely to be forthcoming in practice; thus one must discount a certain percentage of the theoretically gain predicted above to allow for this inability to realise a 3-dimensional antenna steering system.

4.4.2 Directional antenna application to hybrid networks

All the theoretical performance analysis cited above has been carried out in the context of *intra-mesh* traffic, having links formed between node-pairs and with uniform random distribution of Nodes and traffic flows. In particular, this scenario leads to a random distribution of the orientation of radio links and hence a random distribution of interference paths – both of which are conducive to improvements through the use of directional antennas, as discussed above.

However, the situation for a mesh system handling extra-mesh services via Access Points, or intra-mesh traffic via Relay Nodes to a backbone mesh network, is different. When handling such traffic links are aligned roughly radially about the fixed nodes (Access Point or Relay Node) and therefore neighbouring mobile nodes have their antennas aligned radially inwards to the fixed node as illustrated in Figure 25. This means that the arrangement of antennas is almost pathological in that it will tend to maximise the chances of interference between neighbouring nodes and so the benefit from directional antennas will be diminished.

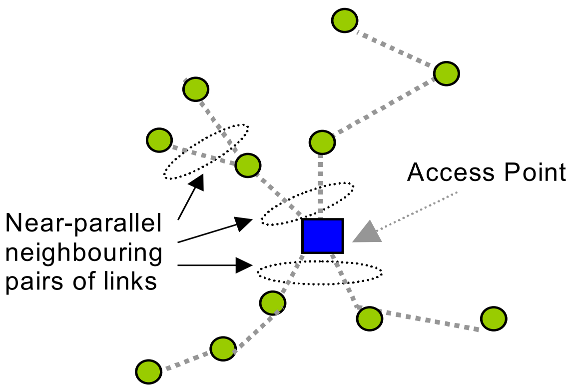


Figure 25: Interfering Neighbours in a Hybrid Network

4.4.3 Conclusions

The conclusions may be summarised as follows. The overall conclusion is that directional antennas do not confer significant benefits on mobile meshes at the frequencies of interest to this report.

- Network capacity can be improved through the use of directional antennas. However for handheld devices the extent of directionality (beam width) is limited due to the product's small physical size in relation to wavelength. But, more significantly, the high side lobes levels associated with such small antennas (coupled with degradations due to body effects etc.) severely limit the improvements in spatial reuse that would otherwise be possible.
- Primarily because of the side lobe problem, the potential gain in spatial reuse offered by directional antennas diminishes as the propagation law increases (e.g. the benefits are less significant in an urban environment), however in such an environment spatial reuse is already quite high so spectral efficiency high. The corollary to this is that the benefits of directionality are more apparent in a low propagation law environment and it is in such an environment that spectral efficiency is low (because of low spatial reuse.
- In the 4^{th}-5^{th} power law environment likely for urban applications: theoretical gains of circa x3 are likely to reduce to x2 because of antenna / RF practicalities, and are likely to be further reduced to possibly little more that x1.5 when one has factored in MAC & Control issues.
- If directionality also provides directional-gain compared to the omni-antenna (i.e. if directionality can be achieved without significant loss of antenna efficiency) then two other attributes can be capitalised on: either,
  - Power saving in the transmit amplifier, or
  - Higher signal level on the link, so enabling higher transmission rate within the same bandwidth, so reducing the time/bandwidth of spectral resource utilised on the link.

  Note that it is possibly disadvantageous to use antenna gain to extend range (hop-length) for a given transmission rate, because this will increase the footprint of interference without reducing the time/bandwidth of spectral resource utilised on the link.
- When concentrated around an access point, the arrangement of directional antennas is almost pathological in that it will tend to maximise the chances of interference between neighbouring nodes and so the benefit from directional antennas will be diminished

It is noted that the above is expected to be equally valid for cellular and other point to multipoint point systems.

4.5 Hypothesis Testing ... that meshes could improve spectrum utilisation

This hypothesis relates to the wider issue of spectrum utilisation, rather than simple specific spectrum efficiency. It suggests that the spectrum may be better utilised by having short line-of-sight (LoS) mesh links use 'less precious' spectrum e.g. up to 6 GHz, hence leaving the present cellular frequencies for best use where they are most needed, for longer-range PMP propagation environments.

From the work of this report four key factors point to mobile mesh networks offering opportunities for use of higher frequency bands:

i. They are not spectrally efficient in comparison to current cellular systems operating in the 2GHz region (see section 4.4). Thus they might use less commercially-precious spectrum.

ii. To achieve useful per-user throughput the relaying capacity of mesh nodes needs to be high (e.g. the need for high value of W in Section 4.2). Thus meshes need access to large allocations of bandwidth.

iii. Throughput latency remains a significant issue with any multi-hopping network. However, to some extent this can be reduced by using a wide bandwidth air interface whereby shorter frame structures in a TDMA architecture can be used. This again points to a wide bandwidth channel allocation and hence high carrier-frequency band operation.

iv. The potential of increased end-to-end throughput by using multi-hop vs. single hop is only realised when there is high propagation path loss – in particular when there is high clutter loss due to in-building penetration and the "urban canyon" effect (see section 4.3). Again these are prevalent at high carrier frequency. In addition to the above, we show later (section 6.3.1) that multi-hopping as an adjunct to PMP coverage offers the potentials of increasing service availability (i.e. reduced coverage-outages) and extending the range of high data-rate services – as an alternative to increasing the density/capability of the fixed infrastructure.

Whilst this report concludes that mobile mesh networks may not be spectrally efficient in terms of the traditional bps/Hz/km^2 metric they offer certain benefits to spectrum planning on a regional basis and to co-existence with like networks:

a) With regard to spectrum sharing on a regional basis; the emissions footprint from a mobile mesh network has sharp fall-off due to the low antenna height and Tx power of terminals (in comparison to the elevated cellular basestation. Hence spectrum use is more contained spatially. This facilitates re-use in adjoining areas (e.g. access mesh can have a frequency re-use of unity).

b) The necessary ad-hoc networking aspects of mobile mesh, such as dynamic channel allocation, collision avoidance and dynamic routing, etc. help mesh networks to co-exist with like and unlike networks – albeit at the expense of spectral efficiency per se. This points to the use of shared (i.e. licence-exempt spectrum) where service-level guarantees can be sufficiently relaxed.

In summary, meshes could improve the utilisation of spectrum.

4.6 Summary of Capacity Constraints

The hypotheses testing resulted in the following broad conclusions, subject to the caveats and assumptions in the text of each discussion.

capacity self generation	FALSE	for a pure mesh. If a hybrid mesh with access points is considered, then with sufficient planned access points, the capacity may be made to scale.
spectral efficiency	FALSE	A practical mesh does not intrinsically have a higher spectral efficiency than cellular. Meshes can however be very good at providing efficient coverage in certain situations.
significant directional antenna benefits	FALSE	Directional antennas will generally out-perform omni-directionals, but antenna performance will be limited by physical size constraints at <5GHz operation, and the complexity of the MAC to control such antennas may be high. Overall, the realisable benefit is not likely to be significant
spectrum utilisation	TRUE	It is possible that the use of meshes will allow less precious spectrum to be utilised (e.g. up to 6GHz), as a consequence of combating high clutter loss and high path loss with diverse-routing and multiple short hops

Additionally, it will also be shown later, that

- The routing overhead for mobile ad hoc networks can be very high. As many routing control packets as user traffic packets might be needed. This would generate a halving of overall system efficiency.
- The degree of mobility able to be supported by a mesh is a subject for continued further work.
- Meshes yield efficient coverage in certain scenarios: Hopping around corners is perhaps the greatest benefit offered by a mesh.

These issues are addressed in section 6.

5 Susceptibility and Co-existence Issues (Ofcom Work Item 5)

5.1 Introduction to susceptibility

There is a large variety in the types of interference which may affect a radio network and these may be categorised in many ways. Figure 26 shows a simple breakdown of interference categories that will be used in discussing the effect of interference on a mesh system.

The distinction is first made between self-interference and external interference. In the former case the modulation, time and frequency characteristics of the interfering signal are the same as those of the wanted signal and the interference is commonly described as co- and adjacent-channel interference. External interference refers to all other forms of interference and will usually have modulation, time and frequency characteristics that are significantly different to those of the wanted signal. The interference commonly arises from unwanted or spurious transmission effects.

Self-interference may result from the components of a receiver's own mesh or from the components of another mesh that uses the same architecture and protocols (e.g. two competing service providers). In the first of these, the elements can in principle exchange information which allows the source of the interference to be influenced, thus reducing it at source. All other types of interference are uncontrollable and must be dealt with through the appropriate design of the radio network, its architecture and its protocol.

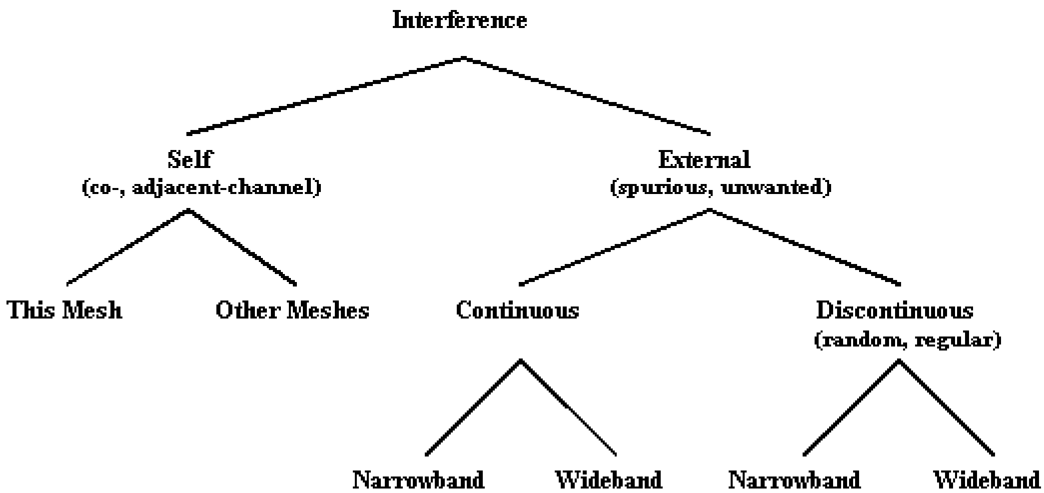


Figure 26: Interference Types

External interference can be characterised in a number of ways, but its time and frequency characteristics are of particular significance. A key aspect of the time characteristics is the duration of the interference; continuous or discontinuous. The latter includes both random interference (e.g. from a microwave oven) and regular interference (e.g. from a DECT cordless phone). The key aspect of the frequency characteristics is the occupied bandwidth and in particular whether the bandwidth interference is small (narrowband) or large (wideband) compared to that of the wanted signal.

Note that no distinction is made between real interference and virtual interference. The former is RF energy that appears in the frequency band occupied by the wanted signal. The latter is RF energy that appears in other frequency bands but which nonetheless affects receiver performance in the wanted band because of the practical limits on the performance of mixers and filters.

This section covers susceptibility from two viewpoints:

- susceptibility due to PHY and MAC issues
- susceptibility due to transport and routing issues

Note that this study was restricted to susceptibility only (i.e. not emissions).

5.2 Susceptibility to interference - PHY and MAC

The susceptibility of a radio network to interference depends on a number of factors and there are many elements of the radio network's design that affect its resistance to interference (*hardness*). The scope of this section covers only the low-level elements that are commonly part of the Physical (PHY) and Medium Access Control (MAC) layers of a design, but higher layers also play an important part (see section 5.3).

Figure 27 shows the key elements in the design of a radio network which affect its susceptibility to interference. The elements can be either dynamic or static in the sense that they may or may not change their characteristics in response to interference problems. The dynamic elements tend to be those in the PHY layer although the impact of the changes may be felt in the higher layers (e.g. changes in throughput).

Even in carefully planned cell-oriented hierarchical networks, the control of dynamic elements in the design is never a simple issue. In mesh systems stable control is even more difficult to achieve and there is a risk of avalanche-like failure of significant areas of the mesh. For example a system which relies on increasing the transmit power to overcome interference in a link between two nodes will, in doing so, increase the self-interference affecting adjacent nodes, which will in turn then need to increase their power etc. As another example consider throughput reduction as a technique to overcome interference – as the throughput of a link is reduced alternative routes through the mesh must be found and so the effect spreads.

Hierarchical or centralised control algorithms can help in preventing these effects in a mesh system, but their presence introduces overhead in the form of control data that must traverse the mesh. The propagation times through the mesh will also negatively affect the performance of such algorithms with the result that localised control centres might be needed.

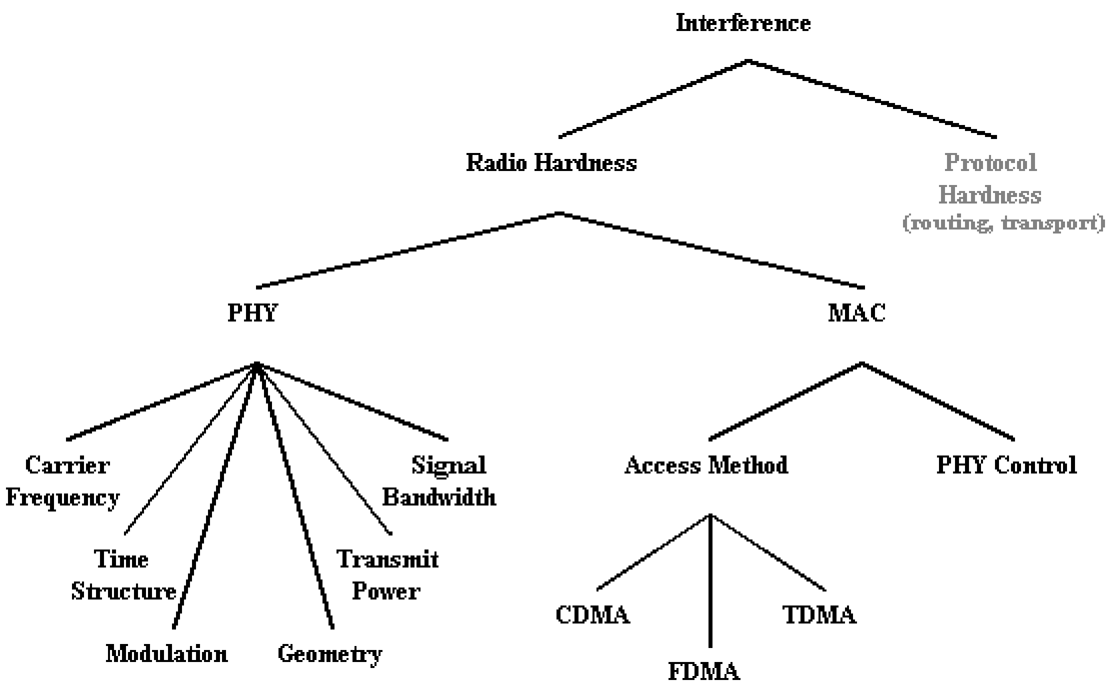


Figure 27: Elements Affecting Interference Susceptibility

With respect to interference susceptibility, the choices made in the design of a radio network must acknowledge the fact that the only way of countering the effects of interference is to ensure that a receiver receives the wanted signal with enough energy to overcome the interference signal. There are however a large number of methods that can be used to achieve this which fall into one or more of the following categories:

- Avoiding interference using time diversity, frequency diversity and spatial diversity, both individually and in combination. For example frequency hopping is a form of frequency diversity that can be used to avoid narrowband interference and to convert continuous interference into discontinuous interference so that time diversity can also be used.
- Rejecting interference. Receivers can use antenna beam-steering to reduce the energy from a few a single-source interferers and phase continuous signals can allow multi-user detection techniques to be used to reject co-channel interference. (But neither scheme is practical for mobiles, see section 4.4).
- Increasing transmission energy. The higher energy may be achieved by transmitting with a higher power spectral density, for a longer period of time or using a wider bandwidth.

In mesh systems, particularly mobile or ad hoc systems, the last of these is least preferable. This is because increasing the transmission energy of the signal means either increasing the self-interference effects on other users of the system (transmit power increased) or reducing the throughput of a link (transmit time increased).

The effect of the MAC layer design on the interference susceptibility is twofold:

- The algorithms in the MAC control have a particular significance in the performance of the

PHY layer elements. However the PHY elements themselves are not mesh specific, hence are covered in Appendix E.

- The implementation of any shared access method for a radio network has an overall negative rather than positive impact on its susceptibility to interference. Multiple access, by its nature, shares the available resources between nodes in a system and as a consequence increases the self-interference affecting a network. In a mesh system in particular a poor choice can result in increased susceptibility and the potential for catastrophic failure of the network. The common multiple access methods are described in section 5.2.1.

5.2.1 Multiple access control

In the following sections the three basic multiple access techniques, Code Division Multiple Access (CDMA), Frequency Division Multiple Access (FDMA) and Time Division Multiple Access (TDMA) including Carrier Sense Multiple Access (CSMA), are described. Although they are described separately, most practical radio network designs use a combination of these. The medium is controlled by the medium access controller (MAC). The MAC protocol itself may be centralised or distributed, plus it may be 'polite' or non-polite in its method of accessing the medium. This aspect of the MAC protocol has a large impact on susceptibility, see 5.2.2.1.

5.2.1.1 Code Division Multiple Access

CDMA is an access method which is used with Direct-Sequence Spread-Spectrum (DSSS). Ideally the use of orthogonal spreading sequences allows signals that are super-imposed in time and frequency to coexist without interfering with each other. In practice such ideal sequences do not exist.

There are sets of sequences that are perfectly orthogonal as long as they are synchronised in time and these can be used effectively to carry multiple data streams transmitted by the same element. However, in a multipath environment delayed versions of the transmitted signal are also received at the destination (unless highly directional antennas are used). Certain receiver designs (e.g. Rake) can take advantage of the delayed versions to increase the energy of the wanted signal, but the unwanted signals that are no longer time-aligned are no longer orthogonal and so act as interference.

If there are multiple active transmitters, the signals at a receiver cannot be synchronised. The signals therefore act as interference to each other, with signals from closer sources having a stronger effect. In the extreme case a signal from a nearby source can result in that from a distant source being completely hidden. To avoid this cellular systems use closed-loop power control to ensure that the power of each signal at a receiver is the same.

An example of the use of CDMA is the 3GPP system. The W-CDMA variant uses orthogonal Hadamard codes so that the base-station's transmissions to the mobiles in a cell are as orthogonal as possible given the propagation conditions. The mobiles use the same codes in their transmissions[15] and are subject to strict closed-loop power control to try and ensure that their signals reach the base-stations with the same power. Complex algorithms in the base-station and in the network balance the power requirements of each mobile with the required throughput. They also balance the power used in neighbouring cells, as this is a major source of interference at a

[15] but since uplink synchronicity cannot be assured, further coding is also used with poor cross correlation properties, in order to enable separation

base-station[16].

For mesh systems, particularly mobile ones, this self-interference is a serious problem. The centralised power control that exists in cellular systems would require significant flow of control data through a mesh and it is doubtful if it could do so fast enough to be of use. High processing gains[17] and advanced Multi-User Detection (MUD) algorithms in receivers might allow a system to be designed without power control, but they would have a significant impact on the throughput and the complexity of elements.

5.2.1.2 Frequency Division Multiple Access

FDMA allocates different carrier frequencies to different RF links. The allocation can be static or dynamic, with the latter combining well with frequency hopping techniques for avoiding interference. Static allocations are generally avoided in mobile networks because a particular frequency may be subject to significant interference or poor propagation conditions. They are however frequently used in carefully planned point-to-point networks.

Dynamic allocation introduces the need for allocation algorithms that generally adopt a TDMA approach. In a hierarchical system a semi-static allocation usually takes place with cells being allocated their own set of frequencies to use. In a mesh system such planning cannot take place and the best recourse is to adopt a frequency hopping approach such as that described in section E.1.

5.2.1.3 Time Division Multiple Access

TDMA allocates different timeslots to different RF links. There are two exclusive techniques:

- Nodes in the network are synchronised and the transmitter and receiver know when a transmission is due to take place. This requires a common timebase which, in a mesh system, may be difficult to achieve without some for of common external reference. Alternatively it requires large guard periods to allow for imperfect synchronisation.
- Nodes in the network are not synchronised and the transmitter transmits at a random time. Collisions may occur, and some form of re-transmission or error correction scheme is required. Refinements include Carrier Sense Multiple Access (CSMA) in which the transmitter first checks that the channel is free. An example is IEEE802.11.

There are a wide range of methods which combine these techniques with varying degrees of time synchronisation and control required. The unsynchronised methods have simpler control algorithms but have the disadvantages that the maximum throughput of a link is reduced and that the receiver must be active for longer periods of time.

5.2.2 Medium access control (MAC)

The MAC protocol layer is part of the data link function and controls access to the physical transmission medium.

5.2.2.1 MAC Politeness issues

Where a MAC operates ‘listen before talk’, it can be described as a ‘polite’ MAC. Such a scheme

---

[16] Typically, for a network of cells with 3 sectors, the interference power from adjacent cells is 0.65 of the power of the wanted signal from the mobiles in a base-stations own cell.

[17] The processing gain of a DSSS signal is the ratio of the signal bandwidth to the data rate.

often fails when operated in the same area as an impolite MAC. Examples of polite and impolite MACs are those of 802.11 and 802.16 respectively, as shown in Figure 28.

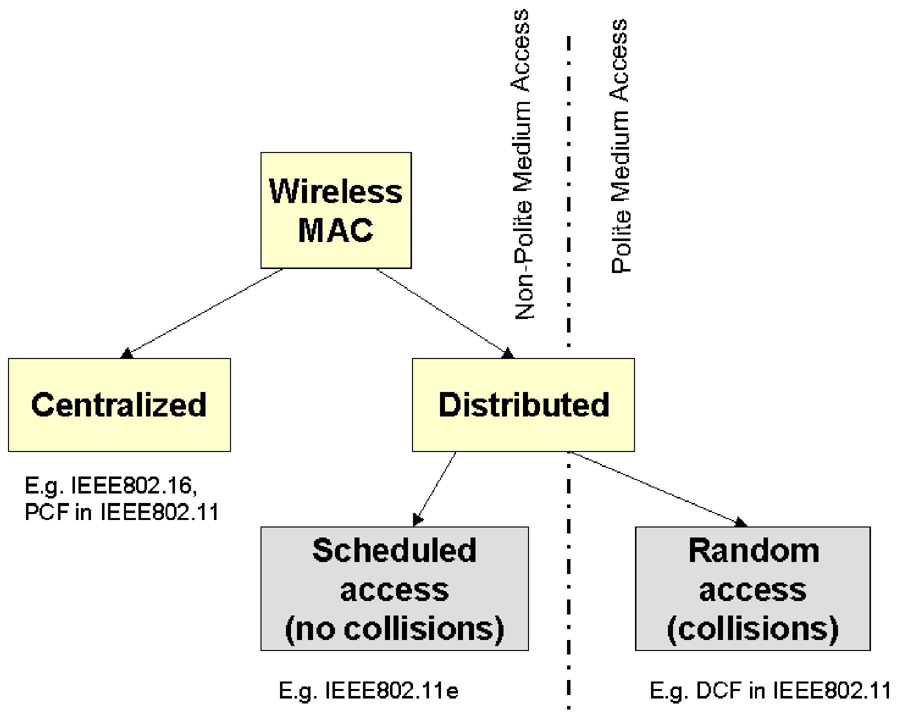


Figure 28 'Politeness' of MAC protocol for medium access

Because of the MAC behaviour, a mesh using a polite protocol will always cede control to a mesh which simply schedules transmissions regardless of any other band users. This is because the back-off algorithm inherent in a CSMA approach will continue to reduce packet size and increase wait time, thus reducing its access to the medium in response to interference, whist the scheduling MAC will continue to fill its time slots with automatic repeats of data which may have been lost in any collisions due to interference. This viscous circle is understood for the case of 802.11 losing out to 802.16 and some steps have been suggested to overcome it, namely the inclusion of quiet periods in the 802.16 scheduling MAC to allow the 802.11 CSMA MAC to transmit, [Li et al 2002].

5.2.2.2 MAC inefficiencies

There are several aspects regarding efficiency which are well known. They include the hidden- and exposed-node problems and the inefficiency of any CSMA protocol in general, due to back-off behaviour. The inefficiency and inequality of the specific 802.11 MAC is also known, where single hops are always unfairly favoured over multiple hop routes. These effects, whilst important, are not within the scope of this report.

5.2.3 Conclusions

This section has reviewed PHY and MAC approaches. Some of the findings of this section are noted to be similar for mesh and non-mesh schemes, such as FHSS being particularly suitable against narrow band interferers. Other general properties were summarised in Table 9. Those conclusions which are mobile mesh specific are listed below:

- The common cellular technique of managing throughput in response to propagation conditions is likely to be much more complex in a mesh where high hop counts occur.
- Current transmit power control schemes use centralised control, whereas in a mesh all control is traditionally expected to be distributed. Distributed power control on a wide scale would be a topic for research.
- Self interference on a mesh employing CDMA would be a serious problem, since it is doubtful whether fast power control for near/far effects could be implemented with a short enough response time over a wide area.
- Ad hoc systems preclude frequency planning by definition, therefore FDMA schemes are disadvantaged in a mesh. (This does not preclude frequency hopping techniques).
- Synchronisation to a common timebase for TDMA[18] will be problematic to implement efficiently in a wide area mesh. Although the use of unsynchronised approaches like CSMA, e.g. in IEEE802.11, avoid this issue, this is at the expense of MAC efficiency and hence overall spectral efficiency.

The effect of MAC politeness, although not mesh specific, is so important that it deserves repeating: Any co-located network using impolite protocols will always take all the available resource away from a network using polite protocols. In other words, shouting loudest will work. The example of 802.16 and 802.11 was given, but Bluetooth and 802.11 could also be an example, where spectrum is shared.

An overall conclusion for meshes is that extra overhead is more likely to be required for meshes to implement the common interference control schemes, since most have a centralised element (e.g. for power control, synchronisation). Whilst this is not to say that meshes will be more susceptible to interference, efficient mesh susceptibility performance may have to rely on layers above the PHY/MAC (see section 5.3) or on radically different approaches such as cognitive radio.

5.3 Susceptibility to interference - transport and routing

Network resistance to node failures is the attraction here. In a manner very similar to the Internet, the mesh may be said to be self healing, given the correct software. If a link drops due to failure or, more likely, due to user mobility patterns, then another link will automatically be brought up, if available. Of course, the success of this plan relies on the balance of system efficiency and the speed and range of user mobility. Unfortunately, re-routing can give rise to problems elsewhere, notably in packet re-ordering (which upsets typical TCP stacks) and varying latency (which leads to issues at the application layer).

5.3.1 Transport

Firstly, traffic types, transport and routing are placed in context.

5.3.1.1 Traffic Perspective and Implications

The ‘traffic model’ used will strongly affect findings, since the range of traffic types is great: Small file transfers can be very tolerant of delay, delay variation, end to end capacity and breaks in transmission (outage), whereas full screen, real time video places very stringent demands on the

---

[18] for variants which require it

same three network parameters.

Currently, many simulations simply use CBR (constant bit rate) traffic. Although simulation results for e.g. routing protocols will include items such as latency and packet loss, and thereby give some idea of how multimedia traffic might fare under the protocol, this is not enough. A further technique is needed to stress the simulation based on the bursty nature of some traffic or the regular periods of very high speed traffic which might be expected in multimedia communications. Such an approach will stress the network as well as look for subsequent multimedia-friendly behaviour of the protocols.

Models of actual traffic loads appear to be largely absent from the present literature.

Further work is required here to move the modelling of performance from average throughput rates (as in the above theoretical analysis and model simulations) towards modelling of performance for actual traffic loads at actual Erlang loadings.

5.3.1.2 Elastic Applications and TCP

Applications may be classified as elastic or inelastic, depending on their tolerance to delay in transmission. Such applications need to be handled differently by the transport protocol.

The majority of traffic in today's Internet is made up of elastic applications, that is applications which are tolerant to variations in the throughput but that typically require 100% packet reliability. Popular examples include email, web access and peer-to-peer file sharing. Because they all use TCP as a transport protocol this makes TCP a crucial protocol/algorithm in the scheme of overall data transfer efficiency.

Historically, TCP was designed to solve the congestion collapse problem in the Internet and is therefore extremely conservative when it sees data loss (it typically reduces to half the sending rate when it detects one lost packet). Since the Internet was, and still is, mostly a wired environment this assumption is almost always correct. In a wireless link, however, this is not the case. Loosing a packet is much more frequent and does not usually imply congestion. TCP will reduce the sending rate unnecessarily, causing lower network efficiency and a degraded user experience. For mobile devices this may not be a problem since they are becoming available with modified versions of TCP that are much less conservative and tolerant to packet loss, e.g. TCP Westwood. Nevertheless there are numerous scenarios (e.g. home Internet access) where the end system is not, and cannot be configured for this - which inevitably causes degradation of service[19].

A similar problem may arise from the increase in delay caused by the larger number of hops in a mesh. Because TCP data rate increases with each reception of an acknowledgement, an increase in delay makes the TCP data rate increase more slowly. This might be a problem for web browsing since each page is typically 5 kbytes and a smaller increase in the total delay of the page transfer may lead to a reduction in user satisfaction. It can also be a problem if high transfer rates are desirable, since TCP is well known to perform badly with high delay-bandwidth-product networks. If adding to the extra delay, the application requires high bandwidth transfer, TCP is proven to be inefficient. Although there are current research proposals to solve this problem, none is satisfactory from the fairness point of view. This is not foreseen to be a problem for any current application today but the near future may bring new applications with higher bandwidth requirements.

Mesh networks where several routes can be used to maximise throughput can cause packet re-ordering since consecutive packets may follow differing paths. Although this does not violate the

---

[19] Whilst some MAC implementations include automatic repeat request (ARQ) which aims to reduce packet loss, the delay caused by this may still cause TCP or other transmission protocols to time-out and thus deduce packet non-delivery.

TCP/IP model it does generate inefficiencies since the majority of practical TCP implementations do not cope well with packet re-ordering. The potential success of mesh networks would have to lead to a redesign of TCP stacks to overcome this issue if it were allowed to develop. – But such a change would be difficult, if not impossible, to propagate into all the very many TCP implementations already deployed.

5.3.1.3 Inelastic Applications and UDP

Inelastic applications, typically interactive real-time ones like audio and video have been gaining popularity in the Internet and causing big concerns on its stability. The recent popularity of Skype (IP telephony), for example, is forcing uncontrolled traffic into the network and some fear that the possibility of congestion collapse in the Internet has thus increased significantly. This is because these applications use UDP as a transport protocol. UDP performs multiplexing (through the UDP port) but does not implement reliability and, more crucially, does not perform any congestion control. Users send at a given rate (in the case of Skype 10kbits/s) and the network is simply expected to cope with this rate.

Although this remains a problem for a TCP/IP network in general, UDP is less affected by the behaviour of wireless mesh networks. Nevertheless, an increase in delay can easily destroy the quality of a phone call since anything above 150ms starts to become problematic. Given the ad-hoc nature of some networks, capacity provision becomes more difficult but this is a problem which should be addressed by the routing protocols in the network layer.

A larger loss rate may also trigger a rethinking in ‘transport’ protocols above UDP such as RTP (Real-time Protocol) which may have to accommodate new error checking functionalities. SCTP (Stream Control Transport Protocol) may also play an important role since it is includes a number of functions for telephony signalling.

It should be noted that a current hot research issue is the interface between applications and the lower layers. New proposed protocols like DCCP (Datagram Congestion Control Protocol) are trying to address this by creating different congestion control profiles that can be selected by the applications. This is more and more important as the heterogeneity of applications in today’s networks is increasing. The impact of mesh networks in current DCCP congestion control profiles is also a open research question.

5.3.2 Routing

If PHY and MAC properties largely determine the scalability of meshes, then the behaviour of the mobile routing protocol largely determines the *degree of mobility* which may be handled, see 6.1.3. Efficient dynamic routing protocols are needed to cope with the frequently changing multi-hop network topology.

The degree of mobility which a routing protocol may handle and how this may be quantified is the subject of intense research at present. Clearly further work is needed.

At the protocol level, route discovery and maintenance is a challenging task. The node connectivity map is very dynamic and must be refreshed regularly. This requires a large number of control packets to be sent. The issue raised by this is one of overhead and it is not uncommon for control packets to outnumber data packets under some conditions. Overhead, along with latency and packet loss rate are normally used as the performance metrics when comparing routing protocols.

5.3.2.1 Proactive and reactive routing in ad hoc networks

There are generally two categories of protocols in ad hoc routing: proactive and reactive, although hybrids do exist in reality.

Proactive routing protocols are based on the 'normal' routing protocols used in wired networks. Algorithms based on distance vector or link state are common. In both cases, there is an attempt to build locally, at each node, a global picture of all routes to all destinations within the network before they are required for use. The routing tables are usually built periodically through the normal operation of the protocol exchanging routing update packets. This has the advantage that the routes are already pre-computed and so packet forwarding can take place as soon as a packet for a particular destination appears at a node. The drawback is that routes may be calculated and re-calculated when they are not actually required, wasting bandwidth and for mobile nodes also wasting battery power, by the sending and receiving of unnecessary routing updates.

Reactive routing builds routes only upon demand and caches route information according to some time-out or 'staleness' policy. Cached routes can be used as required, but if a route is not known then it has to be 'discovered'. This has the advantage that routes are only evaluated when needed, although this approach adds latency to packet forwarding when routes are not known.

Generally, the reactive routing approach is the one that has received most attention in the ad hoc networking community, for the reasons already given.

5.3.2.2 Anatomy of a reactive ad-hoc routing protocol

An ad-hoc routing protocol consists of three main functional components:

- route discovery: this is the part of the protocol that is used to discover routes to destinations within the network
- data forwarding: this is how the data packets are forwarded by the protocol.
- route maintenance: this is how the protocol deals with changes and faults within the network once routes have been established

The design, implementation and interaction of these functional components within the ad-hoc routing protocol greatly affects its performance. The performance considerations encompass several dimensions:

- data forwarding performance: this might typically be measured by evaluating the 'load vs. throughput' performance of the protocol under given network sizes (number of nodes), and various levels of background traffic and source traffic types (CBR, Poisson, on/off, etc)
- protocol overhead: this would be an assessment of what proportion of the available channel capacity is taken up by the operation of the routing protocol. This could need to be considered for the routing protocol in 'isolation' as well as when considered with any radio channel control or media access control considered for a specific use case (e.g. a community area network)
- delay and jitter: under a range of typical scenarios, one could assess end to end delay and jitter across given use cases or scenarios.

The protocol operation of ad-hoc protocols does not lend itself easily to traffic analysis in a general way. For example, a table driven protocol like AODV (adaptive on-demand distance vector) may generate more overhead in routing packets but will have a lower per-data-packet overhead compared to, say, DSR (dynamic source routing). The mobility model, rate of mobility and size of the network will also affect the amount of control information generated by an ad-hoc protocol. Differences in overheads for AODV and DSR are quantified in section 6.1.3.3 for several different situations.

5.3.2.3 Constrained flooding based routing protocols

Even though reactive routing has benefits for ad hoc systems compared to proactive routing there are still issues with the operation and performance of routing protocols that are not prescriptive even within a given scenario.

Table-driven routing protocols (such as AODV) maintain routing tables that are built on demand. Other routing protocols (such as DSR) cache the whole path. There is then a trade-off between the amount of state information held at each node, and the amount of state information that accompanies each packet for the purposes of forwarding. Both approaches have an overhead for route discovery, though the exact nature of this overhead will depend on mobility models, number of nodes in the network and so forth.

However, such protocols also share a common problem; they require the state information for the entire path to be available for a packet to be forwarded. As well as the added latency for (at least) the first packet transmission, this means that the rate of mobility and the mobility model may affect the operation and efficiency of the protocol. Intuitively, we can imagine that if the rate at which topology changes occur is greater than the rate at which route changes can be discovered and the routing protocol stabilise to a new view of the network, then packet forwarding will suffer.

So there is increasing research activity in looking at approaches to ad hoc networking that use constrained flooding approaches, using information about data packet forwarding to neighbours in the network as indicators of the network topology. For example, in a 'backward learning' approach, data packets may be flooded when a route to a destination is not known, but seeing packets from the destination forwarded from a given neighbouring node is a strong hint that that neighbouring node may offer a path back to the destination.

5.4 Other co-existence approaches

Schemes like dynamic frequency selection (DFS) and transmit power control (TPC) are excluded from this report, but the use of higher frequencies (i.e. in less precious spectrum) for extension models has already been raised in section 4.4. Other aspects are briefly listed below, to provide an overview.

5.4.1 Knowledge based approaches

There are several drivers for moving to a knowledge based approach.

- there is no design for interference-free working in the commons approach
- the actual waveforms are unknown in the commons and this can make a difference
- what are the measurement possibilities for interference?

The knowledge required may be either table driven, much like a fixed routing table, or it may be ‘learnt’ using yet-to-be-developed technology.

A cognitive radio is one that can respond to its environment. In a mesh network, particularly a mobile mesh, adaptation of the performance of individual links can be used to trade off throughput, range, power etc such that the overall performance of the mesh is enhanced.

5.5 Summary of susceptibility issues

PHY level issues were found not to be mesh specific, although these are reviewed in Appendix E.

At the MAC level, politeness issues are key, in that an impolite MAC will strongly tend to dominate. Efficiency issues with a distributed MAC, as in a mesh, were raised and the confirmatory example of 802.11 was cited. The distributed nature of a mesh also makes synchronisation issues a challenge, for those schemes which need it..

Open issues in transport are:

- the impact of wireless in TCP
- the impact of high delay on TCP
- the impact of re-ordering on TCP
- the impact of extra delay on the perceived quality of service
- the impact of higher loss rates in the protocol and codec design
- new transport protocols and their relationship with new applications.

With respect to routing, there are three basic problems to be addressed before the promising reconfigurability feature of mesh networking can lead to network resilience for users and their applications. These are:

- the poor understanding of traffic and mobility requirements
- that TCP congestion detection/back-off operation is mismatched to a radio channel
- the effect of re-ordering issues and delay variation.

Whilst mesh networks do have the capability to physically re-route under conditions of interference (for example), the drawbacks of doing so may be felt by the application via effects above the physical layer. Overall, the issues lie within the bounds of access and transport protocols, and specifically with their interaction with the lossy nature of radio transmission and the distributed quality of a mesh network. Many issues would make suitable future work, section 9.

6 The potential case for mobile mesh adoption (Ofcom Work Item 4)

Section 6.1 examines where mesh networking could find application, having acknowledged the capacity and scalability limitations already illustrated by the hypothesis testing of section 4. The potential applicability of mesh networking is next tested via three key questions:

1. Would a mesh enable new services?
2. What degree of mobility could a mesh support?
3. Could a mesh guarantee a quality of service level?

The answers to these questions lead into a critique of several likely adoption scenarios within section 6.3, consisting of the following:

- cellular multi-hopping
- WiFi hotspot extension
- community networking
- home and office indoor networking
- zero- or low-infrastructure environments

This section concentrates mainly on Access Meshes, i.e. those with access to services beyond other user nodes in the mesh.

Wider mesh specific issues including unfamiliar battery exhaustion mechanisms and mesh security approaches are introduced in section 6.4, since they should be considered before any deployment. Finally, section 6.5 provides overall conclusions regarding the potential of adopting mesh networking.

Note that section 7 of the report includes views on the time-scales for mesh adoption in its consideration of how to encourage innovation.

6.1 Remaining key questions for mesh adoption

6.1.1 Review

The report thus far has shown that several idealised expectations of mesh networking do not realistically apply to mobile meshes below 3.5GHz. However, this is far from saying that such meshes will not find application. If meshes were to be deployed, then what could they offer? Three key questions are answered next.

6.1.2 Would a mesh enable new services?

The published literature is not full of ideas for new services which a mesh alone would enable. But this should be taken in context: Published literature has often failed to predict a single killer application, even for the Internet.

Like the internet, meshes are expected to be required to carry a variety of traffic types. All meshes discussed in this report have an underlying attribute of relying heavily on peer to peer working for their basic operation. Comparison with internet traffic types quickly highlights peer to peer file sharing as being similar in its operation, and section 4.2.4.3 showed that localised services would exploit the benefits of mesh very well. If all peer to peer users were within the same mesh and if all sources of traffic were also within the same mesh, then traffic would be localised within that mesh.

Example services could be local file shares of local content e.g. music files, video clips. Privately generated content would be freely distributable this way, but Digital Rights Management (DRM) issues could quash plans for peer-to-peer file sharing with commercial content in the near future, depending on the licensing model. Another issue with peer to peer is the question of what the providers' charging and revenue model will be, since it may be that no traffic leaves the mesh. At the very least some way of measuring the amount of each users' local traffic would be needed, or else a creatively marketed flat rate scheme would need to be offered to users. Finally, of course, there is no guarantee that peer to peer file sharers will all be within the same mesh. The closest potential application would seem to be home or office indoor networking, section 6.3.4.

In summary, it is not clear that mobile meshes will enable any hitherto unattainable services which would create revenue for an operator[20]. Although meshes may affect traffic patterns for existing services in the operator network, that will bring both advantage and disadvantage.

This report thus concentrates on providing already known service types over a mesh.

6.1.3 What degree of mobility could a mesh support?

The examination of the four hypotheses of section 4 used predominantly physical factors within its discussion. Those efficiency conclusions can never be reversed by consideration of higher level factors (transport, routing etc), since they merely add overhead from the viewpoint of the physical layer.

However, more detailed consideration of such higher level factors becomes key if a mesh is to be realistically deployed, so the cumulative, complicating effect on system performance may be assessed. This is the objective of the following discussion on mobility and, in section 6.1.4, of quality of service. These issues are quite complex and interlinked.

The effects of mobility are first put into context.

6.1.3.1 Effects of mobility

A key differentiator of this report is to consider specifically *mobile* ad hoc networks using mesh approaches. Clearly the degree of mobility involved must strongly affect results, since mobility can vary from nomadic, for example a docking laptop on the one hand, to vehicular mobility at high speeds on the other hand. In summary:

- Nomadic mobility has quite tight constraints which ease system design.
- True ad hoc mobility is largely unfettered and is thus the hardest for routing to deal with.

When modelling or analysing mesh networks, the 'mobility model' assumption and parameters used will be key. A high mobility node will be likely to pass by many other nodes and will thus cause many route changes in the mesh over the duration of the communication - and this will stress the routing algorithms most severely. This will lead to overheads via proportionally more control

[20] Whilst applications such as car to car telematics might be a good use for a mesh, this is considered to be beyond the requested scope of this report.

packets and fewer data packets being sent and hence to an overall poorer throughput being seen by all users. Mobility models themselves are expanded upon in Appendix D.

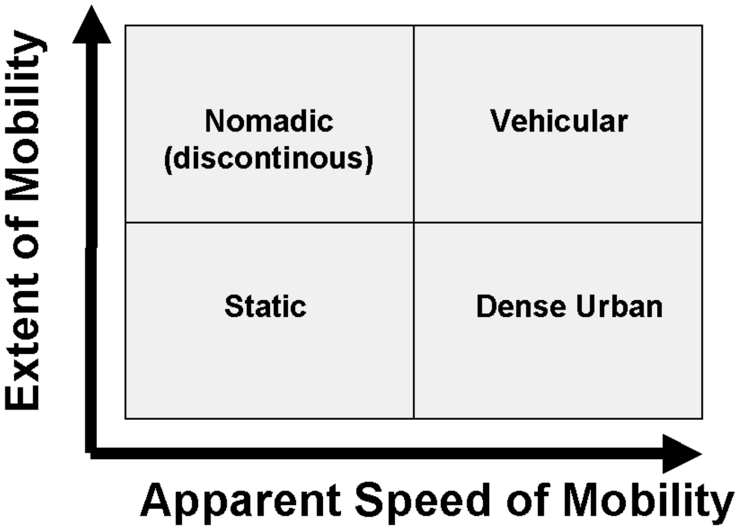


Figure 29 Application scenarios: matrix of range versus apparent speed of node mobility

Figure 29 shows mobility scenarios mapped onto axes of extent and speed of mobility. Note that the x axis is labelled *apparent* speed of mobility – this is to convey the idea that the concept of speed is measured with respect to how many other nodes (which may also be mobile) are passed by the mobile node in question, rather than how fast it moves relative to a fixed point in space. The reason for this is that the number of nodes passed will have direct relevance for the routing algorithm, which will have to discover new routes at a pace related to node speed measured in this way. This is why Dense Urban joins Vehicular as a high speed scenario: In the Dense Urban case, there is a high node concentration: many other mobile nodes will be seen by the user's mobile node and the routing protocol must work quickly to adapt to this (maybe more quickly than the Vehicular case).

Nomadic computing in the IETF mobile-IP sense is a slow, discontinuous process. Although a large range may be involved as a worker travels from fixed location to location, the connectivity expected by the nomadic worker is discontinuous. When travelling, there is no connectivity.

Figure 29 is included here specifically to point out that the critical aspect of dealing with mobility is how quickly the system must adapt (e.g. routing) and this is not always obvious from first impressions of the generic application scenario. This might be borne in mind as the impact of mobility on throughput is discussed in section 6.1.3.2 and routing protocol overhead is discussed in section 6.1.3.3.

<u>The need for new mobility modelling processes</u>

Little is really known about how nodes may move within a mesh, although one likely example is clustering around distributed hotspots. Thus nodes will move as a group around a hotspot. Where there is more than one hotspot, nodes will cluster generally, but individual nodes will move between the two (or more) hotspots. Although Camp et al [2002] (discussed in Appendix D) do cover group mobility models, no model exists for the above example. This is because prior models have assumed the military or emergency services scenarios, where nodes do not swap between hotspots (indeed there are no hotspots), rather they stay grouped in some form and the whole group moves, perhaps amongst other fixed membership groups.

A new mobility model for mesh networks is a ripe area for future work. If real world motion traces of node movements become available, they could be used directly or used to adapt models which could remain computationally attractive whilst still being representative. More information on current modelling approaches can be found in Boukerche and Bononi's [2004] recent review paper.

6.1.3.2 The Impact of Mobility on Throughput: Volatility

One of the generally agreed conclusions from modelling work cited in this report is that:

- For maximising network capacity and per-user throughput it is more beneficial to use many short hops through the mesh (giving high frequency re-use) than it is to use longer hops (lower frequency re-use) to keep the relayed traffic level down.
- However, although this is a reasonably well accepted principle, there are limits to it. As the hop-length reduces and hop-count increases there is increasing volatility in the routes through the mesh and hence increasing overhead for route management and re-transmission of lost packets. In other words, the mesh will eventually break up into pieces due to increased node mobility, which is termed 'partitioning' (e.g. Figure 34).

This volatility is due to:

- Increased dependence on the accurate measurement of interference levels and increased susceptance to short-term changes in these interference levels (as nodes move and/or the radio environment changes).
- Increased probability of link breakages as nodes move.
- Increased susceptance to errors in transmitter automatic power control (APC), (to maintain a target signal-to-interference ratio on each link).

One interpretation of this is to say that the network becomes more "brittle" as its capacity is increased by the spatial diversity techniques inherent in (mobile) mesh networking. Thus in practice it may be the network's stability and the integrity of its quality-of-service (QoS) which are likely to be real limits to its capacity.

Example of network 'brittleness'

One informative example of this "brittleness" is given in a system simulation [Hsieh and Sivakumar 2001] which explores the volatility of routes through a mesh as nodes move. Although this is a specific 802.11-based simulation it does illustrate some interesting characteristics which may be expected to hold for mesh networks in general. This example is examined in detail in the remainder of this section:

The simulation sets all node transmissions to the same power level – regardless of the hop-length between nodes pair. Although this is non-representative of a practical deployment it serves to illustrate the effect of power level (and hence link budget) on system performance.

For a random distribution of static nodes and routes the simulation first determines a necessary minimum power level (defined as MIN) at which the routes are fully connected. Simulations are then run for increased power levels of 2xMIN (= +3dB), 4MIN (+6DB), 16MIN (+12dB) and 256MIN (+24dB). The carrier frequency is 915MHz. The node relay throughput is 2Mbps. The model system is 1500m x 1500m grid in which four network sizes of 50, 100, 200, and 400 Nodes are uniformly distributed. Every Node acts as a Constant Bit Rate (CBR) traffic source.

The impact of node mobility is investigated by introducing the waypoint mobility model (Appendix D). In this simulation the pause period is set to zero in order to simulate continuous motion. Movement speeds are 5, 10, 15 and 20 metres per second.

Throughput: static users

As a benchmark the mean per-user throughput is first characterised for static users. The effect of transmit power on per-node throughput is illustrated in Figure 30. These results are for a 100 node network and an offered traffic rate of 64kbps per node.

It can be seen that as the transmit power is increased above the minimum required for connectivity there is initially little change in the per-user throughput: this is because the network is being run below maximum capacity and so can tolerate the additional mutual interference introduced. As transmit power is increased further the mutual interference starts to dominate and throughput suffers.

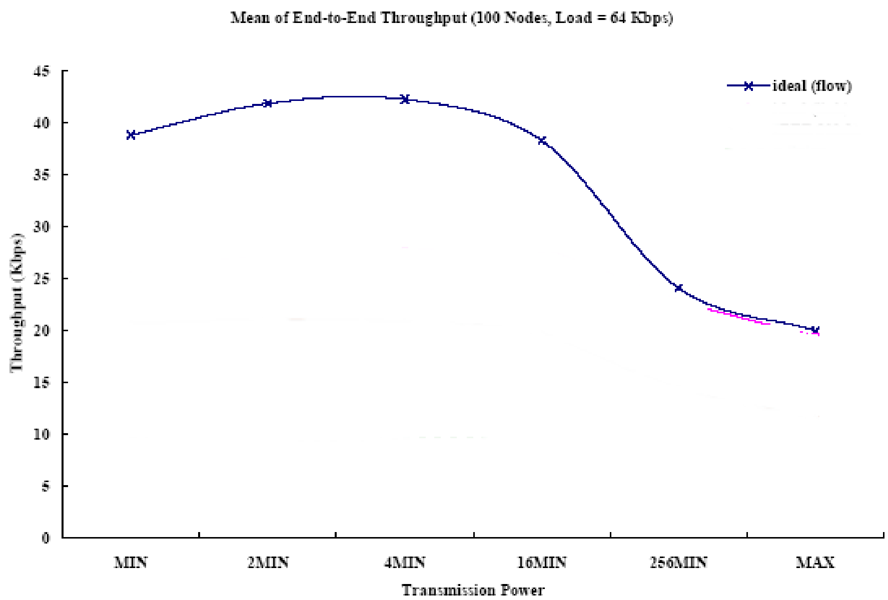


Figure 30: Per-Node throughput as a function of excess Tx power level

Under these static conditions the variation in throughput amongst users is quite low, as indicated by the standard deviation plotted in Figure 31:

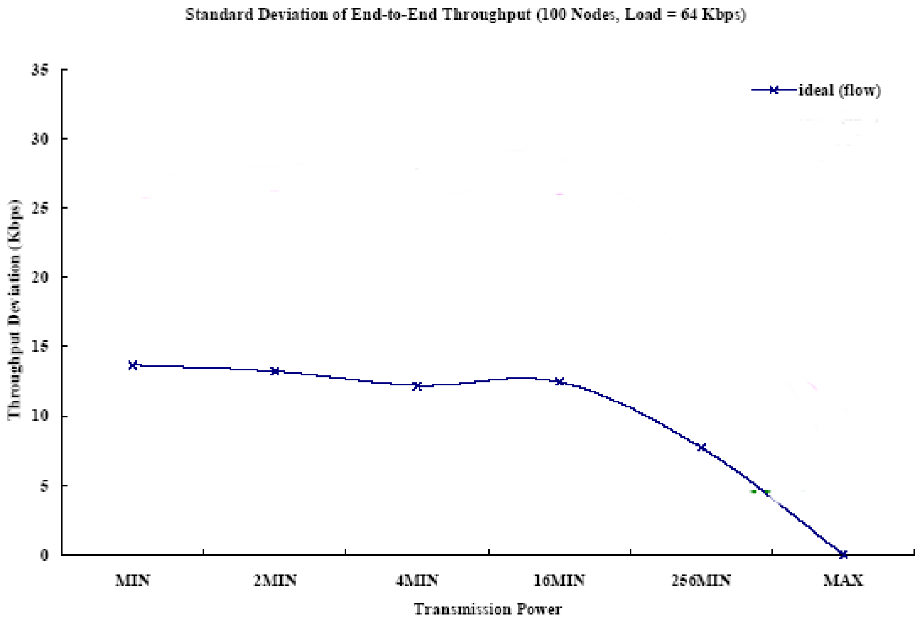


Figure 31: Standard Deviation of throughput vs. Tx power (static users)

Network partitions caused by mobility

The simulation is run for 100 seconds and its status is checked at one second intervals. The percentage of the samples for which the network is partitioned into one or more components is then logged.

Since initially the transmit power level is set at a level sufficient to just achieve connectivity it is unsurprising that any motion will start to cause disconnections. As the transmit power level is increased above the minimum the probability of partitioning decreases to zero as illustrated in Figure 32 for a 100 node system. When it occurs, the effect of partitioning can be to lose key interconnections and thus global re-routing[21] becomes impossible.

[21] i.e. to each and every node in the simulation

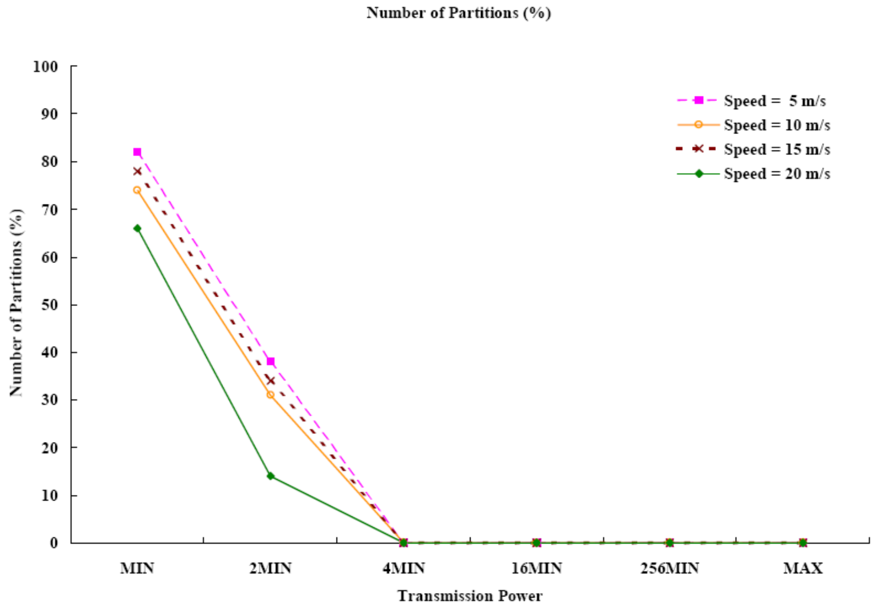


Figure 32: Percentage occurrence of network partitions as a function of excess Tx power level

Demand for path re-routes caused by mobility

Figure 33 presents the percentage of samples for which new paths (routes) have to be computed, as a result of mobility [Hsieh and Sivakumar 2001], and as a function of transmit power level. Three characteristic emerge:

- In the mid-power region, where connectivity is good and per-node throughput is reasonable there is a high probability of need for re-routing.
- Near the minimum power setting there are fewer re-routes implemented however there is a high level of network partitioning (in which state re-routing is not possible so end-to-end links are lost).
- At high power settings there are also fewer re-route occurrence however there are now fewer active routes because capacity is very low (due to the increased level of mutual interference at these higher Tx powers).

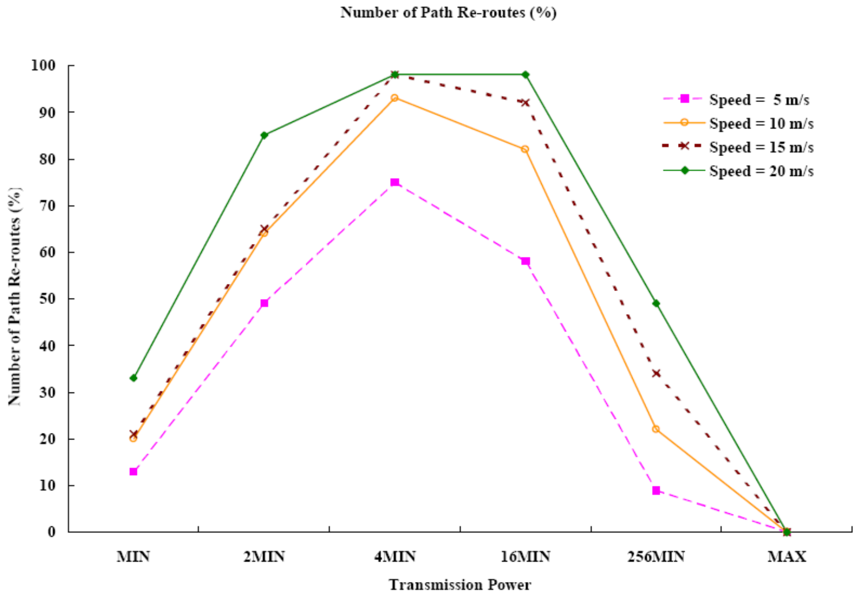


Figure 33: Percentage occurrences of re-route calculations as a function of excess Tx power level and speed of movement

Total Failures caused by Mobility

Figure 34 combines the failures due to partitioning (Figure 32) and path re-routes (Figure 33). This illustrates the overall volatility of the network to increased mobility [Hsieh and Sivakumar 2001].

Only at high Tx power level is there a reduction in the percentage of link failures due to mobility. This is a direct consequence of the fact that the coverage area around each active node is so large that mobility does not cause links to break during the average period of a message transfer. However, at these high power levels the network capacity is very low due to the increased level of mutual interference.

Thus one sees the trade-off between resilience to mobility and network capacity.

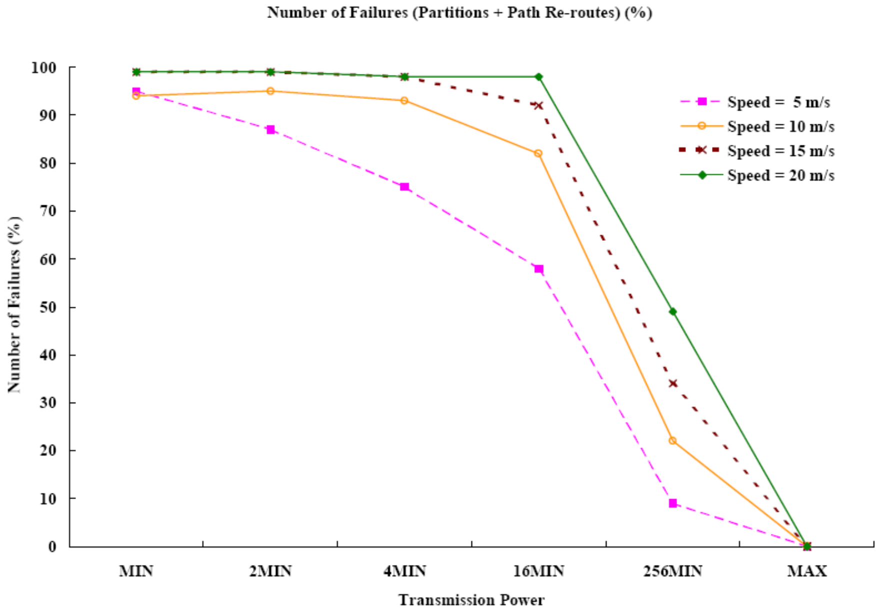


Figure 34: Total percentage routing failures as a function of excess Tx power level and speed of movement

<u>Throughput reduction due to mobility</u>

Whilst re-routing is not as catastrophic as network partitioning, it is a drain on resources as the routing overheads increase. This is illustrated in Figure 35 and Figure 36 which show the *mean* and *standard deviation of* per-node throughput as a function of mobility and excess Tx power level. (These results are for a 100 node network and an offered traffic rate of 50kbps per node)

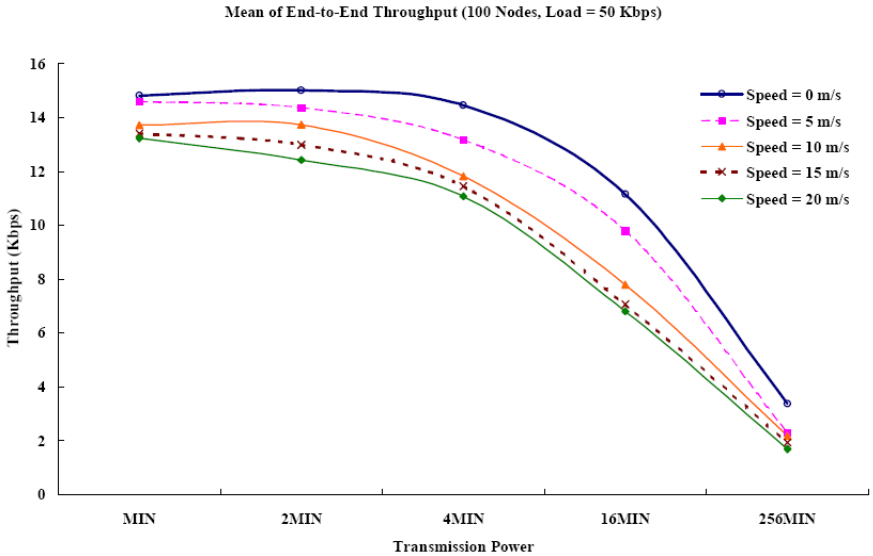


Figure 35: Per-Node throughput as a function of excess Tx power level and speed of movement

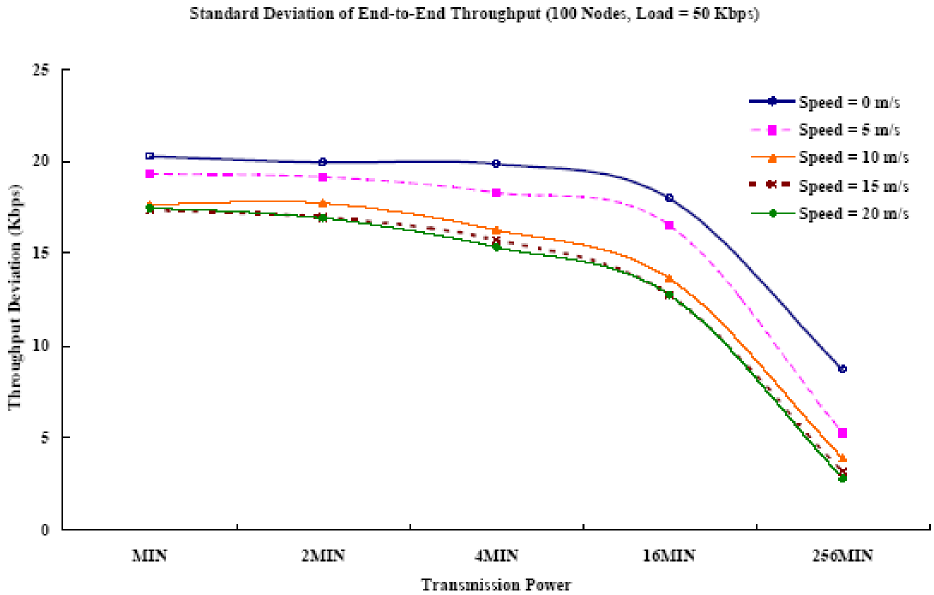


Figure 36: Standard Deviation of throughput vs. Tx power and speed of movement

Figure 35 indicates that the mean throughput achieved is now only about 30% of the offered rate of 50kbps, whereas it was approximately 65% for the static case shown in Figure 30. But furthermore there is now a very significant increase in the variation of the *actual* throughputs achieved by each node, as indicated by the very high standard deviations shown in Figure 36 (compare the static case in Figure 31). The standard deviation is now greater than the mean –indicating that some nodes are possibly not achieving any throughput at all. This is to be expected from the high probability of network partitioning at the lower power levels.

It is clear that throughput has deteriorated rapidly with mobility. Before network connectivity is compromised, routing overheads are seen to rise. The next section looks at what levels of routing overhead and other routing metrics might be expected in normal operation.

6.1.3.3 Routing overheads

A comprehensive overview paper is provided by Das et al [2000] which contrasts DSR and AODV, two popular reactive routing algorithms for ad hoc networking. Examples from this paper are used to bring out the compromise between routing overhead, delay and packet loss. The computational capability a node would require for performing routing calculations are not included, but are not thought to be a major concern, relative to those already required by a node for modulation-related signal processing.

Das et al [2000] tested the following metrics for AODV and DSR:

- packet delivery success rate
- delay
- overhead

Das et al found a wide set of complex results depending on the parameters of the mobility model and which routing algorithm and metric was being tested. For example they found routing overheads ranged from 33% to 75%, see Figure 37 (left hand). Since they are critical to the system, routing packets are always given preference over data, so any routing 'challenge' immediately pushes up the level of overhead.

The mobility model used was random waypoint (see Appendix D): A random, bounded mobile speed and a random, bounded pause was used before the next transmission in order to create different relative speeds between mobiles. Mobile speeds were 0-20 m/s and the area was 1500m x 300m for 50 nodes and 2200 x 600m for 100 nodes. Pause times varied from zero to the full simulation time (900 or 500 seconds). A shorter pause models higher mobility.

CBR traffic was used, always with 512 byte packets. Source-destination pairs were chosen at random and a variable number of pairs and transmission rates were used to load the network. Note that the model was never run with parameters which would break the connectivity of the mesh, unlike 6.1.3.2. Note also that these results are merely examples and are not presented as bounds for each metric, which could actually be much wider.

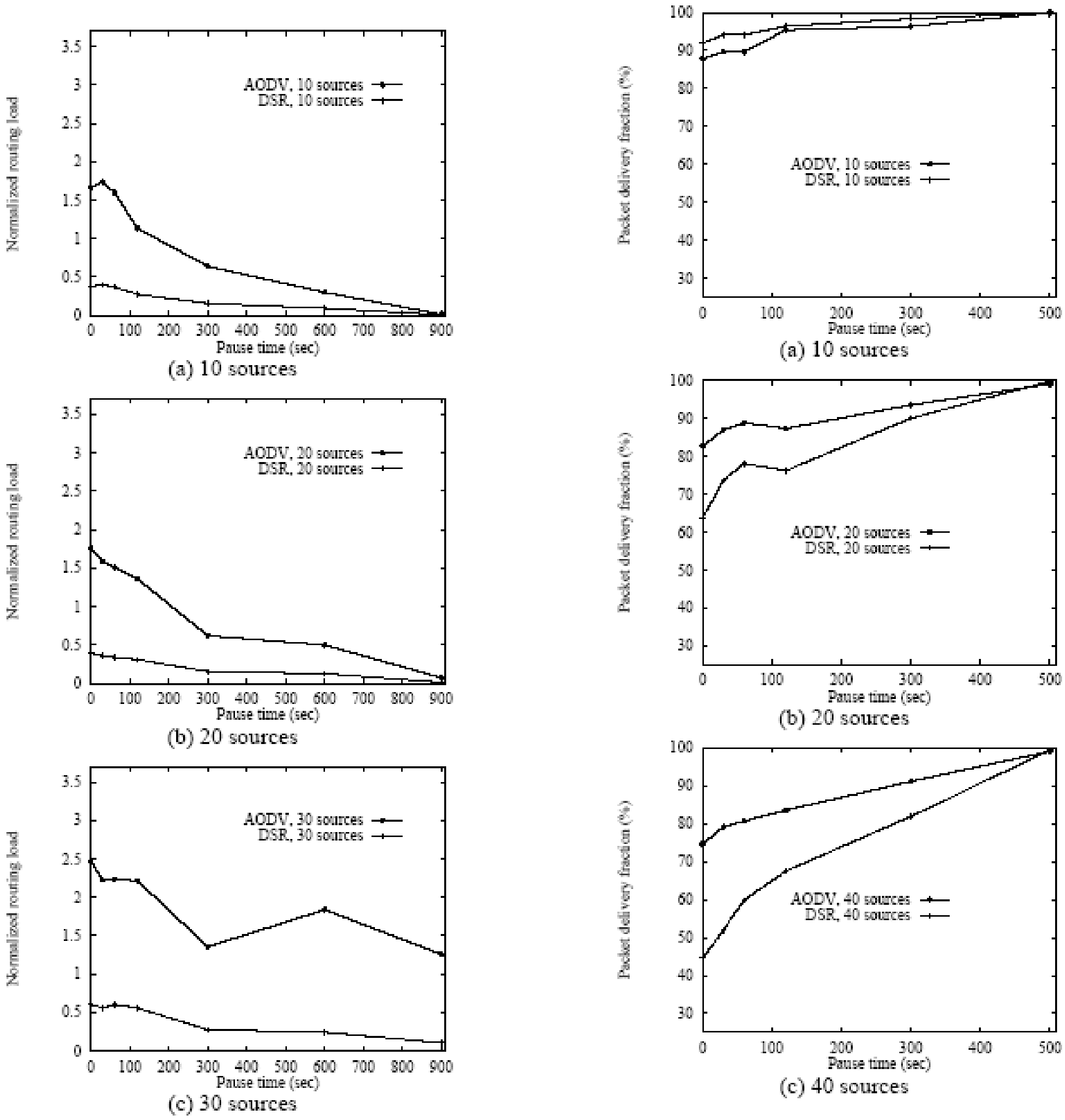


Figure 37 Left: Normalised routing loads for DSR, AODV (0.5=33% overhead, 3=75%) Right: Packet delivery success rate (100% = no dropped packets)

Figure 37 (right hand) shows that at higher network loading (more traffic, more mobility) packet delivery success can be considerably below 100%. Not shown are other graphs which report that packet delay is very dependent on load and protocol. Inelastic applications (e.g. video) may tolerate a few dropped packets well (due to both error control coding and human perception of video), but begin to fail with delay variation. Elastic applications like email require 100% delivery success but are not so sensitive to delay variation.

In summary, neither protocol was found to be better overall in terms of delivery, delay and overhead simultaneously, over all network loads, although DSR was better all round at lower loads and vice versa. This poses an obvious problem for the mesh system designer, who presently does not have the tools needed to please all the stakeholders all of the time. This is an area of intense research at present.

A discussion of routing overheads leads naturally into the wider area of quality of service.

6.1.4 Could a mesh guarantee a quality of service?

Different applications will require different qualities of service and one very important attribute within a quality of service offering will be delay and its variation over time. Applications may be described as elastic or inelastic, meaning they may be tolerant or intolerant of delay (also called latency). Email is elastic as the time taken for delivery is not critical as long as it is within reasonable limits. Much of the LAN based ad hoc networking literature assumes elastic applications are the norm and hence regular store and forward transport will be appropriate. This is unlikely to be true into the future of fixed mobile convergence, where new applications such as video calling or multimedia entertainment will certainly place inelastic demands on the network.

By way of an inelastic application example, whilst not specific to meshes, it is worthy of note that mobile displays are being called the '4th screen' concept (i.e. cinema, TV, PC, mobile). Fox Entertainment has its TV series "24" in 'mobisode' format already. This presently has a 1 minute run time, but is expected to be expanded. The 4th screen is specifically seen as a new entertainment medium which requires bespoke content creation. This year's Edinburgh Film Festival has a 'films for mobiles' component.

Having confirmed the need for quality of service levels, an early question to be asked is whether quality of service is possible at all within an uncontrolled interference environment[22], since unlike cellular, mobile ad hoc mesh networks cannot be planned for a known interference environment *ab initio*, due to user movement. This will be at odds with users' expectations, who will require guaranteed service for inelastic applications like video or anything else intolerant of delay-variation or throughput-variation.

Perhaps because of this basic difficulty, many LAN biased research efforts have concentrated on the notion of best effort delivery, as used in the main on the Internet. Such operation is often called 'store-and-forward', with the main objective being 100% delivery success (but in no way guaranteed, hence the need for transmission control protocol; TCP). Latency is very much a secondary objective with a very low priority: Some commercial latency figures might appear to be driven more by the desire of not overloading storage space queues in transmission equipment, rather than as a service offered to the user. Conversely from the cellular and telecoms world, service level agreements are much more the norm. The concept of service availability is also key.

For whatever reason, the quality of service aspect of meshes does not always receive the attention deserved: The move to being much more reliant on user behaviour for network performance is a situation starkly absent from present mobile networking. Its effects can be disruptive and it is

[22] self-interference from other users is meant here, not interference from outside the mesh. See section 5 for a discussion of susceptibility to outside interference (which would also affect quality of service).

examined next.

6.1.4.1 Dependence of QoS on user behaviour

Some published literature in this field is from the proponents of cellular extension by multi-hopping. This technique will be discussed more closely later in section 6.3.1. However for the time being it is sufficient to appreciate that the technique involves extending the reach of a base station via allowing nodes to act as relays, such that links to the base station may be greater than the one hop of traditional cellular. In this way it is also hoped to increase coverage (by filling in black spots) as well as increase range. As such it is pertinent to multi-hop mesh networking.

Mobility versus connectivity

Royer et al [2001] noted the result of a well known publication which stated that a node with the ‘magic number’ of six neighbours could be shown to possess the optimum trade off between transmission power and self-interference for best throughput [Kleinrock and Syvester 1978]. In other words, with more or less than six neighbours, throughput was shown to fall either due to increased relay hops or bandwidth lost due to the bigger interference footprint of neighbouring nodes. This was derived for fixed networks and Royer at al’s question was whether a similar optimum existed for mobile networks.

Their answer was ‘no’. There is no single optimum power-efficiency trade off for a range of node mobilities, although there is an optimum for each value of node mobility. Essentially as node mobility increases, Royer et al [2001} found that the node density, in terms of connectivity, needed to rise: More mobile nodes need more neighbours. This could be ensured by increasing transmit power or by adding more nodes. The concept is quite easy to rationalise in principle: As nodes move more quickly, the likelihood of any link breakage increases along with their speed (already noted in section 6.1.3.2). In order to have a higher probability of remaining connected by some route, the number of neighbours within radio range must be increased by some method.

The clear issue for the mesh network planner is that network connectivity depends on parameters outside his control – the users. Only by adding some permanent seed or relay nodes can some degree of control be reasserted. This is next examined a little more closely and it is noted that what is really under discussion incorporates the notions of availability and quality of service.

Coverage , availability and quality of service

Firstly Royer et al [2001] showed that connectivity depended on user mobility. More recently Nilsson [2004] showed that connectivity also depends on the traffic level within a multi-hop mesh.

Nilsson extended a similar approach to that of Royer at al, using a similar routing protocol in his model, plus a more modern protocol. In each case, he confirmed the dependence of packet delivery success on user mobility and additionally noted that at low traffic loads packet delivery rate could increase quickly with increasing transmit power (i.e. connectivity), whilst as the traffic load increased, the rate of improvement in packet success rate slowed down. His conclusion was that both predicted traffic levels and node mobility/densities would need to be known at the planning stage, for viable network design.

Lugaro and Wallin [2003] expanded on Lugaro’s earlier work [Lungaro 2003], where notably they added a caveat to his earlier enthusiasm for cellular multi-hopping: They noted that the principle drawback of the scheme was that, due to the ‘partially uncontrolled infrastructure’ (i.e. the user nodes), multi-hopping was not able to guarantee a quality of service. They proposed that a simple solution was the addition of nodes as relays, by the network operator. In fact they proposed a system with three node types: access node, relay node and user node. Their conclusion goes beyond Nilsson’s; knowing traffic and user mobility is still not enough to guarantee quality of service in a mesh of user nodes. Infrastructure must be added. More detail from both Lungaro’s papers is given in the later section 6.3.1.1.

Accepting this conclusion from Lungaro and Wallin, a contemporary paper from Sanzgiri and Royer [2005] suggests a novel way of 'finding' the required quality of service within a mesh by directing the user to a new physical location within the mesh where the desired quality of service is known to be available. A protocol to achieve this was described. This is analogous to walking around in a marginal GSM area, until a stronger signal is found. It is not clear how acceptable this would be to users.

6.1.5 Conclusions

Answering the three key questions within this section provided the following conclusions:

- No evidence was found within the research community to indicate that the introduction of mesh networking would enable new services[23] which only a mesh could provide. The provision of current services via a mesh is thus the focus.
- The degree of mobility a mesh can support should be measured by the number of other user nodes passed rather than by measuring speed relative to a fixed point. This metric begins to quantify the number of re-routes[24] required of the protocol. For large enough mobility, a mesh can break into several disconnected pieces, a process termed 'partitioning'. Before partitioning, an active mobile mesh may require as many routing packets as data packets for its operation. Further quantification of mesh mobility characteristics (as a planner would require) appears to be a large and open research subject well beyond the scope of this report.
- Quality of service within a mesh is user dependent, in terms of both mobility and traffic level. This effectively means that an operator cannot guarantee a quality of service level. There are two ways around this problem:
    1. Add infrastructure, such as fixed nodes
    2. Direct the user to physically relocate to a 'better' part of the mesh, via a protocol.

The key conclusion above is that the **quality of service within a mesh is user dependent, in terms of both mobility and traffic level**. Unless specific additional steps are taken to mitigate this, an operator will be unable to provide service level guarantees. Such mitigation will need to be tuned to the actual mobility and traffic circumstances in each case, if they are known.

6.2 Coverage

Most likely scenarios will exploit either, or both, of

- the coverage, or the
- no-fixed-infrastructure

aspects of meshes, in order to maximise any or all of the metrics of efficiency, economy and effectiveness. Efficiency is taken to mean technical efficiency, economy is a cost based issue and effectiveness is a measure of providing most closely that which is needed.

As has already been said in section 1 of this report, meshes are not a replacement scheme for

---

[23] as defined within the bounds of this report

[24] cellular handovers are a broadly similar concept, in the sense that they create an overhead for a similar reason

existing wireless communications, rather they may expect a place in the wider vision of a future integrated mobile communications architecture. In simple terms they are a 'horses-for-courses' solution. In other words, their appeal will depend on the scenario into which they are deployed, relative to other potential wireless solutions, such as cellular for example.

A submission to the IEEE802.16 committee [Beyer 2002] summarised two scenarios for evaluating meshes, together with a very pragmatic reasoning behind the choice. The submission assumed fixed meshes, but the coverage reasoning applies to any mesh or multi-hop system. Beyer's distinction between the PMP (point to multipoint) approach and the mesh approach is summarised below:

The premise is that real world propagation effects can be well related to the log and normal contributions to the log-normal fading model. Whilst, strictly, this is a probable mis-interpretation of a log-normal function, it does lend a very worthwhile physical viewpoint. The proposition is that PMP copes most easily with the log fading of an open environment, but must use high power to cope with the normal fading environment due to clutter. For mesh it is proposed that the opposite is true, so that mesh deals well with cluttered environments, but not so well with distances.

In short, Beyer's assertion is that meshes are about coverage, since they 'skip around obstacles', as shown in Figure 38.

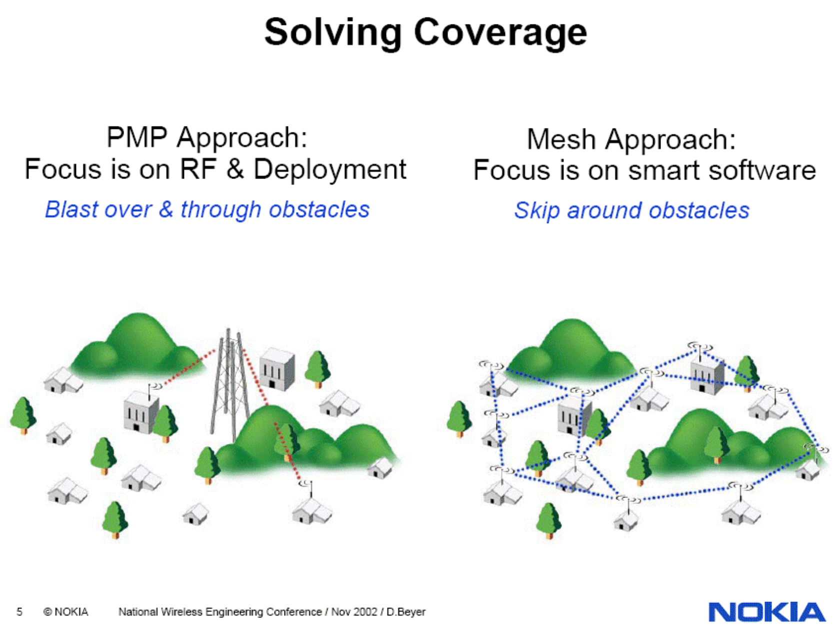


Figure 38 Distinction between PMP and mesh approach in the cluttered environment [Beyer, 2002]

Note that Beyer acknowledges that the success of the mesh does rely on 'smart software': A very similar conclusion has been drawn by this report which has explicitly noted what this smart software needs to do in the routing layers especially (see elsewhere in section 6).

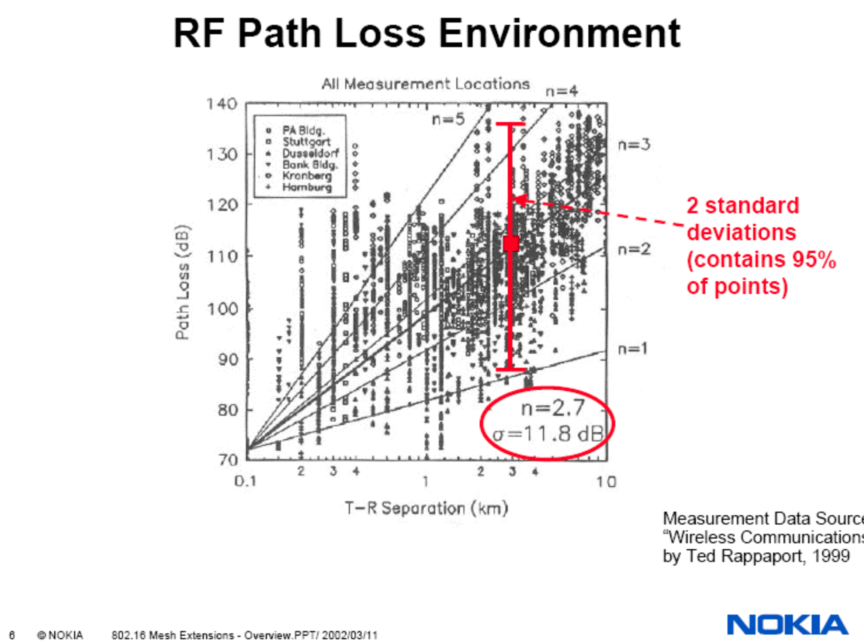


Figure 39 RF path loss, showing log and normal contributions [Beyer 2002]

Figure 39 shows measured signal path loss around several German city locations, as might be observed within a typical cellular deployment. It is the form of the results which is of interest: The distribution may be described as log-normal. The constants n and σ have been drawn on the figure in a best fit approximation to model the measurement distribution. n and σ describe the log-normal path loss following the well known formula:

$$\text{Path loss} = (\text{a constant}) + 10*n*\log(\text{distance}) + X_\sigma$$

n represents the average log fading attributable to distance and X_σ represents the variance of the normal distribution around this average caused by a distribution of cluttered environments.

Referring to Figure 38 and Figure 39 jointly, it may be deduced that for the PMP example, a large variance is bad, since the system must be designed for the worst case, hence a high RF power is needed. This will affect other base stations negatively. In complete contrast, for the mesh system, a large variance can be turned to advantage - if the node has the necessary software intelligence to pick the best link - since the best link may 'skip' around the obstacles over a distribution of RF paths with much less variance. In the mesh case, therefore, lower RF powers are needed to cope with the lower power budget and a smaller interference footprint results.

Figure 40 further illustrates the proposition that meshes avoid the regular PMP problem of dealing with channels with high variance. The mesh deals with what might be called a restricted 'sweet spot' of path losses shown in blue in Figure 40.

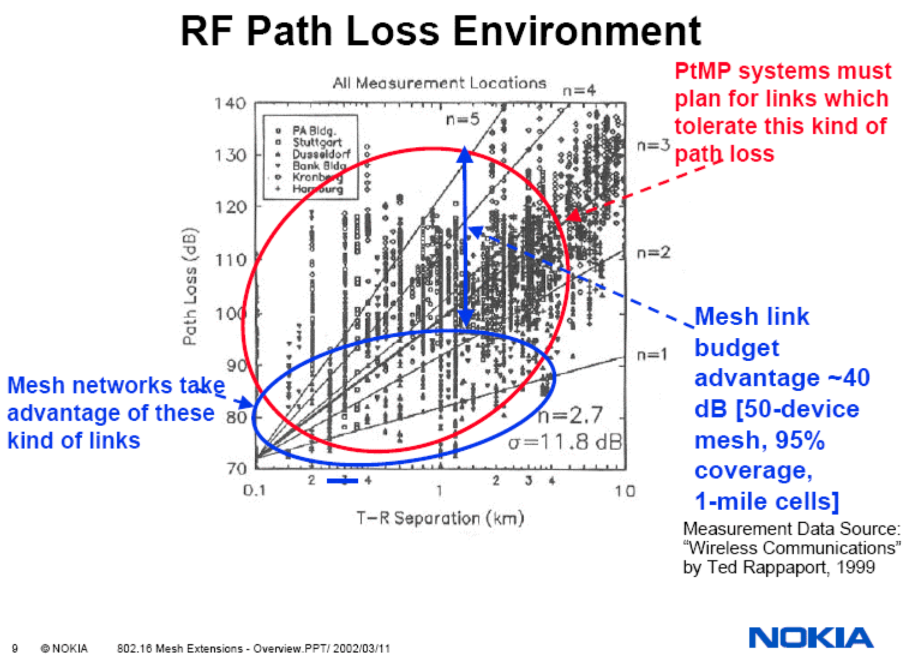


Figure 40 Sweet spot identification of mesh relative to PMP [Beyer 2002]

Next, two scenarios are considered which were specifically chosen to highlight firstly where PMP systems and secondly where mesh systems show clear advantages from the propagation viewpoint alone.

6.2.1 Rural environment with open propagation

This scenario is very simple; all links have low loss and low variance. There is neither advantage nor disadvantage in running a mesh system from a propagation point of view.

6.2.2 Urban environment with obstacles

This scenario is also very simple; the majority of links have a high average loss and a large variance. One extreme example would be the so-called Manhattan model (see Appendix D); a grid of US city streets. Shadowing and thus variance are very high. The proposed advantage of a mesh here is to hop around obstacles. This would be more power efficient than the PMP case which would need to have sufficient power to cope with the worst case. Of course it should be mentioned that there is a middle ground – cellular picocell deployment would be an alternative step to offering what meshes could offer. One tacit assumption is that mesh node density is high enough that mesh connectivity is maintained. In the early stages of mesh deployment, this situation would have to be ensured by a deployment of seed nodes by the operator, although these could also double up as external infrastructure access points, which are likely to be permanently needed in any case.

Esseling et al [2002] compare meshing around obstacles via a 'forwarder' to gain an advantage in throughput, see Figure 41. This clearly supports the assertions of Beyer discussed above.

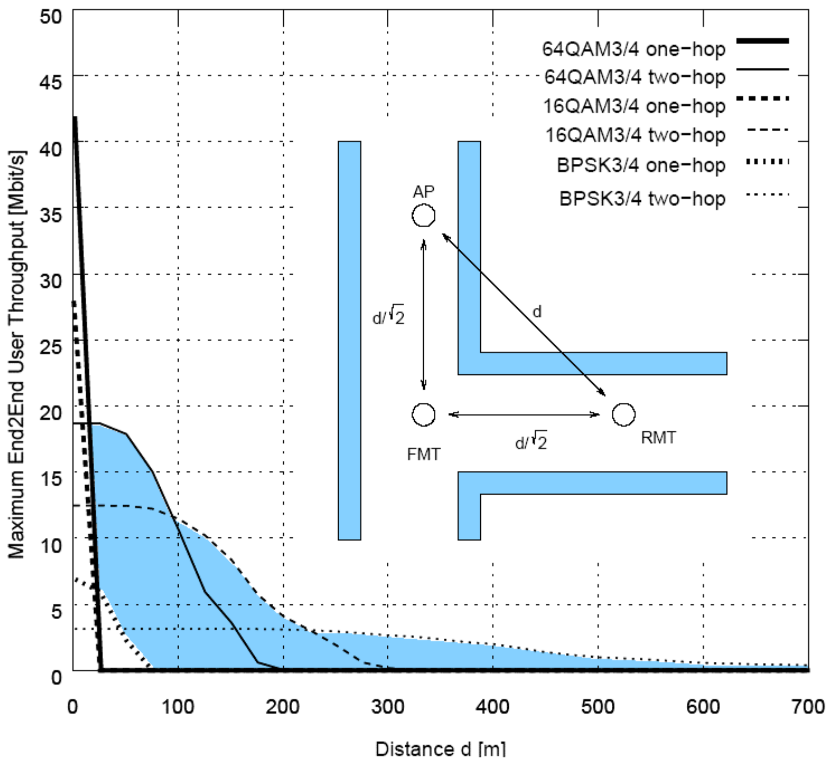


Figure 41 use of a 'forwarder' to 'skip' around obstacles as in a mesh network [Esseling et al 2002]

Figure 41 may be understood as follows: The situation being modelled is as drawn in the figure inset within the graph. The objective is to communicate from the AP (access point) to the RMT (remote mobile terminal), which is sited around a corner. This may be done in one of two ways:

- as a single hop, distance d, 'directly'
- as a two-hop, twice distance d/√2, 'hopping' around the corner via the FMT (forwarding mobile terminal)

Additionally, the effect of adapting the modulation scheme is also shown.

In Figure 41, the potential gains to be had in all the two-hop cases are shown within the total blue shaded area of the graph which is made up of the gains from each two-hop case listed in the graph legend. The white area enclosed by the axes and the blue area is the single hop performance, which is very limited as might be expected.

The results of Figure 41 relate to extending the infrastructure mode of HiperLAN/2 in the 5GHz band, however the implications of this simulation are of a general nature and thus may be applied to other hopping situations, such as a mesh network.

6.2.3 Extension to a mixed environment

Clearly in the real world, real scenarios are in between the extremes outlined above. The best solution thus depends on the detail of the situation. In general however, meshes should show benefits where the problem faced is one of coverage due to a cluttered or shadowed environment. Most often this would be expected to be a dense urban area. Potentially, meshes could be deployed as ‘hot zones’ in a sea of PMP.

Hopping around corners or obstacles is perhaps the greatest benefit offered by a mesh.

The next section looks at several likely scenarios in order to illustrate the points above.

6.3 Likely adoption scenarios

To summarise once more, at this point in the report it has been shown that practical mobile meshes are not primarily chosen for spectral efficiency or self generation of capacity. However, meshes do have other benefits. Section 6.2 provided an introduction to how meshes offer coverage benefits, which is possibly a mesh’s major attribute. The following sections cover potential applications, comprising

- cellular multi-hopping
- WiFi hotspot extension
- community networking
- home and office indoor networking
- zero- or low-infrastructure environments

6.3.1 Cellular multi-hop

At this point, it is pertinent to review a corollary to the Access Mesh. This is the prospect of multi-hopping between hand-held/mobile terminals within some future cellular network. Within the cellular industry this proposed architecture is called ‘cellular with multi-hopping’. The concept is one of using the relaying function to either

1. extend the range of higher bandwidth services, or
2. increase coverage availability within a cell.

Each of these offers the potential benefit of reduced infrastructure costs via fewer cell sites. An additional bonus, given the discussion within section 4.5 of this report, must also surely be the potential to establish cellular-like services at much higher frequencies than the present 2GHz band. Contiguous cellular coverage with the present model is highly unlikely at higher frequencies because of reduced link budget and higher clutter losses, whereas multi-hopping via relay-terminals could potentially offer this.

For example, in the case of 3G, with its higher transmission rate capability, the poor availability of high data rate services over a cell’s coverage area is an acknowledged weakness. Operators are unlikely to invest in the necessary additional infrastructure to remedy this over their entire service areas, but would welcome a more cost-effective solution.

Firstly the mechanisms and benefits of cellular multi-hop are covered, followed by the major disadvantage of increased reliance on user behaviour for system performance, along with an approach to mitigation.

6.3.1.1 Enhancing coverage availability

Considered here the potential improvements in coverage availability achieved through hopping via other mobiles, and/or via fixed relay nodes. A useful analysis of coverage enhancement is given in a comprehensive thesis by Lungaro [2003].

Lungaro's analysis is useful and relevant in that the air-interface is loosely based on 3G Cellular; 5MHz channel bandwidth and 20MHz total bandwidth at 2GHz. However, one weakness is that it adopts a very low value of differential link-budget between basestation-mobile and mobile-mobile links[25]. Nonetheless this does not undermine the conclusions since, whilst this makes Lungaro's work less suitable for comparison with a traditional cellular network, it does have relevance to the performance of an Access Mesh or WLAN-Hotspot network in which lower gain (lower height) base station installations would likely be used.

Lungaro sets a coverage availability design-target of 95% (i.e. 5% outages) and concludes that around 2-3 hops are sufficient to achieve this coverage availability whilst nominally halving base station density.

Two routing strategies are studied:

- Minimum Power Based Routing (PBR) which tends to pick multiple short hops to conserve RF power, and
- Minimum Hop Algorithm (MHA) which tends to choose the route with the least number of hops. With MHA the average number of hops to achieve this coverage target is near constant at around 2, as illustrated in Figure 42.

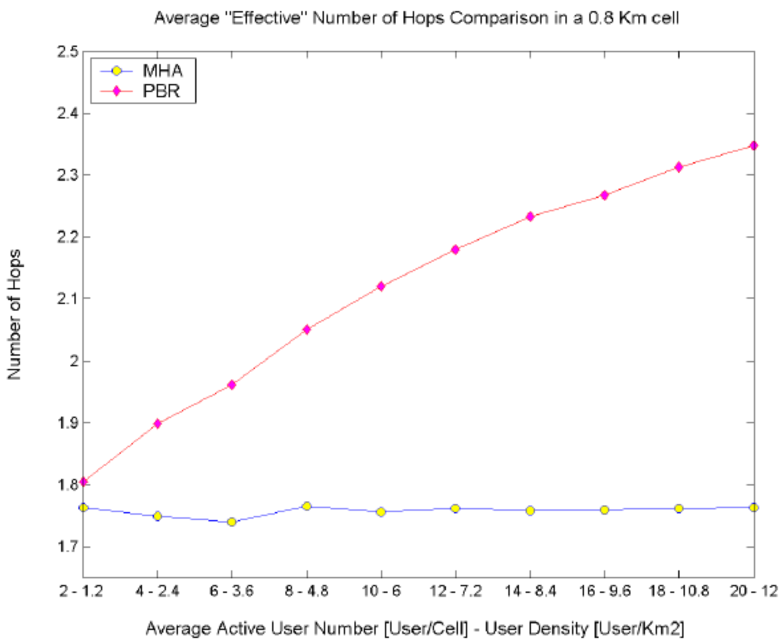


Figure 42 Number of hops vs. routing strategy [Lungaro 2003]

The coverage-availability improvement is illustrated in the graph of Figure 43. This shows the relatively small number of active relay-mobiles per cell which are necessary to support multi-hopping sufficient to achieve a given outage probability for a number of different base station densities.   Three cell radii of 0.4, 0.6, 0.8km at circa 2GHz are used.

[25] All modelling is carried out for 1watt Tx feed powers, 8.2dBi basestation antenna gain and 2.2dBi mobile antenna gain; which implies only x1.5 range differential.

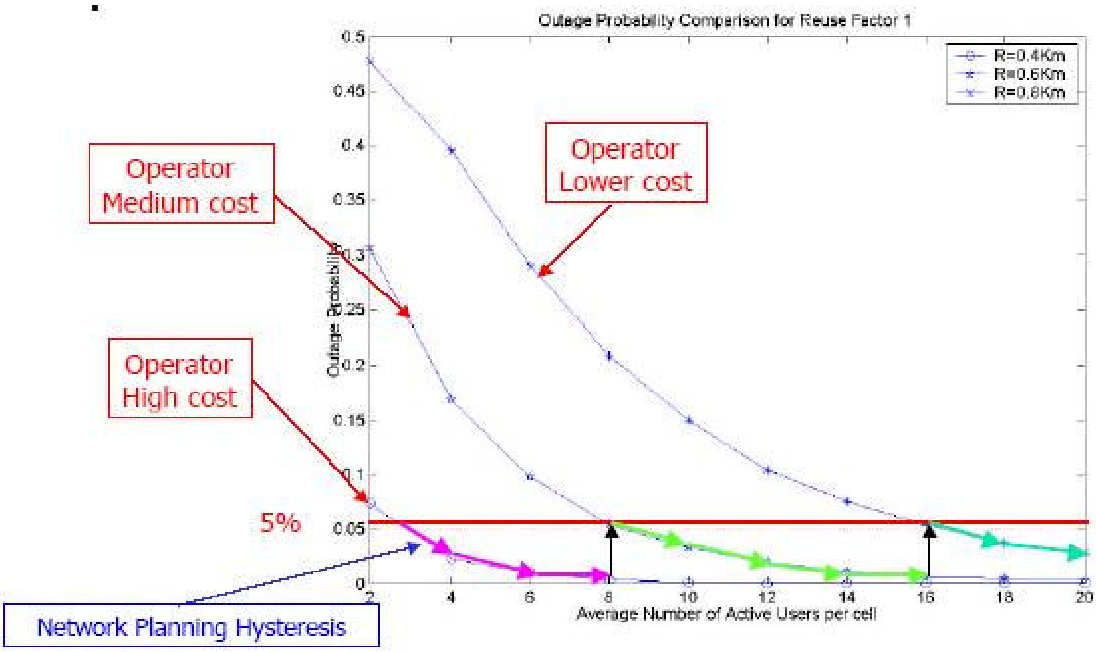


Figure 43 Relay node requirements to achieve coverage outage target [Lungaro]

6.3.1.2 Cell Boundary Extension

A further perspective on the above is the increase in usable cell radius that can be achieved by multi-hopping. Lungaro [2003] presents the results of comprehensive traffic simulation modelling on a TDMA network and concludes that of the order of x3 cell radius extension is achievable. An example is given in Figure 44, which, for a specific set of propagation and link-budget criteria[26], indicates an increase of cell radius from about 0.27km to about 0.82km.

---

[26] 2GHz band, 5MHz BW scenario, 1W Tx feed power, 8dBi base antenna, 2dBi mobile antenna, comprehensive theoretical path-loss model, propagation law = 3.6 + log normal fading. Again the very small differential in base- and mobile-antenna gains is noted.

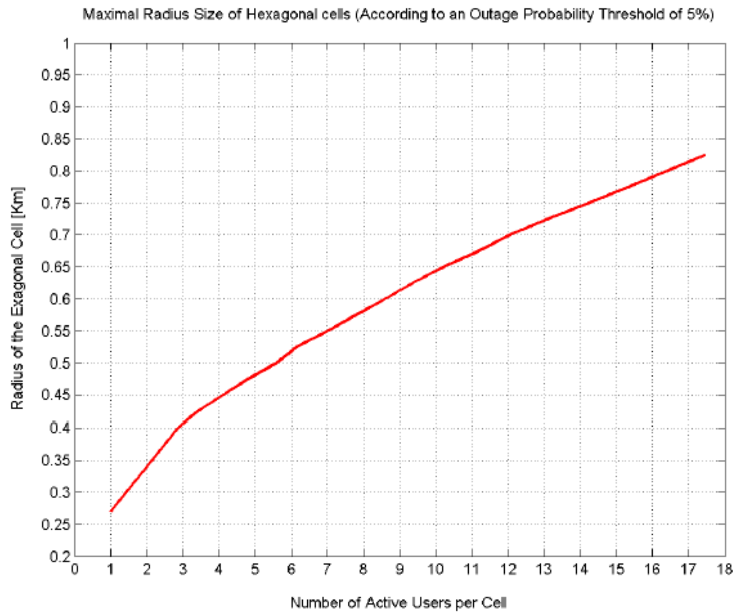


Figure 44 Maximum cell radius retaining 95% coverage availability [Lungaro 2003]

An aspect the above modelling does not consider is the limit caused by high traffic loading on a large network, whereas a major limit to range extension will be the finite throughput capacity of the mobiles/relay-nodes in direct range of the Basestation – which must carry all the traffic of the extended cell coverage.

By considering this aspect, an upper bound may be derived for cell extension.

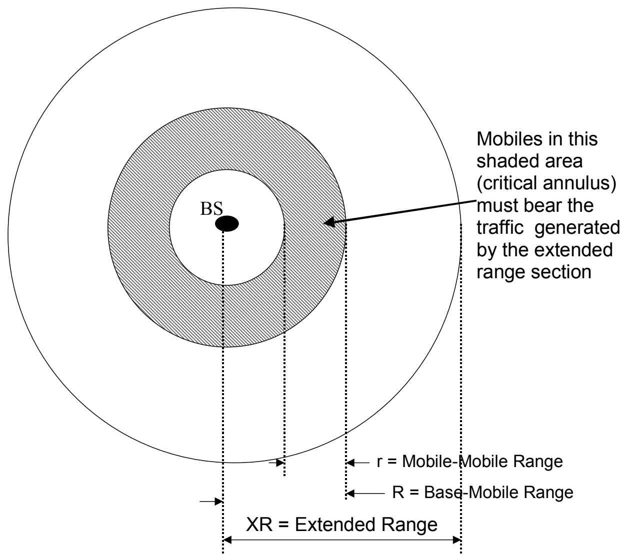


Figure 45 Cell boundary extension footprint

Traffic to mobiles outside the nominal range of the Basestation or Access Point is relayed via mobiles within its range. Therefore, all of this outer traffic must pass through mobiles in an area describing an annulus at the extremity of the basestation's range; the width of this annulus being the mobile to mobile range, see Figure 45. Assuming that the entire unused throughput of a mobile is consumed by acting as a relay then, from a consideration of the traffic concentration effect by taking a ratio of areas, the maximum range extension that can occur[27] can be shown to be given by the expression:

$$E * \frac{X^2 R^2 - R^2}{R^2 - (R-r)^2} \leq (1-E)$$

where E = offered traffic per user (Erlang)

R = Mean Basestation to mobile communication range.

r = Mean mobile to mobile communication range

X = Cell boundary extension factor

Note that range extension scales inversely with the square root of the traffic loading. If we define the ratio of Basestation and mobile ranges as ξ=R/r then this expression becomes :

$$E(X^2 - 1)\frac{\xi^2}{2\xi - 1} \leq (1-E)$$

By way of illustration, consider present, typical mobile-phone usage: Assume a busy hour loading of about 0.04 Erlang. An approximate range for ξ is 3 to 6 assuming 13-20dB additional link margin between Basestation-Mobile and Mobile-Mobile links and for propagation law in the range of 3 to 4. But a key assumption is that all of the mobiles in the critical annulus are able to connect to mobiles at distance r or less; i.e. that at least for the first hop or two outwards nodes fall into quite distinct annuli of width r. This is not realistic: in practice there will be a large overlap between rings to cope with fluctuations in user-density and path loss. The degree of overlap is a matter of debate, but we suggest it could be of the order of 50%. Each annulus would thus cover the outer half of the inner neighbour and the inner half of the outer neighbour. This would halve the effective radius of the mobile/mobile range, thereby doubling the value of ξ.

Figure 46 illustrates the above for a range of ξ and E : there is a clear tendency towards a value of approximately 2.

[27] assuming perfect load balancing, even utilisation, uniform user density, etc.

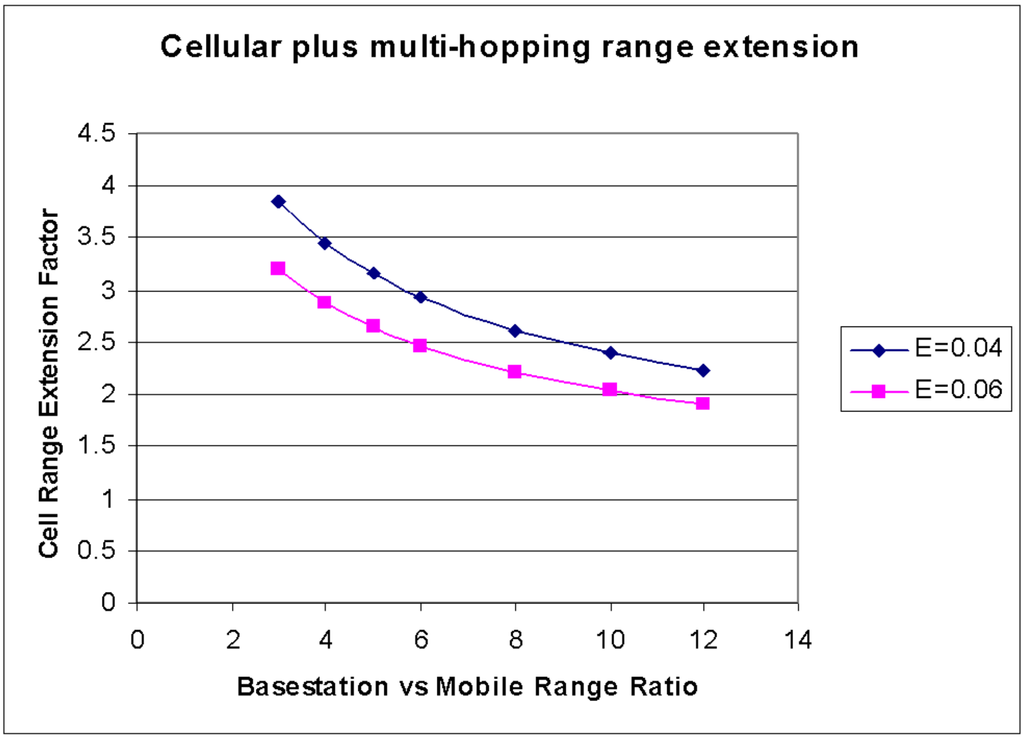


Figure 46 Cellular range extension through multi-hopping (upper limits)

Note that for the simulation of Lungaro cited above with only 6dB differential between basestation and mobile antenna gains and a 3.6 propagation power law, the corresponding value of ξ is 1.5. Then from Figure 46 the upper bound on cell extension is approximately x4. This compares broadly with Figure 44 from Lungaro which by extrapolation appears to be asymptotic towards x4 as the number of users increases[28].

6.3.1.3 The main drawback of cellular multi-hopping

In a later paper [Lungaro and Wallin 2003] Lungaro adds a critical caveat to his work on cellular multi-hopping:

> "The partial uncontrollability of the infrastructure, represented by the 'relaying' users and their density constitutes a significant threat, especially in an environment exposed to large user density fluctuations."

Lungaro immediately suggests that, by adding fixed router nodes (to the user nodes and access points already present), the situation may be improved. At the time of writing, Lungaro would

---

[28] The preceding analysis will be affected by further real world effects: For example, another assumption made in the above was uniform traffic loading and uniform user density. In practice the traffic load generated by users varies around the mean so to be able to offer a certain level of availability the calculations need to allow for a certain amount of headroom. The amount of headroom depends on the number of users involved, since larger numbers reduce the relative size of the fluctuations, and the availability offered. Furthermore there will be variations in the density of nodes in the critical annulus. More headroom will be required to handle these extra variations for a given availability. Detailed analyses of these are beyond the scope of this report, but it is suggested that the maximum achievable range extension, due to traffic throughput considerations only, is unlikely to exceed a factor of approximately 2. This suggests a quartering of the required base station density.

have been unaware of Nilsson's additional conclusion which showed that connectivity also depends on the traffic level within a multi-hop mesh [Nilsson 2004].

Lungaro showed that by adding relatively small numbers of additional routers, coverage outage may be reduced. However that was true only where a minimum number of users (greater than the number of routers added) remained present. Clearly the elimination of user dependence is only partial in this example. Presumably adding more routers would shift the trade-off, but introduce a higher cost. Given that cost reduction via reducing infrastructure is the very basis for cellular multi-hop, such an additional router approach would appear to be in danger of becoming self defeating.

Lungaro and Wellin [2003] also clearly show that quality of service is maximised when hop count is minimised. As such, multi-hop cellular will always be inferior in this regard.

6.3.1.4 Conclusion on Cellular with Multi-hopping

Promised benefits of cellular multi-hopping include:

- A reduction of the link-budget planning margin for log normal fading which traditionally burdens cellular networks. This reduced planning margin promises reduced inter-cell interference, so improving frequency reuse, and so improving spectral efficiency of the cellular network via multi-hopping.
- A range extension of around x2 – x3 could be achieved under certain circumstances. This represents a significant reduction in basestation density. The extent to which the cell boundary can be extended by multi-hopping depends on the ratio of basestation/mobile antenna gains and heights and on the throughput capacity of the relay nodes.

But the technique is not without drawbacks which include:

- Quality of service (QoS) issues such that the operator cannot guarantee service levels, as described in section 6.1.3. This may be combatted by adding fixed infrastructure, but that conflicts with the aim of cellular multi-hop, which is cost reduction via infrastructure reduction.
- An unavoidably increased latency as a result of multi-hopping. This will also affect QoS, see also section 6.1.3

Drawing on the above and the earlier results regarding availability and mobility in section 6.1.3, the overall conclusion is that the multi-hop technique is useful in certain circumstances and would best be implemented by including fixed relay nodes.

6.3.2 WLAN hotspot extension

This is very similar to multi-hop cellular in its basic aim. However there are two potential reductions in service constraints, which may make WiFi hotspot extension more easily deployable than multi-hop cellular. These are:

1. A perceived lesser requirement for the prime quality of service parameters of delay and delay variation. This is an assumption based on the present traffic types for WLAN having a greater proportion of applications which place only elastic demands on the network (e.g. email, web browsing). Such an assumption may not hold well into the future, as cell phones become more like laptops and laptops begin to be used for VOIP etc.).
2. The perceived greater amenability of WLAN users to follow directed quality of service - where the network directs them to alter their physical location to somewhere better suited (see earlier, section 6.1.4.1).

6.3.3 Community networking

Where a remote community has no broadband connection, installing just one connection and sharing it can be the most effective solution. The user nodes are meshed together without infrastructure and one mesh node is connected to the broadband backhaul.

Locust World is a good example of shared ADSL services in community networks: Remote communities can use the Mesh to share a single expensive internet link, like a satellite or leased line, among enough users to make the service affordable. A T1 or Satellite connection is often out of reach of individual small businesses and personal users, but if there is enough local interest then their purchasing power can tip the balance and help to provide excellent value within the local community. Such Locust World based systems are appearing because traditional cost and service inhibitors are controlled by a 'community'. This is a good example of achieving new spectrum utilisation by recognising that certain deployments have very different 'business cases' than is typical. Spectrum efficiency and other engineering measures of mesh networks are of relatively little importance to these cases. The potential capacity of the radio network is so much greater than the ADSL limit at the gateway Access Points that it is not an issue for network design.

6.3.4 Home and office indoor networking

This is essentially similar to WLAN hotspot extension with respect to the requirements for mobility, lack of infrastructure, and the need for high bandwidth. But additionally the naturally closed user group transmissions will play to the best attributes of a mesh (see section 6.1.2 which discusses potential new mesh based services).

More and more devices in the home are expected to gain a wireless data communications capability through Bluetooth, WiFi, ZigBee or UWB. Whether they can communicate adequately to control point directly or will be part of an ad hoc or mesh arrangement is unclear. As propagation within buildings is quite variable the use of mesh approach to relay signals seems a good option. Although some devices are mobile and self powered, the fixed mains powered devices are best for the relay nodes.

There is a new activity in IEEE802.11 aimed at meshing with an indoor, ad hoc focus.

6.3.5 Zero- or low-infrastructure environments

The lack of infrastructure of a pure mesh leads to very quick commissioning and de-commissioning. This will be less true for a Access Mesh where access points are needed, although the Access Mesh is not intended for the type of emergency operation which requires very quick set up and tear down. An Access Mesh will still be quicker to install than any new wired system or PMP, so still has clear benefits for example in a green field business site such as an industrial park which is waiting for permanent infrastructure, or a temporary conference event where permanent infrastructure is not preferred. In each case meshing could offer both lower capital expenditure due to the lack of infrastructure installation and lower operating expenditure, since core network charges will not be incurred for any peer to peer traffic component.

6.4 Wider practical issues for deployment

There are number of problems related to mesh operation which are new and beyond those encountered with PMP or cellular[29]. These include:

- Initial roll-out
- Battery life
- System reliance on node behaviour (see earlier section 6.1.4.1)
- Network management and billing
- Security and Trust

6.4.1 Initial mesh roll-out

When rolling out a new network, seed nodes may be needed [Beyer 2002]. This situation has already been seen in practice with fixed meshes [Radiant 2005] and is beginning to be seen in reports of US 802.11-based mesh deployments which use relay-only nodes mounted on street corners or intersections). The implication is that some infrastructure will be needed from day one (and likely for longer if QoS is desired since it depends on user node density/mobility and traffic, see section 6.1.4.1). These infrastructure nodes may not all need data interconnect via a wired backbone, but they will need power supplied e.g. potentially via solar cells in the right conditions. This kind of infrastructure may also alleviate selfish node problems, see 6.4.2.

6.4.2 Battery life

Battery life is a major issue. Whilst the mesh nodes are transmitting at lower powers due to reduced link losses, their duty cycle of transmission is increased due to the relaying requirements. Nodes very close to an Access Point will suffer much more than nodes at the edge of a mesh. This are would benefit from future technical work.

If a mesh user behaves selfishly by removing their node from the mesh, perhaps to conserve battery life, then the whole mesh will suffer from this loss of connectivity. This is called 'the selfish node problem' and is nothing less than a risk to network viability. User reward systems have been proposed, but are presently very inefficient with respect to network resources [Salem et al 2003].

Relay exhaustion occurs when a node's battery power fails due to use. Routing whilst taking into account the battery level of the nodes in the path appears to be well beyond any property of routing protocols under current discussion for communications networks, although it is theoretically possible[30].

6.4.3 Network management and billing

This subject is strictly outside the scope of this report, but the following list is included for consideration:

- Is there a need to identify and measure all traffic – or just charge flat rate for simplicity? Does this raise inter-operator issues? For example billing will be complicated where several meshes exist from different operators. If a subscriber uses more than one operator's mesh to reach the destination node, then questions arise as to how the revenue should be allocated amongst those operators. This will be further complicated if the operators use different billing strategies e.g. per bit vs. flat rate, vs. application-based.

---

[29] note that co-existence issues are dealt with in section 5

[30] It is a common theme of sensor networks, but these are uniquely delay tolerant networks and are not covered by this report.

- Identifying traffic and managing the user behaviour for non- selfishness is very complex, with a large overhead.
- Closed user groups could help circumvent the selfish node problem, charging and security, but the concept of closed groups is an anathema to a true ad hoc network.
- Existing mesh roll-outs have limited billing - all known examples are single private LAN based for the primary use of internet access.

6.4.4 Security and trust

Security means both control over which nodes are allowed to join a mesh, but also integrity of the message (and its declared end points) during transit.

Again, this subject is strictly outside the scope of this report, but the following bulleted list is included for consideration:

- Security issues arise when

  a new user wishes to join a trusted mesh, and when

  user traffic must transit a third party device via hopping.
- Attacks can also be denial of service (DoS) via connection overload or via targeting battery life exhaustion in portables.

In summary, the challenge of security in an ad hoc environment is large and an open research issue. There are all the usual wireless problems and more, since the whole concept of ad hoc networking is contrary to the usual security approaches of access control etc.

6.5 Summary of considerations for mobile mesh deployment

- Five likely scenarios for mesh adoption were found, which matched the benefits offered by mesh.
- User behaviour (mobility, traffic levels) directly affect mesh performance and hence the ability to offer guaranteed service levels. The addition of infrastructure is necessary to attend to this problem.
- Mesh specific, practical issues remain in the areas of roll-out, battery life depletion patterns and security.

Note that section 7 includes views on the time-scales for mesh adoption in its consideration of how to encourage innovation.

7 How to encourage innovation (Ofcom Work Item 8)

7.1 Introduction

Innovation is not just the introduction of new techniques and technologies; it is also about doing things in a new way. To encourage innovation in any field it is necessary to understand the enablers and barriers in order to see what becomes possible if some can be changed. The previous sections of this study have addressed and commented on technology aspects of mesh networks and identified the properties which make them unsuitable for certain applications yet suitable for others. Achievement of the best overall result will also require consideration of regulation, creation of standards, policies concerning harmonisation effects, etc. from the point of view of how they might encourage or discourage the complex innovation process, Figure 47.

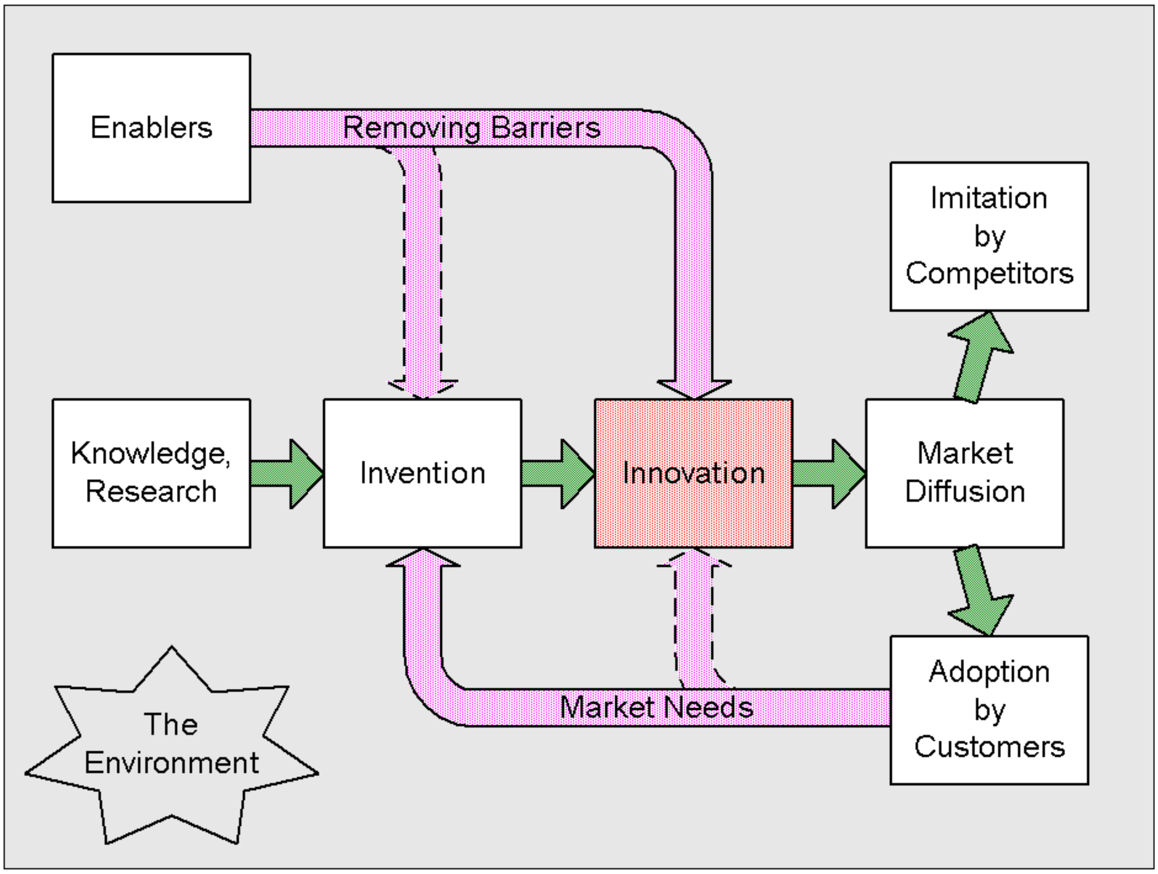


Figure 47 Innovation is part of a larger process within a given environment

Figure 47 shows that necessity is not only the 'mother of invention', but also of innovation. In other words the market drives the need process.

The discussion is structured into five categories

- A definition of the desired goals of innovation
- Consideration of enablers for innovation
- Consideration of barriers to innovation

- An example of how a start-up mesh company might get to market
- An example of how an existing cellular operator company might add mesh capability

7.2 Innovation goals - what, where and how?

It is important to be clear on the goal and the general strategy to achieve it. This report takes the goal as achieving the deployment of Access Mesh networks in particular, i.e. those with access points to external content. Moreover the significant markets are taken to be those already identified in the scenarios of section 6.3. These were:

- cellular multi-hopping
- WiFi hotspot extension
- community networking
- home and office indoor networking
- zero- or low-infrastructure environments

However many parts of this section would be applicable to other scenarios.

The strategic 'how' factor is next addressed via an evaluation of the enablers and barriers, within the specific context of our stated goal.

7.3 Enablers for innovation

Enablers are discussed under the following groupings:

- Technology advances
- Human factor aspects
- Investment and pricing
- Technology understanding

7.3.1 Technology enablers

A number of advances are required in technology areas, as follows:

7.3.1.1 High efficiency air-interface (PHY layer).

The need here is to achieve high throughput radio interface building on the broadband approaches currently taken in the IEEE802 series of Standards. Specifically, multipath-resilient techniques such as OFDM require further development and analysis for (low power) mobile implementations.

The effect of directional antennas has been considered earlier in this report. Due to the severe limitations on antenna size for portable devices the opportunity for high directionality and/or diversity are substantially less in mobile than fixed mesh networks. Nevertheless finite gains in spectral efficiency and quality of service may be possible. Substantial further work is required in this area but not limited to just the antenna technology, but more importantly its control within the network. Incorrect control of directionality can actually lead to a reduction in network performance. Certainly the MAC protocol must be very carefully defined in this respect.

A cognitive radio is one that can respond to its environment. In a mesh network, particularly a mobile mesh, adaptation of the performance of individual links can be used to trade off throughput, range, power etc such that the overall performance of the mesh is enhanced.

General improvements in radio technology will continue to benefit mesh networks. In particular techniques which increase battery life, such as higher efficiency transmit amplifiers and advances in power efficient signal processing algorithms, will particularly benefit mobile meshes.

7.3.1.2 System modelling tools

In order to develop appropriate new MAC and routing protocols access is required to more advanced system modelling tools than are currently available. Much of today's work in this field employs simulations based on the present IEE802 series of standards, due to expediency. Many researchers have ready access to simulators and equipment that use 802 protocols and thus choose to use them rather than endeavouring to develop systems optimised for meshes. To some extent this may reflect an interesting difference in mind set between industry groups: on the one hand groups such as MANET are considering meshes essentially for LANs for which mobility and ease of use are paramount virtues and spectral efficiency may be sacrificed, whilst on the other hand commercial operators with a fixed spectrum allocation will take great pains to achieve the maximum spectral efficiency and quality of service. Within these simulators are requirements for more sophisticated representations of user-mobility and traffic flows.

Currently, many simulations simply use CBR (constant bit rate) traffic. Models of actual traffic loads appear to be largely absent from the present literature. Although simulation results for e.g. routing protocols will include items such as latency and packet loss, and thereby give some idea of how multimedia traffic might fare under the protocol, this is not enough. Something is needed to stress the simulation based on the bursty nature of some traffic or the regular periods of very high speed traffic which might be expected in multimedia communications. Such an approach will stress the network as well as look for subsequent multimedia-friendly behaviour of the protocols. Further work is required here to move the modelling of performance from average throughput rates (as in the above theoretical analysis and model simulations) towards modelling of performance for actual traffic loads.

7.3.1.3 High efficiency multiple access schemes.

The need here is for robust channel acquisition and collision-avoidance mechanisms that can maximise mesh efficiency and minimise signalling overheads. We have stated earlier in this report that the present industry-focus on ad-hoc networking protocols is non-ideal for high-QoS/high-efficiency mesh networking.

7.3.1.4 Routing and transport for real time mesh applications

As for MAC schemes, the present industry focus on WLAN type ad hoc networks will not necessarily help solve issues for real-time traffic.

7.3.1.5 Mobility modelling processes

Little is really known about how nodes may move within a mesh, although one likely example is clustering around distributed hotspots. Thus nodes will move as a group around a hotspot. Where there is more than one hotspot, nodes will cluster generally, but individual nodes will move between the two (or more) hotspots. Although Camp et al [2002] (discussed in Appendix D) do cover group mobility models, no model exists for the above example. This is because prior models have assumed the military or emergency services scenarios, where nodes do not swap between hotspots (indeed there are no hotspots), rather they stay grouped in some form and the whole group moves, perhaps amongst other fixed membership groups.

A new mobility model for mesh networks is a ripe area for future work. If real world motion traces of node movements become available, they could be used directly or used to adapt models which are computationally attractive whilst still being representative.

7.3.1.6 Propagation

For input to these models one requires better representation of propagation characteristics. Much of the current work (especially in academia) employs mathematical representations of mean path loss and log-normal fading. But so much of this is founded on earlier path loss measurements for PMP networks in which the base station is elevated, whereas more real data is required for low antenna height peer-to-peer communications[31] –with attendant higher mean path loss and log-normal standard deviation clutter loss.

Exploiting spatial variations in propagation is also of interest. This includes diversity, MIMO, steerable antennas etc. In addition such information is required at higher frequency bands than current cellular systems if one is to open up the higher frequency bands (e.g. circa 6GHz) to mesh networking (see section 4.5).

7.3.2 Human factor aspects

Communication systems are used by people who all have expectations. Mesh may not satisfy those expectations due to its technical limitations. There are two important mobile mesh specific considerations:

7.3.2.1 Selfish node behaviour (battery life)

A mesh specific problem is the power consumption of a battery powered device. Because mesh nodes are required to relay traffic which may not be their own, especially when close to an AP, more demands are made on battery performance. A situation may occur where a user who is not generating his own traffic may nonetheless experience high battery drain due to relay. A further problem could arise depending on the user's response to such a situation. If he turns off his node when he is not using the device, then he will compromise the mesh for all other users. Clearly some provision needs to be made to encourage good group behaviour from individual users.

User-concerns in the areas of security, trust, reliability, billing and management may also impact user behaviour.

General improvements in radio technology will continue to benefit mesh networks. In particular techniques which increase battery life such as higher efficiency PAs and more efficient modulation methods (energy/bit) will particularly benefit mobile meshes.

A different example comes from the community network (section 6.3.3). This has been a feasible proposition because several barriers to mesh implementation and operation are overcome. Selfish behaviour is less likely thus ensuring critical relays are present, individual traffic demands are moderated for the good of the community, costs are reduced by voluntary efforts, free site rentals, etc.

7.3.2.2 Quality of service expectations

In the UK, the major cellular operators all use GSM and UMTS for their 2nd and 3rd generation mobile technology. There is continual speculation in the media that there will be a profusion of

[31] a future, proposed Ofcom study

multi-mode handsets which will add WiFi access to their service offering. Their technical capability to do this is clear and with T-Mobile already offering a separate extensive WiFi hot spot service the business aspects of the differing services are understood.

However in talking to these operators it became clear that there were worries about combining services which offered very different performance guarantees. Some operators believe that they have established a particular service expectation for their ‘brand’. So for a cellular operator this means that there are levels of service in terms of coverage in its extent and strength and completeness in covered areas, availability, dropped calls (including calls with handovers), consistency, etc. The list of parameters continues to grow with data and IP based calls such that delay, latency, throughput need to be at or above the ‘brand’ level on a regular basis. The operator with a strong brand is also keen on a vertical business such that every part of the service delivery and content is under his control. This includes the ‘walled garden’ approach to services rather than allowing customers unfettered access to third party material.

On the other hand the WiFi hotspot service provider offers a service some aspects of which attract lower expectations but yet are still acceptable to their users. This arises from previous experience of the Internet as a ‘best efforts’ service and the assumed coverage area of a WiFi radio hotspot being conservative.

With these ‘mindsets’, how likely is it that mesh networks will be considered as a suitable technology for delivering part of their service? Clearly if mobile mesh techniques are to form part of a cellular network the operator must ensure that customers’ expectations are not disappointed by variability or lower than expected performance levels. It is sometimes considered better not to offer coverage rather than patchy coverage and when rolling out a network sites are only introduced when they can offer contiguous coverage. On the other hand when an operator is struggling for coverage in a notionally ‘covered’ area (e.g. indoors) any means of improving the signal strength would be welcome.

The service which is expected by a user is in line with that which the operator hopes his brand is offering, as explained in the previous section. Depending on the service type this is either to ‘Telco quality’ or is very much ‘best efforts’. However these are not set in stone and over time users are becoming less tolerant to service imperfections. On the other hand, increasing voice over IP experience may challenge these expectations.

Many customers of mobile cellular operators buy terminals from them. In the UK these are subsidised but lower prices require them to make more commitment to the network through contracts etc. WiFi users have generally bought their own unsubsidised terminals as laptop computers or PDAs with the WiFi function being included so that it has no impact on the cost of them acquiring a radio terminal. It may be that this cost model also affects user expectations and behaviour. The advantage cellular operators have is that they have more control over terminals and can favour devices which are capable of mesh operation.

7.3.3 Investment and pricing

The investment of public finances offers a low cost, low risk enabler to an innovative company. Some risk and cost are borne by the government. The DTI make digital strategy awards as part of Government policy to overcome the ‘digital divide’. Radio could be included in such awards but the policy is not directed at specific technology solutions. Clearly any investment would assist the roll out of a mesh network.

With respect to innovative spectrum pricing, this could be a mechanism by which a regulator might incentivise users to move to mesh networks. This could be seen as compensation for all mesh nodes always being needed in the role co-operative relay nodes for all traffic. Hence this tackles the selfish user problem of section 7.3.2.1 at a higher level.

7.3.4 Technology understanding - operator exploitation

The rate and extent of technology understanding is very relevant. For example work towards this report found that understanding of mesh technology and application performance varied between operators[32]

This highlighted that there is little consistency in what operators understand by the terms mesh networks, ad hoc networks, MANET etc. To some the terms are almost interchangeable, whilst to others they represent subsets and variations of one another. There are also claims which are made as general attributes of mesh networks which only apply in more limited cases e.g. when directional antennas are used.

Furthermore the question of what is a mobile mesh network again elicited a quite varied response from the industry. What is meant by a mobile network is clear to those who take commercial cellular networks as the benchmark, but even if there is a lessening in the degree of device portability, the permitted speed at which terminals can move, limitations in handovers, etc,, many users will still consider it to be a mobile service. For example, Centrino is positioned in the marketplace as a 'mobile' technology for notebooks and laptops, but in communication terms it is simply a wireless LAN offering portability and no automatic handover. In the limit, if any devices in a network are self-contained and portable, then the term mobile might be used by some people.

The most likely reasons for an operator to use mesh technology will be if it reduces costs or speeds up roll out. The traditional approach of building infrastructure is to install more and more base stations each with their own backhaul transmission requirements. Because a mesh network can extend coverage and enhance coverage quality through relaying and re-broadcasting through a mesh, coverage can be extended without the need for acquiring new separate transmission circuits, see section 7.4.1.3. Backhaul transmission can be a major cost an operator and a time consuming part of the network build. In the case of picocells, the backhaul is a disproportionate part of the total cost.

7.4 Removing barriers to innovation

7.4.1 Regulatory

Regulations may of course be an enabler of innovation, but here we consider regulations from the perspective of removing barriers.

7.4.1.1 The model

That command and control is slow and market forces are faster is well accepted. However, like all processes of change, the transition could be a bumpy ride for all concerned, and may be a reason why time-scales are found to be slower than expected.

Ofcom's policy on spectrum management is based on the 'Cave Report' [Cave 2002] as developed further in the Spectrum Framework Review [Ofcom 2004, 2005a]. Ofcom has changed its approach from 'Command and Control' to the use of 'Market Mechanisms' plus some license exempt use, primarily for short range applications. With the 'Market Mechanisms' approach, Ofcom is still responsible for partitioning the spectrum and making it available, an auction being

[32] All UK operators were informally interviewed during this project.

the preferred route, ownership and use is then driven by market forces. Secondary trading is expected and spectrum should be acquired by those who value it most highly. Over time spectrum blocks may be sub-divided or aggregated.

Definition of the liberalisation process

Trading and liberalisation raises issues to do with radio interference. Owners of spectrum need to know what are their ' spectrum usage rights' as this translates into the Quality of Service that they can provide. The issues are:

- Spectrum quality
- Levels of real interference
- Perceived interference
- Protection / adjacent channel ratios
- The affects on interference of new RF structure & usage levels

This leads to the need for carefully defined spectrum usage rights

Clarity of rights and guarantees of rights

The concept of technology-neutral spectrum usage rights has been introduced. The current usage of an allocation is referred to as specific spectrum usage rights. Further restrictive spectrum usage rights are permitted provided that neighbouring users do not suffer increased interference, but these rights are probably too restrictive to allow efficient use. Hence it is proposed that there are new specific spectrum usage rights by agreement with neighbours possibly involving some compensation, Ofcom [2004].

However getting an effective process is still an ongoing issue and Ofcom have commissioned a study[33] to define further the detailed options in implementation of technology-neutral spectrum usage rights. Until this and a subsequent consultation have been completed it is not fully clear that the addition of mesh techniques to an existing approved service technology is not an issue. However Annex H concluded that serving an area by greater densities of base stations was not considered to be an issue. Hence it would seem that using a mesh approach would be an acceptable change within the existing spectrum usage rights of a licensee.

7.4.1.2 Time-scales

The currently used mechanisms for spectrum management are a contributing factor to the long lead times from innovation to market in wireless technologies and systems. Innovation is stifled by traditional spectrum management methods due to time scale for changes to happen being very many years. Unlicensed WLAN bands have improved this situation, but do not offer a quality, efficient option.

Ofcom is introducing spectrum trading & liberalisation in a phased manner. To date there has been little trading and the liberalisation process is still being developed.

The current limitations are due to:

- Insufficient information for traders
- International barriers
    - Harmonisation

[33] by Aegis, Indepen and Transfinite

- Standards - both *de facto* and *de jure*
- Anti-competitive practices
  - It is hoped that competition law is adequate
  - Legacy holdings can obstruct change
- Legacy usage difficult to change
- Paired bands sometimes needed

7.4.1.3 Access/backhaul

Telecoms operators generally treat access and backhaul requirements separately. However because mobile-mesh networks are a means of providing extended coverage areas from a limited number of gateways or Access Points, they provide networks which are providing both the mobile access and a backhaul function. This is an interesting point from a radio spectrum perspective as the mesh network is replacing some point to point or point to multipoint systems which would traditionally might used higher frequency microwave bands. The following section addresses this combined access/backhaul role.

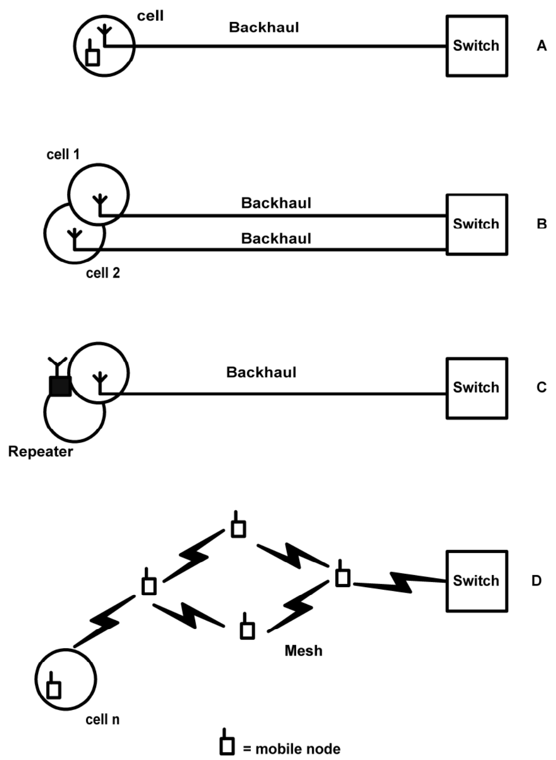


Figure 48 Four different ways of arranging backhaul

In Figure 48 (A), the base station (BS) creates a coverage area, or cell, in which a mobile station (MS) can communicate with the BS. The backhaul to the switch uses fixed infrastructure which may be line or radio (and is generally microwave point to point).

In Figure 48 (B) a second cell has been added to extend coverage, again this cell requires backhaul to the switch.

Figure 48 (C) provides the same coverage by using a repeater but the backhaul from this cell now effectively includes a radio link to the first BS using spectrum from the mobile access spectrum resource. With a mobile mesh network, the concept of RF backhaul extends further, see Figure 48 (D).

This is of interest when the practical construction of radio networks is considered. The provision of backhaul to cell sites is a major factor in roll-out with significant implications in terms of cost, time-scale, site selection and site approval. Indeed the backhaul issues can be dominant in all these aspects. Sites with an existing point of presence to high bit rate bearers may be favoured rather than those providing the most appropriate access network coverage.

Of increasing interest is picocell deployment where due to the cells being of low capacity, the costs of backhaul can be a much larger part of the total cost than is the case for macrocells. Various point-to-multipoint schemes have been proposed as a solution but mesh networks are a clear competitor in such cases.

In transferring some of the backhaul function to the mesh, some of the potential access network capacity is being used instead for the backhaul role. The alternative routing options in the mesh also add to backhaul resilience, the equivalent separate backhaul network might use a ring structure or alternative bearers to protect links but the solution would vary depending on how much traffic was dependent on a particular path. In the picocell case nothing more than a simple star network would be likely.

7.4.1.4 WT act issues

7.4.1.4.1 Elevated relay operation

When used in a mesh network, radios are effectively acting as both infrastructure (relays) and customer equipment (access). For a TDD system there is no physical differences to the transmissions from the radio whether it is at the end or part of a chain of radio transmissions. The only consideration might be that if the radio were regarded as part of the operator's infrastructure its use would be subject to any planning approvals for the site from which it transmits. This should not be an issue from the point of view of power level but the regulations[34] [5] say

> ''5.1 The full site clearance procedure is triggered when a radio installation satisfies any of the following three conditions:
>
> • the radio equipment is capable of transmitting power levels greater than or equal to 17 dBW ERP (Effective Radiated Power), unless specifically mentioned below
>
> or
>
> • the maximum height of the antenna or its supporting structure exceeds 30 metres above ground level (AGL),
>
> or
>
> • the new installation increases the height of an existing (site cleared) structure by 5 m or more.
>
> 5.2 Systems operating below 17 dBW ERP do not need to follow the full site clearance procedure provided they do not breach the height criteria. Operators are nevertheless encouraged to notify the secretariat of such systems for entry onto the sites database, to help in evaluating prospective sites and addressing mast height safety concerns. Such notifications are not subject to approval. ''

So use as a relay from an elevated position is questionable despite the fact that the device could transmit legitimately from the same location as an end terminal.

[34] for fixed nodes

However, all the 802 standards (and DECT, PHP, etc.) have established the case for terminals operating on common frequencies in both directions, so in these cases the only issue may be that the spectrum needs to be assigned on this basis.

7.4.1.4.2 uplink/downlink

For an FDD system there is a further issue in that the operator has a licence to transmit on the down link frequencies and the customer equipment is approved for transmissions on the uplink frequency. When acting as a relay, customer equipment then requires some licence extension or some means by which it is considered to be being operated by someone as the agent and under the control of the operator, fully in line with the provisions of the appropriate Licence under the WT Act.

7.4.1.4.3 Fixed/mobile

Although it is Ofcom's intention to liberalise the use of spectrum there is a legacy position which inhibits this process. This has been partly recognised by deferring the start of trading in the mobile bands. It is also of note that the auction for the 3.4GHz band in the UK (now operated by UK Broadband) was for a nation-wide licence but only for a non-mobile service. The market they have addressed is Fixed Wireless Access (FWA) but the equipment is to the 3G UMTS TDD standard. This is a fully mobile standard even though it is being used for the fixed market only.

What may be regarded as a grey area is the portable market. Some forms of Fixed Wireless Access (FWA) require the customer equipment (CPE) to be specially sited often with outdoor antennas which are directional and mounted in elevated positions. In such cases professional installation is generally used. However if the signal strength is high enough it is possible to enable a connection to be established with a small self-installed wireless modem. The impact of this is that a portable service becomes possible as connection can be established anywhere within the coverage area with such a wireless modem. If handover were allowed it would have to be considered as a mobile service.

On this basis the service is of a form between full cellular radio and WiFi (with hotspots). It allows for portability rather than mobility, although if the technology is a mobile standard such as 3G UMTS TDD, the CPE may work even if it is moving rapidly.

Regulation has stopped cellular operators from offering a fixed bypass service and limited the degree of mobility/portability from non-mobile operators.

The mesh networking scenario we see as having most promise is the Access Mesh, in which the mobile user-node density may be supplemented by additional fixed relay nodes.

At the regulatory level, it is necessary to understand whether there are restrictions on fixed access on such a mobile network and how portability is to be regarded. For example if a fixed relay node is added to a mobile network, its secondary use as an access point (fixed) might currently be illegal.

7.4.1.5 Availability of non-operational licenses

Non-operational licences are better than trials in unlicensed spectrum if real intention is licensed spectrum. Ofcom should continue its policy of allowing trials through Non-operational Development Licences (previously T&D Licences).

7.4.2 Standards

7.4.2.1 Interoperability

Standards are needed for interoperability, but as noted in section 2.2.6, interest and progress is disappointing. Only IEEE802.16e (mobile WiMax) seems to have a clear work item concerned with mesh networking. Even within this community there was some concern that adding a mesh capability might undermine QoS claims for the standard.

7.4.2.2 Harmonisation

Although harmonisation requires standards, standardisation is also worthwhile in non-harmonised markets. Standardisation helps prevent excessive fragmentation and variants in candidate technologies but permitting more than one solution can lead to competition benefits which complete harmonisation might stifle.

To encourage a market in spectrum, the intention of Ofcom is to place as few constraints as possible on any spectrum and allowing Liberalisation of its use but there are problems with moving from the current position.

Spectrum is planned at world, European and national level so the UK is bound by various agreements which allocate blocks of spectrum to particular usage. European standardisation is based on harmonisation principles and this has led to some great successes such as GSM but also many unsuccessful outcomes which have resulted in some bands being dedicated to failed initiatives. It is a slow process to change but the UK intent is to limit rigid spectrum planning to essential points only such as avoiding interference with neighbouring countries and allowing pan-national services such as satellite.

Some spectrum is licence-exempt, this accounts for 7% and Ofcom do not expect this to be increased significantly. Mesh networking, like all networks which desire to offer quality of service guarantees, are not seen as drivers for licence-exempt spectrum, by this report.

7.5 Summary and Time-scales

There seem to be few regulatory objections to mesh networks particularly with a more liberalised regime in which differences between fixed and mobile regulation are reduced.

However, the following points need to be addressed:

- eliminate remaining regulatory blocks to mobile mesh networking
- add mesh capability to relevant radio standards
- ensure sufficient and appropriate spectrum is available
- allow trials through Non-operational Development Licences (previously T&D Licences).

In addition it may be necessary to find means to modify user behaviour if this acts as a deterrent to successful mesh network operation. Such a task is very difficult to quantify.

Table 7 attempts to summarises the relative time scales for the aspects of innovation as discussed in this section.

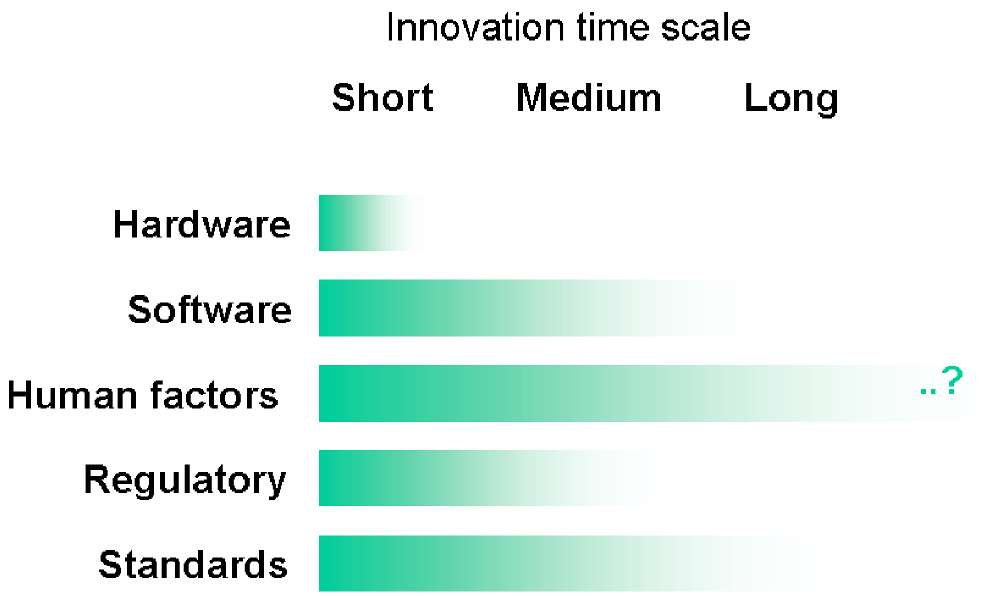


Table 7 Relative time scales for the components of innovation

8 Conclusions

The conclusions of this report are as follows – specifically in the context of mobile mesh networks operating below 3.5 GHz:

1. Mobile meshes do not self-generate capacity regardless of subscriber-base, unless fixed infrastructure is added in the form of relay nodes, a hierarchical network, and/or access points.
2. Mobile meshes are unlikely to prove more spectrally efficient than cellular networks in the general case – primarily because of the need to relay traffic through the network in multiple hops and the substantial protocol overheads required for mobility.
3. Mobile meshes will not significantly benefit from directional antennas for the case of handhelds below 3.5GHz – due to the physical limitations on antenna size vs. wavelength
4. Mobile meshes could offer improved spectrum utilisation through their potential for employing higher frequency allocations, compared to cellular requirements.
5. Access meshes are the likely form in which mobile mesh technology will be adopted – in order to support extra-mesh traffic flows to/from other networks, e.g. the Internet .
6. The routing overhead for mobile mesh networks can be very high. For example, as many routing control packets as user data packets might be needed to cope with mobility. This would force a halving of overall system efficiency.
7. The degree of mobility able to be supported by a mesh is a current research topic. Mesh performance depends on user behaviour in a manner quite unlike cellular. The addition of infrastructure could reduce the dependence of system performance on user mobility and traffic, thus enabling service guarantees to be made.
8. Meshes offer efficient radio coverage in certain scenarios: Hopping around corners or obstacles in cluttered environments is perhaps the greatest benefit offered by a mesh.
9. There is poor predictability associated with mobile mesh deployment due to the transitory nature of its users. Additionally, there is a critical mass needed for roll-out, some of which may need to be provided via the deployment of fixed seed nodes.
10. User expectations in terms of security, multimedia performance and battery life will need to be assessed and carefully balanced with what a mobile mesh could provide. This could have an important, but hard-to-predict influence on time scales for mobile mesh adoption.
11. In general it is not the maturity of electronic hardware which is limiting mobile mesh networking. The elements of transport, routing, medium access and cross-layer protocol co-operation, plus modelling software are notably less mature.
12. Upgrading a mesh for higher performance requires each and every node to be upgraded; it cannot be done piecemeal as it can for cellular. In this respect the performance and capability of user terminals must be sufficiently well scoped to cover the service life of the network.
13. The total capacity of any mesh radio node must always exceed the capacity presented to the node user, in order to allow for the relay overhead which is intrinsic to mesh operation.
14. More likely adoption scenarios are suggested to be:
    - cellular multi-hopping

- WiFi hotspot extension
- community networking
- home and office indoor networking
- zero- or low-infrastructure environments

15. Mobile mesh networking, other than for personal area networks, is not seen as a driver for licence-exempt spectrum, on quality of service grounds.
16. Mobile meshes do have a role as part of a future integrated mobile network deployment.

9 Suggestions for future work

Suggestions for future work have been made throughout the body of the report. Table 8 summarises the main areas raised and the corresponding section of the report.

Future work item	See report section
'Novel' modulation, e.g. UWB	4.2
Link adaptation, cross layer metrics	4.3
Utilisation of higher frequencies for meshing	4.5
Multi-hop access meshes for real time applications	5.3, 6.1
Quality of service provisioning within a mesh architecture	6.1
Mobility models	6.1
Security for mobile, ad hoc networks	6.3
Power consumption issues for mobile mesh nodes	6.4
User expectations and behaviour vs. mesh capabilities	7.3
Technology factors	7.3

Table 8 Suggestions for future work

Appendix A Definitions

access mesh	a development of a pure, isolated mesh to include connection to other networks such as the internet via access points. Access meshes are expected to be the dominant form of mesh deployed for public use. The density of Access Points has a large effect on mesh properties.
access points	see access mesh
allocation	relates to allocating spectrum for a service type and is usually internationally agreed, cf. assignment
arbitrary network	as defined by Gupta and Kumar [2000], see also random network
assignment	relates to assigning a pre-existing allocation to an operator and is usually nationally agreed, cf. allocation
hop	a point to point connection, with no intermediate nodes
link	a point to point connection
network capacity	the average total traffic rate (bps) that can circulate within the network. It is the aggregate of the *per-user throughput*
node transport capability	the average rate at which data can be passed through a node (from Rx to Tx)
path	equivalent to *route*
per-user average throughput	The average data rate that can be sourced/sinked from/to the user of a node
per-user traffic throughput	the data rate and associated Erlang loading that can be sourced/sinked from/to the user of a node.
pure mesh	an isolated mesh
random network	as defined by Gupta and Kumar [2000], see also arbitrary network
raw transmission rate	the raw on-air bit-rate (bps) over a radio *link*
relay nodes	mesh internal infrastructure connection points
route	a point to point connection description via intermediate network nodes
seed node	a node or nodes deployed on day one to ensure connectivity via guaranteeing a minimum density of nodes
spectrum management models	There are three main spectrum management models: • Command and control. Licensed e.g. cellular • Commons. Unlicensed, e.g. WLAN • Market. Secondary Trading. Flexible, 'change of use' allowed, probably via the regulator.
transmission bandwidth	the spectrum bandwidth (Hz) required to accommodate the *raw transmission rate*

Appendix B Glossary and Abbreviations

3G	3^{rd} generation mobile, generic
AAA	access, authentication and authorisation (server)
ADSL	asymmetric digital subscriber line
AODV	adaptive, on-demand distance vector (routing protocol)
APC	automatic power control
APC	automatic power control
BER	bit error rate
CBR	constant bit rate
CDMA	code division multiple access
CPE	customer premises equipment
CSMA/CA	carrier sense multiple access/collision avoidance
CSMA/CD	carrier sense multiple access/collision detection
DFS	dynamic frequency selection
DSR	dynamic source routing (routing protocol)
DTI	department of trade and industry (UK)
ERP	effective radiated power
FDD	frequency division duplex
FDMA	frequency division multiple access
FHSS	frequency hopping spread spectrum
FWA	fixed wireless access
GPS	global positioning system
GSM	2^{nd} generation mobile (Europe etc, not North America)
IEEE	institute of electrical and electronic engineers (publish IEEE802.x etc)
IETF	internet engineering task force (responsible for TCP/IP, mobileIP etc)
ISP	internet service provider
LoS	line of sight (for RF path)
MAC	medium access control
MANET	mobile, ad hoc networking (also MANet, a specific IETF group)
MIMO	multiple in, multiple out (an RF diversity technique)
MUD	multi-user detection (CDMA systems)
OFDM	orthogonal frequency division multiplexing
PAN	personal area network

PDA	personal digital assistant
PHY	physical layer
PMP	point to multi-point
QoS	quality of service
RADIUS	remote access dial in user service
RF	radio frequency
RTS/CTS	ready/clear to send (handshaking protocol)
Rx	receive, receiver
SIR	signal to interference ratio
TCP/IP	transmission control protocol/internet protocol
TDD	time division duplex
TDMA	time division multiple access
TPC	transmit power control
Tx	transmit, transmitter
UMTS	universal mobile telecommunication service
UWB	ultra wide band
VBR	variable bit rate
VOIP	voice over internet protocol
WLAN	wireless local area network
ZigBee	a personal area network, based on IEEE802.15

Appendix C References

Aegis Systems Ltd 2000 "*Examination of issues related to spectrum efficiency of point-to-multipoint and mesh multimedia wireless system architectures proposed for 40.5-43.5 GHz*" , Radiocommunications Agency Report Ref. 1205/AE/MWS2/R/3, April 2000.

Arpacioglu and Zygmunt 2004 Arpacioglu O, Zygmunt J.H : “On the Scalability and Capacity of Wireless Networks with Omnidirectional Antennas”. *Third International Symposium on Information Processing in Sensor Networks* (IPSN 2004), Berkeley, California (April 27–28, 2004)

Basagni et al 2004 Basagni S, Conti M, Giordano S and Stojmenovic I (Eds.), “*Mobile Ad Hoc Networking*”, IEEE Press, Wiley, 2004.

Bensal and Liu 2003 Bansal N and Liu Z,. “Capacity, delay and mobility in wireless ad-hoc networks”, *Proc. of IEEE Infocom*, March 2003

Beyer 2002 Beyer D, “Wireless Mesh Networks for Residential Broadband”, *National Wireless Engineering Conference*, San Diego, Nov 2002.

see also:

Beyer D, van Waes N, Eklund C, “*Tutorial: 802.16 MAC Layer Mesh Extensions Overview*” IEEE802.16 contribution, 2002, available via <http://www.ieee802.org/16/tga/contrib/S80216a-02_30.pdf>.

Boppana and Zheng 2005 R. V. Boppana and Z. Zheng, "Designing Ad Hoc Networks with Limited Infrastructure Support," *IEEE Consumer Communication and Networking Conference (CCNC)*, January 2005.

Boukerche and Bononi 2004 Boukerche, A and Bononi, L, “Simulation and Modelling of Wireless, Mobile and Ad Hoc Networks”, in Basagni et al (Ed.), “*Mobile Ad Hoc Networking*”, IEEE Press, Wiley, 2004.

Camp et al 2002 Camp T, Boleng J, and Davies V, “A Survey of Mobility Models for Ad Hoc Network Research” *Wireless Communication & Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, vol. 2, no. 5, pp. 483-502, 2002.

Cave 2002 Cave M, *“Review of Radio Spectrum Management - An independent review for Department of Trade and Industry and HM Treasury”*, 6 March 2002. Available via http://www.ofcom.org.uk.

Das et al 2000 Das S, Perkins C, and Royer E. “Performance comparison of two on-demand routing protocols for ad hoc networks” *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, March 2000. Tel Aviv, Israel

Ephridemes 2002	Ephremides T, “A Wireless Link Perspective in Mobile Networking", *ACM Mobicom 2002* keynote speech, available at <http://www.acm.org/sigmobile/mobicom/2002/program/>
ESA 1999	EUROPEAN SPACE AGENCY, *“Interference Susceptibility of Selected Bandwidth-Efficient Modulation Schemes (SFCG Action Item AI 18-17/2, Tasks 17/2-1 to 17/2-5)”*, 7-15 September, 1999.
Esseling et al 2002	Esseling N, Weiss E, Kramling A, Zirwas W, “A Multi Hop Concept for HiperLAN/2: Capacity and Interference”, *Proc. European Wireless 2002*, Vol. 1, p.p. 1-7, Florence, Italy, 02/2002.
Golmie et al 2003	Golmie, N, Rebala,, O, Chevrollier, N, “Bluetooth Adaptive Frequency Hopping and Scheduling”, *MILCOM 2003 - IEEE Military Communications Conference*, no. 1, pp 1138-1142, Oct 2003.
Grossglauser and Tse 2001	Grossglauser M, Tse D, “Mobility increases the capacity of ad hoc networks” *Proceedings of IEEE INFOCOM '01*, Anchorage, Alaska, April 2001.
Gupta and Kumar 2000	Gupta P and Kumar P R, "The capacity of wireless networks", *IEEE Trans. Information Theory*, vol 46, pp. 388-404, Mar. 2000.
Gupta et al 2001	Gupta P, Gray R, and Kumar P R, “*An experimental scaling law for ad hoc networks*” Technical report, University of Illinois at Urbana Champaign, available at <http://decision.csl.uiuc.edu/#prkumar>, May 2001.
Hekmat and Mieghem 2004	Hekmat R, Mieghem P, "Interference in Wireless Multi-hop Ad-hoc Networks and its Effect on Network Capacity," *Wireless Networks*, vol. 10 (4), pp. 389-399. Jul. 2004
Hsieh and Sivakumar 2001	Hsieh H-Y and Sivakumar R, “Performance Comparison of Cellular and Multi-hop Wireless Networks: A Quantitative Study”, *Proceedings of ACM SIGMETRICS*, June 2001.
Huang et al 2002	Huang Z, Shen C, Srisathapornphat C, and Jaikaeo C, “ “Topology Control for Ad hoc Networks with Directional Antennas” *Proceedings of 11th International Conference on Computer Communications and Networks (ICCCN 2002)*, Miami, Florida, October 14-16 2002.
Jones et al 2003	Jones S, Levine P and Rickman N, “*The Economics of Radio Spectrum*”, Ofcom report, 2003.
Jun and Sichitiu 2003	Jun J and Sichitiu M L, "The Nominal Capacity of Wireless Mesh Networks," *IEEE Wireless Communications*, Vol. 10, pp. 8-14, Oct. 2003.
Li et al 2001	Li J, Blake C, Couto D D, Lee H and Morris R, "Capacity of ad hoc wireless networks", *Proc. 7th ACM Int. Conf. Mobile Comput. Netw. (MobiCom'01)*, 2001, pp. 61-69.
Li et al 2002	Li L, Sarca O, Selea R, *“A solution/scheme for coexistence*

with 802.11 in the LE bands", submission IEEE C802.16a-02/74, IEEE 802.16 Broadband Wireless Access Working Group, July 02, 2002.

Liu et al 2003 Liu B, Liu Z, Towsley D: "On the Capacity of Hybrid Wireless Networks" *IEEE InfoCom* 2003.

Lungaro 2003 Lungaro P, *"Coverage and capacity in hybrid multi-hop ad hoc cellular"*, Master of Science Thesis, TRITA-S3-RST-0315, Radio Dept. KTH, Stockholm, 2003.

Lungaro and Wallin 2003 Lungaro P, and Wallin E, "Coverage, Capacity and QoS Tradeoffs in Hybrid Multi-Hop Ad Hoc Cellular Access Systems", *Proceedings of the Affordable Wireless Services and Infrastructure, 1st Annual Workshop*, Tammsvik, Bro, Sweden, 3-4 June 2003.

Martin et al 1999 Martin, W L, Yan, T-Y, Gray, A, Lee, D K, *"CCSDS – SFCG; EFFICIENT MODULATION METHODS STUDY AT NASA/JPL; PHASE 4: INTERFERENCE SUSCEPTIBILITY"*, NASA/JPL, 12 October 1999.

McCune 2000 Earl McCune, "DSSS vs. FHSS narrowband interference performance issues", *RF Design*, September 2000.

MeshDynamics 2005 "*High Performance Mesh for Dense Metro WiFi* ", MeshDynamics white paper available at <http://www.meshdynamics.com/wPresentations.html>

Negi and Rajeswaren 2004 Negi R, Rajeswaren A, "Capacity of power constrained ad-hoc networks," *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

Nilsson 2004 Nilsson A, "Performance analysis of traffic load and node density in ad hoc networks", *Proc. European Wireless (EW2004)*, Barcelona, Spain, February 24-27, 2004.

Ofcom 2004 *Spectrum Framework Review*, Ofcom, 23 November 2004.

Ofcom 2005a *Spectrum Framework Review Statement*, Ofcom, 28 June 2005

Ofcom 2005b *UK Radio Site Clearance Procedure*, OfW 191 Version: 1, Ofcom, January 2005

Plextek 2005 Plextek project Ackerman for Ofcom, "*Development of Smart Antenna Technology, WP3 = Beam forming review"* Ofcom ITT reference C31400/004, 2005

Radiant 2005 "*How does **MESH**WORKS™ deliver connectivity that current technologies do not?*", available via <http://www.radiantnetworks.com/meshworks/how.asp>

Ramanathan 2001 Ramanathan R, "On the performance of ad hoc networks with beamforming antennas", *ACM MobiHoc*, Long Beach, CA, Oct. 2001.

Royer et al 2001 Royer E M, Melliar-Smith P M, Moser L E, "An analysis of the optimum node density for ad hoc mobile networks", *Proc. IEEE Comms*, June 2001.

Salem et al 2003	Salem N, Buttyan L, Hubaux J, Jakobsson, M. ”A Charging and Rewarding Scheme for Packet Forwarding” *ACM MobiHoc*, Annapolis, Maryland, USA, 2003.
Sanzgiri and Royer 2005	Sanzgiri K, Belding-Royer E, "Leveraging Mobility to Improve Quality of Service in Mobile Networks", *Proc. MobiQuitous 2004*, Boston, MA, August 2004.
Sharma and Mazumdar 2004	Sharma G, Mazumdar R R, “On Achievable Delay/Capacity Trade-offs in Mobile Ad Hoc Networks,” *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Cambridge, UK, March 2004.
Shepard 1996	Shepard T , A Channel Access Scheme for Large Dense Packet Radio Networks”, *Proceedings of ACM SIGCOMM'96*, August 1996, Stanford University, California.
Spyropoulos and Raghavendra 2003	Spyropoulos A, Raghavendra C S, “Capacity bounds for ad-hoc networks using directional antennas”, *Proc. IEEE 2003 International Conference on Communications (ICC)*, May 2003.
Yi et al 2003	Yi S, Pei Y, Kalyanaraman S, “On the capacity improvement of ad hoc wireless networks using directional antennas”, *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, Annapolis, Maryland, USA, June 01-03, 2003.
Zander 2004	Zander J, “*The Dynamic Spectrum Access project: Scenarios and Concepts*” Workshop on Dynamic Spectrum Access, Kista Science Tower, Stockholm, available via <http://www.wireless.kth.se/projects/DSA>, May 3rd, 2004.

Plextek

Appendix D Mobility models

The aim of using a mobility model is to reflect as accurately as practicable the real conditions themselves. One way to do this is to use motion traces, which are logs of real life node movements over a representative period of time. There are not many such logs available for use even with established cellular schemes and none are known to this report which cover mesh environments. The focus then must move to synthetic models. Such a model will deal with a number of nodes and may include such parameters as speed and direction of movement, the ability to pause at some locations and a bound to the model area. The models available are mostly fairly simple to implement, since they are intended for use in simulators were a tractable run time is expected. It is probably the case that present models err on the side of simplicity at the expense of realism. On the other hand, moving too close to the actual environment requires a very specific model – which may then not be adequately representative of all environments. The choice of models is thus a subject which needs to be understood, in order to interpret specific protocol and other simulation results for wider contexts.

Camp et al [2002] review 12 different mobility models which have been applied to mesh simulations at various points in the published literature. Their work is an often quoted indication that the choice of model alone can strongly affect the results when testing the exact same routing protocol. For the purposes of this report three models are noted as being appropriate:

D.1 The Random Waypoint mobility model

This model has a base of randomly distributed destinations to which any node may move. The node will move with a random speed and will pause for a random amount of time at each point, before moving on again. The model input parameter are the number of nodes, the bounds on both speed and pause time and the physical boundaries of the model area. Figure 49 shows the pattern of a single node in this model. The axes are in units of distance.

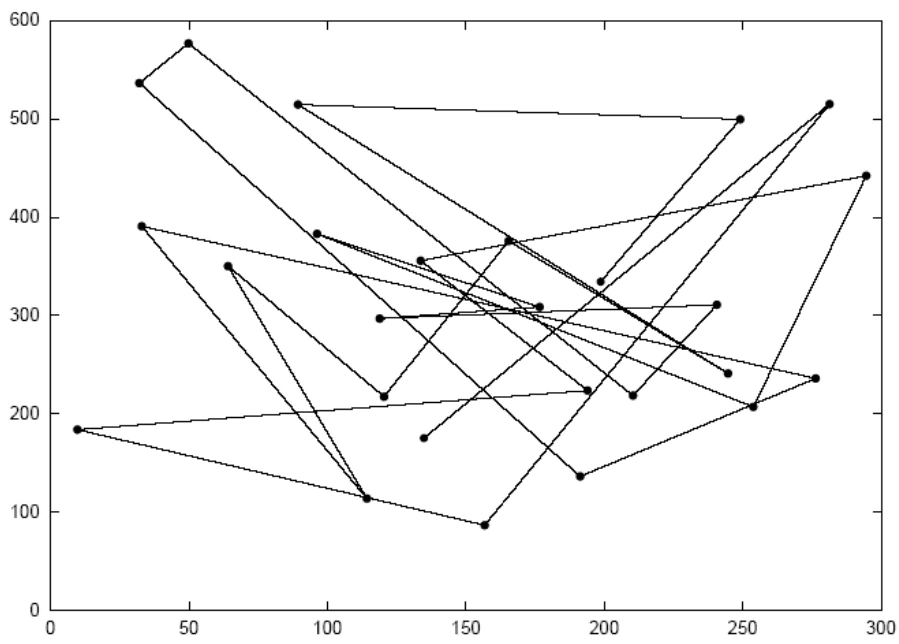


Figure 49 Random Waypoint model (from Camp et al.)

It is a matter of some current debate as to whether this model tends to concentrate the nodes at the centre of the model over time. If it does, the effect is thought to be quite light, but for the purposes of this report, this behaviour is interesting as it is reminiscent of the clustering effect which users may exhibit within a mesh, depending on the scenario. This fact will be referred to when simulation results using this model are interpreted. Like all the models, the pattern in figure 2 can be scaled by changing parameters, even to the point of having quite a static distribution.

D.2 The Random Direction mobility model

Any tendency to concentrate nodes at the centre of the model area is avoided in this model, by design. Nodes travel all the way to the boundaries of the model before setting off again in another direction, with another speed. At this level it is often called the billiard ball model. However pauses may be added, so that nodes spend time at the boundary before moving off again, in order to mimic the pauses seen in real world user behaviour, see Figure 50. Of course not all users will really pause at the boundaries of a given area. Despite this, the model with pauses is useful in that it has been found to create a multiple hop situation more often than not, precisely due to pausing at the boundaries (maximising inter-node distance). This makes the model useful for modelling multi-hop networks. It is also a tractable model to use in a simulation environment.

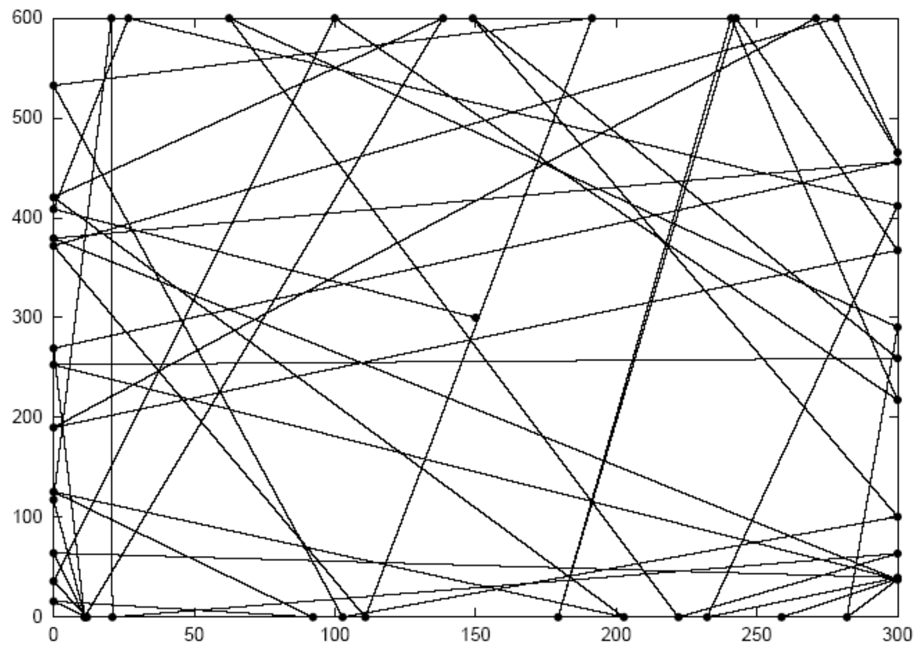


Figure 50 The Random Direction mobility model (from Camp et al.)

D.3 The City Section or Manhattan mobility model

The approach here is very direct. A dense urban environment with a regular grid of streets is assumed, hence the term Manhattan. Nodes are constrained to move within the streets and may be made to pause at street corners to improve coverage around the corners. The inputs to the model are the street length, width and spacing. Such a model was used in the assessment of the 3GPP and GSM coverage quality.

Despite the appeal of a direct analog of the real situation, this model has not seen much application in the study of mesh networks. It may be surmised that this is due to the very specific nature of the model, in other words how are the street dimensions to be picked? If a typical case is taken, how useful is that in reality, where a spread occurs? **This report suggests that such models are worthy of further investigation.**

Appendix E Modulation and multiple access

Table 9 lists the key PHY layer elements in a radio network design and describes their static/dynamic nature in common systems. It also describes the ways in which the control of the element can counter interference.

	Static	Dynamic	Interference avoidance	Interference rejection	Transmit Signal Strengthening
Carrier Frequency					
Allocation	✓		✓		
Sub-bands	✓		✓		
Hopping	✓	✓	✓		✓
Orthogonal Frequency Division Multiplexing	✓	✓	✓		✓
Signal Bandwidth					
Narrowband	✓		✓		
Wideband	✓				✓
Time Structure					
Interleaving	✓		✓		
Throughput	✓	✓			✓
Modulation	✓	✓	✓	✓	
Transmit Power		✓			✓
Geometry	✓	✓	✓	✓	✓

Table 9: Key PHY layer elements, their static/dynamic nature in common radio network designs and their use in affecting interference susceptibility.

E.1 Carrier Frequency

Allocation

The frequency bands occupied by a radio network influence the susceptibility of radio network design to both self-interference and external interference. Ideally a frequency allocation is large in the sense the total bandwidth permits many channels, isolated in the sense that is separated in frequency from potential external interference, and dedicated solely to the radio network.

However there are many other issues, not least existing allocations and legislation, which govern the allocation of frequency spectrum to a particular use. Thus there is frequently little freedom available and the design of a radio network's architecture must begin with the assumption that a particular allocation has been imposed. The allocation can be made to appear more ideal through appropriate selection of the design parameters.

Sub-bands

If the frequency allocation permits many channels, sub-bands may be defined which allow independent networks to co-exist with reduced self-interference. Consider, for example, the assignment of sub-bands within the GSM allocation to different service providers. Such a strategy works equally well for both mesh and cellular systems in reducing self-interference between networks. However it requires a high degree of co-ordination and hence is only practical if the number of networks that must co-exist is small.

Hopping

Frequency hopping can be used to reduce the impact of self-interference and external-interference when the bandwidth of the interference is less than that spanned by the frequency hopping. When used for this purpose it is one of the several techniques that are known as Spread Spectrum (Frequency Hopping Spread Spectrum – FHSS).

There are a two of ways in which frequency hopping can be used:

- Through the use of orthogonal hopping sequences. Transmitters in the radio network use hopping sequences that are guaranteed to never result in two devices using the same frequency, or frequencies close enough to result in adjacent-channel interference, at the same time.

  The GSM system uses orthogonal hopping sequences within a cell. The base station defines one or more sets of frequencies, each used by a group of mobiles. The mobiles use the frequencies in the same order but each has an allocated offset in the sequence.

- Through the use of non-orthogonal hopping sequence. The hopping sequences are designed so that the likelihood of co-channel or adjacent-channel interference is low. This allows error correction schemes to compensate for the effects of the self-interference.

  The Bluetooth system uses this approach. A pseudo-random sequence defined by the master in a piconet uses up to 79 frequency channels. The sequence used by a master depends on its unique Bluetooth Address and so independent piconets use different sequences.

Frequency hopping sequences do not need to be static. A good example is the introduction of Adaptive Frequency Hopping in version 1.2 of the Bluetooth standard specifically to reduce the effects of interference between Bluetooth and 802.11 derivatives occupying the same unlicensed band. The 79 channels used by the frequency hopping is reduced at run-time[35] if measurements show that particular channels are subject to persistent interference or fading effects [Golmie et al 2003].

Frequency hopping methods are attractive because they not only provide protection from interference effects but also from fading effects. Their use does however have a number of implications on the design of the radio network:

[35] 'run-time' means during normal operation. Provision must have been made at design time for the flexibility to allow this.

- Some form of error correction or acknowledgement/repetition protocol is required. Frequency hopping alone does not protect against interference effects. It attempts to effectively shorten the duration of interference so that the performance of error correction and message acknowledgement/repetition protocols is improved.
- There must be guard periods in the on-air time structure that allow the frequency to be changed. This reduces the system throughput and introduces phase discontinuities that may affect the demodulation performance.
- The channel characteristics at two frequencies are not necessarily the same. This is an advantage when frequency hopping is being used to mitigate fading effects but is of no benefit if the purpose is to avoid interference. Some form of overhead is typically required to allow the channel characteristics on a new frequency to be quickly estimated.
- Time synchronisation of the elements in the radio network is required so that frequency hops occur at the same time and the elements must agree on a sequence. This imposes the needs for some form of master/slave structure to the network. This might be an ad-hoc structure such as occurs in Bluetooth piconets or a formal cellular structure. Nevertheless the structure is required and for mesh networks the question of the extent[36] of the relationship must be carefully considered.

Orthogonal Frequency Division Multiplexing

This is a multi-carrier technique in which the data is transmitted on a number of sub-carriers. These sub-carriers are synchronised in time and phase so that they are orthogonal and so do not result in adjacent-channel interference in each other. The performance of each sub-carrier is therefore independent of the presence of the others.

The susceptibility of the method to interference is essentially dependent on the performance of the sub-carriers. However, since the data is distributed across the sub-carriers, the effective data-rate on each sub-carrier is a fraction of the original data rate. Thus the energy to transmit each bit is distributed across a greater period of time and the susceptibility to discontinuous interference is reduced[37]. Further, if a continuous narrowband interferer is present a relatively simple adaptation algorithm can suppress the use of the affected sub-carriers.

E.2 Signal Bandwidth

When describing the signal bandwidth the terms narrowband and wideband are used. Narrowband is used for signals with a bandwidth that is similar to or only slightly greater than the data rate. Wideband is used for signals with a bandwidth that is significantly greater than the data rate[38].

Narrowband

Narrowband signals perform relatively poorly in the presence of interference of a similar power and bandwidth. However the likelihood that a narrowband interfering signal coincides with the wanted signal can be quite low and it is relatively simple matter to use hopping techniques to avoid the interference.

In the presence of wideband interference such techniques do not work well, but there are

[36] In both the geographic and protocol senses.

[37] When compared to the same modulation scheme running at the data rate (single-carrier equivalent).

[38] Note that FSK/PSK modulation with a high modulation index, or high-order FSK, can result in a wide signal bandwidth with the signal energy concentrated at a few discrete frequencies. Despite this they should be considered as narrowband signals.

nevertheless conditions under which the use of narrowband signals can be advantageous. The fact that the power of the interference is spread across a wide band means that if the data rate is particularly low, the bandwidth of a receiver can be small enough that most of the interference signal's power is rejected. Note however the data rates required for speech are generally too great for this case and so the only real solution to overcome wideband interference is to increase the signal energy by increasing the transmit-power or reducing the throughput.

As long as the system is subjected to only narrowband interference, the advantages and disadvantages to using a narrowband signal in a mesh system are the same as those in a cellular system. The relatively poor performance with respect to interference is balanced by the relative ease with which hopping techniques can be used to avoid the interference. However, if wideband interference is likely to arise then the fact that the only real solutions are to increase the transmit-power or reduce the throughput is a problem for mesh systems, as discussed next:

Wideband

In general fact that the signal energy is distributed over a wide bandwidth means that the susceptibility to interference is low. However there are a number of types of wideband signal, each with their own characteristics:

Direct-Sequence Spread-Spectrum

Direct-Sequence Spread-Spectrum (DSSS) achieves a wide bandwidth by replacing each data bit with a sequence of chips that are modulated at a higher rate. The ratio of the chip-rate to the data-rate is the processing gain.

There are various sequences that can be used, with different properties. Generally the decision as to the type of sequence to be used is based on the multiple-access possibilities of DSSS (see section 5.2.1.1). There is little difference in the susceptibility to interference.

As discussed by McCune [2000], the de-spreading process essentially converts a narrowband interferer into a noise-like signal at the input to the decision-making elements of a receiver. An uncorrelated wideband interferer will also result in a noise-like signal. Thus the impact of both type of interferer is similar to that of white noise of the same power. This makes DSSS one of the best schemes for dealing with moderate levels of interference (as long as the processing gain is reasonably high).

In a mesh system in which moderate levels of interference are expected, DSSS is therefore an attractive option. However, at higher interference levels the performance degrades quite quickly and the only recourse is to increase the transmit-power or reduce the throughput, both methods that can cause problems in a mesh system, as already noted.

Frequency-Hopping Spread-Spectrum

Frequency-Hopping Spread-Spectrum (FHSS) uses narrowband modulation but achieves a wide bandwidth by changing the carrier frequency at regular intervals. There are two classes, fast and slow FHSS. In the former the rate at which the frequency is changed is higher than the data rate and in the latter it is lower.

The susceptibility of FHSS to wideband interference is similar and since FHSS has no processing gain it is worse than DSSS. Narrowband interference is different however, as follows:.

FHSS is highly resistant to narrowband interference. As the interference power increases their performance degrades up to the point when the frequencies affected by the interference are completely blocked, but no further. Thus FHSS can handle high-power narrowband interference better than DSSS.

The key difference between fast and slow FHSS is that the blocking of slow FHSS frequencies results in the corruption of several data bits so that other schemes must be used to compensate for the effect (re-transmission, error correction coding etc.), resulting in a loss of throughput. With fast FHSS, blocked frequencies affect a fraction of a bit and so long as the number of blocked frequencies remains small the demodulation will be unaffected.

In a mesh system that is likely to be affected by narrowband interference, FHSS is an attractive option. The loss of throughput in slow FHSS associated with the need to deal with lost data can be a problem which would make fast FHSS the more attractive of the two. However fast FHSS has greater overheads associated with synthesiser switching times so there may be little effective difference.

E.3 Time Structure

Interleaving

Interleaving is a diversity technique which can be used to counter discontinuous interference. The premise is that a non-zero error rate can be tolerated if the errors in the data are evenly distributed in time. For example some forward error correction techniques used by higher layer protocols can typically handle much higher error rates if the errors are so distributed.

The technique is most efficient when the design takes account of the capabilities of the higher layer protocols. Its principal disadvantage is that it introduces delay in the transmission process. In a multi-hop mesh system this delay may become unacceptable in, for example, telephony applications.

Throughput

The information throughput is one of the key elements in a radio network’s design and it is directly related to the interference to which the radio network is subject. By reducing the throughput the design can be made more resistant to interference because the energy that can be used to transmit the information bits increases as the bit-rate is reduced. It is however rare that the conditions under which a radio network will operate can be predicted and so the ability to vary the throughput at run-time is usually included in the design.

In the physical layer there are three methods that can be used to vary the throughput at run-time:

- An adaptive modulation scheme, in which the on-air modulation-rate changes, may be used. For example the sub-carriers of the IEEE802.11a system are modulated with using binary or quadrature phase shift keying (BPSK/QPSK), 16-quadrature amplitude modulation (QAM), or 64-QAM.
- A repetition method may be used to achieve a fixed on-air modulation-rate. The system is designed for a maximum throughput and as the throughput is reduced the information bits are repeated to maintain the same rate at the modulator. The method has the advantage of being simple to implement, but the steady-state conditions may themselves introduce problems and the spreading method is preferable.
- The spreading method achieves the same aim, a fixed on-air modulation-rate, as the repetition method. The difference is that the information bits are not simply repeated, but are replaced by a sequence of modulation-bits. The pattern of modulation-bits may be chosen to have different properties:
    - Pseudo-random sequences spread the bit energy across the channel bandwidth and form the basis of the Direct Sequence Spread Spectrum (DSSS) methods. By spreading the energy across the channel bandwidth the susceptibility to narrowband interference is reduced.
    - Orthogonal sequences allow multiple information bits to be transmitted at the same time

and in the same channel. These form the basis of the Code Division Multiple Access (CDMA) access method.

The 3GPP system uses orthogonal sequences to allow use of the CDMA access method and scrambles the result to spread the bit energy across the channel bandwidth.

Varying the throughput to match the propagation conditions is a technique used in most cellular designs. The management of the throughput can be centralised in the base station and generally a point-to-point connection will include at most 2 RF hops so the negotiation of the throughput can be quite simple.

In meshes the management of throughput could be considerably more difficult. The variable number of RF hops and the possibility that this might change during a call means that the negotiation process could become very complex.

E.4 Modulation

The modulation scheme used in a system affects the susceptibility to both self-interference and external interference. The effects are a result of both the immediate characteristics of the modulation and the characteristics of the practical transmitter and receiver designs that must be used in the real world.

In the case of self-interference the spectrum of modulated signal is the dominant factor as it affects the potential for adjacent-channel interference. Unfiltered modulation schemes such as FSK that have a broad spectrum not only have high out-of-band power levels but also require wide bandwidth receivers to recover the signal energy. In non-linear receivers this means that adjacent channel interference (or indeed any off-frequency interference) may, through intermodulation effects, be converted into co-channel interference.

The differences between more bandwidth efficient modulation methods are more subtle. The power spectral density of MSK, for example, has a wider central-lobe than that of QPSK and BPSK but it has lower side-lobes, which in principle should mean that it would offer better adjacent-channel performance. However all these modulation schemes have most of their energy concentrated in the central lobe so that additional channel filtering can be included in the transmitter without significantly degrading their performance.

The susceptibility of different modulation schemes to external interference is just as complex an issue. Even studies [Martin et al 1999, ESA 1999] which examine similar modulation schemes and similar interference effects can have different results and conclusions. In the two cases cited, the differences arise from the assumptions made about the nature of the receiver designs.

Overall it is impossible to identify one modulation scheme that is better than all others with respect to interference susceptibility. Other aspects of a radio network’s design will have a much greater impact.

E.5 Transmit Power

Increasing the power used to transmit a signal is a simple method of increasing the signal energy to overcome interference. However there are constraints that result from practical design considerations and from standardisation limits. The spectrum assigned to a particular use invariably has an associated maximum transmit power.

Further, increasing the power of the signal also increases the interference that it causes. Thus the self-interference of a system is increased, and in a mesh this may lead to a catastrophic instability as each element increases its power to overcome the interference caused by its neighbours. While it is possible to control the power used by the elements in a system so that this does not happen, some form of centralised control is required. The additional control flow through the mesh decreases user throughput.

E.6 Geometry

The geometry of the transmitters and receivers used in a system can be used to reduce the susceptibility to interference. Techniques such as antenna diversity, polarised antennas, directional antennas, beam-steering and sectoring are all used in cellular systems and fixed mesh systems.

However, such methods rely on the elements in a system having a relatively stable spatial position and orientation. In mobile meshes such stability cannot be assumed to exist and so only those methods that can adapt quickly are likely to be of use.

Appendix F International Conference Publications

F.1 IEE 3G and Beyond (short paper)

Efficient Mobile Mesh Networking: Testing Scalability Hypotheses

S.G.Methley*, M. Crisp*, J. Newman*, P.Ramsdale§, M. Rio¶, S.Bhatti¶, A.Atefi †

*Plextek Ltd, UK, {sgm,mc,jn}@plextek.com, §STA, UK, peter.ramsdale@spectrade.co.uk,
¶UCL, UK, m.rio@ee.ucl.ac.uk, s.bhatti@cs.ucl.ac.uk, †Ofcom, UK, ahmad.atefi@ofcom.org.uk

Keywords: mobile mesh, cellular multi-hopping, efficiency, quality of service, spectrum management

Abstract

This paper examines four scalability hypotheses of interest for mobile meshes via the following questions:

'Do meshes self-generate capacity as new nodes join?'

'Are meshes more spectrally efficient?'

'Do directional antennas confer significant benefits for hand-helds below 3.5GHz?'

'No' is the answer because these hypotheses, whilst having a theoretical basis, can be shown to rely on inappropriate real world assumptions. However the following hypothesis is found to be true:

'May meshes improve spectrum utilisation?'

1 Introduction

The UK Office of Communications (Ofcom) recently commissioned a consortium of industry and academia to investigate the reality of mobile meshes in the bands below 3.5GHz [1]. Such an activity is termed 'sensemaking' by strategists, where the aim is to establish an initial position despite confusing evidence: Ofcom wished to examine the validity of the many competing mesh performance claims in the literature, since subsequent strategic and economic analysis could develop important policy conclusions from such technical claims.

In this paper we attempt to summarise several of the main points of a larger investigation [1]. The core approach begins via an examination of assumptions made by key papers in the literature – and establishing their relevance to mobile meshes under 3.5GHz. This focus is key to the paper's findings. Whilst we do not wish to overstate the case, the results are not all as might be expected from mesh 'folklore'.

Next the real-world potential benefits of multiple hopping and antenna directionality are evaluated from first principles. Finally we note that 'less precious' spectrum e.g. up to 6GHz could usefully be utilised by mesh systems.

2 Hypotheses - Capacity and Scalability

Ofcom wished to test the following widely proclaimed benefits of multiple hop mesh networks:

- capacity self-generation
- spectral efficiency
- omni-vs.-directional antenna benefits
- spectrum utilisation

2.1 Hypothesis Testing ... "that customers self-generate capacity"

There would be huge attractions to having 'self-generation of capacity' in a radio network. Notably, that the network is self-sustaining and that it could avoid the so-called 'tragedy of the commons' (the exhaustion of network resources due to over-use).

We believe misinterpretation of some published work may have led to several unfortunate myths concerning 'self-generation of capacity'. Four published approaches are reviewed below and, whilst each presents a coherent argument based on its stated assumptions, it will be shown that those assumptions do not translate well to practical applications. The four approaches examined are:

Approach	Assumption Challenged
Grossglauser and Tse [2]	Unbounded delay
Gupta and Kumar [3]	Strict localisation of traffic
Shepard [4]	Unbounded spectrum
Negi and Rajeswaren [5]	Unbounded spectrum

<u>Grossglauser and Tse [2]</u>

This paper was taken as the starting point for an economics paper [6] which postulates many benefits if a 'tragedy of the commons' could thereby be avoided.

The model [2] specifically uses the mobility of nodes to act as intermediate 'couriers' of data between source and destination. Datagrams are passed from source nodes to near neighbours and delivery occurs when the courier nodes encounter the target recipients. Under this idealised model the per-node throughput remains constant, i.e. such a network is fully scalable in terms of capacity.

However, a clear consequence of this model is that the end-to-end packet delivery delay is related to the transit time of nodes moving throughout the area covered by the mesh. Statistically the mean delivery time is of order of 2d/v where d is the diameter of the mesh network and v the mean velocity of nodes within it. In a practical situation, the courier nodes may *never* encounter the recipient, in which case traffic is never delivered. The authors accept that this is clearly not acceptable for voice, or other real-time communications, and so direct the concept to non-critical store-and-forward messaging applications. It seems that this caveat may often be missed.

Although therefore limited in application in its basic form, we suggest the technique might be enhanced to reduce the transport delay and increase the probability of message delivery by nodes retaining a database of all other nodes they have had contact with and so selecting courier(s) on the basis of those that have had recent contact with the recipient.

<u>Gupta and Kumar [3]</u>
Their key conclusion is that capacity is shared amongst mesh nodes such that the upper bound for the average throughput $\lambda(n)$ obtainable by each node for a randomly chosen destination is of order of $c2W/\sqrt{(n.\log n)}$ bits/sec for the defined Random Network with the Physical Model. Thus the per-user throughput decreases with increasing node population.

Other authors, e.g. [8], have suggested other dependencies on the order of proportionality with n, but all models agree that average per-user throughput diminishes towards zero as the number of nodes increases, thus the mesh network does not scale indefinitely (and hence does not self-generate capacity).

It is interesting to consider what parameters, if any, might be changed to avoid this demise. Using a model [9] the dependencies on system parameters can be logically and simplistically stated as:
average throughput $\lambda(n)$ is proportional to functions of (γ, W, G/β, $1/L$, $1/r$, A, and $1/n$)

where γ = propagation attenuation law, W = channel transmission rate, G = channel processing gain, β = required signal to noise ratio, L = mean end-to-end path length, r = mean per-hop link length, A = area covered by network, n = number of nodes

This implies that unless one or more of the parameters grows with *n* then per-user throughput will be asymptotic to zero:

- *W* cannot grow arbitrarily large because of thermal noise constraints and limits on transmission power.
- *G/β* depends on the properties of the communication system and increasing it generally makes it necessary to decrease *W*.
- Reducing hop length *r* (e.g. by constraining transmit power) increases spatial re-use but at the expense of increased hop-count and hence increased relay traffic. It transpires [3, 9] that the preference is to reduce *r* to increase spatial re-use. But there is a limit here in that if r is too small then the network can become disconnected, i.e. minimum r is related to the inverse of node density (*A/n*).
- In random traffic flow models with uniform node density the mean end-to-end communication path length, *L*, is assumed to grow with coverage area *A* (*L* proportional to $\sqrt{A}$). This reduces capacity because of increased hop count. Thus, if one could conceive of services with more localised traffic (e.g. amongst localised communities) then *A/L* will increase more rapidly with increasing *A*. This will help to improve scalability.
- The remaining parameter that might scale with n is the area *A*. [9] suggests that three factors are required to achieve a non-zero throughput with increasing n: (i) the attenuation law γ needs to be greater than 3, (ii) the hop count *H* needs to be independent of n, (iii) area, *A*, needs to increase with n (i.e. the node *density* needs to be nearly constant or reducing with increasing *A*).
  However, (iii) requires that as the subscriber base increases those subscribers spread themselves out more thinly. It is not easy to see on what basis this might happen in any practical deployment.
- The propagation attenuation law γ strongly influences the above conclusions. A higher attenuation factor γ will permit higher throughput capacity [3, 9].

From the above list of options, one can see that there appears to be very little prospect of avoiding the asymptotic reduction in per-user throughput with increasing subscriber base. The analysis of [3] and others assumes a random association between source and destination nodes. Thus path lengths range from nearest neighbour (one-hop) to the full diameter of the area covered (many hops), and so, as the network size increases geographically and/or in terms of node-density, the number of hops per path must increase. This is one of the primary factors which cause the reduction in capacity with increasing number of nodes.

It is clear, then, that if traffic flows were more localised amongst neighbouring nodes, regardless of the geographic size of the network, then the number of hops per path would not increase *pro rata* with size and so the network would scale better, but we wonder how such a situation could be guaranteed in a real world deployment.

<u>Shepard [4]</u>
This paper has a relatively ‘out-of-the-box’ approach in suggesting a mesh in which collisions are not fatal for the MAC. It sees multiple concurrent transmissions as a signal-to-noise issue, rather than a requirement to back off and try again. It does this by using spread spectrum transmission, hence multiple transmissions simply raise the noise floor, as in any CDMA system. A complete theory is proposed to enable meshes to scale to millions of nodes. The problem is that it is extremely spectrally inefficient, due to the large processing gain required and in any case the predicted throughput of a large mesh is still only in the several kb/s range.

Negi and Rajeswaren [5]
A broadly similar approach with some similar problems is that of using "infinite" spectral bandwidth, for example in the ultra wide bandwidth (UWB) sense.

2.2 Hypothesis Testing ... "that mobile meshes are more spectrally efficient"

One of the traditionally used scenarios for suggesting that mesh operation into an Access Point might be more spectrally efficient than a PMP cell is the concept that increased throughput can be achieved over a series of short hops rather than one long hop. We shall demonstrate that this is only true for an idealised single-path scenario, and is diminished by the dissimilar antenna gains of Access Points and mobiles.

For the case of hopping between nodes of like type: If two hops of roughly equal length replace a single hop as shown in Figure 1 then:

- only half the time-bandwidth product of spectral resource is available for each hop, and this acts to reduce the delivered data rate by a factor of 2
- but as each hop is half the length of the original link, the link budget is improved. This improvement can be used to improve spectral efficiency either by increasing the transmission rate on each hop or reducing the transmit power. For example, in a third-law propagation environment the link budget is improved by x8 (~9dB); this would permit a four-fold increase in transmission rate by changing from QPSK to QAM64. Alternatively, with spread-spectrum the coding gain could be reduced to realise a similar increase in transmission rate.

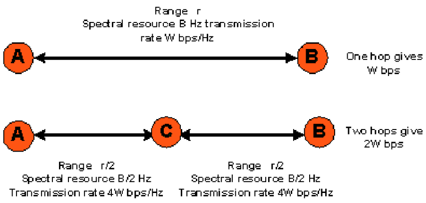


Figure 1 Two-hop vs. one-hop rate improvement between mesh nodes

This example implies that twice as much data can be transferred using two shorter hops: i.e. spectral efficiency is doubled. But this only prevails when the path length is exactly halved. If instead there is asymmetry in the two-hop path lengths then the link-budget gain in the longer hop will diminish and so the higher rate becomes unsupportable. This "sweet spot" in the path length split is illustrated in the graph of link budget in Figure 2.

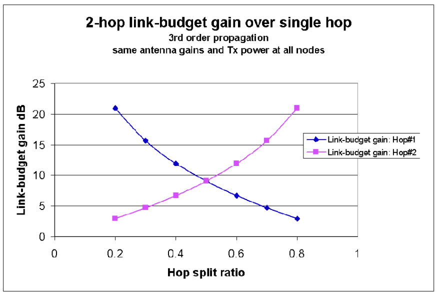


Figure 2: Two-hop link budget gain over single hop

But the comparative performance is further eroded for the case of multi-hopping into a mesh Access Point or cellular base station as represented in Figure 3.

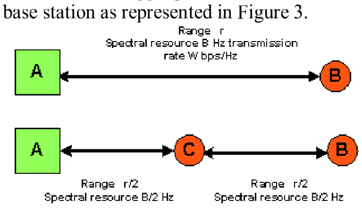


Figure 3: Two-hop vs. one-hop into high gain Access Point

The hop(s) between mobiles lack the higher antenna gain and height of the link into the Access Point (item **A** in Figure 3). Due to this imbalance the "sweet spot" no longer occurs at the 50:50 path-length split. The graph of Figure 4 illustrates this for the case when the Access Point antenna gain is just 13dB above the mobile nodes' gain – the "sweet spot" has moved to approximately 75:25 path length ratio and the optimal link budgets on the two hops are only about 4dB above the single-hop case. With this small link-budget gain the transmission rate might be little more than doubled. Thus the best case throughput rate of this two-hop route is roughly the same as the single-hop route.

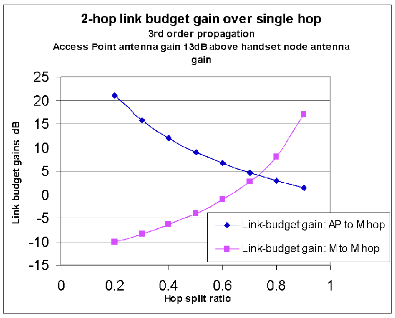


Figure 4: Two-hop vs. one-hop link budgets with high antenna gain

A further implicit assumption in the above simplified analysis is that the multi-hop path length is the same as the single hop length. In practice this may not be the case; nodes will be unevenly distributed and routes may circumvent building and terrain clutter. The detrimental effect of increased route length is illustrated in the graph of Figure 5 which illustrates the reduction in link budget gain at the "sweet spot" of Figure 4 as the route length is increased

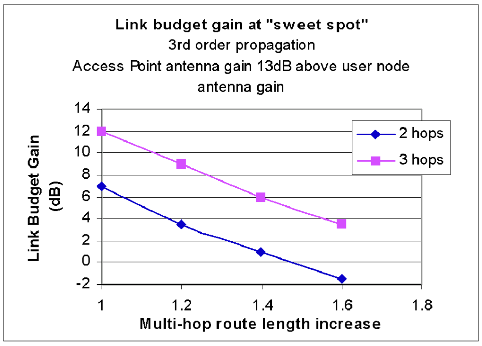


Figure 5 "Sweet Spot" link budget gain vs. extension in total route length

2.3 Hypothesis Testing ... "that directional antennas confer significant benefits for mobile mesh networks below 3.5GHz"

A starting point in the analysis is to consider an idealised antenna having negligible side lobe responses. This can be represented by the "flat top" model – where the antenna beam in the azimuth (horizontal) plane is represented as an arc of a circle subtending an angle equal to the 3dB beam width of a polar response. This leads to a simplistic interfering / non-interfering alignment of beams as illustrated in Figure 6:

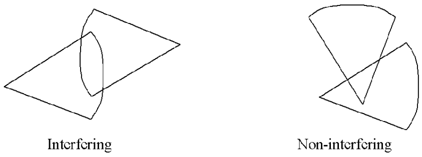


Figure 6: Interference Model for Directional Antennas

For a network of randomly deployed nodes equipped with such antennas, the theoretical upper limit on the improvement of throughput capacity is as large as $4\pi2/\alpha\beta$ [10] (where α and β are the beam widths of the transmit and receive antennas respectively). However, for any practical antenna, and more so for mobile/hand-held products in the bands of interest here (0.5-3.5 GHz), there will be a finite side lobe response which will seriously erode the gains anticipated.

The key manifestation of this finite side lobe response in the network is to extend the interference boundary around nodes [10]. The physical extent of this boundary is governed also by the attenuation factor of the propagation environment. If an antenna has a mean side lobe level which is κdB below the main beam then, in a propagation environment with attenuation rate γ (i.e. path loss proportional to (range)$^{\gamma}$), the differential coverage range, Δ_R, between main beam and side lobe is given by:

$$\kappa = 10.\ \gamma.\log(1/\Delta_R) \qquad (1)$$

It is postulated, from practical work at Plextek and data from the antenna-supply industry, that for mobile/hand-held products operating below approximately 6GHz the side lobe response is unlikely to be more than about 10dB-15dB below the main beam. So, taking a likely figure for side lobe level of κ=13dB, in a fourth-law propagation environment Δ_R is only 0.5. Thus, the interference boundary for the side lobes is only half that within the main beam.

Considering the case of 90° beam widths with -13dB side lobes this implies a capacity gain in the region of x3.3, compared to a theoretical gain of x16 for the zero-side lobes case. This illustrates the detrimental effect of finite side lobe levels.

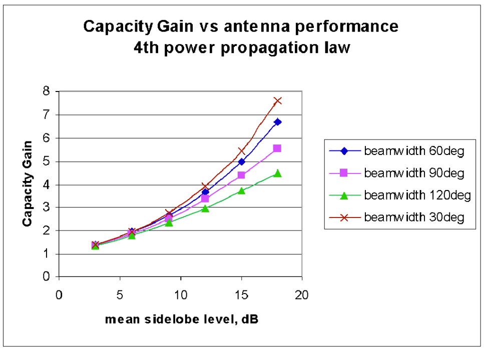


Figure 7: Theoretical Capacity Gain vs. Antenna Performance

Figure 7 illustrates that the capacity gain factor is a more sensitive function of side lobe level than it is of beam width. Furthermore, as beam width is reduced the side lobe level dominates performance, thus indicating that there is little benefit in decreasing beam width without equal attention to reducing side lobe levels, which returns us to the practical barriers first stated.

2.4 Hypothesis Testing ... "that meshes could improve spectrum utilisation"

This hypothesis relates to the wider issue of spectrum utilisation, rather than simple specific spectrum efficiency. It suggests that the spectrum may be better utilised by having short line-of-sight (LoS) mesh links use 'less precious' spectrum e.g. up to 6 GHz.

In [1] three key factors point to mobile mesh networks offering opportunities for use of higher frequency bands:

i. They are not necessarily more spectrally efficient than current cellular systems operating in the 2GHz region (cf. 2.2). Thus they might usefully be allocated less commercially precious spectrum.

Plextek

ii. To achieve useful per-user throughput the relaying capacity of mesh nodes needs to be high (a corollary of 2.1). Thus meshes need access to large allocations of bandwidth.
iii. The potential of increased end-to-end throughput by using multi-hop vs. single hop is best realised when there is a high propagation path loss at the chosen frequency of use.

Conclusions

The main contribution of this paper is the rationalisation and clarification of many competing mesh performance claims within the literature for the specific case of mobile meshes below 3.5GHz. This is important since such technical claims can form the basis of economic and policy planning. We both examined existing work and used analysis from first principles.

From the hypotheses tested we conclude that mesh subscribers cannot self-generate capacity at a rate sufficient to maintain a target level of per-user throughput regardless of network size and population. One way scalability could be achieved is by providing additional capacity in the form of an access network so forming an "Access Mesh". We further conclude that there remain fundamental tradeoffs between throughput, capacity and delay, which cannot be dismissed easily. Thirdly, whilst network capacity can be improved through the use of directional antennas, for handheld devices the extent of directionality is limited since the high side lobes levels associated with such small antennas severely limit the improvements in spatial reuse that would otherwise be possible. Finally we note that spectrum utilisation could be improved by operating meshes within higher, less precious spectrum.

In summary, we still believe that whilst mobile meshes do not live up to all the claims in the literature, they can be very benficial in the area of coverage extension. We suggest that meshes are equally applicable to extending the coverage of WLAN hotspots as they are to cellular multi-hopping, within certain limitations, and should be seen as integral to any 4G or 'beyond 3G' vision. They should also find application in home and office indoor networking and community networks.

Acknowledgements

The authors would like to thank William Webb of Ofcom for his encouragement and colleagues at Plextek (Graham Maile, Bill Garland, Peter Massam, David Cox), Ofcom (Stephen Jones) and UCL (Lionel Sacks, John Mitchell) for their inputs to the project.

References

[1] Methley S G, Crisp M, Newman J, Rio M, Bhatti S, Ramsdale P, *"Study of Efficient Mesh Networks"*, Ofcom Contract Ref. C31400/001, Ofcom, to be published.

[2] Grossglauser M, Tse D, "Mobility increases the capacity of ad hoc networks" *Proceedings of IEEE INFOCOM '01*, Anchorage, Alaska, April 2001.

[3] Gupta P and Kumar P R, "The capacity of wireless networks", *IEEE Trans. Information Theory*, vol 46, pp. 388-404, Mar. 2000.

[4] Shepard T , A Channel Access Scheme for Large Dense Packet Radio Networks", *Proceedings of ACM SIGCOMM'96*, August 1996, Stanford University, California.

[5] Negi R, Rajeswaren A, "Capacity of power constrained ad-hoc networks," *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[6] Jones S, Levine P and Rickman N, *"The Economics of Radio Spectrum"*, Ofcom report, 2003. Also *Fifth Wireless World Conference*, 15-16 July 2004, University of Surrey, Guildford, UK.

[7] Sharma G, Mazumdar R R, "On Achievable Delay/Capacity Trade-offs in Mobile Ad Hoc Networks," *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Cambridge, UK, March 2004.

[8] "*High Performance Mesh for Dense Metro WiFi*", MeshDynamics white paper available at <http://www.meshdynamics.com/wPresentations.html>

[9] Arpacioglu O, Zygmunt J.H : "On the Scalability and Capacity of Wireless Networks with Omnidirectional Antennas". *Third International Symposium on Information Processing in Sensor Networks* (IPSN 2004), Berkeley, California (April 27–28, 2004)

[10] Yi S, Pei Y, Kalyanaraman S, "On the capacity improvement of ad hoc wireless networks using directional antennas", *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, Annapolis, Maryland, USA, June 01-03, 2003.

Plextek

F.2 IEEE DySpan (long paper)

Efficient Mobile Mesh Networking:

Attractions, Myths and Techno-Economic Roadmap to Successful Commercial Innovation

S G Methley, M Crisp, J Newman
Plextek Ltd
Cambridge, UK
{sgm, mc, jn}@plextek.com

P A Ramsdale
STA
Cambridge, UK
peter.ramsdale@spectrade.co.uk

M Rio, S Bhatti
University College
London, UK
{m.rio@ee,s.bhatti@cs}.ucl.ac.uk

A Atefi
Ofcom
London, UK
ahmad.atefi@ofcom.org.uk

***Abstract*— This paper examines four scalability hypotheses of interest for mobile meshes via the following questions:**

'Do meshes self-generate capacity as new nodes join?'

'Are meshes more spectrally efficient?'

'Do directional antennas confer *significant* benefits for *hand-helds* below 3.5GHz?'

'No' is the answer because these hypotheses, whilst having a theoretical basis, can be shown to rely on inappropriate real world assumptions. However the following hypothesis is found to be true:

'Meshes may improve spectrum utilisation'.

Importantly however, there remain properties of meshes which make them uniquely attractive, such as coverage extension. However this raises a further question over the ability of a mobile mesh architecture to provide a guaranteed quality of service.

Finally, the wider aspects of commercial innovation are considered.

Keywords- mobile mesh, cellular multi-hopping, efficiency, quality of service, spectrum management

I. Introduction

The UK Office of Communications (Ofcom) recently commissioned a consortium of industry and academia to investigate the reality of mobile meshes in the bands below 3.5GHz [1]. Such an activity is termed 'sensemaking' by strategists, where the aim is to establish an initial position despite confusing evidence: Ofcom wished to examine the validity of the many competing mesh performance claims in the literature, since subsequent strategic and economic analysis could develop important policy conclusions from such technical claims.

Perhaps the largest initial attraction of mobile meshes is that they can be entirely unplanned in pure form. This is useful to the military and to disaster recovery teams who neither need infrastructure access for content nor want to rely on its presence for operation. It is far less clear what these benefits could lend to the roll-out of a mass market mesh network. On the other hand, to a service provider or regulator, the lure of a network which promises no planning phase must be high and thus must merit investigation.

It has often been said, as if it were a truism, that meshes increase capacity. The reasoning is usually along the lines of *each new user brings additional capacity to the mesh*, or *each new user effectively becomes a base station*. This paper critically examines such statements and aims to separate the reality from a 'something for nothing' type of mythology.

Nonetheless we find that, ultimately, meshes do retain some strongly attractive features, notably in the area of coverage, where they offer complementary performance to that of cellular systems. It is for this reason that meshes should find application in some scenarios, as part of a larger picture of mobile access technology.

In this paper we attempt to summarise several of the main points of a larger investigation [1]. The core approach begins via an examination of assumptions made by key papers in the literature – and establishing their relevance to mobile meshes under 3.5GHz. This focus is key to the paper's findings. Whilst we do not wish to overstate the case, the results are not all as might be expected from mesh 'folklore'.

Next, the real-world potential benefits of multiple hopping and antenna directionality are evaluated from practical first principles. Following these analyses we support meshing as a technique whose prime advantage is coverage extension. However we note that there are important caveats for service quality due to a performance which is dependent on the uncertainty in user movements. Finally we introduce five scenarios where we believe meshes can be innovative.

Remaining enablers and barriers to success are evaluated in the contexts of technology, human factors and regulatory framework.

II. Hypotheses - Capacity and Scalability

Ofcom wished to test the following widely proclaimed benefits of multiple hop mesh networks:

- capacity self-generation
- spectral efficiency
- omni-vs.-directional antenna benefits
- spectrum utilisation
- coverage extension

The first four are dealt with in this section, whilst coverage is dealt with in the following section.

A. *Hypothesis Testing ... "that customers self-generate capacity"*

There would be huge attractions to having 'self-generation of capacity' in a radio network. Notably, that the network is self-sustaining and that it could avoid the so-called 'tragedy of the commons' (the exhaustion of network resources due to over-use).

We believe misinterpretation of some published work may have led to several unfortunate myths concerning 'self-generation of capacity'. Four published approaches are reviewed below and, whilst each presents a coherent argument based on its stated assumptions, it will be shown that those assumptions do not always translate well to practical applications. The four approaches examined are:

Approach	Assumption Challenged
Grossglauser and Tse [2]	Unbounded delay
Gupta and Kumar [3]	Strict localisation of traffic
Shepard [4]	Unbounded spectrum
Negi and Rajeswaren [5]	Unbounded spectrum

1) *Grossglauser and Tse [2]*

This paper was taken as the basis for an economics and regulatory policy paper [6] which postulates many benefits if a 'tragedy of the commons' could be avoided by using mobility itself to increase the capacity of a network.

The model [2] specifically uses the mobility of nodes to act as intermediate 'couriers' of data between source and destination. Datagrams are passed from source nodes to near neighbours and delivery occurs when the courier nodes encounter the target recipients. Under this idealised model the per-node throughput remains constant, i.e. such a network is fully scalable in terms of capacity.

However, a clear consequence of this model is that the end-to-end packet delivery delay is related to the transit time of nodes moving throughout the area covered by the mesh. Statistically the mean delivery time is of order of 2d/v where d is the diameter of the mesh network and v the mean velocity of nodes within it. In a practical situation, the courier nodes may *never* encounter the recipient, in which case traffic is never delivered. The authors accept that this is clearly not acceptable for voice, or other real-time communications, and so direct the concept to non-critical store-and-forward messaging applications. It seems that this caveat may often be overlooked.

Although therefore limited in application in its basic form, we suggest the technique might be enhanced to reduce the transport delay and increase the probability of message delivery by nodes retaining a database of all other nodes they have had contact with and so selecting courier(s) on the basis of those that have had recent contact with the recipient.

2) *Gupta and Kumar [3]*

Their key conclusion is that capacity is shared amongst mesh nodes such that the upper bound for the average throughput λ(n) obtainable by each node for a randomly chosen destination is of order of W/√(n.logn) bits/sec for the defined Random Network with the Physical Model. Thus the per-user throughput decreases with increasing node population.

Other authors, e.g. [8], have suggested other dependencies on the order of proportionality with n, but all models agree that average per-user throughput diminishes towards zero as the number of nodes increases, thus the mesh network does not scale indefinitely (and hence does not self-generate capacity).

It is interesting to consider what parameters, if any, might be changed to avoid this demise. Using a model [9] the dependencies on system parameters can be logically and simplistically stated as:

average throughput λ(n) is proportional to functions of (*γ, W, G/β, 1/L, 1/r, A,* and 1/*n*)

where γ = propagation attenuation law, W = channel transmission rate, G = channel processing gain, β = required signal to noise ratio, L = mean end-to-end path length, r = mean per-hop link length, A = area covered by network, n = number of nodes

This implies that unless one or more of the parameters grows with n then per-user throughput will be asymptotic to zero:

- *W* cannot grow arbitrarily large because of thermal noise constraints and limits on transmission power.
- *G/β* depends on the properties of the communication system and increasing it generally makes it necessary to decrease *W*.
- Reducing hop length *r* (e.g. by constraining transmit power) increases spatial re-use but at the expense of increased hop-count and hence increased relay traffic. It transpires [3, 9] that the preference is to reduce *r* to increase spatial re-use. But there is a limit here in that if r is too small then the network can become

disconnected, i.e. minimum r is related to the inverse of node density (*A/n*).

- In random traffic flow models with uniform node density the mean end-to-end communication path length, *L*, is assumed to grow with coverage area *A* (*L* proportional to $\sqrt{A}$). This reduces capacity because of increased hop count. Thus, if one could conceive of services with more localised traffic (e.g. amongst localised communities) then *A/L* will increase more rapidly with increasing *A*. This will help to improve scalability.

- The remaining parameter that might scale with n is the area *A*. [9] suggests that three factors are required to achieve a non-zero throughput with increasing n: (i) the attenuation law γ needs to be greater than 3, (ii) the hop count *H* needs to be independent of n, (iii) area, *A*, needs to increase with n (i.e. the node *density* needs to be nearly constant or reducing with increasing *A*). However, (iii) requires that as the subscriber base increases those subscribers spread themselves out more thinly. It is not easy to see on what basis this might happen in any practical deployment.

- The propagation attenuation law γ strongly influences the above conclusions. A higher attenuation factor γ will permit higher throughput capacity [3, 9].

From the above list of options, one can see that there appears to be very little prospect of avoiding the asymptotic reduction in per-user throughput with increasing subscriber base. The analysis of [3] and others assumes a random association between source and destination nodes. Thus path lengths range from nearest neighbour (one-hop) to the full diameter of the area covered (many hops), and so, as the network size increases geographically and/or in terms of node-density, the number of hops per path must increase. This is one of the primary factors which cause the reduction in capacity with increasing number of nodes.

It is clear, then, that if traffic flows were more localised amongst neighbouring nodes, regardless of the geographic size of the network, then the number of hops per path would not increase *pro rata* with size and so the network would scale better, but we wonder how such a situation could be guaranteed in a real world deployment.

3) Shepard [4]

This paper has a relatively 'out-of-the-box' approach in suggesting a mesh in which collisions are not fatal for the MAC. It sees multiple concurrent transmissions as a signal-to-noise issue, rather than a requirement to back off and try again. It does this by using spread spectrum transmission, hence multiple transmissions simply raise the noise floor, as in any CDMA system. A complete theory is proposed to enable meshes to scale to millions of nodes. The problem is that it is extremely spectrally inefficient, due to the large processing gain required and in any case the predicted throughput of a large mesh is still only in the several kb/s range.

4) Negi and Rajeswaren [5]

A broadly similar approach to Shepard [4] with some similar problems is that of using "infinite" spectral bandwidth, for example in the ultra wide bandwidth (UWB) sense.

B. Hypothesis Testing ... "that mobile meshes are more spectrally efficient"

One of the traditionally used scenarios for suggesting that mesh operation into an Access Point might be more spectrally efficient than a regular PMP cell is the concept that increased throughput can be achieved over a series of short hops rather than one long hop. We shall demonstrate that this is only true for an idealised single-path scenario, and is diminished by the dissimilar antenna gains of Access Points and mobiles.

For the case of hopping between nodes of like type: If two hops of roughly equal length replace a single hop as shown in Figure 1 then:

- only half the time-bandwidth product of spectral resource is available for each hop, and this acts to reduce the delivered data rate by a factor of 2

- but as each hop is half the length of the original link, the link budget is improved. This improvement can be used to improve spectral efficiency either by increasing the transmission rate on each hop or reducing the transmit power. For example, in a third-law propagation environment the link budget is improved by x8 (~9dB); this would permit a four-fold increase in transmission rate by changing from QPSK to QAM64. Alternatively, with spread-spectrum the coding gain could be reduced to realise a similar increase in transmission rate.

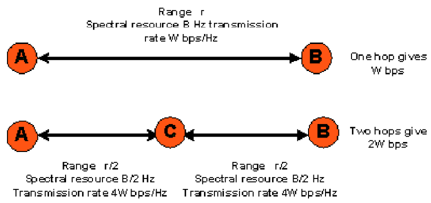


Figure 1 Two-hop vs. one-hop rate improvement between mesh nodes

This example implies that twice as much data can be transferred using two shorter hops, i.e. spectral efficiency is doubled. But this only prevails when the path length is exactly halved. If instead there is asymmetry in the two-hop path lengths then the link-budget gain in the longer hop will diminish and so the higher rate becomes unsupportable. This "sweet spot" in the path length split is illustrated in the graph of link budgets in Figure 2.

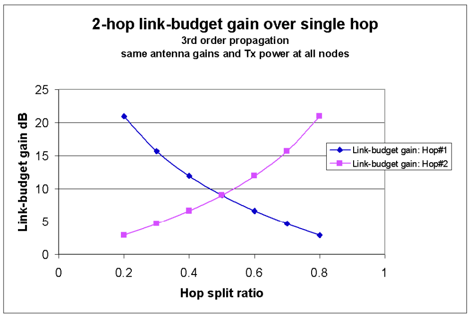


Figure 2: Two-hop link budget gain over single hop

But the comparative performance is further eroded for the case of multi-hopping into a mesh Access Point or cellular base station as represented in Figure 3.

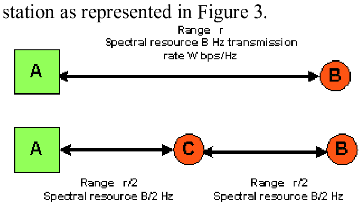


Figure 3: Two-hop vs. one-hop into high gain Access Point

The hop(s) between mobiles lack the higher antenna gain and height of the link into the Access Point (item **A** in Figure 3). Due to this imbalance the "sweet spot" no longer occurs at the 50:50 path-length split. The graph of Figure 4 illustrates this for the case when the Access Point antenna gain is just 13dB above the mobile nodes' gain – the "sweet spot" has moved to approximately 75:25 path length ratio and the optimal link budgets on the two hops are only about 4dB above the single-hop case. With this small link-budget gain the transmission rate might be little more than doubled. Thus the best case throughput rate of this two-hop route is roughly the same as the single-hop route.

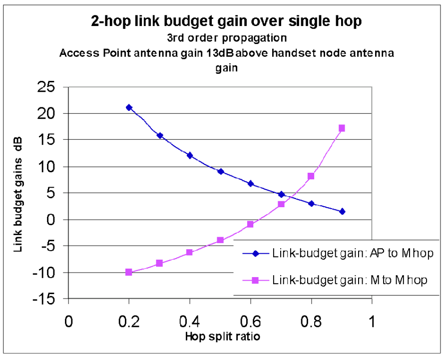


Figure 4: Two-hop vs. one-hop link budgets with high antenna gain

A further implicit assumption in the above simplified analysis is that the multi-hop path length is the same as the single hop length. In practice this may not be the case; nodes will be unevenly distributed and routes may circumvent building and terrain clutter. The detrimental effect of increased route length is illustrated in the graph of Figure 5 which illustrates the reduction in link budget gain at the "sweet spot" of Figure 4 as the route length is increased

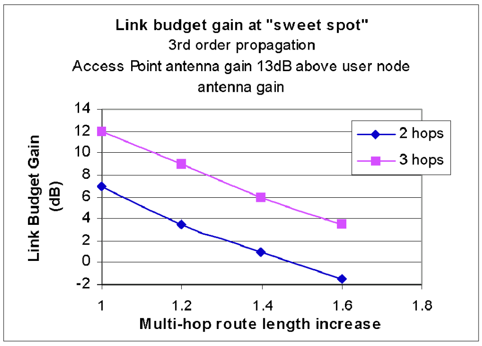


Figure 5 "Sweet Spot" link budget gain vs. extension in total route length

C. *Hypothesis Testing … "that directional antennas confer significant benefits for mobile mesh networks below 3.5GHz"*

A starting point in the analysis is to consider an idealised antenna having negligible side lobe responses. This can be represented by the "flat top" model – where the antenna beam in the azimuth (horizontal) plane is represented as an arc of a circle subtending an angle equal to the 3dB beam width of a polar response. This leads to a simplistic interfering / non-interfering alignment of beams as illustrated in Figure 6:

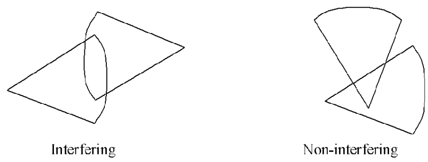


Figure 6: Interference model for directional antennas

For a network of randomly deployed nodes equipped with such antennas, the theoretical upper limit on the improvement of throughput capacity is as large as $4\pi2/\alpha\beta$ [10] (where α and β are the beam widths of the transmit and receive antennas respectively). However, for any practical antenna, and more so for mobile/hand-held products in the bands of interest here (0.5-3.5 GHz), there will be a finite side lobe response which will seriously erode the gains anticipated.

The key manifestation of this finite side lobe response in the network is to extend the interference boundary around nodes

[10]. The physical extent of this boundary is governed also by the attenuation factor of the propagation environment. If an antenna has a mean side lobe level which is κdB below the main beam then, in a propagation environment with attenuation rate γ (i.e. path loss proportional to $(\text{range})^{\gamma}$), the differential coverage range, Δ_R, between main beam and side lobe is given by [10]:

$$\kappa = 10.\ \gamma.\log(1/\Delta_R) \quad (1)$$

It is postulated, from practical work at Plextek Ltd and data from the antenna-supply industry, that for mobile/hand-held products operating below approximately 6GHz the side lobe response is unlikely to be more than about 10dB-15dB below the main beam. So, taking a likely figure for side lobe level of κ=13dB, in a fourth-law propagation environment Δ_R is only 0.5. Thus, the interference boundary for the side lobes is only half that within the main beam.

Considering the case of 90° beam widths with -13dB side lobes this implies a capacity gain in the region of x3.3, compared to a theoretical maximum gain of x16 for the zero-side lobes case. This illustrates the detrimental effect of finite side lobe levels.

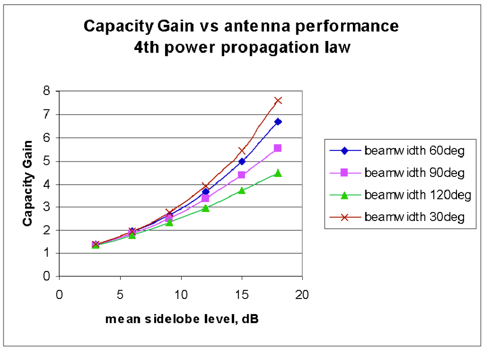


Figure 7: Capacity gain vs. antenna performance

Figure 7 illustrates that the capacity gain factor is a more sensitive function of side lobe level than it is of beam width. Furthermore, as beam width is reduced the side lobe level dominates performance, thus indicating that there is little benefit in decreasing beam width without equal attention to reducing side lobe levels, which returns us to the practical barriers first stated.

This capacity gain performance is also a function of the propagation environment. The range difference, Δ_R, between main beam and side lobes decreases with increasing propagation attenuation law and so the benefit of side lobe attenuation diminishes. This follows from the premise that, for a given density of nodes, the ratio of the number of nodes residing inside the main beam coverage area to the number residing in the side lobe coverage area diminishes with increased propagation law. This effect is illustrated in Figure 8 for a beam width of 90°. From this it can be seen that the benefit from antenna directionality decreases with increasing propagation attenuation factor.

Note that the capacity gain curves in Figure 8 are normalised to the omni-directional antenna case for each respective propagation law. Thus the curves do not imply that the network has the same capacity for an omni-directional antenna, independent of propagation law; there are scale-factors to be applied to the vertical axis for each curve. In fact the corollary to this is that the high attenuation environments enable greater spatial reuse and hence higher spectral efficiency than a low attenuation environment [3, 9]. However, the low attenuation environment will reap more benefit from the use of directional antennas, because of reduced interference over the longer propagation ranges.

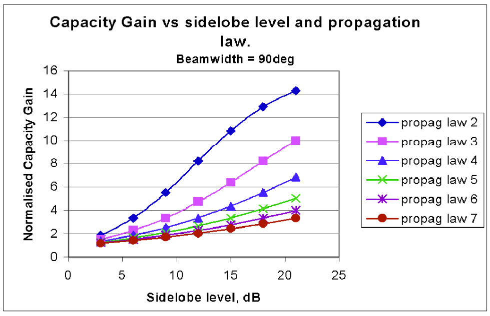


Figure 8 Capacity gain vs. propagation attenuation factor

As a final aside, another way of visualising the potential benefits of directional antennas is to consider that a mesh may be made physically or logically. Physical meshes are those made by physical level constraints, for example by directional antennas or perhaps by constraint of the signal path by terrain or medium. The wired internet is clearly a perfect physical mesh in that transmitting on one link does not interfere with any others. On the other hand a logical mesh is configured above the physical layer. There is not necessarily any physical constraint to a station's neighbours imposed by the system. Omni-directional antennas in an open field could be connected as a logical mesh, although their interference footprint would clearly be quite different from that of a physical mesh and the full benefits of physical meshing should not be expected. (The unrealised benefit would be that the omni-directional antennas could equally be re-configured as e.g. point to multi-point, in the logical sense, if that were ever required.) The relevance of this discussion to the preceding argument is that it is how closely a true physical mesh may be approached which is important, be that via the degree of antenna directionality available, or otherwise.

D. Hypothesis Testing … "that meshes could improve spectrum utilisation"

This hypothesis relates to the wider issue of spectrum utilisation, rather than simple specific spectrum efficiency. It suggests that the spectrum may be better utilised by having short line-of-sight (LoS) mesh links use 'less precious' spectrum e.g. up to 6 GHz.

In [1] three key factors point to mobile mesh networks offering opportunities for use of higher frequency bands:

i. They are not necessarily more spectrally efficient than current cellular systems operating in the 2GHz region (cf. I.B). Thus they might usefully be allocated less commercially precious spectrum.
ii. To achieve useful per-user throughput, the relaying capacity of mesh nodes needs to be high (a corollary of II.A). Thus meshes need access to large allocations of bandwidth.
iii. The potential of increased end-to-end throughput by using multi-hop vs. single-hop is best realised when there is a high propagation path loss at the chosen frequency of use (cf. I.B, II).

III. MESH COVERAGE

One clear claim for mesh networking is that it helps coverage by allowing multiple hops around obstacles. This is of most use in a cluttered environment, where a powerful base station would otherwise have to be used to combat the large variance in path loss. Figure 9 [11] compares meshing around obstacles via a 'forwarder' to gain an advantage in throughput.

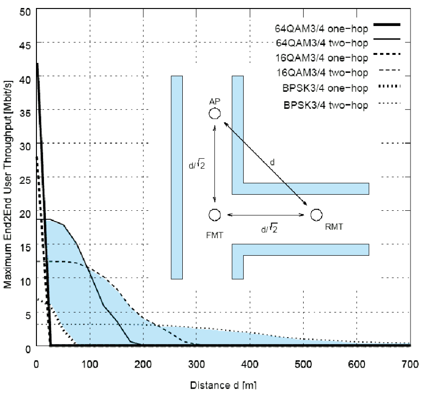


Figure 9 Use of a 'forwarder' to 'skip' around obstacles as in a mesh network

Figure 9 presents a specific argument for hopping around obstacles, rather than multi-hopping *per se*, and may be understood as follows: the situation being modelled is as drawn in the figure inset. The objective is to communicate from the AP (access point) to the RMT (remote mobile terminal), which is sited around a corner. This may be done in one of two ways:

- ∅ as a single hop, distance d, 'directly'
- ∅ as a two-hop, twice distance $d/\sqrt{2}$, 'hopping' around the corner via the FMT (forwarding mobile terminal)

Additionally, the effect of adapting the modulation scheme is also shown.

The potential gains to be had in all the two-hop cases are shown within the total blue shaded area of the graph which is made up of the gains from each two-hop case listed in the graph legend. The white area, enclosed by the axes and the shaded area, is the single hop performance, which is very limited as might be expected. Although the results of Figure 9 relate to extending the infrastructure mode of HiperLAN/2 in the 5GHz band, the implications of this simulation are of a general nature and thus may be applied to other hopping situations, such as a mesh network.

Clearly, in the real world, propagation scenarios may be more or less extreme than Figure 9. The best system solution thus depends on the target environment. In general however, meshes should show benefits where the problem faced is one of coverage due to a cluttered or shadowed environment. Most often this is associated with a dense urban area. Potentially, therefore, meshes could be also deployed as 'hot zones' in a sea of cellular, due to their complementary coverage characteristic. We would define this as an 'access mesh', but this might also be included under the collective of 'cellular with multi-hopping'.

Hopping around corners or obstacles is perhaps the greatest benefit offered by a mesh.

IV. COULD A MESH GUARANTEE A QUALITY OF SERVICE?

Accepting the limitations of II.A to I.D, we have stated that we still believe that meshes have much to offer from a coverage point of view. Hence the next question to ask is whether quality of service is possible within a mesh, despite user movement. Poor quality of service would be at odds with users' expectations, as they will require guaranteed service for inelastic applications like video or other applications intolerant of delay-variation or throughput-variation.

Royer et al [12] noted the result of a well known publication which stated that a fixed node with the 'magic number' of six fixed neighbours could be shown to possess the optimum trade-off between transmission power and self-interference for best throughput [13]. The question was whether a similar optimum carried across to mobile networks.

The answer was 'no'; there is no single optimum power-efficiency trade off for a range of node mobilities, although there is an optimum for each value of node mobility. Essentially, as node mobility increases it was found that the node density, in terms of connectivity, needed to rise. Otherwise, the mesh could eventually break apart or 'partition' [14]. More mobile nodes need more guaranteed neighbours. The clear issue for the mesh network planner is that network connectivity, hence quality of service, depends on parameters outside his control – the users.

More recently Nilsson [15] showed that connectivity also depends on the traffic level within a multi-hop mesh. His conclusion was that both predicted traffic levels and node mobility/densities would need to be known at the planning stage, for viable network design.

Lungaro and Wallin [16] expanded on Lungaro [17], where notably they added a caveat to his earlier work on mesh-like cellular multi-hopping. They noted that the principal drawback of the scheme was that, due to the 'partially uncontrolled infrastructure' (i.e. the user nodes), multi-hopping was 'not able to guarantee a quality of service'. They proposed that a simple solution was the addition of nodes, as relays, by the network operator. In fact they proposed a system with three node types: access node, relay node and user node. Their conclusion goes beyond Nilsson's. Knowing traffic and user mobility is still not enough to guarantee quality of service within a mesh of user nodes - infrastructure must be added.

We note that the addition of some infrastructure is necessary in any case where access beyond the mesh is required, e.g. to the Internet. We expect such access meshes will dominate completely over isolated pure meshes for public deployments. The fixed part of an access network could be arranged to provide (i) seeding of a new deployment (to ensure connectivity whilst user number ramps up), (ii) reduction of network dependence on user mobility and traffic levels, as well as (iii) access to the wider Internet.

V. Innovation

The previous sections of this paper have addressed and commented on technology aspects of mesh networks and identified the properties which make them unsuitable for certain applications yet suitable for others. Achievement of the best overall result will also require consideration of regulation, creation of standards, development of policies concerning harmonisation effects, etc., each from the point of view of how they might encourage or discourage the complex innovation process, Figure 10.

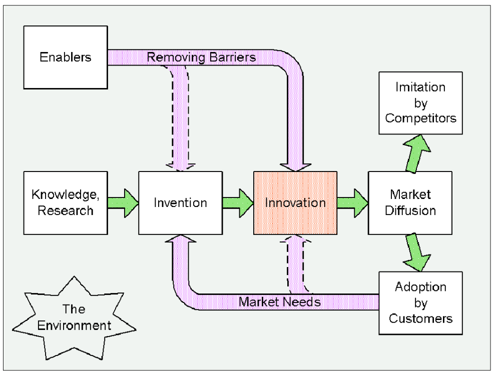


Figure 10 Innovation as part of a larger process

Innovation is not simply the introduction of new techniques and technologies; it is also about doing things in a new way. To encourage innovation in any field it is necessary to understand all the enablers and barriers in order to see what becomes possible by forcing change. Figure 10 shows that necessity is not only the 'mother of invention', but also of innovation. In other words the market drives the need.

The discussion is structured into three categories:

- a definition of the desired end result of innovation
- consideration of enablers for innovation
- consideration of barriers to innovation.

A. *Innovation - What, Where and How?*

It is important to be clear on the goal and the general strategy to achieve it. This paper takes the goal as achieving the deployment of access mesh networks in particular, i.e. those with access points to external content. Moreover the likely scenarios are taken to be those identified in [1]. These are:

- cellular with multi-hopping
- WLAN hotspot extension
- community networking, e.g. [22]
- home and office indoor networking
- zero-infrastructure environments

The latter two scenarios may not be familiar: for home and office networking we observe that a closed user group, peer to peer network within a cluttered environment is the requirement and this plays to the strengths of a mesh. By zero-infrastructure environments, we mean those places where infrastructure is yet to be installed or where a permanent infrastructure is not needed, e.g. a new technology park or a temporary conference venue.

The strategic 'how' factor is addressed via an evaluation of the enablers and barriers, within the specific context of our stated goal. We believe that the very creation of the likely scenario list above already shows how mesh networks can be strongly innovative in removing barriers. What follows identifies those barriers which may remain and the further enablers which need consideration.

B. *Enablers for Innovation*

Enablers are discussed under the following groupings:

- technology
- human factor aspects
- technology understanding.

1) *Technology*

In Figure 10, these include the remaining inventions necessary to feed into the innovation process. In some cases the inventions are very difficult technically and hence a solution is sought but not yet available. In other cases the need for the invention may not yet be clear. The physical layer is fairly well served by existing technology, although there is substantial scope for inclusion of the emerging cognitive radio technology. But substantial further work is required in the fields of MAC and routing protocols to improve the efficiency of carrying delay sensitive traffic. Alongside these are requirements for more advanced system modelling tools for mobility and traffic flows.

2) Human factor aspects

Some user behaviour and expectations will affect the successful uptake of mesh networking. A mesh specific problem is the power consumption of a battery powered device. Because mesh nodes are required to relay traffic which may not be their own, especially when close to an access point, more demands are made on battery performance. A situation may occur where a user who is not generating his own traffic may nonetheless experience high battery drain due to relaying the traffic of others. Further problems could arise depending on the user's response to such a situation. If the user turns off his node when not using the device, then the mesh availability may be compromised for many other users. Clearly some provision needs to be made to encourage good group behaviour from individual users. Reward schemes have been suggested [18], but may carry a large traffic overhead.

General improvements in radio technology will continue to benefit mesh networks. In particular, techniques which increase battery life, such as lower voltage rails, higher efficiency PAs, and more efficient modulation methods (energy/bit), will particularly benefit mobile meshes.

3) Technology understanding

Both users and operators need to understand the technology in terms of what it could do for them. Currently, much of this thinking may be dominated by individual perception.

In the UK, the major cellular operators use GSM and UMTS for their 2nd and 3rd generation mobile technology. Multi-mode handsets which add WLAN access to the service offering are becoming available. In the UK, one operator already offers a separate, extensive WLAN hotspot service, and the business aspects of the differing services are beginning to be understood.

However there remain concerns about combining services which offer very different performance guarantees. Some operators believe that they have established a particular service expectation for their 'brand'. So, for a cellular operator, this means that there are levels of service in terms of coverage, availability, dropped calls (including calls with hand-overs), consistency, etc. The list of parameters continues to grow with data and IP based calls, such that delay and throughput, for example, also need to be at or above the 'brand' level on a regular basis.

On the other hand the WLAN hotspot service provider offers a service, some aspects of which attract lower expectations, but yet is still acceptable to their users. This arises from experience and tolerance of the Internet as a 'best efforts' service and the assumed coverage area of a WLAN radio hotspot being conservative.

With these mindsets, how likely is it that mesh networks will be considered as a suitable technology for delivering part of a service? Clearly, if mobile mesh techniques are to form part of a cellular network, the operator must ensure that customers' expectations are not disappointed by variability or lower than expected performance levels (section III). It is sometimes considered better not to offer coverage rather than to offer a poor service and, especially when rolling out a network, sites are only introduced when they can offer contiguous coverage.

An enabler for innovation would be to first evaluate these 'soft' aspects of rolling out a mesh based service.

C. Removing barriers to innovation

1) Spectrum management model

Ofcom's policy on spectrum management is based on the Cave Report [19] as developed further in the Spectrum Framework Review [20, 21]. Ofcom has changed its approach from command-and-control to the use of market mechanisms plus some licence exempt use. With the market mechanisms approach, Ofcom is still responsible for partitioning the spectrum and making it available - an auction being the preferred route. Ownership and use is then driven by market forces. Secondary trading is expected and spectrum should be acquired by those who value it most highly. Over time spectrum blocks may be sub-divided or aggregated.

This encompasses trading and liberalisation activities and raises issues of radio interference. Owners of spectrum need to know what their spectrum usage rights are, as this impinges on the quality of service which they can provide. The issues are:

- spectrum quality
- levels of real interference
- levels of perceived interference
- protection / adjacent channel ratios
- the effects on interference of new RF structure & usage levels.

The above leads to the need for carefully defined spectrum usage rights: The existing usage of an allocation is referred to as specific spectrum usage rights. Further rights, called restrictive spectrum usage rights, are permitted provided that neighbouring users do not suffer increased interference, although initially these rights may be too restrictive to allow efficient usage. Hence it is proposed that new specific spectrum usage rights are developed together by agreement with neighbours, potentially involving compensation elements.

Ofcom's new policy is designed to speed innovation.

2) Legacy licence effects

When used in a mesh network, nodes are effectively acting as both infrastructure (relays) and customer equipment (access).

In a TDD system there is no physical difference in the transmissions from the node, regardless of whether it is at the end or within part of a chain of radio transmissions. However, such systems may fall victim to older legislation concerning the physical elevation of infrastructure transmitters. In short it may transpire that different rules could apply to the same radio, depending on whether it is said to be functioning as a user or backhaul node.

For an FDD system there is a further issue in that the operator has a licence to transmit on the down-link frequencies and the customer equipment is approved for transmissions on the up-link frequency. When acting as a relay, customer equipment then requires some licence extension or some means by which it is considered to be being operated by someone as the agent and under the control of the operator, fully in line with the provisions of the appropriate legislation.

These examples are from the UK, but similar circumstances may prevail elsewhere.

D. Summary

To precis, we note specifically that Ofcom are implementing a move away from the old, slow command-and-control method of spectrum management towards market mechanisms. This should be expected to enable faster innovation.

We note also that the softer issues of user perception should not be ignored when planning a service and that radio deployment approval processes need to be updated to cope with the capabilities of mesh.

VI. CONCLUSIONS

The main contribution of this paper is the rationalisation and clarification of many competing mesh performance claims within the literature for the specific case of mobile meshes below 3.5GHz. This is important since such technical claims can form the basis of economic and policy planning. We both examined existing work and used analysis from practical first principles.

From the hypotheses tested we conclude that mesh subscribers cannot self-generate capacity at a rate sufficient to maintain a target level of per-user throughput, regardless of network size and population. We further conclude that there remain fundamental trade-offs between throughput, capacity and delay, which cannot be dismissed easily. Thirdly, whilst network capacity can be improved through the use of directional antennas, for handheld devices the extent of directionality is limited since the high side lobes levels associated with such small antennas severely limit the improvements in spatial re-use that would otherwise be possible. Finally we note that spectrum utilisation could be improved by operating meshes within higher, less precious spectrum, e.g. up to 6GHz.

Noting that coverage extension is a likely application for mesh, we note that the quality of service within a mesh is user dependent, in terms of both mobility and traffic level. Unless specific additional steps are taken to mitigate this e.g. via additional fixed nodes, an operator will be unable to provide service level agreements. Such mitigation will need to be tuned to the actual mobility and traffic circumstances in each case, if they are known. The provision of appropriate additional infrastructure could also be arranged to provide scalability, so forming a viable access mesh.

Innovation could be encouraged by attending to technology enablers, which consist of many protocol and modelling aspects. The softer issues of user perception and behaviour become relevant for some unique mesh system properties which are not present under regular cellular systems. Barriers to innovation which can be removed are the tight, slow command-and-control model of spectrum management in favour of market mechanisms. Unforeseen barriers may also exist within radio legislation which predates mesh networking proposals.

In summary, we still believe that whilst mobile meshes do not live up to all the claims in the literature, they can be very beneficial in the area of coverage extension. In fact, hopping around corners or obstacles is perhaps the greatest benefit offered by a mesh. We suggest that meshes are equally applicable to extending the coverage of WLAN hotspots as they are to cellular multi-hopping, within certain limitations, and should be seen as integral to any 4G or 'beyond 3G' vision. They should also find application in home and office indoor networking and community networks.

ACKNOWLEDGMENTS

The authors would like to thank William Webb of Ofcom for his encouragement and colleagues at Plextek (Graham Maile, Bill Garland, Peter Massam, David Cox), Ofcom (Stephen Jones) and UCL (Lionel Sacks, John Mitchell) for their inputs to the project.

REFERENCES

[1] Methley S G, Crisp M, Newman J, Rio M, Bhatti S, Ramsdale P, *"Study of Efficient Mesh Networks"*, Ofcom Contract Ref. C31400/001, Ofcom, to be published.

[2] Grossglauser M, Tse D, "Mobility increases the capacity of ad hoc networks" *Proceedings of IEEE INFOCOM '01*, Anchorage, Alaska, April 2001.

[3] Gupta P and Kumar P R, "The capacity of wireless networks", *IEEE Trans. Information Theory*, vol 46, pp. 388-404, Mar. 2000.

[4] Shepard T, A Channel Access Scheme for Large Dense Packet Radio Networks", *Proceedings of ACM SIGCOMM'96*, August 1996, Stanford University, California.

[5] Negi R, Rajeswaren A, "Capacity of power constrained ad-hoc networks," *Proc. IEEE INFOCOM*, Hong Kong, China, Mar 2004.

[6] Jones S, Levine P and Rickman N, *"The Economics of Radio Spectrum"*, Ofcom report, 2003. Also *Fifth Wireless World Conference*, 15-16 July 2004, University of Surrey, Guildford, UK.

[7] Sharma G, Mazumdar R R, "On Achievable Delay/Capacity Trade-offs in Mobile Ad Hoc Networks," *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, Cambridge, UK, March 2004.

Plextek

[8] "*High Performance Mesh for Dense Metro WLAN*", MeshDynamics white paper available at <http://www.meshdynamics.com/wPresentations.html>

[9] Arpacioglu O, Zygmunt J.H : "On the Scalability and Capacity of Wireless Networks with Omnidirectional Antennas". *Third International Symposium on Information Processing in Sensor Networks* (IPSN 2004), Berkeley, California (April 27–28, 2004)

[10] Yi S, Pei Y, Kalyanaraman S, "On the capacity improvement of ad hoc wireless networks using directional antennas", *Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, Annapolis, Maryland, USA, June 01-03, 2003.

[11] Esseling N, Weiss E, Kramling A, Zirwas W, "A Multi Hop Concept for HiperLAN/2: Capacity and Interference", Proc. European Wireless 2002, Vol. 1, p.p. 1-7, Florence, Italy, 02/2002.

[12] Royer E M, Melliar-Smith P M, Moser L E, "An analysis of the optimum node density for ad hoc mobile networks", *Proc. IEEE Comms*, June 2001.

[13] Kleinrock L, Silvester J, "Optimum transmission radii for packet radio networks or why six is a magic number", *Proc. IEEE Nat. Telecoms Conf.*, pp4.3.1-5, Birmingham, Alabama, Dec. 1978.

[14] Hsieh H-Y and Sivakumar R, "Performance Comparison of Cellular and Multi-hop Wireless Networks: A Quantitative Study", *Proceedings of ACM SIGMETRICS*, June 2001.

[15] Nilsson A, "Performance analysis of traffic load and node density in ad hoc networks", *Proc. European Wireless (EW2004)*, Barcelona, Spain, February 24-27, 2004.

[16] Lungaro P, and Wallin E, "Coverage, Capacity and QoS Tradeoffs in Hybrid Multi-Hop Ad Hoc Cellular Access Systems", *Proceedings of the Affordable Wireless Services and Infrastructure, 1st Annual Workshop*, Tammsvik, Bro, Sweden, 3-4 June 2003.

[17] Lungaro P, "Coverage and capacity in hybrid multihop ad hoc cellular", Master of Science Thesis, TRITA-S3-RST-0315, Radio Dept. KTH, Stockholm, 2003.

[18] Salem N, Buttyan L, Hubaux J, Jakobsson, M. "A Charging and Rewarding Scheme for Packet Forwarding" ACM *MobiHoc*, Annapolis, Maryland, USA, 2003.

[19] Cave M, "Review of Radio Spectrum Management - An independent review for Department of Trade and Industry and HM Treasury", 6 March 2002. Available via http://www.ofcom.org.uk.

[20] Spectrum Framework Review, Ofcom, 23 November 2004. Available via http://www.ofcom.org.uk.

[21] Spectrum Framework Review Statement, Ofcom, 28 June 2005. Available via http://www.ofcom.org.uk.

[22] Lad M, Bhatti S, Kirstein P, *"Challenges: Opportunities For Coalition-Based Community Networking"*, submitted to the 11th Annual International Conference on Mobile Computing and Networking. Paper: 1568957932, 2005

Appendix G Comparison Guidelines for Fixed and Mobile Meshes

Ofcom commissioned two mesh reports which were completed independently, one focussed on fixed meshes and one on mobile meshes. This appendix contains some guidelines for those readers who wish to compare the details of each report's findings. The two reports were:

- SES-2004-1a Fixed Meshes, led by CCLRC. This study focused on fixed mesh networks operating at frequencies in the range 5.5-10GHz, but with some reference to systems operating up to 42GHz.
- SES-2004-1b Mobile Meshes, led by Plextek Ltd. This study focused on mobile mesh networks operating at frequencies below 3.5GHz, but with some reference to operation up to circa 6GHz.

G.1 Guidelines for Comparison of Results

Detailed comparison of results should be made carefully and may not be valid in some cases. This is primarily because:

- The fixed mesh study primarily relates to providing fixed wireless access (FWA) services between one or more Points of Presence (POPs) and fixed customer premises equipment (CPE) having roof-top mounted antennas. Furthermore, by considering only frequencies above 5.5GHz it relates to radio paths which are predominantly line of sight. On this basis the fixed-PMP benchmark will have relatively poor coverage to customer premises due to shadowing. For example, the report cites ITU data on LOS statistics suggesting circa 50% probability of coverage in rural environments and 30% in urban environments.
- In contrast to the above, the mobile mesh study has, as its PMP benchmark, a standard cellular network (e.g. GSM or 3G) in which a fixed base station (POP-equivalent) serves mobile subscriber terminals. Such a network has a very high coverage availability to subscribers of circa 95%.

Thus the benchmark PMP case for the fixed study is not the same as the benchmark PMP case for the mobile study and is therefore not valid for direct comparison purposes.

In addition to the PMP baseline/frequency-of-use difference, there is also the aspect of mobility itself, which leads to a collection of issues specifically for the mobile mesh, for example route discovery and maintenance is a dynamic problem. However, attempting a comparison by simply ignoring the mobility overheads is to miss the point; a 'frozen mobile' mesh is not directly comparable to a fixed mesh (e.g. in terms of efficiency), not least since different design constraints will already have been employed in the two approaches.

Where the study areas of the reports overlapped, they were in agreement. Examples include:

- Coverage is the great benefit offered by meshes
- Seed nodes are needed for initial roll-out of service
- Routing can be problematical to achieve efficiently
- Latency can quickly become an issue for real time traffic
- If practically available, antenna directionality is helpful

G.2 Summary

The two mesh reports are in agreement with respect to their findings. Nevertheless, care should be taken with the details of a direct comparison.