UNIVERSIDAD DE CANTABRIA

DEPARTAMENTO DE INGENIERÍA DE COMUNICACIONES

TESIS DOCTORAL

CONTRIBUTION TO THE CROSS-LAYER OPTIMIZATION OF INTRA-CLUSTER COMMUNICATION MECHANISMS IN PERSONAL NETWORKS

Autor: *Luis Sánchez González*

Director: *Luis Muñoz Gutiérrez*

Santander, diciembre 2008

Contribution to the Cross-Layer Optimization of Intra-Cluster Communication Mechanisms in Personal Networks

Tesis que se presenta para optar al título de

Doctor por la Universidad de Cantabria

Autor: Luis Sánchez González

Director: Luis Muñoz Gutiérrez

Programa Oficial de Posgrado en Tecnologías de la Información y Comunicaciones en Sistemas de Telecomunicación

Departamento de Ingeniería de Comunicaciones

Escuela Técnica Superior de Ingenieros Industriales y de Telecomunicación

Universidad de Cantabria

Diciembre 2008

RESUMEN

En el futuro, los dispositivos digitales formarán parte del entorno en el que las personas se desenvuelvan, participarán en nuestros objetivos y necesidades y nos ayudarán a "*hacer más haciendo menos*". A través de los dispositivos portátiles o aquellos que estén embebidos a nuestro alrededor el usuario será capaz de interactuar con el futuro universo de servicios e infraestructuras ubícuas. El principal paradigma que se seguirá se basa en que este universo estará centrado en el usuario ya que éste demandará los servicios que más le convengan en cualquier momento y lugar, todo ello preservando nuestra privacidad y seguridad. Este nuevo concepto no sólo se aplica a un entorno de ocio sino que en el campo profesional las redes inalámbricas de próxima generación permitirán incrementar nuestra productividad, reduciendo el peso de tareas repetitivas, poniendo a nuestra disposición la información relevante en el momento adecuado y según sean las necesidades particulares del usuario en ese momento y permitiéndonos trabajar con otras personas independientemente de donde se encuentren. En particular, se intuye que las redes de próxima generación se diseñen de forma que aglutinen todos los servicios disponibles a través de los diferentes sistemas que existan de forma que las posibles deficiencias de alguno de ellos se vean compensadas por otro. Lo que se pretende conseguir es que el usuario pueda disfrutar en todo momento y lugar de los servicios que desee sin que ello suponga un esfuerzo.

Este concepto implica diferentes retos tecnológicos y la integración de múltiples sistemas. Dentro de estos retos tecnológicos esta Tesis aborda los siguientes: soporte de la heterogeneidad en lo referente a las tecnologías de acceso radio que existen y que eventualmente aparecerán en el futuro y que coexistirán en un mismo terminal; desarrollo

de técnicas de optimización basadas en la cooperación entre diferentes capas de la pila de protocolos; implementación de estrategias de selección de la red que mejor pueda soportar un determinado servicio ante la posibilidad de utilización de múltiples tecnologías de acceso; optimización del uso de recursos energéticos en las comunicaciones dentro de la red; protección de la privacidad y la seguridad de las comunicaciones personales del usuario.

Desde el punto de vista de las aportaciones, en esta Tesis se ha contribuido mediante el diseño, implementación y validación de una serie de técnicas de optimización de las comunicaciones en redes de dispositivos móviles basadas en información intercapa. Para ello, se propone una arquitectura de protocolos novedosa que permite soportar la heterogeneidad en términos de tecnologías de acceso dentro del mismo terminal. Igualmente, se desarrollan una serie de técnicas basadas en optimización intercapa que permiten una gestión eficiente de los recursos disponibles así como una mejora en el uso de la energía. Finalmente, se implementan los mecanismos de autenticación y encriptación que permiten asegurar las comunicaciones dentro de la red. El mayor avance respecto del estado del arte se centra en habilitar al usuario para que utilice de manera transparente, eficiente y segura los dispositivos que tiene a su alrededor independientemente de la heterogeneidad que presenten y sin requerir de un conocimiento intensivo de la tecnología. El usuario podrá comunicarse haciendo un uso óptimo de los recursos a su alcance sin preocuparse de tener que gestionarlos él mismo.

Todas estas contribuciones se sustentan en una estructura de documento compuesta por seis capítulos. Así, tras introducir los objetivos y motivación del trabajo, en el Capítulo 2 se incluye el estado del arte relativo a los desarrollos realizados. Fundamentalmente este capítulo resume cuáles son los mecanismos que hay actualmente propuestos como solución a los retos anteriormente descritos.

En el Capítulo 3 se describe la arquitectura de protocolos en la que se basan los trabajos realizados en esta Tesis. El concepto de aislar las capas superiores de la pila de protocolos de las tecnologías de acceso subyacentes se consigue a través de una Capa de Convergencia Universal (UCL, en sus siglas en inglés). Se ha generado una arquitectura capaz de gestionar de manera inteligente y transparente para el usuario y la red todos los interfaces de acceso de los que disponga un dispositivo. Igualmente, se ha descrito en detalle las funcionalidades de los módulos que forman parte de la UCL. El diseño y la especificación esta arquitectura así como de los bloques funcionales que la componen son la primera contribución que se hace en esta Tesis. La UCL supone el marco en el que el resto de técnicas de optimización que se presentan han sido desarrolladas.

El Capítulo 4 recoge los resultados de las evaluaciones realizadas en relación a los mecanismos de selección dinámica de interfaz implementados. El diseño, implementación y validación de estos mecanismos supone la segunda contribución en esta Tesis. El empleo de técnicas de optimización basadas en información procedentes de diferentes capas de la pila de protocolos es la base de los dos mecanismos que se han propuesto. El primero de ellos se basa en la selección dinámica de la tecnología de acceso a utilizar para obtener un rendimiento óptimo del sistema. La segunda estrategia de optimización consiste en el uso

simultáneo de varias tecnologías de acceso para conseguir una mejora en las prestaciones de la red. Aparte de la optimización en cuanto al rendimiento en términos de ancho de banda y calidad de servicio, se ha evaluado la mejora de la eficiencia energética conseguida gracias a las soluciones propuestas. Los resultados obtenidos permiten concluir que las propuestas realizadas en el marco de esta Tesis representan una optimización tanto en parámetros de calidad de servicio como en la eficiencia energética del sistema.

Los mecanismos de descubrimiento y autenticación de dispositivos así como los encargados de asegurar la privacidad de las comunicaciones se evalúan en el Capítulo 5. Se trata de la última contribución incluida en la Tesis. Al contrario que las anteriores acepciones de las redes de área personal, el concepto de red de dispositivos personales considerado no se limita al puro ámbito de la cobertura radioeléctrica sino que se extiende por el concepto de pertenencia a un usuario. En este sentido las relaciones de confianza entre los dispositivos de una persona se han utilizado para crear una red segura en la que sólo los terminales que tengan esa confianza puedan participar y que lo hagan manteniendo en todo momento la privacidad de sus comunicaciones. La evaluación llevada a cabo permite cuantificar el impacto que tienen estos mecanismos sobre el rendimiento del sistema así como proceder a compararlo con la solución más utilizada hoy en día. Los resultados que se derivan a raíz del análisis llevado a cabo no hacen sino confirmar que las prestaciones de las técnicas propuestas son superiores a las del resto de alternativas analizadas.

Por último, en el Capítulo 6 se resumen las principales contribuciones y conclusiones a las que se llega en esta Tesis así como se presentan las posibles líneas de trabajo que se abren tras su finalización.

En los diferentes anexos se describe el marco en el que las soluciones propuestas se han implementado sobre plataformas reales lo cual constituye otra de las grandes aportaciones realizadas en esta Tesis.

ABSTRACT

In the future, computation will be human-centred: it will enter the human world, handling our goals and needs and helping us to do more by doing less. Next generation wireless systems should provide the user access with a broad range of services in a transparent way, independently of user location, by making the technology invisible and embedded in the natural surroundings. New systems will boost our productivity. They will help us automate repetitive human tasks, control a wide range of physical devices in our environment, find the information we need (when we need it, without obliging us to examine thousands of search-engine hits), and enable us to work together with other people through space and time.

The achievement of this paradigm led to the identification of a set of optimizations in intra-cluster communications that were needed to fully support it. Firstly, heterogeneity will be a fundamental characteristic of next generation wireless communications since more and more personal devices are equipped with multiple network access technologies so that the user can have access to the different services that the different operational environments provide. However, Next Generation Networks (NGN) will comprise such a diverse number of possibilities that the users cannot be expected to take technical decisions on their own. It is necessary to provide mechanisms that intelligently select the optimal available access network based on context information such as user preferences, power consumption, link quality, etc. Finally, users need to trust the system that supports their personal communications. Within a personal network the most confidential information might be exchanged and the user need to be sure that this will never be

disclosed. If the system fails in these features, NGN in general and PNs in particular will never happen.

This Thesis has contributed with the development of the mechanisms that tackle the abovementioned challenges. The design and specification of a convergence framework, the so-called Universal Convergence Layer (UCL), has been the first topic addressed. This framework aims to manage all the network access interfaces with which a device is equipped so that they can be transparently used by upper layers as if the node were equipped with a single access technology. On the other hand, the UCL enables the cross-layer optimization paradigm. Its privileged location within the protocol stack gives the UCL the possibility to support both bottom-up and top-down information flow. In this sense, two different solutions based on cross-layer optimization have been proposed to enhance the performance and energy efficiency of the system. Finally, the UCL also plays a key role in security issues as an enabler for providing link-layer security mechanisms that ensure data confidentiality and integrity, authenticity and non-repudiation. The techniques implemented for node authentication combined with traffic encryption in ad-hoc networks have been thoroughly assessed and have demonstrated their appropriateness.

The biggest advance in the state-of-the-art comes from enabling the user to have easy, affordable and seamless control of their devices over heterogeneous communications networks. They are empowered to communicate efficiently and securely with their selected interaction groups, no matter what kind of access is available for them to use.

All these contributions are structured as follows: After introducing the main objectives and the motivation of the work at hand in Chapter 1, a review of the current state-of-the-art is provided in Chapter 2. Fundamentally, this chapter will cover the solutions that are currently proposed for handling the problems described above.

In Chapter 3, the architecture of the framework proposed to optimize intra-cluster communications will be presented. The concept of isolating the upper-layers from underlying wireless technologies, thus providing real multi-mode, can be achieved by introducing the UCL. The high-level protocol architecture of the UCL will be presented and the main building blocks of the architecture will be described along with the related mechanisms implemented within them. The main work developed as part of this Thesis has been to provide detailed understanding of the UCL architecture and its building blocks not only limited to the specification of the architecture and functionalities but also by implementing a prototype over real platforms.

Chapter 4 will further develop the mechanisms implemented within the UCL to improve the system performance. The focus has been placed on two algorithms that are based on cross-layer optimization. The first one deals with the selection at run-time of the most appropriate wireless interface to be used in order to improve the system performance. The second one leverages the striping concept in order to exploit all the network interfaces available. The main aim of this chapter is to present the benefits in terms of performance optimization and power efficiency obtained through intelligent selection of the most

appropriate wireless interface and in particular to comment on the results derived from the experimental and analytical validation carried out.

Neighbour discovery and authentication and intra-cluster communications privacy protection mechanisms will be presented in Chapter 5. In contrast to other descriptions of cluster or Personal Area Network that limit the concept to a matter of radio coverage, the concept of cluster proposed in Personal Networks architecture relies on the trust relationships established between the cluster constituents. The analytical and experimental validation, which also includes comparison analyses with competing solutions, presented in this chapter, allow quantifying the impact of the proposed techniques on the system performance.

Finally, Chapter 6 summarizes the main contributions of the thesis and concludes this work. An overview of opened research opportunities of interest will also be presented.

The implementation framework and the system in which the UCL has been integrated are briefly described in the Annexes.

TABLE OF CONTENTS

ACRONYMS

2G	Second Generation
3DES	Triple Data Encryption Standard
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
AC	Always Cheapest
ACK	Acknowledgement
AES	Advanced Encryption Standard
AH	Authentication Header
AI	Ambient Intelligence
AIPN	All-IP Networks
AN	Ambient Network
ANS	Ad hoc Network System
AODV	Ad hoc On demand Distance Vector protocol

AP	Access Point
API	Application Programming Interface
ARI	Ambient Resource Interface
ARP	Address Resolution Protocol
ARQ	Automatic Repeat Request
AS	Authentication Server
ASI	Ambient Service Interface
aT	Advanced Terminal
BER	Bit Error Rate
BGP	Border Gateway Protocol
BNEP	Bluetooth Networking Encapsulation Protocol
BO	BackOff
BRAN	Broadband Radio Access Networks
bT	Basic Terminal
BW	Bandwidth
CA	Certificate Authority
CALA	Context Access Language
CAN	Community Area Network
CBC	Chain Block Chaining Mode
CDMA	Code Division Multiple Access
CHAP	Challenge Handshake Authentication Protocol
CMG	Context Management Gateway
CMN	Context Management Nodes
CPFP	Certified Private Personal Area Network Formation Protocol
CPU	Central Processing Unit
CRC	Cyclic Redundancy Check
DES	Data Encryption Standard
DIFS	DCF Interframe Space
DLC	Data Link Control
DNS	Domain Name System
DoS	Denial of Service

DSR	Dynamic Source Routing
EAP	Extensible Authentication Protocol
EAP	Extensible Authentication Protocol over Secure Remote Password
EAP-TLS	Extensible Authentication Protocol over Transport Layer Security
EAP-TTLS	Extensible Authentication Protocol over Tunnelled Transport Layer Security
ECB	Electronic Codebook Mode
ECN	Explicit Congestion Notification
ESP	Encapsulating Security Payload
ETCP	Extended Transport Control Protocol
ETSI	European Telecommunications Standards Institute
ETX	Expected Transmission Count Metric
FEC	Forward Error Correction
FER	Frame Error Rate
FIFO	First In First Out
FTP	File Transfer Protocol
GI	Generalized Identifier
GUI	Graphical User Interface
GW	Gateway
HI	Host Identity
HIP	Host Identity Protocol
HiSWANa	High-Speed Wireless Local Area Network
HLP	Host Layer Protocol
HSDPA	High Speed Downlink Packet Access
HTTP	HyperText Transfer Protocol
HW	Hardware
IANA	Internet Assigned Numbers Authority
ICMP	Internet Control Message Protocol
ID	Identifier
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force

IKE	Internet Key Exchange
IP	Internet Protocol
IPSec	IP Security
IrDA	Infrared Data Association
IRTF	Internet Research Task Force
ISAKMP	Internet Security Association and Key Management Protocol
ISP	Internet Service Provider
IST	Information Society Technologies
ITU	International Telecommunication Union
IV	Initialization Vector
L2CAP	Logical Link Control and Adaptation Protocol
LAN	Local Area Network
LCD	Liquid Crystal Display
LEAP	Lightweight Extensible Authentication Protocol
LIN6	Location Independent Networking for IPv6
LKM	Loadable Kernel Module
LMSK	Link Master Session Key
MAC	Medium Access Control
MADM	Multiple Attribute Decision Making
MAGNET	My Adaptive Global NETwork
MANET	Mobile Ad hoc NETwork
MBWA	Mobile Broadband Wireless Access Networks
MHTP	Multi-Homing Transport Protocol
MIT	Massachusetts Institute of Technology
MMAC	Multimedia Mobile Access Communication
MOO	Multiple Objective Optimization
MOPED	Mobile Grouped Device
MSMP	MAGNET Service Management Protocol
MTM	Medium Time Metric
MTU	Maximum Transmission Unit
NAS	Network Access Server

NAT	Network Address Translation
NDISC	Neighbour Discovery
NDMA	Network-assisted Diversity Multiple Access
NGN	Next Generation Networks
NGWS	Next Generation Wireless Systems
NIST	National Institute of Standards and Technology
OFB	Output Feedback Mode
OS	Operating System
OSI	Open System Interconnection
P2P	Peer-to-Peer
PACWOMAN	Power Aware Communications for Wireless Optimised Personal Area Network
PAN	Personal Area Network
PC	Personal Computer
PCBI	Protocol Control Block Identifier
PDA	Personal Digital Assistant
PDU	Protocol Data Unit
PEAP	Protected Extensible Authentication Protocol
PEP	Performance Enhancing Proxy
PER	Packet Error Rate
PFP	Private Personal Area Network Formation Protocol
PGP	Pretty Good Privacy
PHY	Physical Layer
PILC	Performance Implications of Link Characteristics
PKC	Public Key Cryptography
PKI	Public Key Infrastructure
PLCP	Physical Layer Convergence Procedure
PMH	Personal Mobile Hub
PMK	Primary Master Key
PN	Personal Network
PNDS	Personal Network Directory Service

PN-F	Personal Network Federation
POS	Personal Operating System
P-PAN	Private Personal Area Network
PPP	Point-to-Point Protocol
PUCC	Peer-to-Peer Universal Computing Consortium
QoS	Quality of Service
RADIUS	Remote Authentication Dial-In User Service
RAM	Random Access Memory
RAN	Radio Access Network
RC4	Rivest Cipher 4
RD	Radio Domain
RFC	Request For Comments
RSA	Rivest Shamir and Adleman
RTO	Retransmission TimeOut
RTT	Round Trip Time
S/MIME	Secure / Multipurpose Internet Mail Extensions
SA	Security Association
SCMF	Secure Context Management Framework
SCTP	Stream Control Transmission Protocol
SHA	Secure Hash Algorithm
SIFS	Short Inter-Frame Space
SK	Session Key
SMN	Service Management Nodes
SNR	Signal to Noise Ratio
SOHWNE	Service Oriented Heterogeneous Wireless Network Environment
SRP	Secure Routing Protocol
SRR	Surplus Round Robin
SSL	Secure Socket Layer
SW	Software
SYN	SYNchronization
TCP	Transmission Control Protocol

UCL	Universal Convergence Lauer
UDP	User Datagram Protocol
UIA	User Information Architecture
UIP	Unmanaged Internet Protocol
UMTS	Universal Mobile Telecommunications System
UPN	Universal Personal Networking
UWB	Ultra Wide Band
VoIP	Voice over IP
VPN	Virtual Private Network
WAN	Wide Area Network
WCETT	Weighted Cumulative Expected Transmission Time
WEP	Wired Equivalent Privacy
WiFi	Wireless Fidelity
WIG	Wireless Interworking Group
WPAN	Wireless Personal Area Network
WWAN	Wireless Wide Area Network
WWRF	Wireless World Research Forum
XOR	eXclusive OR

REFERENCES

[1] L. Muñoz, R. Agüero, J. Choque, J. A. Irastorza, L. Sánchez, M. Petrova, P. Mähönen, "Empowering Next-Generation Wireless Personal Communication Networks", IEEE Communications Magazine, vol. 42, nº5, pp 64-70, May 2004

[2] I.G. Niemegeers, S. Heemstra de Groot, "From Personal Area Networks to Personal Networks: A user oriented approach", Journal on Wireless and Personal Communications, vol. 22, nº 2, pp 175-186, August 2002.

[3] I.G. Niemegeers, S. Heemstra de Groot, "Personal networks: Ad hoc distributed personal environments," Proceedings from the 1st Annual Mediterranean Ad Hoc Networking Workshop, September 2002.

[4] M. Alutoin, K. Ahola, S. Lehtonen, L. Sánchez, J. Lanza, J. Hoebeke, G. Holderbeke, I. Moerman, M. Girod-Genet, W. Louati, "Self-organisation and mobility in Personal Networks", Proceedings from the 8th International Symposium on Wireless Personal Multimedia Communications, pp. 318–322, September 2005.

[5] M. Alutoin, S. Lehtonen, L. Sánchez, J. Lanza, J. Hoebeke, G. Holderbeke, I. Moerman, D. Zeghlache, W. Louati, "Towards Self-organising Personal Networks", Proceedings from the 1st ACM Workshop on Dynamic Interconnection of Networks pp. 12–16, September 2005.

[6] L. Muñoz, L.Sanchez, J. Lanza, M. Alutoin, S. Lehtonen, D. Zeghlache, M. Girot Genet, W. Louati, I. Moerman, J. Hoebeke, G. Holderbeke, M. Ghader, M. Jacobsson, "A proposal for Self-Organizing Personal Netowrks", Proceedings from the WWRF meeting 15, December 2005.

[7] FP6-IST-IP-507102 'My personal Adaptive Global Net' IST-MAGNET project. www.ist-magnet.org

[8] E. Gustafsson, A. Jonsson, "Always best connected", IEEE Wireless Communications, vol. 10, nº 1 , pp. 49–55, February 2003.

[9] IEEE 802.15 Working Group for WPAN, http://www.ieee802.org/15/

[10] IETF Mobile Ad hoc NETworks (MANET) working group, http://www.ietf.org/html.charters/manet-charter.html

[11] IETF Zero Configuration Networking (Zeroconf) working group, http://www.zeroconf.org/

[12] UPnP™ Forum, www.upnp.org

[13] M. Petrova, et al., "MAGNET Deliverable 2.1.2: Overall secure PN architecture", October 2005.

[14] S. Deering, R. Hinden, "RFC 2460: Internet Protocol, Version 6 (IPv6) Specification", Network Working Group of the Internet Engineering Task Force, December 1998.

[15] R. Braden, "RFC 1122: Requirements for Internet Hosts -- Communication Layers", Network Working Group of the Internet Engineering Task Force, October 1989.

[16] S. Mirzdeh et al., "MAGNET Deliverable D4.3.2: Final version of the Network-Level Security Architecture Specification", March 2005.

[17] E. Kovacs et. al., "MAGNET Beyond Deliverable D1.1.1: MAGNET System Specification", January 2008.

[18] UMTS Forum, Report on Candidate Extension Bands for UMTS/IMT-200 Terrestrial Component, Report No. 7, 1999.

[19] J. D. Day, H. Zimmermann, "The OSI reference model," in Proceedings of IEEE, vol. 71, nº. 12, pp. 1334–1340, December 1983.

[20] M. Ibnkahla, "Signal Processing for Mobile Communications Handbook", CRC Press, pp. 28-1 – 28-26, August 2004.

[21] V. Srivastava, M. Motani, "Cross-layer design: a survey and the road ahead," IEEE Communications Magazine, vol. 43, nº 12, pp. 112–119, December 2005.

[22] S. Shakkottai, T. S. Rappaport, P. C. Karlsson, "Cross-layer design for wireless networks," IEEE Communications Magazine, vol. 41, nº 10, pp. 74–80, October 2003.

[23] X. Lin, N. B. Shroff, R. Srikant, "A tutorial on cross-layer optimization in wireless networks," IEEE Journal on Selected Areas in Communications, vol. 24, nº 8, pp. 1452–1463, August 2006.

[24] B. Zhao, M. C. Valenti, "Practical relay networks: A generalization of hybrid–ARQ," IEEE Journal on Selected Areas in Communications, vol. 23, nº 1, pp. 7–18, January 2005.

[25] S. Udani, J. Smith, "Power Management in Mobile Computing (A Survey)", http://www.cis.upenn.edu/~udani/papers.html. University of Pennsylvania. 1996

[26] M. Stemm, R. H. Katz, "Measuring and reducing energy consumption of network interfaces in hand-held devices", IEICE (Institute of Electronics, Information and Communication Engineers) Transactions on Fundamentals of Electronics, Communications and Computer Science, Vol.E80-B, nº8, pp.1125–1131, August 1997.

[27] A. Chandraskasan, R. W. Broedersen, "Low Power Digital CMOS Design", Kluwer Academic Publishers, June 1995

[28] M. Zorzi, R. R. Rao, "Energy constrained error control for wireless channels", IEEE Personal Communications Magazine, vol. 4, nº 6, pp. 27–33, December 1997.

[29] P. Lettieri, C. Fragouli, M.B. Srivastava, "Low power error control for wireless links", Proceedings from the 3rd annual ACM/IEEE international conference on Mobile computing and networking, pp.139–150, September 1997.

[30] M. Woo, S. Singh, C.S. Raghavendra, "Power Aware routing in mobile ad hoc networks", Proceedings from the 4th annual ACM/IEEE international conference on Mobile computing and networking, pp.181–190, October 1998.

[31] J-H. Chang, L.Tassiulas, "Energy conserving routing in wireless ad-hoc networks", Proceedings from the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 22–31, March 2000.

[32] R. Ramanathan, R. Rosales-Hain, "Topology Control of Multihop Wireless Networks using Transmit Power Adjustment", Proceedings from the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 404–413, March 2000.

[33] W. Arbough, N. Shanker, Y. Wan, "Your 802.11 Wireless Network has no Clothes", available at http://www.cs.umd.edu/waa/wireless.pdf

[34] J. Khan, A. Khwaja, "Building Secure Wireless Networks with 802.11", Wiley, January 2003.

[35] L. Sanchez, J. Lanza, L. Muñoz, J. Perez, "Enabling Secure Communications over Heterogeneous Air Interfaces: Building Private Personal Area Networks", Proceedings from the 8th International Symposium on Wireless Personal Multimedia Communications, pp. 1963–1967, September 2005.

[36] M. Presser, R. Tafazolli, István Z. Kovács, D. Dahlhaus, J. Farserotu, F. Platbrood, L. Sanchez, K. Schoo, "MAGNET 4G Personal Area Network Air-Interfaces for Personal Networks", Proceedings from the 13th IST Mobile & Wireless Summit Communications Summit, vol. 1, pp. 169–175, June 2004.

[37] L. Sanchez, J. Lanza, L. Muñoz, "Experimental Assessment of a Cross-Layer Solution for TCP/IP Traffic Optimization on Heterogeneous Personal Networking Environments", Lecture Notes in Computer Science (Vol 4217), pp. 284–296, September 2006.

[38] J. Lanza, L. Sánchez, L. Muñoz, "Performance Evaluation of a Cross-layer based Wireless Interface Dynamic Selection on WPAN/WLAN Heterogeneous Environments: An Experimental Approach", Proceedings from the 6th International Workshop on Applications and Services in Wireless Networks, pp. 139–146, May 2006.

[39] L. Sanchez, J. Lanza, L. Muñoz, "Self-Configuring Private Personal Area Networks: The first stage towards Personal Networking", Proceedings from the 8th International Symposium on Wireless Personal Multimedia Communications, pp.1983–1987, September 2005.

[40] L. Sánchez, J. Lanza, L. Muñoz, K. Ahola, M. Alutoin, "Securing the communication in Private Heterogeneous Mobile Ad-hoc Networks", To appear on Wireless Personal Communications Journal, Springer, To appear on 2009.

[41] J. Lanza, L. Sánchez, L. Muñoz, "Experimental Comparison of Two Solutions for Securing Heterogeneous Ad-Hoc Network Communications", Proceedings from the 11th International Symposium on Wireless Personal Multimedia Communications, September 2008.

[42] F. Teraoka, M. Ishiyama, M. Kunishi, A. Shionozaki, "LIN6: A Solution to Mobility and Multi-Homing in IPv6", Internet-Draft, Work in progress, August 2001.

[43] M. Py, "Multi Homing Translation Protocol (MHTP)", Internet-Draft, Work in progress, November 2001.

[44] C. Huitema, "Multi-homed TCP", Internet-Draft, expired, May 1995.

[45] R. R. Stewart, Q. Xie, K. Morneault, C. Sharp, H. J. Schwarzbauer, T. Taylor, I. Rytina, M. Kalla, L. Zhang, V. Paxson, "RFC2960: Stream Control Transmission Protocol", October 2000.

[46] P. Nikander, J. Lundberg, C. Candolin, T. Aura, "Homeless Mobile IPv6", Internet-Draft, Work in progress, February 2001.

[47] R. Moskowitz, "Host Identity Payload and Protocol", Internet-Draft, Work in progress, November 2001.

[48] H. Adiseshu, G. Parulkar, G. Varghese, "A reliable and scalable striping protocol", Proceedings from the ACM SIGCOMM, pp. 131–141, August 1996.

[49] C. de Launois, B. Quoitin, O. Bonaventure, "Leveraging network performances with IPv6 multihoming and multiple provider-dependent aggregatable prefixes", The International Journal of Computer and Telecommunications Networking, vol. 50, nº 8, pp. 1145–1157, June 2006.

[50] B. Traw, J. Smith, "Striping within the network subsystem," IEEE Network, vol. 9, nº 4, pp. 22–32, July 1995.

[51] D. Phatak, T. Goff, J. Plusquellic, "IP-in-IP tunneling to enable the simultaneous use of multiple IP interfaces for network level connection striping," Computer Networks, vol. 43, nº 6, pp. 787–804, December 2003.

[52] H.-Y. Hsieh, R. Sivakumar, "A transport layer approach for achieving aggregate bandwidths on multihomed mobile hosts," Proceedings from the 8th Annual International Conference on Mobile Computing and Networking, pp. 83–94, September 2002.

[53] H. Sivakumar, S. Bailey, R. Grossman, "PSockets: The case for application-level network striping for data intensive applications using high speed wide area networks," Proceedings from International Conference for High Performance Computing, Networking, Storage, and Analysis, November 2000.

[54] V. T. Raisinghani, S. Iyer, "Cross-Layer Design Optimizations in Wireless Protocol Stacks", Computer Communications, vol. 27, pp. 720–724, October 2003.

[55] S. Shakkottai, T. S. Rappaport, P. C. Karlsson, "Cross-Layer Design for Wireless Networks," IEEE Communications Magazine, vol. 41, no. 10, pp. 74–80, October 2003.

[56] Z. Ji et al., "Exploiting Medium Access Diversity in Rate Adaptive Wireless LANs", Proceedings from the 10th ACM Annual International Symposium on Mobile Computing and Networks, pp. 345–359, October 2004.

[57] G. Xylomenos, G. C. Polyzos, "Quality of Service Support over Multi-service Wireless Internet Links", Computer Networks, vol. 37, nº 5, pp. 601–615, November 2001.

[58] G. Dimic, N. D. Sidiropoulos, R. Zhang, "Medium Access Control — Physical Cross-Layer Design", IEEE Signal Processing Magazine, vol. 21, nº. 5, pp. 40–50, September 2004.

[59] L. Tong, V. Naware, P. Venkitasubramaniam, "Signal Processing in Random Access," IEEE Signal Processing Magazine, vol. 21, nº 5, pp. 29–39, September 2004.

[60] Q. Liu, S. Zhou, G. B. Giannakis, "Cross-Layer Combining of Adaptive Modulation and Coding with Truncated ARQ Over Wireless Links," IEEE Transactions on Wireless Communications, vol. 3, nº 5, pp. 1746–1755, September 2004.

[61] J. Border, M. Kojo, J. Griner, G. Montenegro, Z. Shelby, "Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations", IETF Network Working Group, June 2001.

[62] E. M. Royer, S. J. Lee, C. E. Perkins, "The effects of MAC protocols on ad hoc networks communication", Proceedings from the IEEE Wireless Communications and Networking Conference, vol. 2, pp. 543–548, September 2000.

[63] O. Arpacioglu, T. Small, Z. J. Haas, "Notes on Scalability of Wireless Ad Hoc Networks", <draft-irtf-and-scalability-notes-00.txt>, August 2003.

[64] G. Carneiro, J. Ruela, M. Ricardo, "Cross-Layer Design in 4G Wireless Terminals", IEEE Wireless Communications Magazine, vol. 11, nº 2, pp. 7–13, April 2004.

[65] A. Doufexi, S. Armour, P. Karlsson, M. Butler, A. Nix, D. Bull, J. McGeehan, "A comparison of the HIPERLAN/2 and IEEE 802.11a wireless LAN standards", IEEE Communications Magazine, vol. 40, nº 5, pp. 172–180, May 2002.

[66] T. S. Rappaport, A. Annamalai, R. M. Buehrer, W. H. Tranter, "Wireless communications: Past events and a future perspective", IEEE Communications Magazine, vol. 40, nº 5, pp. 148–161, May 2002.

[67] J. McNair, F. Zhu, "Vertical Handoffs in Fourth-Generation Multinetwork Environments", IEEE Wireless Communications Magazine, vol. 11, nº 3, pp. 8–15, June 2004.

[68] J. Ylitalo, T. Jokikyyny, T. Kauppinen, A. J. Tuominen, J. Laine, "Dynamic Network Interface Selection in Multihomed Mobile Hosts", Proceedings from the 36th Hawaii International Conference on System Sciences, January 2003.

[69] E. Bircher, T. Braun, "An Agent-Based Architecture for Service Discovery and Negotiations in Wireless Networks", Proceedings from the 2nd International Conference on Wired/Wireless Internet Communications, February 2004.

[70] H. J. Wang, R. H. Katz, J. Giese, "Policy-enabled Handoffs Across Heterogeneous Wireless Networks", Proceedings from the 2nd IEEE Workshop on Mobile Computing Systems and Applications, February 1999.

[71] L.-J. Chen, T. Sun, B. Chen, V. Rajendran, M. Gerla, "A Smart Decision Model for Vertical Handoff", Proceedings from the 4th International Workshop on Wireless Internet and Reconfigurability, May 2004.

[72] J. Murphy, L. Murphy, "Bandwidth Allocation By Pricing In ATM Networks", IFIP Transactions C-24, pp. 333–351, Broadband Communications II, March 1994.

[73] M. Angermann, J. Kammann, "Cost Metrics For Decision Problem In Wireless Ad Hoc Networking", Proceedings from the IEEE CAS Workshop on Wireless Communications and Networking, September 2002.

[74] G. Le Bodic, J. Irvine, D. Girma, J. Dunlop, "Dynamic 3G Network Selection for Increasing the Competition in the Mobile Communications Market", Proceedings from the 52nd IEEE Vehicular Technology Conference, vol. 3, pp. 1064–1071, September 2000.

[75] M. Falkner, M. Devetsikiotis, I. Lambadaris, "An Overview of Pricing Concepts for Broadband IP Networks", IEEE Communication Surveys and Tutorials, pp. 2–13, September 2000.

[76] X. Wang, H. Schulzrinne, "An Integrated Resource Negotiation, Pricing, and QoS Adaptation Framework for Multimedia Applications", IEEE Journal on Selected Areas in Communications, Special Issue on Internet QoS, vol. 18, nº 12, pp. 2514–2529, December 2000.

[77] S. Das, H. Lin, M. Chatterjee, "An Econometric Model for Resource Management in Competitive Wireless Data Networks", IEEE Network Magazine, vol. 18, nº 6, pp. 20–26, November/December 2004.

[78] V. Gazis, N. Houssos, N. Alonistioti, L. Merakos "On the Complexity of Always Best Connected in 4G Mobile Networks", Proceedings from the 58th IEEE Vehicular Technology Conference, October 2003.

[79] R.G Gallager, "Energy limited channels: Coding, Multi-access, and Spread Spectrum", Proceedings from the Conference Information Science and Systems, pp. 372, March 1988.

[80] H. El Gamal, et al, "A new approach to layered Space-time Coding and Signal Processing", IEEE Transactions on Information Theory, vol. 47, nº 6, pp. 2321–2334, September 2001.

[81] A. Spyopoulos, et al, "Energy Efficient Communications in Ad hoc network using Directional Antennas", Proceedings from the 21st IEEE Conference on Computer Communications, pp. 220–228, June 2002.

[82] S.L. Wu, Y. C. Tseng, J. P. Sheu, "Intelligent Medium Access for mobile Ad hoc Network with Busy Tones and Power control", IEEE Journal on Selected Areas on Communication, vol. 18, nº 9, pp. 1647–1657, September 2000.

[83] S.Kandukuri et al., "Power controlled Multiple Access in Wireless Communication Networks", Proceedings from the 18th IEEE Conference on Computer Communications, pp. 386–395, March 2000.

[84] S. Singh, M. Woo, C.S. Raghavendra, "Power-Aware Routing in Mobile Ad Hoc Networks", Proceedings from the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking, October 1998.

[85] M.B.Pursley, et al., "Energy-efficient Routing in Frequency-hop Network with Adaptive Transmission", Proceedings from the IEEE Military Communications Conference, pp.1409–1413, October 1999.

[86] S. Narayanaswamy, et al., "Power control in ad hoc networks: Theory, Architecture, Algorithm and Implementation of the COMPOW Protocol", Proceedings from the European Wireless Conference, pp.156–162, February 2002.

[87] K. Chandran et al., "A Feedback-based scheme for improving TCP performance in ad hoc wireless networks", IEEE Personal Communications, vol. 8, nº 1, pp. 34–39, February 2001.

[88] J.-P. Ebert, B. Stremmel, E. Wiederhold, and A. Wolisz, "An Energy-efficient Power Control Approach for WLANs", Journal of Communications and Networks (JCN), vol. 2, nº. 3, pp. 197-206, September 2000.

[89] S-B. Lee, G. Seop, et al., "Improving UDP and TCP Performance in Mobile ad hoc networks with INSIGNIA", IEEE Communication Magazine, vol. 39, nº 6, pp. 156–165, June 2001.

[90] B. Guttman, E. Roback, "An Introduction to Computer Security: The NIST Handbook", Special Publication 800-12, October 1995

[91] N. Borisov, I. Goldberg, D. Wagner, "Intercepting Mobile Communications: The Insecurity of 802.11", Proceedings from the 7th Annual International Conference on Mobile Computing and Networking, July 2001.

[92] A. Freier, P. Karlton, P. Kocher, "The SSL Protocol Version 3.0", November 1996.

[93] S. Blake-Wilson, M. Nystrom, D. Hopwood, J. Mikkelsen, T. Wright, "RFC 3546: Transport Layer Security (TLS) Extensions", June 2003.

[94] T. Dierks, E. Rescorla, "RFC 5246: The Transport Layer Security (TLS) Protocol Version 1.2", August 2008.

[95] B. Ramsdell, "RFC 3851: Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification", July 2004.

[96] S. Garfinkel, "PGP: Pretty Good Privacy", O'Reilly & Associates, December 1994.

[97] S. Kent, K. Seo, "RFC 4301: Security Architecture for the Internet Protocol", December 2005.

[98] S. Kent, "RFC 4302: IP Authentication Header", December 2005.

[99] S. Kent, "RFC 4303: IP Encapsulating Security Payload (ESP)", December 2005.

[100] C. Kaufman, Ed., "RFC 4306: Internet Key Exchange (IKEv2) Protocol", December 2005.

[101] L. Zhou, Z. J. Haas, "Securing ad hoc networks", IEEE Network, vol. 13, nº 6, pp. 24–30, November/December 1999.

[102] H. Yang, H. Luo, F. Ye, S. Lu, L. Zhang, "Security in Mobile Ad Hoc Networks: Challenges and Solutions", IEEE Wireless Communications Magazine, vol. 11, nº 1, pp. 38–47, February 2004.

[103] Y.-C. Hu, A. Perrig, D. B. Johnson, "Ariadne: A secure on-demand routing protocol for ad hoc networks", Proceedings from the 8th ACM International Conference on Mobile Computing and Networking, September 2002.

[104] M. Zapata, N. Asokan, "Securing Ad Hoc Routing Protocols", Proceedings from the 1st ACM workshop on Wireless Security, pp. 1–10, September 2002.

[105] Y. Hu, D. Johnson, A. Perrig, "Sead: Secure Efficient Distance Vector Routing for Mobile Wireless Ad Hoc Networks", Proceedings from 4th IEEE Workshop on Mobile Computing Systems and Applications, June 2002.

[106] M. Guerrero, "Secure ad hoc on-demand distance vector (SAODV) routing", August 2001. INTERNET-DRAFT draft-guerrero-manet-saodv-00.txt.

[107] P. Papadimitratos, Z.H. Haas, "Secure Routing for Mobile Ad Hoc Networks." Proceedings from the Communication Networks and Distributed Systems Modeling and Simulation Conference, January 2002.

[108] S. Marti, T.J. Giuli, K. Lai, M. Baker, "Mitigating Routing Misbehavior in Mobile Ad Hoc Networks", Proceedings of 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 255–265, August 2000.

[109] H. Yang, X. Meng, S. Lu, "Self-Organized Network Layer Security in Mobile Ad Hoc Networks", IEEE Journal on Selected Areas in Communications, vol. 24, nº 2, pp. 261–273, February 2006.

[110] B. Awerbuch et al., "An On-Demand Secure Routing Protocol Resilient to Byzantine Failure", Proceedings from the 1st ACM Workshop on Wireless Security, pp. 21–30, September 2002.

[111] http://www.wireless-world-research.org/

[112] Ambient Networks (AN), http://www.ambient-networks.org/.

[113] IST PACWOMAN - Power Aware Communications for Wireless Optimised Personal Area Networks, http://www.imec.be/pacwoman/Welcome.shtml.

[114] F. Louagie, L. Muñoz, S. Kyriazakos, "Paving the Way for the Fourth Generation: A New Family of Wireless Personal Area Networks", Proceedings from the 12th IST Mobile & Wireless Communications Summit, June 2003.

[115] MyNet, http://projects.csail.mit.edu/nrcc/mynet-uia.php.

[116] F. Kaashoek, R. Morris, "User-Relative Names for Globally Connected Personal Devices", Proceedings from the 5th International Workshop on Peer-to-Peer Systems, February 2006.

[117] B. Ford, "Unmanaged Internet Protocol: Taming the Edge Network Management Crisis", Proceedings from the 2nd Workshop on Hot Topics in Networks, November 2003.

[118] Universal Computing Consortium (PUCC), http://www.pucc.jp/.

[119] K. Braun, J. Grollman, M. Horn, H. Raffler, W. Thulke, W. Weigel, "Universal Personal Networking", Proceedings from the 2nd International Conference on Universal Personal Communications, October 1993.

[120] The Siemens LifeWorks Concept, White Paper, http://www.siemensenterprise.com/attachments/2gip/LifeWorksWhitePaper.pdf, Accessed March 2008.

[121] D. Husemann, C. Narayanaswa, M. Nidd, "Personal Mobile Hub", Proceedings from the 8th IEEE International Symposium on Wearable Computers, October 2004.

[122] R. Kravets, C. Carter, L. Magalhaes, "A Cooperative Approach to User Mobility", ACM Computer Communications Review, vol. 31, nº 5, pp 57–69, October 2001.

[123] P. Debaty, D. Caswell, "Uniform Web Presence Architecture for People, Places, and Things", IEEE Personal Communications, vol. 8, nº 4, pp. 46–51, August 2001.

[124] P. Maniatis, M. Roussopoulos, E. Swierk, K. Lai, G. Appenzeller, X. Zhao, M. Baker, "The Mobile People Architecture", ACM Mobile Computing and Communications Review, vol. 3, nº 3, pp. 36–42, July 1999.

[125] 3rd Generation Partnership Project (3GPP), "Service requirements for Personal Network Management (PNM) - Stage 1", Technical Specification, 3GPP TS 22.259 V8.3.0 (2006-06), March 2007.

[126] D. M. Kyriazanos et. al., "MAGNET Beyond Deliverable D4.3.2: Specification of user profile, identity and role management for PNs and integration to the PN platform", March 2007.

[127] J. Daemen, V. Rijmen, "FIPS PUB 197: Advanced Encryption Standard (AES)", National Institute of Standards and Technology (NIST), November 2001.

[128] J. Reynolds and J. Postel, "RFC 1700: Assigned Numbers", October 1994.

[129] T. Narten, E. Nordmark, W. A. Simpson, "RFC 2461: Neighbor Discovery for IP Version 6 (IPv6)", December 1998.

[130] G. Krishnamurthi, "Requirements for CAR Discovery Protocols," Internet-Draft, Work in progress, May 2002.

[131] IST-GOLLUM - Generic Open Link-Layer API for Unified Media access. http://www.ist-gollum.org/

[132] L. Sanchez, J. Lanza, R. Olsen, M. Bauer, M. Girod-Genet, "A Generic Context Management Framework for Personal Networking Environments", 3rd Annual International Conference on Mobile and Ubiquitous Systems, pp. 1–8, July 2006.

[133] Federal information processing standards publication (fips 197), "Advanced Encryption Standard (AES)", 2001. http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

[134] A. Rawat, P.D. Vyavahare A. K. Ramani, "Evaluation of Rushing Attack on Secured Message Transmission (SMT/SRP) protocol for Mobile Ad-Hoc Networks", Proceedings from the International Conference on Personal Wireless Communications, January 2005.

[135] G. Stoneburner, A. Goguen, A. Feringa, "Risk Management Guide for Information Technology Systems", Recommendations of the National Institute of Standards and Technology, July 2002

[136] Y-C. Hu, A. Perrig, D. B. Jonson, "Rushing Attacks and Defense in Wireless Ad Hoc Network Routing Protocols", Proceedings from the 2nd ACM Workshop on Wireless Security, December 2003.

[137] Y-C. Hu, A. Perrig, D. B. Johnson, "Packet Leashes: A Defense against Wormhole Attacks in Wireless Ad Hoc Networks", Proceedings from the Joint Conference of the IEEE Computer and Communications Societies, April 2003.

[138] D. S. J. De Couto, D. Aguayo, J. Bicket, R. Morris, "A High-Throughput Path Metric for Multi-Hop Wireless Routing", Proceedings from the 9th ACM International Conference on Mobile Computing and Networking, September 2003

[139] B. Awerbuch, D. Holmer, H. Rubens, "High Throughput Route Selection in Multi-Rate Ad Hoc Wireless Networks", Proceedings from the 1st Working Conference on Wireless On-demand Network Systems, January 2004.

[140] R. Draves, J. Padhye, B. Zill, "Routing in Multi-radio, Multi-hop Wireless Mesh Networks", Proceedings from the ACM International Conference on Mobile Computing and Networking, September 2004.

[141] J.P. Pavon and S. Choi, "Link adaptation strategy for IEEE 802.11 WLAN via received signal strength measurement", Proceedings from the 38th Annual IEEE International Conference on Communications, vol.2, pp.1108–1113, May 2003.

[142] K. Balachandran, S. R. Kadaba, and S. Nanda "Channel Quality Estimation and Rate Adaptation for Cellular Mobile Radio", IEEE Journal on Selected Areas in Communications, vol. 17, nº 7, pp. 1244–1256, 1999

[143] L. Muñoz, M. Garcia, J. Choque, R. Aguero and P. Mähönen, "Optimizing internet flows over IEEE 802.11b wireless local area networks: A performance-enhancing proxy based on forward error correction", IEEE Communication Magazine, vol. 39, nº 12, pp. 60–67, December 2001.

[144] J. Arauz, P. Krishnamurthy, "Markov modeling of 802.11 channels", Proceedings from the 58th IEEE Vehicular Technology Conference, October 2003.

[145] Low Power Advantage of 802.11a/g vs. 802.11b, Texas Instruments White Paper SPLY006 - December 2003.

[146] Power Consumption and Energy Efficiency Comparisons of WLAN products, Atheros Communications White Paper.

[147] M. García, R. Agüero, L. Muñoz, P. Mahonen, "Behavior of UDP-Based Application over IEEE 802.11 Wireless Networks," Proceedings from the 12th International Symposium on Personal, Indoor and Mobile Radio Communications, vol. 2 , pp. 72–77, October 2001.

[148] I. D. Chakeres, E. M. Royer, C. E. Perkins "Dynamic MANET On-demand Routing Protocol", IETF Internet Draft, draft-ietf-manet-dymo-03.txt, Work in Progress, October 2005.

[149] T. Clausen, C. Dearlove, P. Jacquet, "The Optimized Link-State Routing Protocol version 2", IETF Internet Draft, draft-ietf-manet-olsrv2-02, Work in Progress, June 2006.

[150] Iperf. http://dast.nlanr.net/Projects/Iperf

[151] C. Perkins, E. Belding-Royer, S. Das, "RFC3561: Ad hoc On-Demand Distance Vector (AODV) Routing", IETF Network Working Group, July 2003.

[152] L. Sanchez, J. Lanza, L. Muñoz, J. Perez, "Enabling Secure Communications over Heterogeneous Air Interfaces: Building Private Personal Area Networks", 8th International Symposium on Wireless Personal Multimedia Communications, pp. 1963–1967, September 2005.

[153] IEEE 802.11a, Part 11: Wireless U N Medium Access Control (MACJ and Physical Layer (PHYJ specifications High Speed Physical Layer in the 5 GHz Band, IEEE Standard for Information technology, 1999.

[154] IEEE 802.11b, Wireless LAN MAC and PHY specifications: Higher speed Physical Layer (PHY) extension in the 2.4 GHz band, IEEE Standard for Information technology, 1999.

[155] Specification of the Bluetooth System – Core vol.1 v1.1, www.bluetooth.com

[156] J. Haarsten, "Bluetooth - the universal radio interface for ad-hoc, wireless connectivity", in Ericsson Review, vol. 3, 1998.

[157] H. Niedermayer, A. Klenk, G. Carle, "The networking perspective of security performance - a measurement study", Proceedings from the 13th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems, March 2006.

[158] P. Papadimitratos, Z.J. Haas, "Secure Message Transmission in Mobile Ad Hoc Networks", Elsevier Ad Hoc Networks Journal, vol. 1, nº 1, pp. 193–209, March 2003.

[159] M. Alutoin et. al., "MAGNET Deliverable D2.4.3: Refined Architectures and Protocols for PN Ad-hoc Self-configuration, Interworking, Routing and Mobility Management", December 2005.

[160] V. Kawadia and P. R. Kumar, "A cautionary perspective on cross-layer design," IEEE Wireless Communications, vol. 12, no. 1, pp. 3–11, February 2005.

[161] OpenSSL project, http://www.openssl.org/

[162] L. Sánchez, J, Lanza, J. Hoebeke, I. Moerman, M. Alutoin, K. Ahola, J. Jaen-Pallares, M. Girod-Genet, M. Bauer, J. Zeiss, "Assessing Personal Networks on a pan-European Testbed", Proceedings from the ICT Mobile and Wireless Communications Summit, June 2008.

[163] D. M. Kyriazanos, M. Argyropoulos, L. Sánchez, J. Lanza, M. Alutoin, J. Hoebeke and C. Z. Patrikakis, "Overview of a Personal Network Prototype", IEC Annual Review of Communications, vol. 59, pp. 521–534, December 2006.

[164] J. Hoebeke, G. Holderbeke, I. Moerman, W. Louati, W. Louati, M. Girod-Genet, D. Zeghlache, L. Sánchez, J. Lanza, M. Alutoin, K. Ahola, S. Lehtonen, J. Jaen-Pallares, "Personal Networks: From concept to a demonstrator", Proceedings from the 15th IST Mobile & Wireless Summit Communications Summit, June 2006

[165] M. Alutoin, K. Ahola, S. Lehtonen, J. Paananen, "Personal Network Directory Service", Telektronikk Journal, vol. 103, nº 1, pp. 85–92, March 2007.

[166] R. L. Olsen, M. Bauer, L. Sanchez, J. Lanza, "Self Organisation of Context Agents in Personal Networks and Federations", Proceedings from the 10th International Symposium on Wireless Personal Multimedia Communications, December 2007.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION AND OBJECTIVES

The purpose of this chapter is to provide the initial problem definition and describe the framework addressed in the thesis. Besides, the main contributions of the work carried out will be presented. Additionally, it will provide an overview of the structure of the thesis. Initially, the main challenges motivating the development of this work will be introduced. The Personal Networking paradigm will be the focus of the work although the thesis will also contribute to the optimization of future wireless communication scenarios in general. The particular problems to be tackled in the thesis and the scope of the solutions developed will be described. Finally, an outline of the thesis structure is provided.

1.1 MOTIVATION AND BACKGROUND

For forty years, computer systems have catered to machines. Although they purport to serve people, in fact, they have forced people to serve them. They have been difficult to use. They have required us to interact with them on their terms, speaking their languages and manipulating their parts. They have not been aware of our needs or even of our existence.

In the future, computation will be human-centred: it will enter the human world, handling our goals and needs and helping us to do more by doing less. Computation will be pervasive, like batteries, power sockets, and the oxygen in the air we breathe. Configurable generic devices, either handheld or embedded in the environment, will bring computation to us, whenever and wherever we need it. As we interact with these "anonymous" devices, they will adopt our information personalities. They will respect our desires for privacy and security. Mobile users demand access to high-speed data real- and non-real time multimedia services from Next-Generation Wireless Systems (NGWS) anywhere and anytime. New systems will boost our productivity. They will help us automate repetitive human tasks, control a wide range of physical devices in our environment, find the information we need (when we need it, without obliging us to examine thousands of search-engine hits), and enable us to work together with other people through space and time.

Next generation wireless systems should provide the user access with a broad range of services in a transparent way, independently of user location, by making the technology invisible and embedded in the natural surroundings. Reaching this goal requires efficient cooperation between heterogeneous networking technologies and different protocols. Wireless personal networks are an integral part of such an emerging heterogeneous infrastructure. It is highly desirable, and, in fact, necessary due to economical constraints, to incorporate the present wireless systems in building the new paradigm.

A large number of wireless access technologies are envisaged to coexist in future wireless communication spaces, so the necessary methods for them to interwork seamlessly have to be deployed. In this sense, the corresponding Medium Access Control (MAC) and link layer protocol(s) should be accessed from upper layer protocols and applications for control purposes, in a generic way, independently of the type of technology that is being used (in the same way higher layer protocols and applications access the underlying protocol stack through the socket interface for data purposes).

Taking into account the situation defined in the previous section many initiatives have been triggered to develop a solution that takes into consideration the user requirements and provides the necessary technology to support the future networking paradigms. The work developed herein has its roots in the architecture I co-designed in [1]. This is an architecture suited for present and future personal mobile communications networks and services. This was part of the seed that was further developed into the concept of Personal Networks.

1.1.1 Overview of Personal Networks

Personal Networks (PN) [2]-[3], are based on taking the concept of pervasive computing and combining it with strong user focus. A PN is a collection of one's most private devices, referred to as personal devices/nodes, that forms a virtual network where neighbouring personal devices organize themselves in clusters which are in turn interconnected over the Internet. From a technical point of view, the PN is seen to consist of devices sharing a common trust relationship. Security and privacy are the fundamental properties of the PN, as well as its ability to self-organize and adapt to mobility and changing network environments [4]-[6].

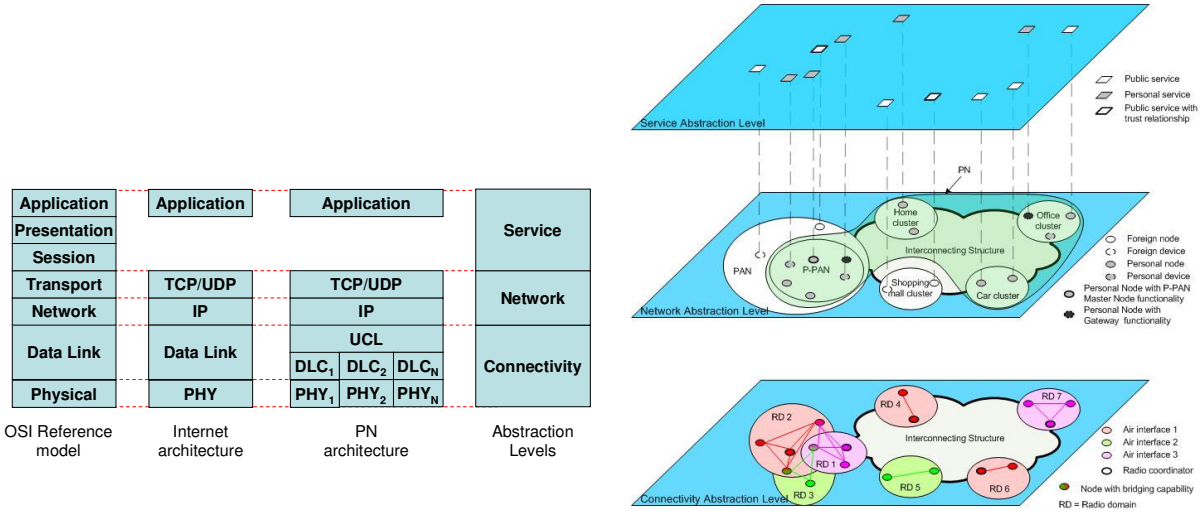
The IST project MAGNET [7] philosophy is that Personal Networks (PNs) will support the users' professional and private activities, without being obtrusive and while safeguarding privacy and security [8]. A PN can operate on top of any number of networks that exist for subscriber services or are composed in an ad hoc manner for this particular purpose. These networks are dynamic and diverse in composition, configuration and connectivity depending on time, place, preference and context, as well as resources available and required, and they function in cooperation with all the necessary and preferred partners.

In contrast to other initiatives that explore fields such as wireless personal area networking [9], mobile ad hoc networks [10] or self-configuration [11]-[12] in isolation, focusing on optimizing the characteristics of each field without taking the others into account. The PNs require an integrated approach that copes with the different connectivity, networking and service requirements in order to accomplish the aforementioned vision of an autonomous and self-organized secure network serving the user for private and professional activities.

A typical misunderstanding comes when PNs are compared with the widely extended concept of Personal Area Networks (PANs). The concept of a PN goes beyond the concept of a PAN. The latter refers to a space with a small coverage area (less than 10 m) around a person where ad-hoc communication occurs, e.g., using Bluetooth or IEEE 802.15.3. These are intended to interconnect portable and mobile computing devices such as PCs, Personal Digital Assistants (PDAs), peripherals, cell phones, and consumer electronics. PNs extend the local scope of PANs to a global one by addressing virtual personal environments that span a variety of infrastructure- as well as ad-hoc networks. PNs are very much centred on a person and his/her needs. They are dynamic in composition, configuration and connectivity depending on the time, place and circumstances, the resources required and the partners one wants to interact with.

Besides the personalization and privacy requirements that are imposed on the Personal Networking paradigm, self-configuration and heterogeneity support are the main cornerstones for supporting this concept.

As shown in Figure 1-1, the architecture defined within MAGNET presents a layered view where three abstraction levels have been identified. This approach enables the detachment of the different requirements and challenges that need to be tackled in each of the different abstraction levels.



Matching of the abstraction with the OSI model

Graphical representation of the abstraction

Figure 1-1: Illustration of abstraction levels of a Personal Network.

Going from the bottom up, the first level is the Connectivity Level, which can be roughly mapped onto OSI layers 1 and 2. Here the devices are organized in Radio Domains (RD). It is important to note that a node can belong to multiple RDs since it can be equipped with multiple interfaces of different access technologies.

The Network Level, consisting of OSI layers 3 and 4, is placed above the Connectivity Level. The Cluster and the PN are defined at this level. The Private Personal Area Network (P-PAN) is the cluster around the user. The PN further extends the P-PAN concept as a collection of all "my active personal nodes" both remote and in the vicinity of the user. The personal nodes in the PN are grouped into clusters such as: the P-PAN itself, Home cluster, Office cluster, etc. The communication among the different clusters happens via the interconnecting structure over secure dynamic tunnels. The important point in this architecture is the strong focus on the long-term trust concept. Only nodes that are able to establish long-term trust can be part of the user's P-PAN/PN.

In order to reflect the provision and usage of services in the P-PAN/PN concept, a Service Level is defined above the Network Level, which comprises the remaining OSI layers 5, 6 and 7. It contains all the services offered in the nodes/devices in the Network Level. Personal services are offered and used only by personal nodes in PN sense. This implies that these services can be used only if the long-term trust relationship is established.

Once the protocol architecture for establishing PNs is presented, Figure 1-2 shows the high-level network architecture of a PN.

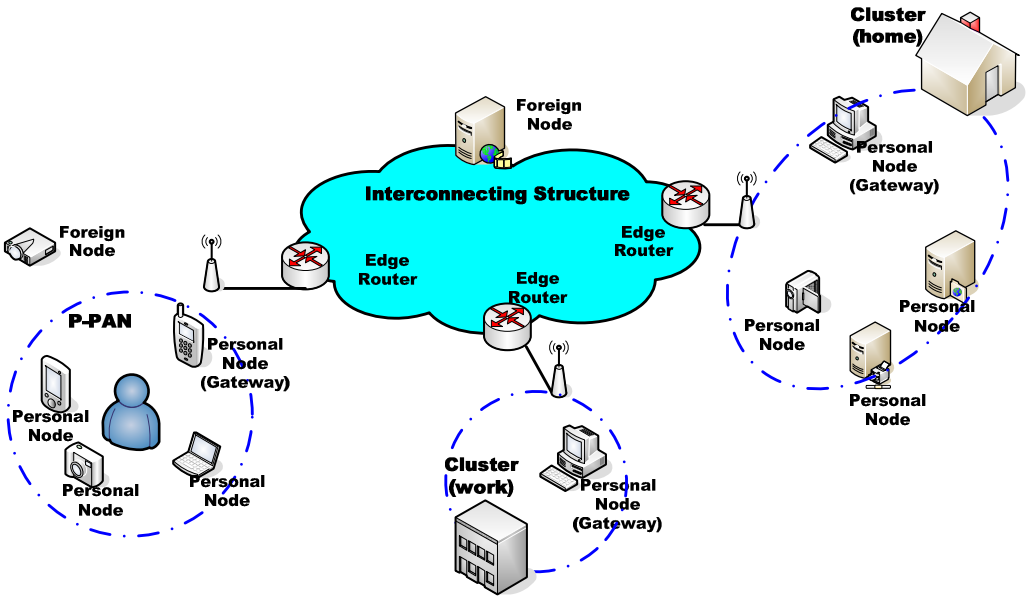


Figure 1-2: PN high-level network architecture

As shown in Figure 1-2, the PN consists of clusters of personal nodes. One of the clusters is special because it is located around the user. The clusters are connected to each other via an interconnecting structure, which is likely to be infrastructure based. In order to protect the privacy of the user and the integrity of the PN, security measures are used to encrypt the user's data when it is sent outside of the device, i.e. using a wireless medium or the infrastructure. The user can reach all of his or her devices using a variety of underlying networking technologies, which are invisible to the user. The user only sees the services that are available in the PN and on foreign nodes that have been made available to the user.

In the subsequent sections the aspects of the Personal Network concept that relate to the work presented in this thesis will be briefly described.

1.1.1.1 Terminology

The terminology used in this chapter and the rest of this document has been introduced in the MAGNET architecture deliverable [13]. For the sake of completeness of this thesis, it is included here. Some new terms that are used throughout the thesis are explained next. Further terms are introduced when used in various chapters.

Device — Any communicating entity.

Node — A device that implements IPv6 [14] and/or IPv4 [15].

Personal Node — A node related to a given user or person with a pre-established trust attribute. Such a node is typically owned by the user in the PN concept. However, any node exhibiting the trust attribute can be considered as a personal node. For instance an arbitrary node can be perceived as a personal node as long as it has been imprinted with the common trust attribute essentially

	defining a fully trusted group of nodes. These attributes are typically cryptographic keys with a permanent (as long as it is not cancelled, redefined or revoked) trust relationship.
P-PAN	Personal nodes are organized in *Clusters* which are a dynamic collection of personal nodes and devices that can communicate among themselves without having to use any other node, except other personal nodes. A Private Personal Area Network or P-PAN is the cluster around the PN owner. The privacy in a P-PAN (as well as in any other cluster) is guaranteed by mandating a mutual trust relationship between every pair of nodes in a cluster. A P-PAN is often referred to as a personal bubble.
Personal Network	A Personal Network (PN) includes the P-PAN and a dynamic collection of remote personal nodes and devices in clusters that are connected to each other via Interconnecting Structures.
Gateway Node	A personal node within a cluster that enables connectivity to nodes and devices outside the Cluster.
Interconnecting Structures	Public, private or shared wired, wireless or hybrid networks such as a UMTS network, the Internet, an intranet or an ad hoc network.
Foreign Node	A node that is not personal and cannot be part of the PN. Foreign nodes can either be trusted or not trusted. Whenever trusted, they will typically have an ephemeral trust relationship with a node in a PN.
Trust Relationship	A trust relationship is established when two parties communicate and determine with a degree of certainty each other's credentials to set up a secure communication channel using encryption mechanisms. When devices and nodes want to establish a secure communication channel, they build a trust relationship by whatever means possible.
Imprinting	A procedure to bootstrap a trust relationship between two nodes that basically consists of an authenticated key exchange.

Figure 1-3 indicates how the PN entities defined in this section relate to each other. Note that, the P-PAN is an example of the Cluster concept. The difference between the P-PAN and the rest of the PN Clusters is that the user is in the surroundings. This difference only becomes important at the Service Abstraction Level. Therefore, both terms can be used indistinctively when connectivity and network-level mechanisms are being discussed.

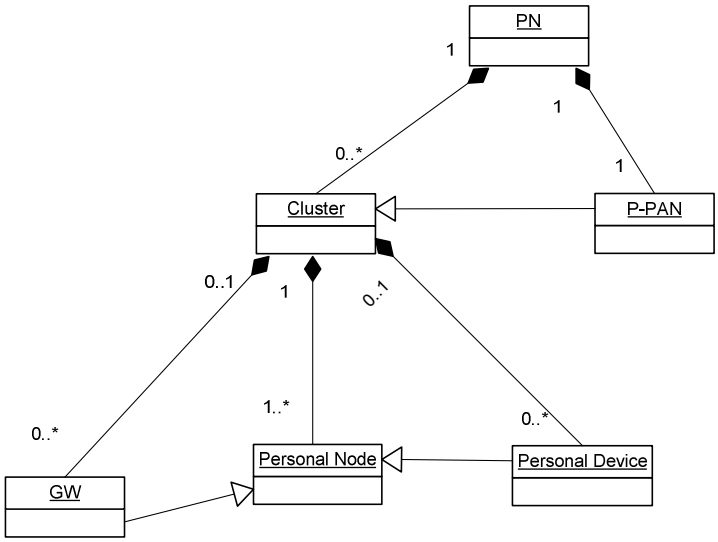


Figure 1-3: Relationship among main entities of Personal Networks

1.1.1.2 Cluster and P-PAN Formation and Characteristics

Before any specific description of the PN self-configuration mechanisms at the abstraction levels can be presented, a number of basic security notions and concepts must be introduced since privacy and security are the key features that govern the formation of the PN. The PN architecture relies on the notion of long-term and short-term trust relationships. The long-term trust, which could also be perceived as permanent trust, is used to establish a strong security association or relationship between the nodes and devices of the PN. Long-term secrets, in fact cryptographic keys, are used to form the trust among the PN constituents, and especially the P-PAN/Cluster components.

These trust relationships are intended to be used between personal nodes owned by the same user. That is to say, the design is based on node ownership, which is an easily understood concept for end users. This is crucial since the end-user understanding of the trust relationship model influences the security of PNs. A lack of understanding of how this works and what consequences it has can jeopardize the security of that person's PN. Nevertheless, while the design is made with ownership in mind, there is nothing in the technical solution that will prevent a user from using the trust relationships in different ways. Someone can create long-term trust relationships between nodes of a family for instance.

The long term trust keys are used as a basis to establish communications between PN nodes. The process of inserting a given secret in a device or node is referred to as imprinting a device [16]. The goal of imprinting is to exchange the pair-wise keys that will be used afterwards as the basis for deriving the actual session keys used for protecting any communication between that particular pair of nodes.

Therefore, when introducing a new device to the PN, this device will be paired with at least one other device participating in the PN and thus trusted by the other personal nodes. During this procedure the new device will securely exchange a long term pair-wise key with the personal node. This key will be referred to as the PN master key in this thesis.

As a result of the pairing procedure, the peers derive a long-term shared key that is subsequently used to secure the communication between them. Each device must store this information securely in the form of a device record. A peer record contains the following information: (1) Peer identifier – a unique identifier associated with the device; (2) PN key – the shared secret derived from the pairing process.

In contrast to other descriptions of cluster or Personal Area Network [10] that limit the concept to a matter of radio coverage (e.g. 10m range), the concept of cluster proposed in this architecture stands on an opportunistic, distributed, multihop and proactive approach based on the trust relationships established between the cluster constituents. Furthermore, it copes with heterogeneity support, dynamic adaptation, infrastructureless environment survival and privacy requirements imposed by the P-PAN concept. Clusters are dynamic in nature. Nodes can be switched off or they can become available as well as roaming and showing up in a different cluster. Clusters can split when a person takes some of the nodes and leaves the rest behind. Likewise, clusters can be merged when a person arrives home and his/her P-PAN merges with the home cluster. Potentially, there is no limit on how large a cluster can grow, neither in terms of number of nodes nor hops. However, typically we expect clusters to have a small number of nodes and a limited geographical span, because of the way they will be deployed. In this sense, the clusters will be as large as possible (as long as a new personal node or device is reachable through a PAN air interface, the cluster will add a new wireless hop to its structure), adding new personal nodes and devices as soon as they appear in the cluster surroundings.

1.1.1.3 PN Organization

In order to form the PN and perform inter-Cluster communication over a fixed infrastructure, four requirements need to be fulfilled. First of all, the clusters need to have access to the fixed infrastructure through one or multiple Gateway Nodes (GW). Secondly, once access to the fixed infrastructure is available, the clusters need to be capable of locating each other. Thirdly, once they have located each other, they should establish tunnels between them. Last but not least, once the PN has been formed, it should be able to maintain itself in view of dynamics in the network. We will now discuss how these requirements lead to a conceptual PN architecture that relies on the concept of a PN Agent.

Connectivity between remote clusters can only be established if they can locate each other. The PN Agent concept has been introduced to assist in this localisation and in the overall PN establishment. The PN Agent could be implemented as part of the user's fixed PN Cluster (e.g. the cluster of nodes around the user's home or office). It can also be implemented as a service under the control of service or network providers.

The PN Agent keeps track of each cluster GW point of attachment. Clusters that have connectivity to the infrastructure need to register themselves with the PN Agent. Based on this information, the PN Agent can inform the other registered clusters about the location of respective PN clusters. This information is indispensable for the creation of the tunnels between the remote clusters. The purpose of the tunnels is twofold. First, they provide a secure means for inter-Cluster communication by shielding the intra-PN communication

from the outside world. Secondly, these tunnels will be established and maintained dynamically, efficiently dealing with cluster mobility.

Establishing and maintaining these tunnels dynamically is based on the same concept since GW nodes maintain their registration updated in the PN Agent which in turn informs the others upon any change that occurs at the point of attachment of any of the registered GWs.

In addition, the PN Agent concept can be extended, meaning that it could provide additional functionalities such as naming, service discovery and foreign communication. The PN Agent offers a good entry point for PN to PN communication. The PN Agent should be considered as a concept rather than as a PN entity, since there may be many different solutions to implement the PN Agent concept.

1.2 PROBLEM DEFINITION AND GOAL OF THE THESIS

Making Personal Networks happen is a challenging task that requires the integration of multiple subsystems dealing with the multiple necessities imposed by the PN concept [17]. This thesis will focus on the optimization of the intra-cluster communications. In this section the main problems on which this optimization should be focused will be described.

1.2.1 Need for multiple interfaces support

In recent years, there has been a process termed "the *deployment stage*" in which multiple wireless access technologies have appeared. Each of them is focused on different operational environments for personal wireless systems that are different in terms of cell sizes, coverage areas and services to be provided. Satellites provide global coverage and are suitable for urban and remote areas with low traffic density, and without access to terrestrial telecommunications networks. Terrestrial macro cells are typically situated in rural or suburban areas with medium traffic density. The cell radius has a range of several tens of kilometres. Micro cells are situated in urban areas with a cell radius of up to 1 km. Traffic density varies from medium to high, and the mobile speed remains moderate. Pico cells are predominantly situated indoors with a cell radius of less than 100 m. Their characteristics include low-speed terminals, medium-to-high traffic density, and wideband or broadband services [18]. The final level of the hierarchy is personal area cells that refer to networks connecting fixed, portable and moving devices operating in a Personal Operating Space (POS). The cell radius is typically up to 10 m around a person, whether stationary or in motion. A typical cell would contain a limited number of devices (e.g. less than 10) with low traffic density and wideband services [9]. For each of these operational environments a different wireless access technology (where not multiple of them) has been deployed, using 2G and 3G networks for the macro cells scenario, IEEE 802.16 for micro cells environments, IEEE 802.11 for pico-cells and Bluetooth and IEEE 802.15 technologies for the personal area cells.

Thus, heterogeneity will be a fundamental characteristic of next generation wireless communications, where a new stage of the technology is foreseen, the "*convergence stage*". Next generation wireless systems are on a continuous search for the *Always Best Connected*

paradigm [8]. More and more personal devices are equipped with many of the aforementioned wireless technologies so that the user can access the different services that the different operational environments provide. Nevertheless, very little attention is paid to these system features when solutions for self-configuring personal networks are proposed. Routing protocols [10] assume that nodes in the network share a common connectivity level, but mobile wireless ad-hoc networks will not only be composed of nodes with quite different capabilities but also from the wireless access technology perspective, multiple standards will coexist. All-IP architectures have usually been proposed to tackle this aspect, putting all the pressure on the IP layer. Nevertheless, it is necessary to first solve the heterogeneity problem by offering to the network layer a single interface, and so hiding the possible underlying complexity.

1.2.2 Need for cross-layer optimization

The traditional protocol stack has been designed to deal with complicated problems by breaking them into smaller parts. It consists of layers whose definitions and tasks are defined explicitly and independently. In other words, each layer is isolated from the others except for providing output to and getting input from adjacent layers [19]. According to the direction of the flow upward/downward, each layer conducts its own task by taking inputs from the layer below/above and conveys the outputs obtained to above/below.

Table 1-1: Major advantages and disadvantages of the traditional layered architecture and their effects

	Explanation	Effect
Advantages		
Modularity	Each layer can be designed independently of others	Simpler design
Standardization	Design only requires the knowledge of explicit definitions and abstractions	Interoperability
Expandability	Layers can be updated, or expanded as long as interfaces are kept	Individual flexibility
Disadvantages		
Ordering	Execution of any process in any layer has to be done after the execution of previous processes in previous layers	• Inefficiency • Latency
Interaction	Due to strict isolation, information cannot cross other layers	• Unawareness • Redundant processes • Sub-optimal performance
Adaptation	In wireless communications, rapid channel variations cannot be responded to immediately	• Decrease in capacity • Sub-optimal performance
Topologies	Some of the network topologies need flexible layer architecture	• Inefficiency

Even though the layered structure overcomes many problems successfully, it is obvious that the stringent architecture creates some problems such as asynchrony and

inefficiencies. A brief list of the major advantages and disadvantages of the traditional layered architecture and their effects is given in Table 1-1.

Migration from a strict layered architecture to a more flexible interactive one has another very strong motivation: wireless networks. Because of the different nature of wireless communications, numerous concepts defined in wired networks need careful re-consideration or even modification, as does the protocol stack. Specifically for wireless communications, due to small-scale fading, wireless channel conditions may change drastically in a very short time [20]. Therefore, in order to take advantage of the periods in which the channel is identified as "good," a flexible design is essential [21]. Large-scale channel variations contribute to the necessity for flexible architecture as well [22]. Interference and time-varying capacity due to multipath, relative mobility, and shadowing are other very crucial parameters that affect wireless networks [23]. Apart from those, new transmission schemes for wireless communications such as relay networks [24] may not be established via a strictly isolated layered architecture [21] and might require a different design.

We have observed that under some circumstances, strictly layered architecture performs inefficiently. An increase in the amount of information flow between layers and a re-organization the processes according to their dependency on each other have been proposed in order to overcome this major problem. All this work has led to a novel concept called *cross-layer architecture*.

1.2.3 Need for a network selection strategy

Accustomed to broadband data networks with flat-rate pricing in wired networks, users have grown dependent on having a high quality broad service range available without caring too much about the cost. The perspective of the desired service priority, quality and budget requirements will change depending on the user profile and current context. Users may be sitting with their laptops in a cafeteria with time to wait for long downloads expecting a high-quality presentation of requested data on a large screen, alternatively they could be rushing down a busy street with low battery power requiring only minimal presentation of their requested data on a handset.

In general, users will seek different features depending on their particular interests and specific contextual situation. If a user considers a particular application too expensive she will be discouraged from using it, and keeps consumption of this wireless data service to a minimum. Consumers choose the products and services that give them the greatest satisfaction. In the world of economics, utility functions are often used to describe this user satisfaction. In our case the utility function depicts the preference relation for the different available service and access options. Users not only want the best obtainable value for their budget in monetary terms; they seek cost effective solutions to meet their communications expectations. Users aim to maximise their utility, and to this end they will compare marginal benefits against marginal costs. As long as marginal benefits exceed marginal costs, and it is within their budget, they will consume the commodity.

However, next generation networks will comprise such a diverse number of possibilities that the users cannot be expected to take the decisions on their own. It is necessary to provide them with mechanisms that intelligently select the optimal available access network based on context information such as user preferences, power consumption, link quality, etc. The network selection strategy has to be designed so that it maximizes the users' communication expectations leading to the Always Best Connected paradigm [8].

1.2.4 Need for energy efficient operation

Although wireless networks have existed for many years already, explicit concern about their energy efficient operation has emerged only recently. It is quite evident that when the power source is either costly or in short supply, energy efficiency is of paramount importance. In some wireless network applications, energy is actually entirely non-renewable and is thus an overriding constraint for the design and operation of the network.

Studies show that the significant consumers of power in a typical personal device are the microprocessor (CPU), liquid crystal display (LCD), hard disk, system memory (RAM), and the wireless network interface card [25][26]. Consequently, energy conservation has been considered in the hardware design of the mobile terminal [27] and in components such as CPUs, disks, displays, etc. Significant additional power savings may result from incorporating low-power consumption strategies into the design of network protocols used for data communications.

The sources of power consumption, with regard to network operations, can be classified as either communication related or computation related. Communication involves usage of the transceiver at the source, intermediate (in the case of ad-hoc networks), and destination nodes. The transmitter is used for sending control, as well as data packets originating at or routed through the transmitting node. The receiver is used to receive data and control packets – some of which are destined for the receiving node and some of which are forwarded. Energy efficiency should attempt to maximize the system throughput while minimizing the total amount of power consumed, in other words, optimizing the time the transmitter or receiver are active.

However, things are not that simple. First of all, if energy efficiency is the only concern in a communication system, one might as well transmit nothing. Energy reserves would thus remain intact perpetually. Clearly communication performance is also of paramount interest. Thus, the choice of how to incorporate energy efficiency in the overall design is far from clear. One approach is to try to minimize energy consumption subject to throughput (or delay) staying above (or below) a certain threshold. Alternatively, one can try to maximize throughput (or minimize delay) per joule of expended energy. Neither of these approaches led to simple precise formulations or easy solutions.

Different solutions have been proposed to provide more efficient energy consumption at different layers. At the link layer, transmissions may be avoided when channel conditions are poor, as studied in [28]. Also, error control schemes that combine Automatic Repeat Request (ARQ) and Forward Error Correction (FEC) mechanisms may be used to conserve

power (i.e. tradeoff retransmissions with ARQ versus longer packets with FEC) as in [29]. Energy efficient routing protocols may be achieved by establishing routes that ensure that all nodes equally deplete their battery power, as studied in [30]-[31] or that avoid routing through nodes with lower battery power. More complex solutions such as [32] exploit cross-layer operation and control the network topology by varying the transmitted power of the nodes so that certain network properties are satisfied.

Cross-layer design is particularly interesting under energy constraints, since not only energy across the entire protocol stack must be minimized, but also system performance must be optimized. While layer-specific solutions might disregard valuable information residing in other layers, cross-layer solutions can exploit global knowledge of the system to provide sub-optimal solutions at each of the different layers that result in optimal solutions at system level.

1.2.5 Need for user privacy protection

Mobile and wireless communication networks have created a major breakthrough in new telecom applications and services. Unfortunately, wireless networking is a double-edged sword. Ever since it introduced new factors such as mobility, the temptation for unauthorized access and eavesdropping has been a reality because an attacker can easily access the transport medium. During the past few years, we have seen that wireless networks are rapidly evolving as an effective medium of communications. At the same time, we have also experienced an enormous increase of data exchange between wired and wireless networks. Although they offer new services like mobility and roaming, wireless networks have introduced new security threats, sometimes referred to as parking lot attacks [33]. In fact, wireless data flows over public airwaves, so a knowledgeable person with a wireless computer and an antenna can gain access to the domain. Moreover, wireless communications have inherited the same wired threats and attacks; especially denial of services (DoS) and Man in the middle attacks [34].

On the other hand, other constraints on wireless environments lie in the scarce radio resources, the high error rates and the limited computational power of mobile devices. This should be taken into consideration whenever we are to implement security protocols for such environments.

Users need to trust the system that supports their personal communications. The most confidential information might be exchanged within a personal network and the user needs to be sure that this will never be disclosed. If the system fails in this feature, Personal Networks will never happen. In order to defend against targeted attackers, wireless networks in general and PNs in particular need to define robust protection mechanisms that allow efficient communication while safeguarding the user communications privacy.

1.3 THESIS' GOALS AND OUTLINE

The thesis aim is to develop the mechanisms that contribute to the optimized communications within clusters of personal nodes. To achieve this goal, the

aforementioned needs must be attended to and the measurement of the enhancements obtained through the adoption of the solutions proposed in this thesis will drive the concept.

In this sense, the thesis will first propose a novel architecture for handling the heterogeneity in terms of wireless access technologies within the same node. This solution will exploit the cross-layer optimization paradigm to enable the dynamic selection of the best available communication interface while assuring energy efficient operation as well as implementing the necessary security mechanisms to protect the communications between nodes within the same cluster.

Although the work described in this thesis has been carried out with a PN scope, the problems defined here and addressed during the thesis development are valid for other 4G wireless systems. Furthermore, the conclusions derived from the analyses done will be centred on the personal networking paradigm in particular, but they can be extrapolated to other future networks concepts.

The rest of the thesis is organized as follows.

Chapter 2 provides the necessary background information, which includes an introduction to different concepts managed during the thesis development and an overview of the current state-of-the-art with particular focus on how it relates to the network scenario envisioned. Fundamentally, this chapter will cover the solutions that are currently proposed for handling the problems described above. Starting from the approaches used to support future heterogeneous scenarios, we will describe how cross-layer optimization mechanisms are being applied as Performance Enhancing Proxies as well as the standardization groups that are working on applying cross-layer interaction mechanisms to ad-hoc networking protocols. We will also describe different models for coordinating the interactions between different layers and the current status of the standardization of the cross-layer optimization design and interoperability. The current research approaches in dynamic network selection will also be presented highlighting the different models applied and the algorithms supporting the decision making process. The fundamentals or power-aware wireless communications will also be addressed in order to present the main attributes considered when power efficiency is to be optimized. Finally, we will provide an overview of wireless security research. We will include the main points of current threat analyses as well as a description of some of the mechanisms currently used to assure communications integrity within ad-hoc networks. The overview will also include a summary of security techniques most commonly used on IP based communications.

In Chapter 3 the architecture of the framework proposed to optimize intra-cluster communications will be presented. The concept of isolating the upper-layers from underlying wireless technologies, thus providing real multi-mode, can be achieved by introducing a Universal Convergence Layer (UCL). It will mainly act as an enabler for backward and forward compatibility by defining a common interface towards the network layer while managing several different wireless access technologies independently of their PHY and MAC layers. Additionally, in a multi-radio architecture the UCL is specified not only to hide the link layer complexity and specifics from the upper layers but also to

facilitate cross layer optimization. Firstly, we will present the high-level protocol architecture of the UCL. The main building blocks of the architecture will be described along with the related mechanisms implemented within them. Then, the core aspects of the implementation carried out will be showcased highlighting the fundamental software architecture of the UCL Linux prototype.

Chapter 4 will further develop the mechanisms implemented within the UCL to support dynamic selection of the most appropriate interface based on cross-layer information. The idea behind these techniques is to take advantage of the heterogeneity instead of thinking of it as an obstacle. In this chapter an analytical assessment of the utility function optimization for best decision making will be presented. This analysis has not only been made looking for optimal bandwidth efficiency (i.e. obtain best possible throughput) but also power efficiency has been studied. Additionally, the results from the experimental validation of the dynamic network selection mechanisms implemented within the UCL will be presented. The main aim of this chapter is to present the benefits in terms of performance optimization and power efficiency obtained through intelligent selection of the most appropriate wireless interface and in particular to comment on the results derived from the experimental validation carried out over the UCL dynamic network selection system implementation.

Neighbour discovery and authentication and intra-cluster communications privacy protection mechanisms will be presented in Chapter 5. In contrast to other descriptions of cluster or Personal Area Network [10] that limit the concept to a matter of radio coverage, the concept of cluster proposed in Personal Networks architecture relies on the trust relationships established between the cluster constituents. The problem of detecting surrounding peers and authenticating their adscription to the same PN is the initial step to be taken to support the cluster formation. The protocol implemented in the UCL to deal with this problem will be described. Furthermore, theoretical and experimental analyses of the impact of this process on the overall cluster formation will be presented. Finally, the mechanism to assure intra-cluster traffic integrity, privacy and origin authenticity will be specified and the performance observed when using different Wireless Personal Area Network (WPAN) technologies will be assessed from both a theoretical and an experimental point of view. The main objective of this chapter is the definition of the techniques that allow personalization, from a security point of view, of the intra-cluster communications and the analytical and experimental validation, which also includes comparison analyses with competing solutions.

Finally, Chapter 6 summarizes the main contributions of the thesis and concludes this work. An overview of opened research opportunities of interest will also be presented.

1.4 MAIN RESEARCH CONTRIBUTIONS

The design and specification of the UCL [35]-[36] is the first contribution made in this thesis and it represents the framework over which intra-cluster communications optimization techniques have been developed. The main work developed as part of this thesis has been to provide detailed understanding of the UCL architecture and its building

blocks not only limited to the specification of the architecture and functionalities but also by implementing a prototype over real platforms. The implementation effort is driven by many. Firstly, it serves as proof-of-concept validation. Additionally, it enables the development of measurement campaigns over real-world scenarios to assess the behaviour of the optimization techniques proposed and compare them both with theoretical analyses and other competing implementations. Finally, the UCL is part of a vertically integrated system for supporting PNs, comprising several other HW and SW components and so its implementation was necessary to provide a full-blown system.

Once the framework for supporting intra-cluster communications optimization techniques was designed and specified, the research and development of specific mechanisms addressing this optimization was tackled. In this thesis, advanced techniques for optimizing intra-cluster communications over heterogeneous wireless scenarios [37]-[38] have been proposed, designed, specified and implemented. The focus has been placed on two algorithms that are based on cross-layer optimization. The first one deals with the selection at run-time of the most appropriate wireless interface to be used in order to improve the system performance. The second one leverages the striping concept in order to exploit all the network interfaces available.

The last contribution made in this thesis is the definition, specification and implementation of the mechanisms that trigger the self-configuration of PN clusters [39] and the procedures for assuring intra-cluster traffic integrity, privacy and origin authenticity. The validation of the techniques proposed and implemented and its comparison with other alternative solutions are also part of the contributions of this thesis [40]-[41].

Concluding, the biggest advance of the state of the art follows from enabling the user to have easy, affordable and seamless control of their devices over heterogeneous communications networks. These are empowered to communicate efficiently and securely with their selected interaction groups, no matter what kind of access is available for them to use.

CHAPTER 2
RELATED WORK AND STATE OF THE ART

The purpose of this chapter is to present a comprehensive and complete overview of previous and current research approaches in the areas that have been addressed in this thesis. The examples pointed out are meant to be representative, not exhaustive.

2.1 HETEROGENEOUS COMMUNICATIONS

Current mobile devices are often equipped with several network interfaces, which may be of different wireless access technologies. Moreover, it is anticipated that some applications and services will be able to adapt to changing access situations. Information on network characteristics should also be available to the applications, so that they have the possibility of adapting to the continuously changing environment. A natural requirement is that the applications running on the user equipment should be network-access independent in the sense that the applications themselves are unaware of the underlying access technology and have an interface toward the IP layer only through an IP-based application programming interface (API). Handling this heterogeneity is one of the main research topics to be tackled in supporting next generation networks. Several strategies have been adopted in order to address this issue. In this section we will present an overview of the main approaches adopted by briefly describing some examples from each of them.

As different wireless network technologies are being deployed at an increasing rate, interworking of these various technologies has become an important issue. Already, mobile Internet hosts are often equipped with several network interfaces or are at least able to connect to such interfaces. These interfaces may use different access technologies such as Bluetooth, WLAN and 3G cellular. For this purpose, a few mobile host multihoming protocols supporting handoffs between interfaces have been proposed. The most advanced protocols are able to move single traffic flows independently of each other.

There are several proposals related to multihoming. In this section, we focus on proposals that are implemented at the network and transport layers because those are transparent to applications.

Additionally, we also present another strategy, denominated striping, that takes advantage of heterogeneity.

2.1.1 Multihoming

Multihoming is widely used in the IPv4 Internet today and is an essential component of service for businesses. In recognition of this, the built-in features of IPv6 make it easier for end-hosts and networks to be multihomed than in IPv4. This, combined with the expected continued growth of the Internet, means that it is likely that multihoming will become an increasingly common phenomenon.

Recently, there have been numerous proposals made within the IETF to allow IPv6 multihoming to prosper without incurring the associated scalability and transport problems. The following section gives a brief summary of some of the approaches to multihoming solutions that have been proposed in the current literature.

2.1.1.1 LIN6

LIN6 [42] provides a solution for mobility and multihoming using LIN6 generalized IDs. The basic idea is that each host has a 64-bit globally unique identifier, called LIN6 ID, which is presented in the IPv6 interface identifier portion of an IP address used by a host.

In addition to this, LIN6 reserves a special IPv6 prefix, called LIN6 prefix, which is not routable. A host can be uniquely named by prefixing its LIN6 ID with the LIN6 prefix, resulting in what is called LIN6 Generalized ID (GI). These generalized IDs are then stored inside the DNS, together with the address of a Mapping Agent. Since the GIs are globally unique and permanent, the communicating hosts use them as endpoint names. The Mapping Agent is asked for the mobile host's current addresses. The host then dynamically translates the prefixes of outgoing and incoming packets, making it possible to use GIs in sockets and real addresses in routing. LIN6 also supports multihoming through enabling a single GI to be associated with several real addresses.

2.1.1.2 Multi Homing Transport Protocol

Multi Homing Transport Protocol (MHTP) [43] has been proposed for IPv6 network layer multihoming. MHTP is targeted for site multihoming, being a feature of routers. It is strongly based on BGP4+ routing information. It can be described as a semi-symmetric, end-to-end, NAT protocol. The main idea behind MHTP is that multihomed traffic is transformed into single-homed traffic at a router close to the source and transformed back into multihomed traffic at the last router, which is the MHTP endpoint - a multihomed site that has been allocated an MHTP prefix. This prefix is a /48 block of multihomed addresses.

Implementing the MHTP in a mobile host would greatly increase the load of the MHTP endpoints because of the amount of signaling. Therefore it is not feasible for host multihoming. It seems that MHTP can only be a mid-point solution for site multihoming.

2.1.1.3 Multihomed TCP

Multihomed TCP [44], also called Extended Transport Control Protocol (ETCP), extends the TCP protocol. The ETCP makes it possible to use a set of IP addresses at both endpoints. Addresses can be updated dynamically during communication, enabling multihomed mobility. The draft also presents an address selection mechanism based on time-stamps. Instead of IP addresses, the ETCP uses a separate 32-bit connection identifier, named Protocol Control Block Identifier (PCBI) to identify connections. Hosts exchange their PCBIs during an ETCP handshake.

The ETCP does not solve fast movement or double jump problems but they are identified in the draft [44]. The protocol does not take a stance on how to avoid PCBI collisions or how the current socket API would be affected.

2.1.1.4 Stream Control Transmission Protocol

Stream Control Transmission Protocol (SCTP) [45] is a reliable transport protocol operating on top of the network layer where both of the endpoints may be presented by multiple IP addresses. The motivation in SCTP to use multihoming is the potentially better survivability of the connection in the presence of network failures. However, the connection management policy in SCTP is restricted, and multihoming is mainly used for redundancy purposes.

A host has one primary address and may have zero or more alternative addresses. The primary address is used by peer hosts as the destination address for all packets in normal data transmission. Alternate addresses are used for retransmitted packets to improve the probability of reaching the remote endpoint. When transmission to a primary address fails several times, packets are transmitted to alternative addresses until the protocol confirms that the primary address is reachable.

SCTP endpoints exchange a list of addresses during the initiation of a connection association. Every endpoint must be able to send messages through any interface that is bound with local address set and receive messages from the address set associated with the remote endpoint.

2.1.1.5 Homeless Mobile IPv6

In Homeless Mobile IPv6 [46] architecture the connections are not bound to interfaces represented by IP addresses, but to hosts themselves which are represented by sets of IP addresses. Every host has a cache consisting of local and foreign host cache entries. A local host cache entry contains a set of local IP addresses. Any of them can be used as a source address for outgoing IP packets. Correspondingly, there is a foreign host cache entry for each of the peer hosts.

2.1.1.6 Host Identity Payload and Protocol

The Host Identity Payload and Protocol (HIP) [47] architecture describes a new space for a name called the Host Identity (HI), which completes the IP and DNS name spaces. The use of HIP requires a special protocol layer called the Host Layer Protocol (HLP) located between the IP and transport layers. Cryptographically generated HI is used to identify a connection, while IP addresses are only used for routing information. This kind of namespace separation allows easier mobility and multihoming implementation and management because the IP addresses can be changed without affecting the connection identification.

The HIP architecture seems to be a good alternative for Mobile IPv6 to implement interface selection policy effectively for mobile multihomed hosts.

2.1.2 Communications striping

Striping is a general technique to aggregate multiple resources for better performance. The aggregation of multiple network interfaces can be done at different layers of the protocol stack. In [48] a general framework for striping as a logical FIFO queue (defined as a channel) at the link, network, or transport layers is proposed.

2.1.2.1 Surplus Round Robin striping

The striping method proposed in [48] is a reliable and scalable striping protocol. It is named Surplus Round Robin (SRR). The SRR method is a general purpose designed. Three key ideas of SRR are load sharing, logic reception, and marker packets. On the sender side, SRR uses a load sharing algorithm to stripe packets over multiple links. Logic reception means that the receiver buffers incoming packets and performs the inverse algorithm to

predict which stripe to receive packets from. Marker packets are used in SRR to perform synchronization recovery at the receiver.

The load-sharing algorithm uses a basic round-robin scheme to switch among stripes for transmission. Starting from the first stripe, after sending some packets on a stripe, the sender switches to the next stripe. This operation is repeated until the sender rolls back to the first stripe. At that time, a round is completed. The load sharing algorithm allows each stripe to share the traffic load equally. In SRR, a counter is kept for each stripe. The counter represents how many bytes can still be sent on the stripe. Whenever a packet is sent on a stripe, the size of the packet is subtracted from the stripe's counter. When the counter becomes negative, the sender cannot send more packets on the current stripe. Therefore, it switches to the next stripe. After a round, a stripe will again receive its turn to send packets. At that time, a fresh quantum value will be added to the stripe's counter. This value represents how many bytes can be sent on a stripe in a round. Figure 2-1 illustrates this operation.

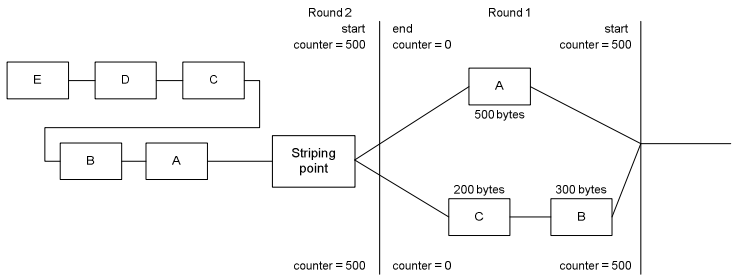


Figure 2-1: The operation of SRR at the sender

At the receiver, packets coming from each stripe are buffered in a per-stripe queue. The receiver simulates the inverse version of the striping algorithm used in the sender to decide which stripe queue to receive packets from. Like the mechanism used in the sender, a counter is kept for each stripe. Whenever a packet is received from a stripe queue, the size of the packet is subtracted from that stripe's counter. When the counter becomes negative, it is refilled with the same quantum value used in the sender, and the receiver switches to the next stripe. Using this design, the receiver can receive packets in their original order without using sequence numbers in these packets. Figure 2-2 illustrates this operation.

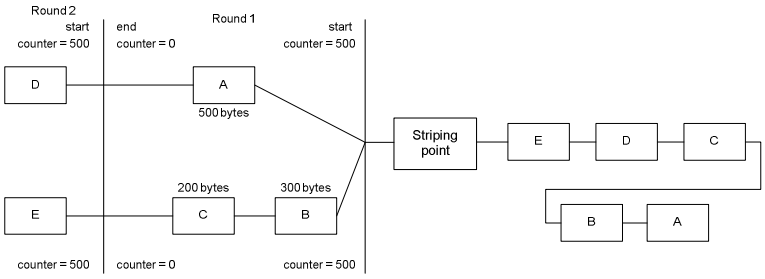
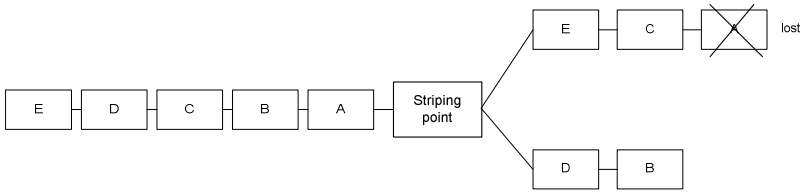


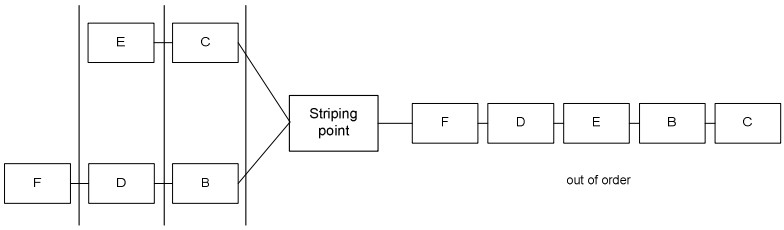
Figure 2-2: The operation of SRR at the receiver

If packets are lost on the stripes, the reception of packets will become out of order. Marker packets are thus introduced to solve this problem. In SRR, the sender periodically sends a special packet named "marker packet" on each stripe. A marker packet is different from

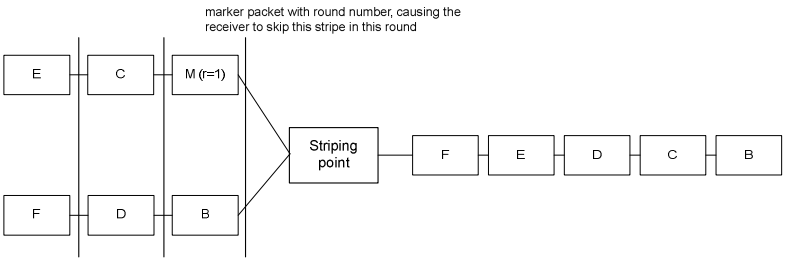
normal data packets. A marker packet contains the information about the current round number. When the receiver receives such a packet, the receiver can synchronize its current round number with the sender. With these marker packets, the receiver can recover from out-of order packets. Figure 2-3 shows this design. In Figure 2-3 (a), the packet labeled 'A' is lost. In Figure 2-3 (b), we can see that the reception in the receiver now becomes out of order. If the marker packet is sent, as shown in Figure 2-3 (c), the lost packet can be detected via the round number carried in the marker packet and packet reordering can be avoided.



(a) on the sender side



(b) on the receiver side



(c) in the receiver side with marker packets

Figure 2-3: Marker packets used in SRR for synchronizing the sender and receiver.

SRR can guarantee FIFO packet delivery order only when no packet is lost. When a data packet is lost, all data packets following it will be received and delivered in the wrong order to the upper layer until a marker packet arrives. When a data packet and a marker packet are both lost, this out-of-order problem will become worse because more data packets will be received and delivered in the wrong order. It is clear that marker packets are more important than data packets. However, in wireless channels, data and marker packets are equally subjected to bit errors.

As mentioned earlier, striping should be transparent to the upper layer. The guarantee of FIFO delivery order is an important issue. Even though SRR supports order recovery via marker packets, the synchronization can only be re-established after future marker packets are received. Before that time, some out-of-order packets will have been handed to the upper layer. The re-establishment of synchronization will be too late for these packets.

Synchronization can be re-established very soon by sending marker packets frequently. However, since marker packets are a bandwidth overhead, a large fraction of bandwidth will be wasted by these marker packets. In addition, SRR has another problem. Since the receiver simulates the algorithm used at the sender, it expects to receive some packets from the current stripe. Until the counter of the current stripe becomes negative, it will not switch to the next stripe. This design will cause blocking if no packet is expected to arrive on the current stripe.

For example, if (1) the link of the current stripe is broken and thus no new packet will arrive soon or (2) a packet is lost on the current stripe but no new packet is expected to arrive soon due to TCP's windowing congestion control, the receiving operation will block forever on the current stripe until some new packets arrive. In the first case, the operation resumption time is when the link becomes re-established, while in the second case it is when the TCP sender times out and resends the lost packet. In either case, during this period the receiver cannot proceed with the reception of more packets from other stripes. It can be seen that in such a situation, the fault tolerance capability inherent in striping is lost. In addition, TCP will perform poorly due to frequent timeouts.

2.1.2.2 Striping at different layers

IPv6 multihoming can be performed at any layer ranging from the network to the application layer, by assigning multiple provider-dependent aggregatable IPv6 prefixes to each site [49]. In general, striping at lower layers leads to a high striping point utilization, and striping at higher layers leads to less head-of-line blocking [50], i.e., a situation where a multihomed connection throughput is limited by that of its slowest path. We now describe the advantages and drawbacks of striping at different layers.

2.1.2.2.1 Striping at link layer

Link layer striping aggregates available physical links into a single communication path. Transmissions are done on a byte-by-byte basis over the physical interfaces, improving the utilization of the links. Byte ordering must be preserved in link layer striping, and padding may be necessary when the number of bytes in a datagram is not a multiple of the number of interfaces. Thus, there may be a significant overhead in striping at the link layer. In addition, IP datagrams may need to be reconstructed before crossing network boundaries, which makes link layer striping only truly useful for local area communications.

2.1.2.2.2 Striping at network layer

Even though network layer striping should make multihoming transparent to the transport and higher layers, network layer striping causes poor TCP performance over heterogeneous paths [51]. While performance can be improved by modifying TCP retransmission timers and window sizes, such modifications essentially require changes at the transport layer, which makes network layer striping a relatively unattractive option.

2.1.2.2.3 Striping at transport layer

Striping IP packets at the transport layer requires a transport protocol that, unlike TCP, can control multiple paths simultaneously. The SCTP protocol [45], for instance, can handle multiple data streams across multiple interfaces. However, SCTP only uses more than one interface in case of failure of the "primary" interface. More recent work on SCTP investigates how one can use multihoming for concurrent transfers. However, the application remains bound to the SCTP semantics.

A few transport protocols, such as pTCP [52], are explicitly designed with transport layer striping in mind. pTCP stripes a connection over a set of (modified) TCP connections (one per interface), and can achieve high throughput aggregation. While pTCP is an excellent option for transport layer striping, the disadvantage of transport layer striping is that it imposes a specific set of transport layer semantics. The semantics, which, in pTCP's case, are similar to those of TCP, may be unsuitable for some applications such as media streaming.

2.1.2.2.4 Striping at application layer

As applications are aware of the characteristics of the data being transferred, application layer striping can theoretically provide fine-grained performance tuning. For instance, an application may achieve high bandwidth aggregation by sending data via multiple sockets on multiple interfaces [53]. However, due to head-of-the-line blocking phenomena at the transport layer [52], application layer striping can also result in throughput well below the capacity of the slowest path when (1) in-order packet delivery is a must, and (2) the underlying physical paths have a wide range of delay-bandwidth products.

2.2 CROSS-LAYER OPTIMIZATION

In the past few years, there has been an avalanche of cross-layer design proposals for wireless networks. A number of researchers have looked at specific aspects of network performance and have presented several cross-layer design proposals, depending on their interpretation of what it implies,. These proposals involve different layers of the protocol stack, and address both cellular and ad hoc networks. As wireless communications and networking have rapidly occupied centre stage in research and development activity in the area of communication networks, the suitability of one of the foundations of networking, namely the layered protocol architecture, is coming under close scrutiny by the research community. It is repeatedly argued that although layered architectures have worked well for wired networks, they are not suitable for wireless networks. Generally speaking, cross-layer design refers to protocol design done by actively exploiting the dependence between protocol layers to obtain performance gains.

We will first introduce a brief overview of the proposed cross-layer design principles that have been predominantly followed in the literature. Afterwards, we will describe some of the efforts made on standardization regarding cross-layer optimization.

2.2.1 A snapshot of cross-layer design proposals

As mentioned above, there are many cross-layer design proposals in the literature. The authors in [54] present a survey of several cross-layer design proposals from the literature

based on the layers that are coupled. Nevertheless, it is more interesting to see how the layers are coupled, in other words, what kind of architecture violation has taken place in a particular cross-layer design. The layered architecture rules have been violated in the following basic ways: creation of new interfaces, merging of adjacent layers, design coupling without new interfaces and vertical calibration across layers.

2.2.1.1 Creation of new interfaces

The layered architecture forbids direct communication between nonadjacent layers; communication between adjacent layers is limited to procedure calls and responses. In this type of cross-layer designs, new interfaces are used for information sharing between the layers at runtime. The architecture violation here is obviously the creation of new interfaces, unavailable in the layered architecture. These interfaces can be divided into subcategories depending on the direction of information flow along the new interfaces:

- upward: from lower layer(s) to a higher layer. For example, the explicit congestion notification (ECN) from the router to the transport layer at the TCP sender can explicitly tell the TCP sender if there is congestion in the network to enable it to differentiate between errors in the wireless link and network congestion [55], or similar upward information flow at the MAC layer (link layer in general) in form of channel-adaptive modulation or link adaptation schemes [56]. The idea is to adapt the parameters of the transmission (e.g., power, modulation, code rate) in response to the channel condition, which is made known to the MAC layer (link layer) by an interface from the physical layer,
- downward: from higher layer(s) to a lower layer. Some cross-layer design proposals rely on setting parameters on the lower layer of the stack at runtime using a direct interface from some higher layer. As an example, applications can inform the link layer about their delay requirements, and the link layer can then treat packets from delay-sensitive applications with priority [57],
- or back and forth: iterative flow between two layers. As an example, we refer to the network-assisted diversity multiple access (NDMA) proposal [58], whereby the physical (PHY) and MAC layers collaborate in collision resolution in the uplink of a wireless LAN system.

2.2.1.2 Integration of adjacent layers

Another way to do cross-layer design is to design two or more adjacent layers together such that the service provided by the new superlayer is the union of the services provided by the constituent layers. This does not require any new interfaces to be created in the stack. Architecturally speaking, the superlayer can be interfaced with the rest of the stack using the interfaces that already exist in the original architecture.

Although it does not explicitly create a superlayer, it is interesting to note that the collaborative design between the PHY and MAC layers (discussed earlier with the NDMA idea) tends to blur the boundary between these two adjacent layers.

2.2.1.3 Design pairing without new interfaces

Another category of cross-layer design involves coupling two or more layers at design time without creating any extra interfaces for information sharing at runtime. While no new interfaces are created, the architectural cost here is that it may not be possible to replace one layer without making corresponding changes to another layer.

For instance, [59] considers the design of a MAC layer for the uplink of a wireless LAN when the PHY is capable of providing multipacket reception capability. Multipacket reception capability implies that the PHY is capable of receiving more than one packet at the same time. Notice that this capability at the physical layer considerably changes the role of the MAC layer; thus, it needs to be redesigned.

2.2.1.4 Adjustment across layers

This refers to calibrating parameters that span across layers. Basically, the performance seen at the level of the application is a function of the parameters at all the layers below it. Hence, it is conceivable that joint tuning can help to achieve better performance than individual settings of parameters — as would happen had the protocols been designed independently — can achieve.

As an example, [60] presents an example where the delay requirement dictates the persistence of link-layer automatic repeat request (ARQ), which in turn becomes an input for deciding the rate selection through a channel-adaptive modulation scheme.

Adjustment can be done in a static manner, which means setting parameters across the layers at design time with the optimization of some metric in mind. This can also be done dynamically at runtime, which emulates a flexible protocol stack that responds to variations in the channel, traffic, and overall network conditions.

2.2.2 PILC (IETF)

The Internet network-layer and transport-layer protocols are designed to accommodate a very wide range of networking technologies and characteristics. Nevertheless, experience has shown that the particular properties of different network links can have a significant impact on the performance of Internet protocols operating over those links, and on the performance of connections along paths that include such links. This is especially relevant to the wireless networking community.

The Internet Engineering Task Force (IETF) Performance Implications of Link Characteristics (PILC) working group in its RFC 3135 [61] discussed the capabilities, limitations and pitfalls of PEPs, that is, active network elements that modify or splice end-to-end flows in an attempt to enhance the performance they attain in the face of particular link characteristics. Emphasis is put on proxies operating with TCP.

Many types of PEPs can be distinguished. Different types of PEPs are used in different environments to overcome different link characteristics which affect protocol performance. Note that enhancing performance is not necessarily limited in scope to throughput. Other performance related aspects, such as usability of a link, may also be addressed.

Layering, *Distribution*, *Implementation Symmetry* and *Transparency* are the main features used to classify PEPs. In principle, a PEP implementation may function at any protocol layer. Nevertheless, link layer mechanisms can be and typically are implemented transparently to network and higher layers, requiring no modifications to protocol operation above the link layer. It should also be noted that some PEP implementations operate across several protocol layers by exploiting the protocol information and possibly modifying the protocol operation at more than one layer.

A PEP implementation may be integrated, i.e., it comprises a single PEP component implemented within a single node, or distributed, i.e., it comprises two or more PEP components, typically implemented in multiple nodes. An integrated PEP implementation represents a single point at which performance enhancement is applied. A distributed PEP implementation is generally used to surround a particular link for which performance enhancement is desired.

A PEP implementation may be symmetric or asymmetric. Symmetric PEPs use identical behaviour in both directions, i.e., the actions taken by the PEP occur independent of which interface a packet is received from. Asymmetric PEPs operate differently in each direction. An asymmetric PEP implementation is generally used at a point where the characteristics of the links on each side of the PEP differ or with asymmetric protocol traffic. For example, an asymmetric PEP might be placed at the intersection of wired and wireless networks. A PEP implementation may also be both symmetric and asymmetric at the same time with regard to different mechanisms it employs. Whether a PEP implementation is symmetric or asymmetric is independent of whether the PEP implementation is integrated or distributed.

Finally, another key characteristic of PEPs is their degree of transparency. PEPs may operate totally transparently to the end systems, transport endpoints, and/or applications involved (in a connection), requiring no modifications to the end systems, transport endpoints, or applications. On the other hand, a PEP implementation may require modifications to both ends in order to be used. In between, a PEP implementation may require modifications to only one of the ends involved. It is sometimes useful to think of the degree of transparency of a PEP implementation at four levels, transparency with respect to the end systems (network-layer transparent PEP), transparency with respect to the transport endpoints (transport-layer transparent PEP), transparency with respect to the applications (application-layer transparent PEP) and transparency with respect to the users.

2.2.3 Interlayer Interactions (IRTF)

The Ad hoc Network Systems Research subgroup belonging to the IRTF is concerned with the study of Ad hoc Network Systems (ANS). Ad hoc networks are complex systems, with cross-layer protocol dynamics and interactions that are not present in wired systems, most prominently between the physical, link and network layers. The IETF community and the wider research community could benefit from research into the behaviour of ad hoc networks that would enable advanced routing protocol development. This research group

will endeavour to develop sufficient understanding in topic areas of interest to enable the desired protocol specification work.

Interlayer interaction is the inter-communication between different layers of the network protocol stack. This subgroup provides interlayer interaction metrics and the related information exchange procedures to enhance performance of the wireless ad hoc networks. It also highlights the importance of interlayer interaction in determining the scalability of a method.

Ad hoc networks are infrastructure-less, multi-hop routing, wireless mobile networks formed spontaneously. Multi-hop routing, random movement of the nodes and other features unique to ad hoc networks results a large control signal overhead for route discovery and maintenance. This is highly unacceptable in bandwidth-constrained ad hoc networks. Usually these devices have limited computing resources and severe energy constraints. Due to these characteristics, there is a lot of research being done in the performance optimization of ad hoc networks. However, most of the research is based on optimization at individual layer.

In [62] authors indicate that optimizing a particular layer might improve the performance of that layer locally but might produce non-intuitive side effects that will degrade the overall system performance.

Even though the scope of the IRTF ANS subgroup is quite broad, with large-scale ad hoc networks, it briefly summarizes some interlayer interaction metrics, points to the information to be exchanged and the benefits of such information exchange between lower layer, routing layer and transport layer. This is useful when designing and standardizing an adaptive architecture that can exploit the interdependencies among link, medium access, network and application protocols. The architecture, where each layer of the protocol stack responds to the local variations as well as to the information from other layers is a major challenge.

Currently ad hoc routing protocols are researched to work mainly on the network layer. They guarantee the independence of the network layer. However each layer needs to do redundant processing and unnecessary packet exchange to get information that is easily available to other layers. This increases control signals resulting in wastage of bandwidth, packet collision, etc. By using interlayer interaction, different layers can share locally available information. This will result in a substantial performance improvement. Any method is scalable with respect to a (parameter, metric, environment) triple [63]. The interlayer interactions are nothing but the interactions among various environmental parameters. These interactions may affect the fairness of the scalability comparison between different methods. Thus, it is important to standardize these parameters as well as the interactions among them. Ad hoc network protocol architecture can give optimized performance by exploiting the environmental parameters information flow across different layers.

Interlayer interaction schemes that can support adaptability and optimization of the routing protocols are needed. We need to treat the entire stack as a single algorithmic

construct in order to improve the performance. Such optimized routing protocol can discover and maintain the routes based on current link status, traffic congestion, signal strength etc. There is need for a general framework to improve performance in scalable ad hoc networks.

Some of the metrics envisaged to support their operation are: (1) Signal strength; (2) Interface identifiers; (3) Link failure detection time; (4) Interference; (5) Congestion; (6) Bandwidth; (7) SNR information; (8) Link type information. Nevertheless, other important aspects such as the ones related to the network nodes computational load conditions are also studied as part of the parameters used in the cost functions used to calculate the optimal routes within the network.

2.2.4 Interlayer Coordination Model

In [64] one possible model for interlayer coordination consists of a set of modules (protocols) connected to a central interlayer coordination manager. The modules display events and state variables to the manager. Events are notifications sent to the manager, such as handover begins or link lost. They are used to trigger, or "wake up," the management algorithms. State variables represent entry points to get/set operations that allow the manager to query or modify the internal state of a protocol/module.

Table 2-1 lists the events and state variables (control points) that need to be displayed by protocols to be used by management algorithms. The problems that have been identified in this document are presented below.

Table 2-1: Interfaces displayed for Interlayer Coordination

Protocol/module	Events	State variables
TCP	Connection initiated Connection terminated Acknowledgement timeout	Congestion window (cwnd) Retransmission timeout (RTO) Round-trip time (RTT) Slow start threshold (SSthreshold)
IP	Handover start Handover end Routing table updated QoS reservation completed QoS Modification	Routing table Security associations List of QoS-enabled flows Per-flow QoS
Link	Link lost Fading start Fading end Corrupt packet received	Capacity (bandwidth), delay, BER Long retry limit Transmission power Noise/interference level

2.2.5 Standardization Status for Cross-layer Design and Interoperability

Wireless networks can be implemented attending to many diverse characteristics. Coverage, mobility support, and throughput differ significantly between different network architectures, and there is no single system that currently stands out as the best solution for mobile and nomadic users. The system performance of future networks will be enhanced by cross-layer design between PHY, MAC and higher layer protocols, as

discussed in earlier sections. However, for a combination of services and applications demanded by users in a variety of environments, each wireless network can provide an important piece in the large mobile multimedia jigsaw puzzle. Standardization efforts are in progress to integrate various architectures.

The important co-design of physical layer, MAC and higher layers has been taken into account in some of the latest standardization issues. 3G standards such as CDMA2000 have been designed with cross-layer issues in mind. The Data Link Control (DLC) and physical layer of BRAN HiperLAN2 have been designed for high throughput, low latency and QoS support [65]. In 3GPP, the enhancement High Speed Downlink Packet Access (HSDPA) with hybrid ARQ and scheduling at the base-station has been introduced for reduced latency. The important topic of cross-layer design has recently been considered in the Study Group on Mobile Broadband Wireless Access Networks (MBWA) within IEEE, with the goal of improving throughput and reducing latency both in downlink and uplink.

Interoperability between standards of distinct networks is crucial both for user adoption and operator management in a wide scale deployment of products and services over wireless networks. At the time of writing, dozens of specialized applications and network protocols are being designed for dual use on IEEE 802.11a/b and cellular/PCS networks [66]. End user equipment for multi-mode terminals, such as PDAs, web phones and smart-cards, are emerging on the market.

Successful integration of existing and emerging wireless systems is a demanding task. Inherently, several standardization and regulatory bodies need to be involved, and the process is made difficult by the status of each specification process. A first WLAN/WWAN standardization initiative was taken by the ETSI BRAN project, which started the UMTS-HiperLAN2 interworking specification work. The scope has been generalized to include all WLAN standards with a loose coupling reference architecture based on the Internet Engineering Task Force (IETF) protocols. The standardization work was adopted by 3GPP, where a phased approach based on six scenarios ranging from common billing and access control toward seamless services has been defined for future releases. A new Wireless Interworking Group (WIG) was established in 2002 as a joint effort by the WLAN Community including ETSI BRAN, IEEE 802.11 and MMAC HiSWANa. The scope of WIG is to solve the interworking issues by defining an interface behind the WLAN AP toward the IP core network.

2.3 RESEARCH APPROACHES TO NETWORK SELECTION

Decision metrics and need for decision policy design are outlined in [67] in the context of vertical handoff for a single mobile user running multiple communication sessions. These metrics are also relevant to the initial network selection decision for any user with a choice in available networks.

Much of the work in Radio Access Network (RAN) selection policy design investigates the access network selection problem as part of the seamless handover venture. In [68] authors look at how to facilitate a user making a network interface selection decision. They concentrate on a possible architecture for the end terminal and not on any particular

strategy, but they do mention an Always Cheapest (AC) network selection strategy. [69] proposes an agent-based architecture with a user agent decision function. Customers compare and select services with the best performance/price ratio, negotiate with providers for offered services and pre-reserve the resources for an agreed price. Details of the exact negotiation terms are not covered but they are based on differently weighted QoS parameters such as delay, bandwidth, packet loss, etc.

In [70] a handover 'policy' for heterogeneous wireless networks, which is used to select the 'best' available network and time for handover initiation is described. It considers the cost of using a particular network in terms of the sum of weighted functions of bandwidth, power consumption and cost. Bandwidth is determined either by using an agent in each RAN which estimates and broadcasts the current network load, or in the case of commercial networks, by the 'typical' value of bandwidth advertised by these networks. The network which is consistently calculated to have lowest cost is chosen as the target network. A randomised waiting scheme based on the impact of the estimated handover delay is used to achieve stability and load-balancing in the system, and to avoid handover synchronisation.

A number of papers, all of which reference [70] use similar cost functions. A smart decision model for vertical handoff is the focus of [71]. The proposed scheme relies on a score function which, is based on functions of allocated bandwidth (transfer completion time), battery power consumption and cost charged by the available networks. The value or benefit function of offered link capacity is a concave increasing function following the economic assumption of diminishing marginal utility, i.e. value will increase for each unit of added bandwidth up to certain point after which the gain in value for extra capacity is marginal as described in [72]. The bandwidth in this case is measured using a probing tool. Both [70] and [71] collect current information on bandwidths available on all local networks – this requires heavy power consumption and introduces a certain delay, factors which should possibly be added to the cost of implementing the suggested strategies. The HTTP handoff decision model presented in [73] and the work in [67] also utilise a cost function which is used to compare available networks from the list of options and establish the network to handoff to according to the importance weightings associated with different metrics.

The user network selection decision strategy will be influenced by the pricing scheme employed in the available networks. [74] looks at networks with an auction based pricing scheme and employs two different strategies based on user preference for low service charge, or user preference for networks with a good reputation. Many new pricing schemes are being proposed for RANs [75], the majority focus on network-centric benefit.

Utility-based functions are commonly used to describe user preference rating relationships for a number of metrics. [76] considers user preferences to be represented quantitatively through a utility function when comparing two congestion pricing schemes. In [77] authors consider users to choose a pricing plan based on their data delay considerations, described by a user utility function. They use their understanding of user behaviour to maximise network gains, however they do not consider a user-centric Service Oriented

Heterogeneous Wireless Network Environment (SOHWNEs), but instead look at efficient network resource management with the goal of reducing customer turnover and maximizing wireless network operator revenue.

In [78] the complexity of being 'Always Best Connected' is analyzed. The decision about which network to select is user-centric. It involves identifying the network, or combination of access networks, from the available candidates which will best satisfy the current user requirements in their current circumstance. The complexity of the access network selection decision problem is mapped onto an NP-hard optimisation problem. Heuristics may be used to solve this combinatorial optimisation problem to produce optimal or near optimal solutions in some input instances.

2.4 POWER-AWARE COMMUNICATIONS ON MULTIHOP MOBILE NETWORKS

Energy efficiency is critical to ensuring scalability, embedding, and portability of emerging computing and communication systems. It is of particular interest in the design of mobile computing systems because of the limitations in energy and power availability.

We will present some of the main approaches that have been followed on the different layers in order to achieve power awareness in mobile ad hoc networks. Different strategies are followed depending on the layer in which the technique is applied.

2.4.1 Power conservation at the Link Layer

The goal of link layer design in ad hoc networks is to achieve rates close to the fundamental capacity limits of the channel while overcoming channel impairments using relatively little energy. The fundamental capacity of a channel dictates the maximum data rate that can be transmitted over the channel with an arbitrarily small probability of error. The pioneering work by Gallager [79] in this area defines reliable communication under a finite energy constraint in terms of the capacity per unit of energy. It indicates that ad hoc networks with finite energy nodes only have a finite number of bits that a given node can transmit before exhausting its energy. Allocating those bits to the different requirements of the network such as information transmission, exchange of routing information, forwarding bits for other nodes, channel estimation, and so on becomes an interesting and challenging optimization problem that clearly requires cross-layer design.

Channel coding can significantly reduce the power required to achieve a given bit error rate and is therefore a common feature in energy constrained link layer design, the study is mainly concentrated on turbo coding and space-time coding [80]. Multiple antennas at the transmitter and receiver have an importatnt role in reducing the required transmitted power. Simulation in [81] has shown a 45% improvement in energy cost for routing is achieved using multiple directional antennas.

2.4.2 Power conservation at the MAC Layer

Firstly, at the MAC layer, unnecessary collisions should be eliminated as much as possible since retransmission incurs power consumption. The collisions introduced by hidden terminals and inefficiencies introduced by exposed terminals are often addressed by a

four-way handshake prior to transmission, as in the 802.11 WLAN protocol. However, this handshake protocol is based on single-hop routing. Another technique to avoid hidden and exposed terminals is busy tone transmission, but this scheme works well in preventing collisions only when a centralized controller can be "heard" by users throughout the network. Hybrid techniques using handshakes, busy tone transmission, and power control are investigated in [82].

Secondly, in another scenario, instead of having the receiver powered on at all times in a wireless network, we could instead broadcast a schedule defining the transmission times for each mobile device, hence allowing mobile devices to switch to standby mode to save power. Moreover, in a single-transceiver system, switching from transmit to receive mode is required to support both uplink and downlink communications. By providing a means for allocating contiguous slots for transmission and reception, power consumed as a result of transceiver mode switching can be reduced. In terms of channel reservation, power may be conserved by supporting the request of multiple slots with a single reservation packet. From another perspective, considerable power savings can be obtained by intelligently turning off radios when they cannot transmit or receive packets.

Thirdly, delay is another problem that should be controlled in ad hoc network. A power control strategy for multiple accesses that takes into account delay constraints is proposed and analyzed in [83]. This strategy optimizes the transmitted power relative to both channel conditions and the delay constraint via dynamic programming. The optimal strategy exhibits three modes: very low-power transmission when the channel is poor and the tolerable delay large, higher power when the channel and delay are average, and very high-power transmission when the delay constraint is strict. This strategy exhibits significant power savings over constant power transmission while meeting the delay constraints of the traffic.

2.4.3 Power conservation at the Network Layer

The most important function of the network layer is to support routing. In ad hoc networks, there are no base stations and each node acts as a router and packet forwarder. Hence, the computation and communication (packet processing, transmission, reception, etc.) load can be quite high. In fact, power control impacts on the routes employed and vice-versa, the power control strategy needs connectivity information that is provided by the routing layer. This mutual dependency motivates the need for a joint solution for power control and routing.

In [84], new power-aware metrics are used for determining routes in wireless ad hoc networks. It was argued that routing protocols that derive routes based on minimizing hop count or delay will result in some nodes depleting their energy reserves faster, causing them to be powered down at an earlier stage. In addition, routing packets through lightly loaded nodes is considered energy conserving since there is less likelihood of contention. Some of the metrics proposed for power-aware routing are: (a) minimize energy consumed per packet, (b) maximize time to network partition, (c) minimize

variance in node power levels, (d) minimize cost per packet, and (e) minimize maximum node cost.

One can also reduce power by improving periodic route updates, as commonly found in table-driven routing protocols. Such protocols result in intensive route updates when the frequency of link changes increases. Mobile agent-based topology discovery techniques can also help to reduce excessive power consumption. Transmitting packets with a more compact header is another power efficient technique. We can see the routing protocol under energy constraints must somehow balance delay constraints, battery lifetime, and routing efficiency. So a cross-layer design-based routing protocol is necessary. One example that combines routing, power control, and adaptive coding to minimize the energy cost of routes can be found in [85]. Another simple but effective protocol called COMPOW is intorduced in [86]. COMPOW protocol simultaneously satisfies the three objectives of maximizing the traffic carrying capacity of the entire network, extending battery life through providing low-power routes, and reducing the contention at the MAC layer. Furthermore, the protocol has the plug and play feature that pro-actively maintains a routing table.

2.4.4 Power conservation at the Transport Layer

In ad hoc networks, frequent topology change may lead to sudden packet losses and delays. Transport protocols such as TCP, which have been designed for reliable fixed networks, misinterpret this packet loss as congestion and invoke congestion control, leading to unnecessary retransmissions and power consumption. To overcome this problem, a feedback scheme is proposed so that the source can distinguish between a route failure and network congestion [87].

Meantime, ARQ is used to provide error control in TCP. ARQ employs error-detection codes to detect errors and request retransmission. Repeated retransmissions consume power and should be avoided as much as possible while still maintaining a certain level of communication performance. Several proposals for power efficient error control schemes have been proposed [88].

The study in [89] reports that the performance of multiple TCP connections or UDP over various ad hoc routing protocol and link states is quite different in fairness coefficient and throughput. So the power control strategy in transport layer must be coordinated with network layer and other lower layers.

2.5 WIRELESS NETWORKS SECURITY

Wireless communications offer organizations and users many benefits such as portability and flexibility, increased productivity, and lower installation costs. Wireless technologies cover a broad range of differing capabilities oriented toward different uses and needs. Wireless Local Area Network (WLAN) devices, for instance, allow users to move their laptops from place to place within their offices without the need for wires and without losing network connectivity. Less wiring means greater flexibility, increased efficiency, and reduced wiring costs. Ad hoc networks, such as those enabled by Bluetooth, allow

data synchronization with network systems and application sharing between devices. Bluetooth functionality also eliminates cables for printer and other peripheral device connections. Handheld devices such as personal digital assistants (PDA) and cell phones allow remote users to synchronize personal databases and provide access to network services such as wireless e-mail, Web browsing, and Internet access. Moreover, these technologies can offer dramatic cost savings and new capabilities in diverse applications ranging from retail settings to manufacturing shop floors to first responders. However, risks are inherent in any wireless technology. Some of these risks are similar to those of wired networks; some are exacerbated by wireless connectivity; some are new. Perhaps the most significant source of risks in wireless networks is that the technology's underlying communications medium, the airwave, is open to intruders, making it the logical equivalent of an Ethernet port in the parking lot.

2.5.1 Security Requirements and Threats

The NIST handbook "An Introduction to Computer Security" [90] generically classifies security threats in nine categories ranging from errors and omissions to threats to personal privacy. All of these represent potential threats in wireless networks as well. However, the more immediate concerns for wireless communications are device theft, denial of service, malicious hackers, malicious code, theft of service, and industrial and foreign espionage.

As is shown in Figure 2-4, network security attacks are typically divided into passive and active attacks. These two broad classes are then subdivided into other types of attacks. All are defined below.

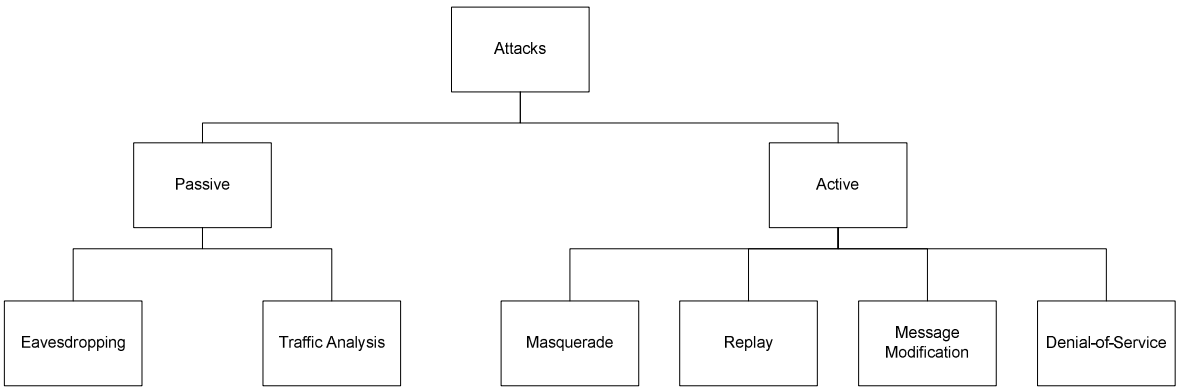


Figure 2-4: Taxonomy of Security Attacks

- Passive Attack—an attack in which an unauthorized party gains access to an asset and does not modify its content (i.e., eavesdropping). Passive attacks can be either eavesdropping or traffic analysis (sometimes called traffic flow analysis). These two passive attacks are described below.
    - Eavesdropping—the attacker monitors transmissions for message content. An example of this attack is a person listening into the transmissions on a LAN between two workstations or tuning into transmissions between a wireless handset and a base station.

  - Traffic analysis—the attacker, in a more subtle way, gains intelligence by monitoring the transmissions for patterns of communication. A considerable amount of information is contained in the flow of messages between communicating parties.

- Active Attack—an attack whereby an unauthorized party makes modifications to a message, data stream, or file. It is possible to detect this type of attack but it may not be preventable. Active attacks may take the form of one of four types (or a combination thereof): masquerading, replay, message modification, and denial-of-service (DoS). These attacks are defined below.
  - Masquerading—the attacker impersonates an authorized user and thereby gains certain unauthorized privileges.
  - Replay—the attackers monitor transmissions (passive attack) and retransmit messages as if they were the legitimate user.
  - Message modification—the attacker alters a legitimate message by deleting, adding to, changing, or reordering it.
  - Denial-of-service—the attacker prevents or prohibits the normal use or management of communications facilities.

Attacks can also be classified according to network protocol stacks. Table 2-2 shows an example of a classification of security attacks based on the protocol stack; some attacks could be launched at multiple layers.

The consequences of these attacks include, but are not limited to, loss of proprietary information, legal and recovery costs, tarnished image, and loss of network service. Nevertheless, the main aspects that should be tackled when designing the threat mitigation plan are the following:

- **Confidentiality** is to keep the information sent unreadable to unauthorized users or nodes. Wireless networks use an open medium, so usually all nodes within the direct transmission range can obtain the data. One way to keep information confidential is to encrypt the data, and another technique is to use directional antennas.
- **Authentication** is to be able to identify a node or a user, and to be able to prevent impersonation. In wired networks and infrastructure-based wireless networks, it is possible to implement a central authority, but this is not always possible in personal networking scenarios. For example, there is no central authority in MANET, and it is much more difficult to authenticate an entity.
- **Integrity** is to be able to keep the message sent from being illegally altered or destroyed in the transmission. When the data is sent through the wireless medium, the data can be modified or deleted by malicious attackers. The malicious attackers can also resend it, which is called a replay attack.

- **Non-repudiation** is related to the fact that if an entity sends a message, the entity cannot deny that the message was sent by it. By producing a signature for the message, the entity cannot later deny the message. In public key cryptography, a node A signs the message using its private key. All other nodes can verify the signed message by using A's public key, and A cannot deny that its signature is attached to the message.
- **Availability** is to keep the network service or resources available to legitimate users. It ensures the survivability of the network despite malicious incidents.
- **Access control** is to prevent unauthorized use of network services and system resources. Obviously, access control is tied to authentication attributes. In general, access control is the most commonly thought of service in both network communications and individual computer systems.

Table 2-2: Security Attacks on Protocol Stacks

Layer	Attacks
Application layer	Repudiation, data corruption
Transport layer	Session hijacking, SYN flooding
Network layer	Wormhole, blackhole, Byzantine, flooding, resource consumption, location disclosure attacks
Data link layer	Traffic analysis, monitoring, disruption MAC (802.11), WEP weakness
Physical layer	Jamming, interceptions, eavesdropping
Multi-layer attacks	DoS, impersonation, replay, man-in-the-middle

2.5.2 Security Mechanisms for Wireless Communications

In the previous section the main requirements and threats for wireless communications have been identified. We will now briefly present the main techniques that are used in order to tackle the aforementioned threats for supporting the communication networks requirements.

2.5.2.1 Node Authentication

When designing an authentication mechanism the main features that need to be pursued are:

- **Mutual authentication.** Is the two-way authentication between the Authentication Server (AS) and the supplicant. The supplicant authenticates the AS to ensure that it is not communicating with an imposter pretending to be the AS. If the authentication method does not enforce mutual authentication, an imposter pretending to be the AS may be able to act as a man-in-the-middle between the

supplicant and the AS and gather private messages from the client (this attack is called the man-in-the-middle attack). Borisov et al. [91] showed that the absence of mutual authentication in Wired Equivalent Privacy (WEP) allowed a myriad of attacks, including man-in-the-middle attacks, which an attacker can use to decrypt WEP encrypted messages.

- **Identity privacy.** Hiding the client's identity (e.g., his username or email address) from eavesdroppers of the authentication process.
- **Dictionary attack resistance**. In a dictionary attack, the victim must have some potentially guessable secret (usually a password or passphrase), and the attacker has access to some data derived from the secret in a known way, typically independent of the context. Thus, the attacker can verify guesses and, if the derivation is independent of context, the attacker can pre-compute a dictionary of likely passwords.
- **Replay attack resistance**. In a replay attack, an eavesdropper records the authentication process of a legitimate client and replays it to gain the access to the network. Note that replay attacks are possible even when the eavesdropper does not know the secret required for the authentication process. An authentication protocol can resist this attack by including a nonce –a timestamp or a sequence number– in the authentication process so that the authenticating parties can detect that a replayed authentication session is not fresh.
- **Derivation of strong session keys**. Any secret is not likely to remain secret indefinitely. If an eavesdropper sniffs many messages encrypted with the same secret, he may eventually be able to derive the secret key from the messages. Authentication protocols should derive a different session key to use for each session's message-protection process. Thus, even if an eavesdropper discovers the secret key, the eavesdropper cannot decrypt the messages from past or future sessions using the stolen key. Moreover, the longer-lived, multi-session secret is only used to derive the session key during the authentication process. As the authentication process is shorter, it is less likely for the attackers to gather enough messages encrypted with the same secret.
- **Fast reconnect**. Unlike a wired network, a wireless one gives the client the freedom to move from one attachment point to another while maintaining his connection. When the client changes location and associates with another AP, the new AP, which did not broker the authentication process, may not be aware of the trust that the supplicant and the previous authenticator established. In such cases, the client may lose connection to the network until he re-authenticates via the new AP.

2.5.2.2 Data Encryption

Cryptography can be directly used to help ensure these security properties:

- **Confidentiality,** preventing open access.
- **Integrity,** preventing unauthorized modification.

- **Authentication**, verification of identity. Sometimes split into: entity authentication and data origin authentication.
- **Non-repudiation,** preventing denial of actions

These properties must be ensured even when another party may eavesdrop or intercept messages. Carefully designed cryptographic protocols help this.

2.5.2.3 IP Security

TCP/IP communication can be made secure with the help of cryptography. Cryptographic methods and protocols have been designed for different purposes in securing communication on the Internet. These include, for instance, the SSL [92] and TLS [93][94] for HTTP Web traffic, Secure Multipurpose Internet Mail Extensions (S/MIME) [95] and Pretty Good Privacy (PGP) [96] for e-mail. Nevertheless, when talking about IP security, the referenced technique involves security at network level. IPsec is designed to protect communication in a secure manner by using TCP/IP. The IPsec [97] protocol is a set of security extensions developed by the IETF and it provides privacy and authentication services at the IP layer by using modern cryptography.

To protect the contents of an IP datagram, the data is transformed using encryption algorithms. There are two main transformation types that form the basics of IPsec, the Authentication Header (AH) and the Encapsulating Security Payload (ESP). Both AH and ESP are two protocols that provide connectionless integrity, data origin authentication, confidentiality and an anti-replay service. These protocols may be applied alone or in combination to provide a desired set of security services for the IP layer. They are configured in a data structure called a Security Association (SA).

The basic components of the IPsec security architecture are explained in terms of the following functionalities:

- Security Protocols for AH [98] and ESP [99].
- Security Associations for policy management and traffic processing
- Manual and automatic key management for the Internet Key Exchange (IKE) [100], the Oakley key determination protocol and Internet Security Association and Key Management Protocol (ISAKMP).
- Algorithms for authentication and encryption

The set of security services provided at the IP layer includes access control, connectionless integrity, data origin authentication, protection against replays and confidentiality. The modularity of the design makes it algorithm independent and permits selection of different sets of algorithms without affecting the other parts of the implementation.

2.5.2.3.1 IPSec Tunnel vs Transport mode

IPSec can be run in either tunnel mode or transport mode. Each of these modes has its own particular uses and care should be taken to ensure that the correct one is selected for the solution:

- Tunnel mode is most commonly used between gateways, or at an end-station to a gateway, the gateway acting as a proxy for the hosts behind it.
- Transport mode is used between end-stations or between an end-station and a gateway, if the gateway is being treated as a host—for example, an encrypted Telnet session from a workstation to a router, in which the router is the actual destination.

As Figure 2-5 shows, basically transport mode should be used for end-to-end sessions and tunnel mode should be used for everything else.

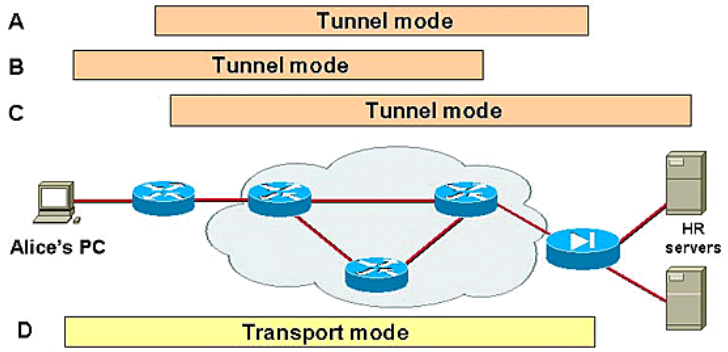


Figure 2-5: Tunnel and transport modes in IPSec

Figure 2-5 displays some examples of when to use tunnel versus transport mode:

- Tunnel mode is most commonly used to encrypt traffic between secure IPSec gateways, such as between the router and Firewall (as shown in example A in Figure 2-5). The IPSec gateways proxy IPSec for the devices behind them, such as Alice's PC and the HR servers in Figure 2-5. In example A, Alice connects to the HR servers securely through the IPSec tunnel set up between the gateways.
- Tunnel mode is also used to connect an end-station running IPSec software to an IPSec gateway, as shown in example B.
- In example C, tunnel mode is used to set up an IPSec tunnel between the router and a server running IPSec software.
- Transport mode is used between end-stations supporting IPSec, or between an end-station and a gateway, if the gateway is being treated as a host. In example D, transport mode is used to set up an encrypted Telnet session from Alice's PC running a Virtual Private Network (VPN) Client software to terminate at the Firewall, enabling Alice to remotely configure the Firewall securely.

2.5.2.3.2 IP Authentication Header

The IP AH is used to provide data integrity and authentication for IP packets. It also provides protection against replays. The AH provides authentication for the IP header, as well as for upper-level protocol (TCP, UDP) data. However, some IP header fields may change in transit and the sender may not be able to predict the value of these fields when the packet arrives at the receiver. Thus, the protection provided to the IP header by AH is somewhat piecemeal.

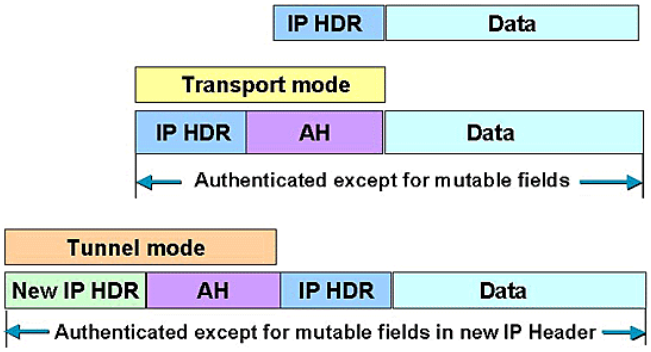


Figure 2-6: AH tunnel versus transport mode.

As is shown in Figure 2-6 AH can be employed in two ways: transport mode or tunnel mode. The transport mode is applicable only to host implementations and provides protection for upper-layer protocols. In the transport mode, AH is inserted after the IP header and before an upper-layer protocol (TCP, UDP or ICMP), or before any other IPsec header that may have already been inserted.

2.5.2.3.3 IP Encapsulating Security Payload

ESP is used to provide confidentiality (encryption), data authentication, integrity and anti-replay service, and limited traffic flow confidentiality. As is shown in Figure 2-7, the ESP header is inserted after the IP header and before the upper-layer protocol header (transport mode) or before an encapsulated IP header (tunnel mode).

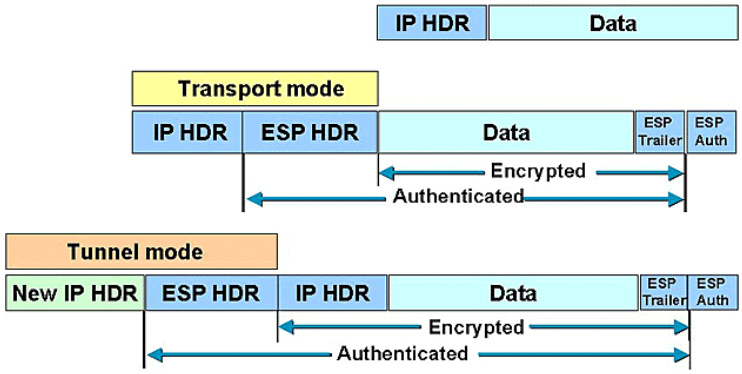


Figure 2-7: ESP tunnel versus transport mode.

In transport mode, the IP payload is encrypted and the original headers are left intact. The ESP header is inserted after the IP header and before the upper-layer protocol header. The upper-layer protocols are encrypted and authenticated along with the ESP header. ESP does not authenticate the IP header itself.

2.5.3 Security in Multi-Hop Routing

In general, the wireless environment is susceptible to security threats. This is even more visible in ad hoc networks as all nodes have access to other user's data while forwarding messages for them. The absence of a trusted third party certification body is one of the main problems for development of appropriate security protocol.

One common feature of most of the existing ad hoc routing protocols is the lack of any built-in security [101]. There are two main approaches to deal with security in ad hoc networks [102]. The first is proactive, using cryptographic techniques to prevent security threats. The second is reactive, detecting threats and reacting accordingly. Most of the protocols for secure routing adopt the proactive approach. The reactive approach is widely used for protection of packet forwarding operations. Complete security solutions do not rely on prevention alone, but also include measures for detecting intrusions and reacting accordingly to avoid persistent adverse effects.

Prevention in mobile ad hoc networks is typically performed by extending the basic routing protocols with cryptographic primitives, such as Hash Message Authentication Codes (HMAC), digital signatures, and hash chains, to authenticate the routing messages. In this way the attacker is prevented from installing incorrect routing states at other nodes. Detection of attacks by discovery and identification of abnormal behaviour mostly takes place either at the end nodes or by neighbouring nodes through overhearing and collaborative consensus. Once the hostile node has been identified, adjustments take place in the routing and forwarding operations. These measures range from avoiding the node in the route to collectively excluding the node from the network.

Regarding prevention, an example of a secure extension for Dynamic Source Routing (DSR) is Ariadne [103] where a one-way HMAC key chain is used for message authentication. For distance vector routing protocols, such as AODV, methods based on a hop count hash chain have been proposed [104]-[106] so that an intermediate node cannot decrease the hop count in a routing update. A hash chain is used for this purpose. As example, consider secure-AODV [106] that assumes that each node has the certified public keys of all network nodes, so that intermediate nodes can validate in-transit routing packets. The basic idea is that the originator of a message appends an RSA signature and the last element of a hash chain. One of the drawbacks of this method is the high processing overhead on the intermediate nodes due to the use of public key cryptography.

A different approach is taken by the Secure Routing Protocol (SRP) [107]. SRP provides correct end-to-end routing information with the only assumption of the existence of a security association between the node initiating the query and the sought destination. SRP provides one or more route replies whose correctness is verified by the route ”geometry” itself, while compromised and invalid routing information is discarded.

As mentioned above, the protection of the exchange of routing messages has to be complemented with measures to ensure that each node correctly forwards packets according to its routing table. This is typically performed by a reactive approach that comprises a detection technique and a reaction scheme.

Detection can be performed by each node in a localised manner by overhearing ongoing transmissions and evaluating the behaviour of its neighbours. Since the accuracy of this method may be affected by interference, channel errors, mobility, and even abuse by malicious nodes, extra measures must be taken. Detection schemes often integrate and refine the results of individual nodes in a distributed manner to achieve consensus among

a group of nodes. Another approach is to rely on explicit acknowledgments from the destination and/or intermediate nodes.

Examples of the first approach are found in [108] and [109]. In [108] a watchdog is proposed to monitor packet forwarding on top of source routing protocols like DSR. Each node may choose the "best" route, comprised of nodes that do not have a history of avoiding forwarding packets along established routes. In [102] a similar idea is followed for distance vector protocols such as ADOV. It adds a next hop field in AODV packets to make nodes aware of the correct next hop of their neighbours. It also protects from other types of attacks such as modification, duplication and jamming of packets. Each independent detection is signed and flooded. The integration of many such results from different nodes can lead to the exclusion of a node from the network.

An example of acknowledgement-based detection is the fault detection mechanism proposed in [110]. The destination sends back acknowledgements to the source for each packet that has been successfully received. A path becomes suspicious when it has dropped more packets than an acceptable threshold. In such a case, the source initiates a fault detection process consisting of a binary search between itself and the destination, which results in the identification of the faulty node.

After detecting a malicious node, actions must be taken in order to protect the network from additional attacks from the same node. This may be done either globally or based on individual decisions at each end host. In global reaction, all nodes in the network react to a malicious node as a whole. The scheme proposed in [109] falls in this category. The detection of a malicious node results in its exclusion from the network by collectively revoking its certificate. In contrast, in the end-host reaction each node makes its own decision about how to react to a malicious node. The scheme proposed in [108] is based on rating the nodes according to their behaviour. Based on the rating, the source always selects the path with the highest average rating.

2.6 OTHER INITIATIVES IN PERSONAL NETWORKING

Many research initiatives have been proposed in the area of personal and wireless communication. In this section, we list some earlier and current work aimed at either analyzing future personal communication requirements or building integrated solutions in the area of ubiquitous computing.

2.6.1 Wireless World Research Forum

The Wireless World Research Forum (WWRF) [111] was founded jointly by Alcatel, Ericsson, Nokia, and Siemens. The forum deals with the definition of research directions within the Wireless World, i.e., wireless communication beyond the third generation. The objectives of the forum are to contribute to a common and comprehensive vision for the Wireless World by concentrating on the definition of the required research, including investigations of pre-regulatory impact. It will also provide a platform for an in-depth discussion of research results and their implications on spectrum and regulatory issues. The forum will actively disseminate Wireless World concepts. It is intended that the

forum's activities constructively contribute to the work of the UMTS Forum, ETSI, 3GPP, IETF, ITU, and other relevant bodies. The forum invites a worldwide participation and is open to all interested parties.

One outcome of this joint forum was The Book of Visions. The idea behind The Book of Visions was to bring together the experts in this field to gather ideas and outline grand visions and challenges for the research and development of future wireless communication systems.

The WWRF has a global scope targeting complete wireless system challenges from the human perspective and future service concepts in new radio interfaces ranging through ad-hoc networks and cognitive wireless systems. It also tackles other more vertical issues such as security and trust or self-organization features. In this sense, Personal Networks and the WWRF user-centred design approach go together on the path towards next generation wireless paradigms.

2.6.2 Ambient Networks

Ambient Networks (AN) [112] was an integrated project sponsored by the European Commission under the Information Society Technology (IST) priority of the 6th Framework Programme. Its main objective was to create network solutions for mobile and wireless systems beyond 3G by enabling scalable and affordable wireless networking while providing rich and easy to use communication services for all.

The project's main objective is enabling seamless interoperation between heterogeneous networks. Ambient Networks aim to establish this interoperation through a common control plane distributed across the individual, heterogeneous networks. This new common control plane functionality can be deployed both as an integral component of future network architectures and as an add-on to an existing, legacy network, enabling legacy interoperability. Ambient Networks enable interoperation of legacy networks by abstracting from the intricacies of legacy connectivity planes and by providing a number of common communication primitives to services and applications. These two "connectivity" abstractions are provided through two interfaces of the Ambient Networks control space: the Ambient Resource Interface (ARI) and the Ambient Service Interface (ASI).

Ambient Networks is more about the composition of networks and the mechanisms needed to support them than about the users' networks themselves. However, these links are still important for personal network communication and will aid in the attainment of ubiquitous networking with QoS-support and reliability. Ambient Networks therefore constitute an important related work since they are developing mechanisms that PNs can implement afterwards to be part of these ambient networks.

Prototyping work has been carried out within the project although there is no single integrated solution but many different developments on the individual frameworks forming the Ambient Networks' project. From radio technology automatic selection (i.e. user can switch between different flavours of 3G systems, WLAN, Bluetooth, or forthcoming 4G systems depending on what is the best network for a particular service or

multimedia content) to quality of service (QoS) sensitive multimedia services or content adaptation upon composition between multiple networks.

2.6.3 IST PACWOMAN

Power Aware Communications for Wireless Optimised Personal Area Network (PACWOMAN) [113] worked not only on networking issues but also on physical technologies such as Ultra Wide Band (UWB). The network architecture was divided into three levels [114]. The first space was the Personal Area Network (PAN), where personal devices can communicate with each other. In the simplest case, the PAN may be a stand-alone network capable of operation independently from other networks. However, due to the very large range in data rates, it is useful to put some hierarchy in this simple network by separating the low rate devices from the high rate devices. The low rate PAN, also referred to as Virtual Device consists of two types of devices: Basic Terminals (bTs) that can be very simple tele-monitoring sensors or actuators and a Master (M) controlling them and acting as gateway towards other Advanced Terminals (aTs) within the PAN. The second level was the Community Area Network (CAN), which consists of nearby PANs belonging to different people that wish to interact with each other. The PACWOMAN concept includes ad-hoc networking in order to support the ability to form networks anytime, anywhere, while maintaining the integrity of the information and applications within an individual personal area space. To complete the picture, the system provides global communication possibilities to the user with acces to classical Wide Area Network systems (wireless or not). To enable this, the communications go through a Gateway. Issues that arise at this level are the end-to-end QoS and security.

The work on PACWOMAN is highly relevant to personal networks and was partially used as a foundation for developing many of the concepts of personal networks. PACWOMAN addresses requirements related to ubiquitous networking and device heterogeneity.

2.6.4 MyNet

The MyNet project [115] is a recently started project. It is a collaboration between Nokia and Massachusetts Institute of Technology (MIT) and aims to study and develop a network architecture, tools and applications for simple, secure, personal overlay networks. The User Information Architecture (UIA) [116] and the Unmanaged Internet Protocol (UIP) [117] are projects within the MIT, which have been merged with the MyNet initiative of Nokia.

UIP combines the self-management of ad hoc networks with the scalability of IP by creating a self-organized overlay network for personal devices. UIA, on the other hand, is intended to allow global interaction and sharing among information devices between persons. The UIA protocols are the foundation upon which the rest of the MyNet project work is layered. The UIA is based on two principles: (a) security is decoupled from physical connectivity; and (b) establishment of trust is based on social connectivity. This is achieved by creating personal and private name spaces. These are simple to use mechanisms that promote social relationships allowing a user to share access to their devices and resources using these name spaces.

These projects stem from the peer-to-peer research community, but are still highly relevant as they focus on many overlapping areas with personal networks. Security, ease of use, and self-organization are also goals for these projects. An interesting viewpoint of this project for personal networking is that the social-networking principle makes end-to-end cryptographic security possible without needing a universal Public Key Infrastructure (PKI). Simple to use mechanisms will be developed that make it natural for users to promote relationships to choose how and with whom to share limited access to their devices and files. Hence they are developing intuitive mechanisms for "introducing" a device into a personal network and for granting some other user limited access to the resources of a personal network; user friendly personal network navigation and management GUIs and tools.

2.6.5 P2P Universal Computing Consortium

The P2P Universal Computing Consortium (PUCC) [118] is a university and inter-industry cooperative project of some Japanese universities and companies active in Japan, such as NEC, Toshiba, and NTT DoCoMo. The target for PUCC is to develop a seamless Peer-to-Peer (P2P) communications technology platform that enables the creation of ubiquitous services between networked devices.

Devices connecting to the Internet utilize the standardized internet protocol stacks, TCP/IP. However today, there are many devices around us equipped with short distance wireless communication capabilities such as Bluetooth, IrDA, Wibree or Zigbee which are used to form PANs. These technologies were not designed based on TCP/IP, but other communication protocols to satisfy specific usage requirements. However, as these connectivity capabilities are already implemented in devices, richer user experience can be achieved by expanding their scope of usage.

PUCC aims to pave the way for development of new applications by maximizing the use of these technology-specific protocols. This is accomplished by utilizing a gateway to connect these devices which are communicating via different bearers and protocols.

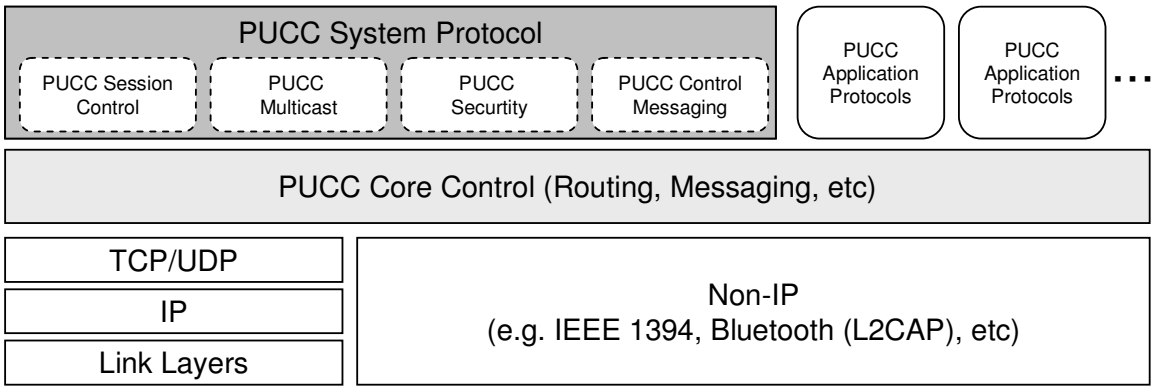


Figure 2-8: PUCC service platform protocol stack

With P2P overlays, they provide seamless communication between IP networks and non-IP networks such as home networks and sensor networks. A service platform provides seamless integration of services and other higher layer functionalities. However, the

network layer is kept as is without any extra support. Figure 2-8 shows the proposed PUCC protocol stack.

PUCC is an organization aiming to overcome these hurdles and provide every device user with an easy way to utilize these device communication capabilities without any special knowledge or skills. For this purpose, industry and academic experts are collaborating in building the solutions. PUCC has already established solutions for several application areas and has succeeded in building functional demonstration systems. Some PUCC members have already begun implementing PUCC solutions into their products.

2.6.6 Other research initiatives

There are numerous other projects that touch on the aspects of personal networks. One such attempt is Universal Personal Networking (UPN) [119]. UPN was a Siemens project in the early 1990s. At that time, WPANs were virtually non-existent and WLAN was very new. However, the aim of UPN was similar to the concept of personal networks, but the existing technologies at that time were a limitation. Hence, UPN focused on infrastructure-support for personal networking, device technology, and user interfaces. Taking place prior to the big break-through of the Internet, they expended a lot of effort on Internet-like techniques and security aspects were neglected. A more recent initiative from Siemens is their LifeWorks [120], which is a visionary concept of a unified communications experience for both business and private users. Under the umbrella of LifeWorks, Siemens develops products that aim at seamless convergence between fixed and mobile networks, new better services for mobile users, and ease of use. It is not only Siemens that works in this direction, but the whole telecommunication industry is showing an increased interest in this area. However, the current focus is more towards businesses and business services.

IBM defined and showcased a concept called Personal Mobile Hub (PMH) [121], which acts as a hub between a PAN and the infrastructure network. It can connect and control the PAN consisting of a person's wireless devices and also interconnect them to servers in the infrastructure. To demonstrate the concept, they developed a health-related application that monitored heart beats and blood pressure and alerted when certain thresholds were exceeded. Furthermore, it could monitor whether a person took his/her medication and if not warn the person and/or caretakers.

In the academic world, it is also worth mentioning the work on personal networking by Robin Kravets' group at University of Illinois at Urbana-Champaign. Among the solutions they worked on, there is one called Mobile Grouped Device (MOPED) [122]. MOPED is a system that represents a person's set of personal devices as one entity towards the Internet using only one single Internet address. That address is given to a proxy node that is always available through the Internet. It is the task of the proxy to keep track of all the other personal devices and how they are connected to the Internet and to each other. Personal devices that can connect directly with each other form what they call components. The components may then connect to the Internet and the proxy. Hence, MOPED provides a technical solution to achieve personal networking and the focus is clearly on addressing, routing, load balancing, and mobility. While they solve many

important aspects, there are many more still left open, such as security, support to higher layers, and direct wireless communication between MOPEDs.

There are many more small projects or specific solutions that target selected areas of personal networks. For instance, HP's CoolTown [123] gives people, places, and things a presence on the web. These dynamic web presences can then be related to each other to form new interesting applications that connects the virtual world with the real world. Standford's Mobile People [124] offers an application-level mobility solution for mobile persons, since it is the people that are the end points and not the devices. It introduces a personal proxy that tracks the person and handles personal-level mobility aspects, including accepting incoming communication on the person's behalf, directing it to the correct device, converting the communication stream if necessary, and protecting the person's privacy. Both CoolTown and Mobile People are completely infrastructure-based and do not consider local communication and many other aspects of personal communication. However, they are useful as parts of the personal networks solution we are aiming for.

It is also worth mentioning that the Third Generation Partnership Project (3GPP) recently started to consider use-cases similar to personal networks in their drive towards All-IP Networks (AIPN) [125]. In fact, they use the term “Personal Networks” for those use-cases, which involve a person with devices in different locations that are interconnected using 3GPP-networks as well as non-3GPP networks.

CHAPTER 3
UNIVERSAL CONVERGENCE LAYER

The concept of isolating the upper-layers from underlying wireless technologies and thus providing real multi-mode can be achieved by introducing a Universal Convergence Layer. The UCL can be seen as a twofold approach. It will mainly act as an enabler for backward and forward compatibility by defining a common interface towards the network layer while managing several different wireless access technologies independently of their PHY and MAC layers. On the other hand, UCL also enables the cross-layer optimisation paradigm. Its privileged location within the protocol stack gives the UCL the possibility to support the information flow both bottom-up (e.g. use of SNR information for enriching the decision-making process in an ad hoc routing algorithm) and top-down (e.g. tuning of MAC parameters depending on the battery status or QoS requirements).

The UCL also plays a key role in security issues as an enabler for providing link-layer security mechanisms that ensure data confidentiality and integrity, authenticity and non-repudiation.

In this thesis, the high-level architecture will be detailed for those components that have been developed in order to make a proof of concept implementation of the UCL. Additionally, this implementation allows us to carry out performance and behavioural analyses based on real testbeds.

The following sections will introduce the architecture designed for the UCL as well as some concepts regarding the technological options chosen to carry out the implementation work. It will also depict the different procedures and data flow of the packets through the UCL.

3.1 HIGH-LEVEL ARCHITECTURE

Figure 3-1 presents the different pieces in which the UCL can be divided. Each of these modules is specialized in providing the different features the UCL offers. This approach allows the easy addition and removal of functionalities depending on the requirements and characteristics of the system on which it will run.

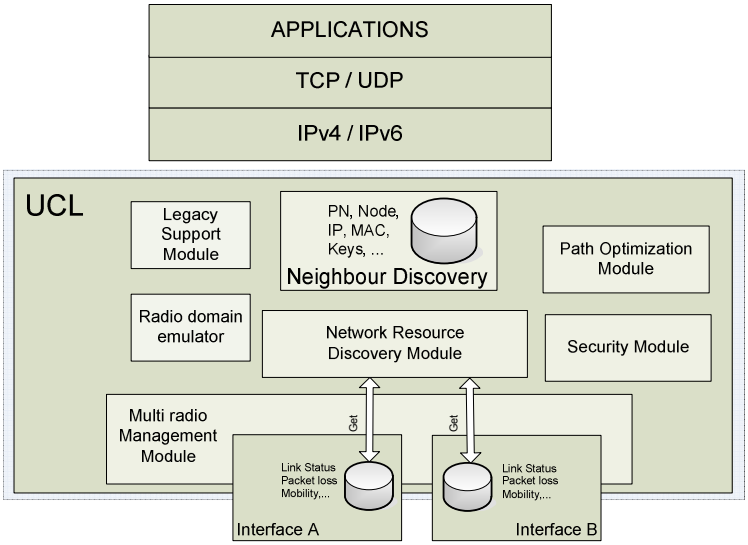


Figure 3-1: Universal Convergence Layer high-level architecture diagram

3.1.1 Multi-radio Management

One of the main objectives of the UCL is to hide the complexity of the available air interfaces and to offer a unique interface to the upper layers. This module will handle this task by discovering and managing the different network resources (set them up, acquire statistics for feeding cross-layer optimization techniques, etc.).

UCL aims at masquerading multihoming by aggregating the different network interfaces (one for each access technology the node is equipped with) on a single interface. By doing this, the IP address of this unique interface become a valid identifier for that host thus alleviating the protocol stack from having to implement multihoming solutions on Layer 3 or 4.

Moreover, UCL provides an overlay Data Link Control (DLC) layer which is set on top of the existing DLC layers of the different access technologies but which does not modify its mode of operation. This way, the UCL can be transparently inserted into the protocol stack since it affects neither the lower nor the upper layers.

On start-up, UCL looks for local wireless network interfaces and incorporates them under its control. Currently, only Bluetooth and IEEE 802.11x are handled when operating in intra-cluster environments, but at any moment more interfaces (WPAN, WLAN or WWAN) can be incorporated both manually and automatically.

3.1.2 Radio Domain Emulator

In [13] a Radio Domain is defined as a collection of nodes/devices with a common radio technology that are controlled by a single MAC mechanism (either centralised or distributed) which can communicate without the need of network layer routing. However, there are wireless access technologies (e.g. Bluetooth when using the BNEP profile) that do not support this definition since it follows a connection-oriented approach that implies the need of establishing a point-to-point connection with a central node. In this sense, all communications are centralized in a node forming a star topology. This module enables the distribution of the packets to the rest of the radio domain in this central node.

This module does not enable inter-radio domain communications as this is achieved by using the ad-hoc routing protocol.

3.1.3 Neighbour Discovery and Authentication

Before delving any deeper into the actual neighbour discovery and authentication mechanisms, it is important to note that the architectural assumption made is that nodes participating in the network have a trust relationship between each pair of them that is established under the supervision of the network owner during the so-called imprinting procedure. The secure cluster formation is based on long-term bilateral shared secrets which are the materialization of these trust relationships. The long-term pair-wise secrets, which are in fact cryptographic keys (also referred to as K_{PN}), are used to form a strong security association between any pair of nodes that are part of the network. The main result of the imprinting is the K_{PN} which is kept together with the public identifier of the node with which this secret is shared. The neighbour discovery and authentication algorithm used in our system relies on the results of the imprinting procedure.

First of all, it is important to note some of the design assumptions that have been made:

- A proactive approach has been selected for forming the cluster and for discovering the peers that become part of it.
- Node discovery is an issue that is resolved at connectivity level. Any node is aware of those nodes and/or devices within the same radio domain.
- The Neighbour Discovery module performs at link layer, so it is only retrieving information about the nodes at a one hop distance.

3.1.3.1 Node Cataloguing

As has already been described, the main characteristic that governs when a node is in or out of the cluster, is the long-term trust relationship established with the other nodes. In this sense, when two nodes meet and discover each other, they can use the shared secrets to verify their membership.

Several relationships can then be identified since not all the nodes around will belong to the same PN:

- FOREIGN: This is a node whose membership has not yet been verified. Thus, it is assumed to be a completely untrusted node and only public communication is allowed with it.
- PERSONAL: This is a node that belongs to the PN. It has already gone through the authentication procedure and its identity has been verified. It has to be noted that personal nodes are treated as foreign nodes until they have proved their belonging to the PN.
- FRIEND: It is a special kind of foreign node with whom a short-term trust relationship has been established. Its identity can be verified when discovering this kind of nodes.

3.1.3.2 Node Advertising

To proactively discover neighbours, each node periodically broadcasts beacon messages advertising its presence. The periodicity of the beacons will be designed depending on how dynamic the cluster is. Context awareness techniques could be applied to set the inter-beacon time. The structure of a beacon is shown in Figure 3-2.

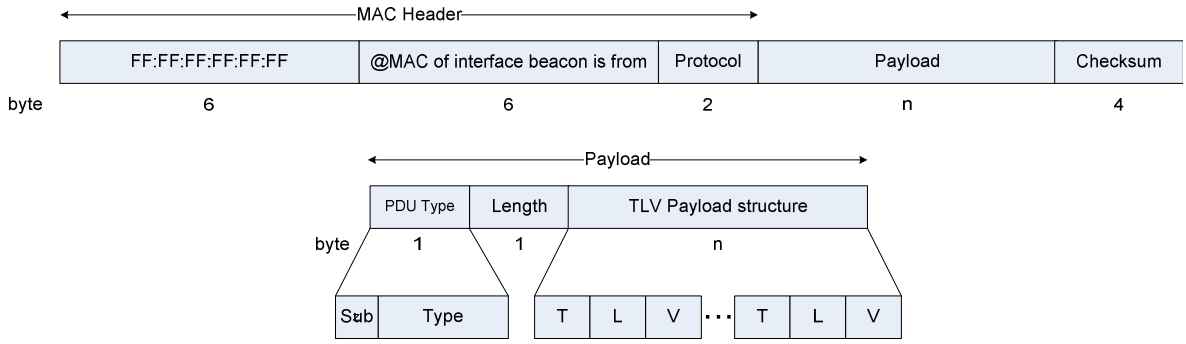


Figure 3-2: Beacon packet format

The payload of a beacon may vary depending on the features of the node. However it is mandatory that the node and PN identifiers are included since these are the indexes used for addressing the corresponding pre-established primary keys. Some of the possible fields are depicted in Table 3-1.

Table 3-1: Beacon fields

Tag	Length	Description
Node Identifier	20 bytes	Cryptographically generated identifier
Node Name	Variable	Friendly name assigned to the node by the user
PN Identifier	20 bytes	Cryptographically generated identifier
PN Name	Variable	Friendly name assigned to the PN by the user

The proposed beacon structure is extensible in order to support future neighbour discovery features. The information that is broadcasted is going to be publicly known, so it

may be required to allow the user to decide on whether she wants to make it public or if it might be only accessible to selected nodes. In this sense, three different visibility levels have been identified:

- Full Visibility: The node advertises itself all the time, and is visible proactively.
- Limited Visibility: The node makes itself available only to other personal nodes by encrypting the beacon with a PN-wide broadcast key. Additionally, it can reactively disclose its presence to non-personal nodes in full visibility if the node or PN is specified (e.g. according to user profile) by responding to the beacon and starting the authentication procedure. Together with the beacon acknowledgement, the node inform the non-personal nodes about its Limited Visibility state in order to make clear to the other node that no beacons will be received from him so other keep-alive mechanism have to be used (e.g. respond to one out of each N beacons).
- Invisible: In this mode the node does not advertise itself even to other personal nodes. Nevertheless, it reactively discloses its presence to other personal nodes (whether in Limited or Full visibility) and it can do the same with non-personal ones (in Full visibility) if specified. When disclosing its presence it also informs about its Invisible state so us corresponding node can adapt the keep-alive mechanism.

3.1.3.3 Node Discovery

Periodic beacon messages are the mechanism used by a node to advertise its presence, so the rest of the peers have to keep track of them in order to not only detect new nodes but also to maintain a complete view of its neighbourhood.

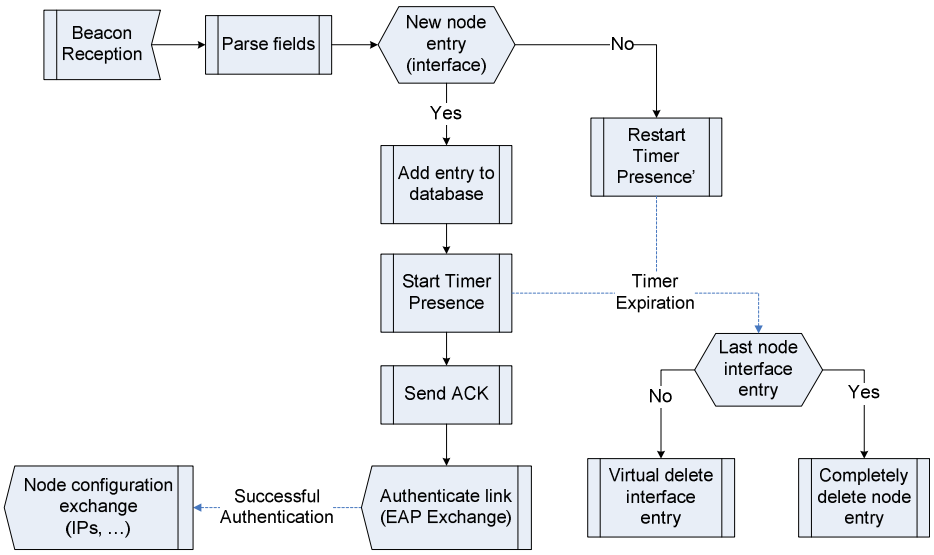


Figure 3-3: Node discovery procedure flow diagram

Upon the reception of a beacon the procedure depicted on Figure 3-3 is triggered. Every time a beacon is received it is checked to see if it belongs to a new neighbour. In that case, it is inserted into the database with all the information that can be extracted from the beacon. By parsing beacon payload fields, data such as the node identifier or node name is

retrieved. In addition to this, the MAC address and link layer interface the beacon has arrived from is registered. An acknowledgement is sent to the beacon before initiating the authentication procedure. This way, both peers recognize each other at the same time without having to wait for the next beacon to be received in the reverse direction.

If the beacon received does not belong to a new neighbour, no acknowledgement is sent and no authentication procedure is triggered. However, its database entry will be updated by reinitialising its expiration timer. Note that there is not one unique expiration timer for each node, but as many as link layer interfaces (peer's MAC addresses) have been discovered. Upon timer expiration, the database entry for this interface is removed. Whenever the last timer associated to a node has expired, the complete node entry is deleted.

The authentication procedure that is run just after a new node is discovered is thoroughly described in the following section. Upon a successful authentication the new peer is correspondingly catalogued (i.e. personal, friend). This implies that a secure communication channel can be established between both nodes. It is then time to securely exchange significant configuration information about the nodes. For instance, IP addresses are considered as private information that has to be safely exchanged. The information sent is encrypted using the previously generated link layer session key (output of the authentication procedure) assuring the confidentiality of the data. Note that this configuration data may be transferred any time if it is needed due to an update in node setup.

3.1.3.4 Authentication

The main characteristic of the neighbour discovery procedure implemented is that it is able not only to find the peers in the node surroundings but also to determine which of them are trusted nodes and consequently with whom communication can be securely established.

Any personal node is imprinted at the very beginning of its life with the rest of nodes belonging to the same user. To imprint a node, the P-PAN Formation Protocol (PFP) is used [126]. As a result of this imprinting procedure, every node shares a secret or key with each of the rest of personal nodes. These keys will be used as the basis for authentication and secure communication between personal nodes at any layer in the network stack. The first step in any communication is to establish a link-layer channel. The neighbour discovery module, after detecting a new neighbour claiming to be one of these personal nodes (by looking at the node and PN identifier included in its beacons), uses the appropriate primary key to derive a session key that secure the newborn link layer channel.

Moreover, since several networking procedures make use of broadcast traffic, this kind of traffic must also be secured. Obviously, the session key cannot be used for protecting the broadcast traffic because it is bilateral. Hence, in addition to the link-layer session keys, each node has a broadcast key for encrypting the broadcast frames. This per-node broadcast key will henceforth be denoted as B_x. Each node has to inform its peers about its

broadcast key. This information is encapsulated and exchanged using the same procedure as used for deriving the session key. This key is used to secure broadcast messages originated by a node. As broadcast messages are intended to be directed to everybody in a certain domain, the broadcast key is distributed amongst its trusted nodes (within the domain for which the key is intended) so that they can decrypt the broadcast packages received.

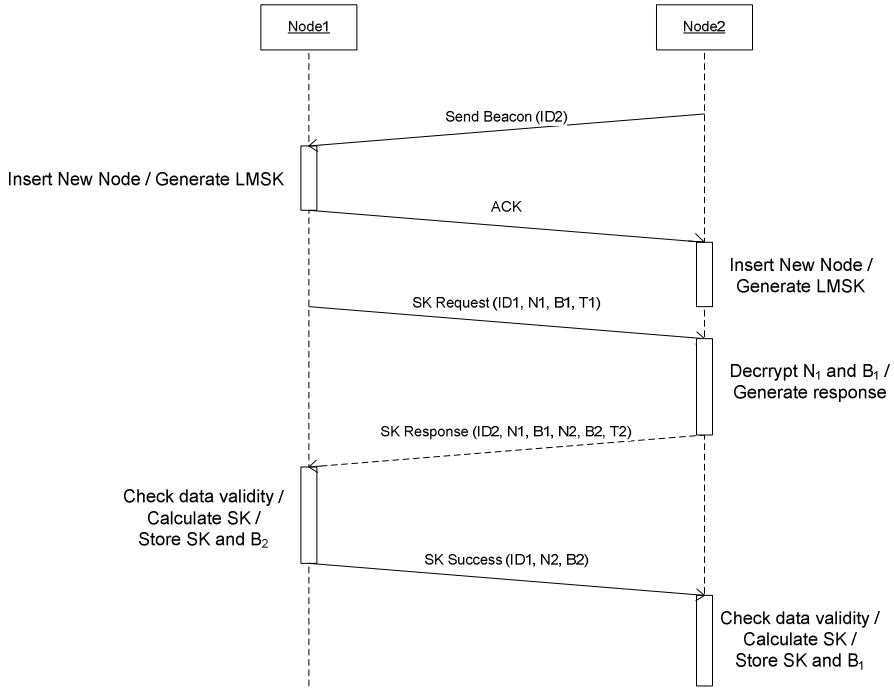


Figure 3-4: Authentication plus Session and Broadcast keys exchange protocol

Figure 3-4 shows the 4-way handshake used for the exchange of the unicast and broadcast session keys. As shown, the keys are derived using the following key exchange protocol. The following notations are used:

- | – concatenation
- HMAC(key, data) – hashing function
- N_X – nonce
- B_X – broadcast key
- E(key, data) – symmetric encryption

Symmetric encryption is done using the Advanced Encryption Standard (AES) [127] cryptographic algorithm with a key length of 256 bits.

1. Node 1 receives a beacon from Node 2
2. Node 1 sends EAP_request ($E(LMSK_{1\_2}, N_1 | B_1 | T_1)$)
3. Node 2 replies with EAP_response ($E(LMSK_{1\_2}, N_1 | B_1 | N_2 | B_2 | T_2)$)
4. Node 1 sends EAP_success ($E(LMSK_{1\_2}, N_2 | B_2)$)

where $LMSK_{1\_2}$ (Link Master Session Key) is calculated as HMAC_SHA_256 (K_{PN}, "MAC_1+MAC_2").

Use of the MAC addresses of the candidate radios in the derivation function ensures that different pairs of hardware adaptors of a radio subsystem share different link keys even for the same pair of devices. This is particularly relevant in the presence of detachable wireless interface adaptors (USB or card based).

The $SK_{1\text{-}2}$ (Session Key) is computed as HMAC SHA-256($LMSK_{1\_2}$, $N_1 \otimes N_2$) and is valid for T_2 seconds ($T_2 \leq T_1$). This procedure is run any time a new neighbour is discovered by a peer and whenever the derived session keys expire.

The actual authentication and session keys exchange procedure has been encapsulated using modified Extensible Authentication Protocol (EAP) where success messages are also authenticated. Details on message format and EAP attributes can be found in Appendix A.

Neighbour authenticity is assured if the session keys' exchange is completed successfully.

3.1.3.5 Configuration information exchange

After the node has been authenticated the node's configuration information is exchanged. This basically means that the IP addresses that correspond to the relationship between the nodes as certified during the authentication phase. In this sense, foreign nodes will only exchange the public addresses used on the access network. On the contrary, when a personal node is discovered the personal IP addresses are exchanged.

It is important to note that since this procedure takes place just after the node has been authenticated, the IP addresses are exchanged encrypted so that only the two peers know the personal IP addresses. This prevents from IP spoofing attacks since the IP addresses from the personal address space are not disclosed to the open air.

The PDU used for the configuration information exchange is the same as for the beacons where the Type on the payload is set to a different value for recognizing them upon reception.

3.1.3.6 Neighbour Database

The information that is retrieved by the Neighbour Discovery module is stored in an internal database and made available to the rest of the system through several interfaces.

As is shown in Figure 3-5 the database is organised starting from the PN to which the neighbouring nodes belong. This first table contains a list of nodes. Each entry in this list contains the information about a node, basically its identifier and the Primary Master Key (PMK – the one exchanged during the imprinting procedure). Besides, a list of the IP and MAC addresses of this node is also included in each node record. For each IP address, the type of address (i.e. IPv4 or IPv6), its value and its category (PERSONAL or PUBLIC) is stored. Finally, for the MAC addresses, not only its value and the network interface to which this entry corresponds are stored, but the session and broadcast keys are stored on each MAC record too. It is important to note that, although a node is univocally identified by its node identifier, provided within the beacon payload, a different authentication

process is performed for each of the air interfaces, through which it is possible to communicate with the neighbour. Thus, there will be different keys with the same node if they belong to more than one radio domain.

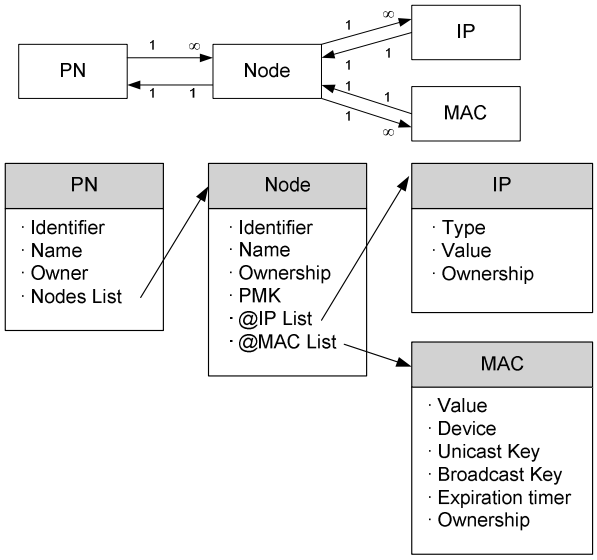


Figure 3-5: Table structure within Neighbour Database

A detailed description follows of each field in the Neighbour Database tables:

- PN
    - Identifier: 20-byte long unique identifier of the PN to which the neighbouring node belongs. For those nodes that do not belong to any PN, an identifier, where all the bits are set to zero, has been reserved.
    - Name: User friendly descriptive name of the PN to which the neighbouring node belongs.
    - Owner: Name or descriptive information of the owner of the PN to which the neighbouring node belongs.

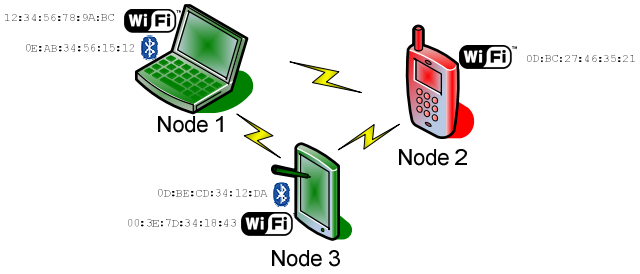
- Node
    - Identifier: 20-byte long unique identifier of the neighbouring node.
    - Name: User friendly descriptive name of the neighbouring node.
    - Ownership: Type of relationship that is shared with the neighbouring node (e.g. PERSONAL, FOREIGN, FRIEND).
    - PMK: 32-byte long Primary Master Key (the one exchanged during the imprinting procedure) shared with the neighbouring node.

- IP
    - Type: IPv4 or IPv6
    - Value: Actual IP address
    - Ownership: Addressing space to which the address belongs (e.g. PERSONAL, PUBLIC). PNs have a private addressing space so that when personal communications are established, IP addresses used are specific for that PN (PERSONAL), but it is also possible to discover the addresses of the neighbouring nodes that they have obtained from the access network (PUBLIC).
- MAC
    - Value: Actual MAC address
    - Device: Network interface which give access to the radio domain shared with the neighbouring node.
    - Unicast Key: 32-byte long session key used for encrypting and decrypting the unicast traffic exchanged with the neighbouring node through the aforementioned device.
    - Broadcast Key: 32-byte long key used for decrypting broadcast messages sent by the neighbouring node through the aforementioned device.
    - Expiration Timer: Time after which the aforementioned keys are no longer valid.
    - Ownership: Type of relationship that is shared with the neighbouring node (e.g. PERSONAL, FOREIGN, FRIEND). In contrast to the Ownership field in the node table, this field is only set to a value different from FOREIGN if the authentication procedure has finished successfully. Once this is completed, MAC ownership can actually change to PERSONAL or FRIEND.

Figure 3-5 also presents the model of the possible relationships between entities in the Neighbour Database. Hence, a node can only belong to one PN, that is, a node record can only be present in the node's list of one PN. In contrast, a PN can have any number of node's records in its nodes list. A similar situation appears in the relationship between IP addresses and Nodes. While a node can have many IP addresses, an IP address can only belong to one node. For the MAC addresses, the relationship is the same. A MAC address record can only be associated with one node while this node can be reached through multiple network interfaces, thus the MAC address list for that node might contain multiple MAC address records.

As can be seen in Figure 3-6, the Neighbour Database will not only contain personal nodes, but also non-personal ones (both belonging to other PNs and those that do not belong to any PN).

Identifier	97d23aafec685728419f4d6db8a6cacb13dcb2e0	5fa5e5269746ecb658e9ba56f899642cabd8e4d8
Name	Node 3	Node 2
Ownership	Personal	Foreign
PMK	f50e38298a4529cf650dc9f2ea6ef9a0742e286e	



Identifier	693459e8a916784118e1d334544176a06e696764		5fa5e5269746ecb658e9ba56f899642cabd8e4d8
Name	Node 1		Node 2
Ownership	Personal		Foreign
PMK	f50e38298a4529cf650dc9f2ea6ef9a0742e286e		
MAC List	Value	12:34:56:78:9A:BC	0D:BC:27:46:35:21
	Device	eth1	eth1
	Ownership	Personal	Foreign
	Unicast Key	1ead8d6ddb612542177a0a2b414e3fb38d1ce0ab954fdb14a328e6b55cdbe851	
	Broadcast Key	2ad8d66ddb612542177a0a2b414e3fb38d1ce0ab954fdb14a328e6b55cdbe8ae	
	Exp. Timer	2558796	
	Value	0E:AB:34:56:15:12	
	Device	eth2	
	Ownership	Personal	
	Unicast Key	387d0ac00b0473218732ecb1bad90098f033492eadccbf980981223462812271	
	Broadcast Key	4354572451512987b07982f70f9ba2987b0ffb2a09fb2a6f2fa2bfae2f2b6b89	
	Exp. Timer	2789678	
IP List	Value	1.2.3.4	192.168.0.112
	Type	IPv4	IPv4
	Ownership	Personal	Public
	Value	192.168.0.100	
	Type	IPv4	
	Ownership	Public	

Figure 3-6: Example of a filled Neighbour Database table

3.1.4 Legacy Support

As already said, one of the key features of the UCL is that it can be transparently inserted into the protocol stack without needing any modification in the rest of the layers (both upper and lower ones). Besides, it is important to note that it is necessary that those nodes that implement the UCL can transparently communicate with those that implement the current TCP/IP protocol stack. A key requirement to seamlessly handle this situation is to support legacy neighbour discovery mechanisms. As presented in the previous section, the UCL has its own neighbour discovery mechanism so that UCL-enabled nodes do not require additional address resolution protocols to discover the MAC addresses of the corresponding nodes. Nevertheless, non-UCL nodes rely on the address resolution mechanisms (Address Resolution Protocol (ARP) for IPv4 and Neighbour Discovery (NDISC) for IPv6) to fill the MAC header of their outgoing packets.

Figure 3-7 presents the ARP and IPv6 neighbour discovery protocol messages' format.

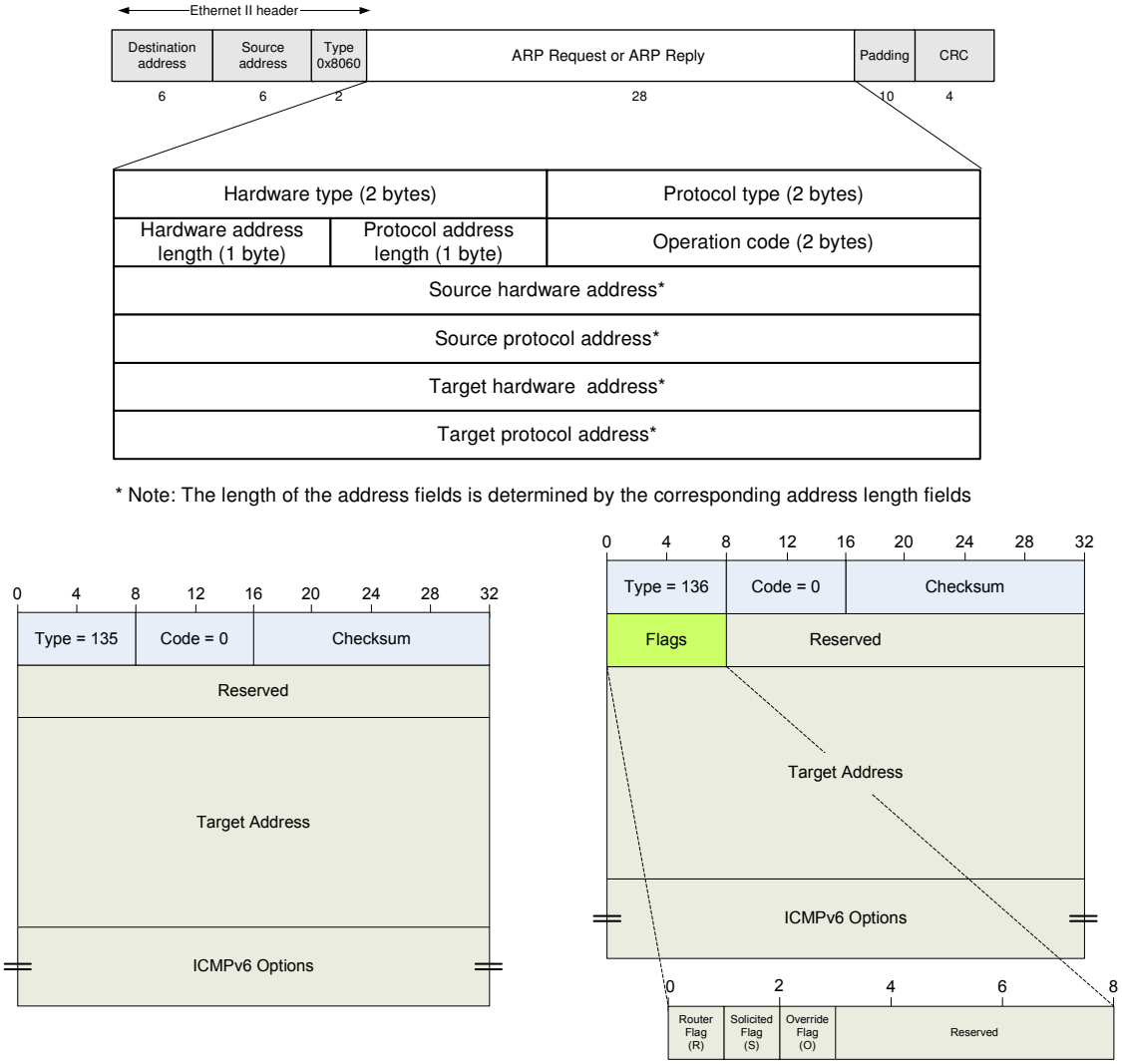


Figure 3-7: ARP and NDISC message format

When ICMPv6 Neighbour advertisement is sent in response to a multicast Neighbour Solicitation, it must contain a Target Link-Layer Address option, which carries the link-layer address of the device sending the message.

As can be seen, both the protocol (i.e. the payload of the messages) and the MAC header for this message contain the MAC address information about the network interface through which the packets are going. However, for UCL-enabled nodes this information corresponds to the UCL MAC address (currently the UCL MAC address is randomly selected from the interfaces controlled) and since the UCL may manage multiple interfaces, when the corresponding node is non-UCL, the address resolution will fail as depicted in Figure 3-8.

To prevent these situations, the UCL legacy support module intercepts those messages that correspond to an address resolution protocol (ARP or NDISC) and performs the necessary changes both on the protocol payload as well as on the packet MAC header so that information contained there corresponds with the outgoing interface. In this way,

when the packet arrives at the non-UCL node, the information corresponds to that of the interface with which it is actually communicating.

Figure 3-8 shows how the legacy support module works over ARP messages. As can be seen, if a node with multiple interfaces (Laptop in the figure) starts an ARP process, the success of the address resolution depends on which MAC address the UCL has taken. If the MAC address selected corresponds to that of the interface that is used to communicate with the legacy node (WiFi in the figure), then the address resolution would work correctly, but if by chance the UCL is assigned with an address from the other interfaces managed, then the address resolution would fail. As can be seen, in this latter case, although the MAC header source address contains the actual MAC address of the outgoing device, the address in the ARP protocol payload is the one assigned to the UCL. The legacy node fetches the MAC address from the ARP protocol payload field and not from the MAC address header. Hence, the wrong MAC header destination address is used in the ARP response and the packet never reaches the other node and the address resolution is not completed. If the legacy node were the one starting the address resolution procedure, the response containing the incorrect UCL MAC address would actually come back to it, but with incorrect information. Thus, communication would be impossible due to MAC addresses mismatch. The Legacy Support module responsibility is to adapt the address resolution protocols' packets as is shown in Figure 3-8. If it detects that the MAC address used as the Source Hardware Address (see Figure 3-7) does not correspond to the one of the outgoing network interface, then it makes the corresponding changes so that the address resolution procedure can finish successfully. ICMPv6 Neighbour Solicitation and ICMPv6 Neighbour Advertisement packets are handled similarly.

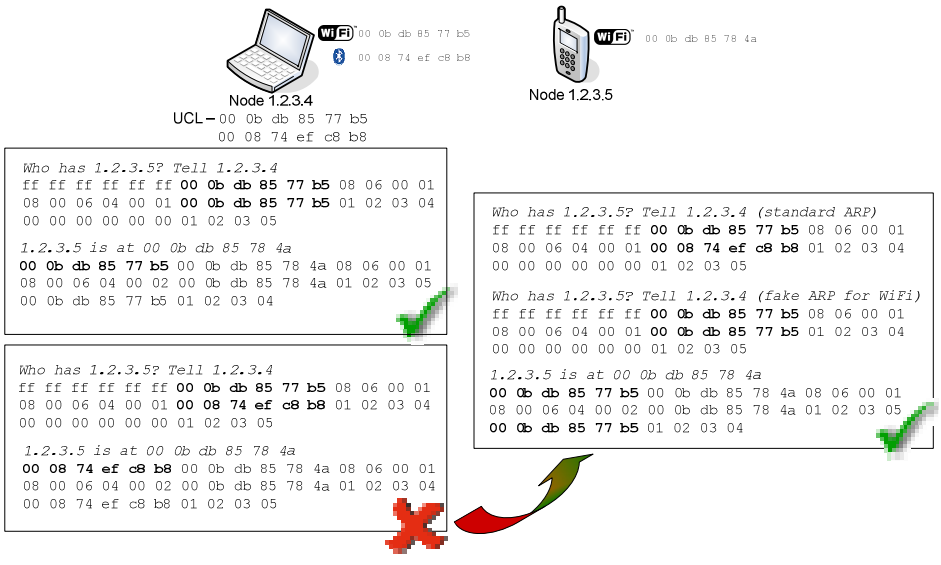


Figure 3-8: Legacy support module operation over ARP messages

As already mentioned, this module assures transparent operation between UCL and non-UCL enabled nodes. Currently only ARP and NDISC messages are handled since they are the ones corresponding to IP protocols but if other legacy address resolution mechanisms

are used, it would not be difficult to adapt the legacy support module to deal with them too.

3.1.5 Network Resource Discovery

This module is fundamental for the cross-layer approach to be taken within the UCL architecture. Normally routing decisions and interface selection are based entirely on IP/network layer information. In host multihoming this is inadequate, since we need to take multiple factors into account when selecting interfaces for outgoing traffic. As a rule, these factors lie outside the IP layer, thus forcing us to break the level hierarchy in order to provide the necessary interface selection functionality.

Some of the areas where the UCL could get valuable knowledge are:

1. Link Layer Information: In wireless networks, signal quality and related metrics play an important role when deciding which interface to use. Though other universal factors must be taken into account, link quality is imperative, since it dictates what quality of service demands and policies can be fulfilled and how much of the theoretical bandwidth is actually available. Moreover, if the link quality is poor, a user with a PDA might not want to use it at all because of the increased power consumption. To be able to make as smooth and intelligent use of the available interfaces as possible, link quality must be constantly monitored and the information must be made available for the network layer and user applications in a form that suits them best. They may be used it in combination with other information to make the best possible proactive routing decisions and interface selections. Most of the currently available wireless network drivers support information gathering and although the information and its presentation are far from uniform, the most important metrics are widely available in some form or another. UCL aims at unifying the access to this information.
2. IP and Transport layer Information: Several attributes can be retrieved from the IP header without looking into the data, e.g., source address and destination address. Some attributes can also be retrieved from IPv6 extension headers. Given that only transport protocols, such as TCP and UDP, can be identified directly from the IP header, the higher level protocols, such as HTTP, can be mapped to port numbers as these are visible in the IPv6 header or transport layer protocol header. For example, in the case of HTTP, the port number could be, e.g., 80 and/or 8080, based on port number assignments by the Internet Assigned Numbers Authority (IANA) [128].
3. Network Originated Information: A service provider may disseminate information about cost, bandwidth and availability of the Internet access. For example, an ISP might offer Internet access through WLAN and Bluetooth within an area. In addition to advertising the default gateway by means of Router Advertisements [129], access routers could also send cost and bandwidth information. The mobile user could then have preferences for connections, such as maximize bandwidth or minimize price, and the host would select the appropriate interface satisfying these

preferences. Disseminating information from the network may be implemented using a new protocol (e.g., CAR [130]) just for this purpose, or Router Advertisements could be extended to carry the information.

4. Information Originated from Users and Applications: Some applications may require certain characteristics from the connections. An application should be able to adapt to the changing network environment and set its own preferences for connections. This can be achieved by introducing a kind of API as defined in [131].

This module can also be the interface that the UCL offers to a Context Management Framework as in [132]. The information retrieved in the Network Resource Discovery module can be exported to this context management and distribution system so that it is available for other system's components to benefit from a more complete picture of the situation.

3.1.6 Path Optimization

The possibility of using different links to the destination allows UCL to intelligently modify the output interface according to the requirements and needs of the system.

Taking into account the destination and locally retrieved information about interfaces and channel status (SNR; available bandwidth, etc.) gathered through the Network Resource Discovery module, many transmission alternatives can be selected. Weighting this information using user profile preferences enables the selection of the most appropriate interface. Amongst the currently available options, the following can be found:

- Traffic striping using at the same time several of the available transmission channels
- Use of the best link on the basis of the SNR, bandwidth, packet loss, etc. statistics retrieved.

3.1.6.1 UCL traffic striping approach

Using a communication system composed of multiple links, with different characteristics and throughputs, not only increases the global transmission throughput but also considerably improves the reliability of the system. However, to be fully efficient, such a system requires particular schemes, called striping schemes, to split traffic over the links. Different proposals for splitting information over several links have been envisaged. In this section two striping strategies are presented for allocating the fragments to the links, which will produce the most optimal system performance to be included in the UCL.

Let t_1 to t_n be the transit times (inversely proportional to the transmission bit rates) of n links L_1 to L_n shared between the two nodes communicating through the UCL. Also let T_{min} be the transit time of the fastest link and T_{max} the transit time of the slowest link.

Figure 3-9 shows the way packets are assigned to the available links with the *sequential strategy*. This method loads the links with fragments taking into account the capacity of the different links and the relationship between these capacities. This strategy gives a number of tokens to each of the available links.

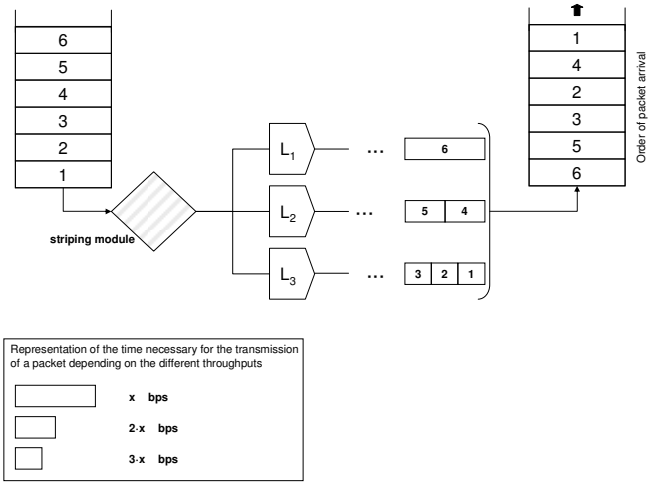


Figure 3-9: Sequential strategy for striping over the available links

The number of tokens assigned to each link depends on the relationships between the binary rates of the different links. They are calculated as $\tau_i = \frac{T_{\max}}{t_i}$. After the tokens have been assigned the algorithm starts fetching for the fastest L_i that already has tokens. Then as many packets are injected as tokens are available in the link. When all the links have consumed the tokens provided, they are reassigned and the process starts again.

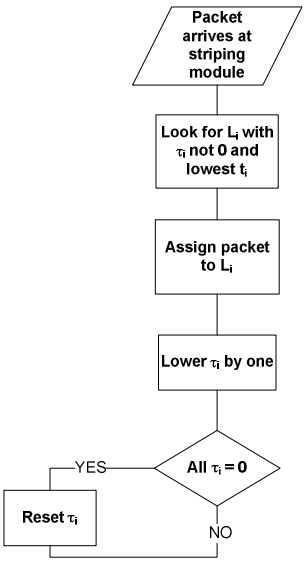


Figure 3-10: Sequential strategy flowchart

The second strategy consists in anticipating the sending order of the fragments which are pending according to the available links in order to minimize the gap in the arrival times between two consecutive fragments. As is shown in Figure 3-11, this strategy allows that the packets to arrive in the correct order to the receiver side.

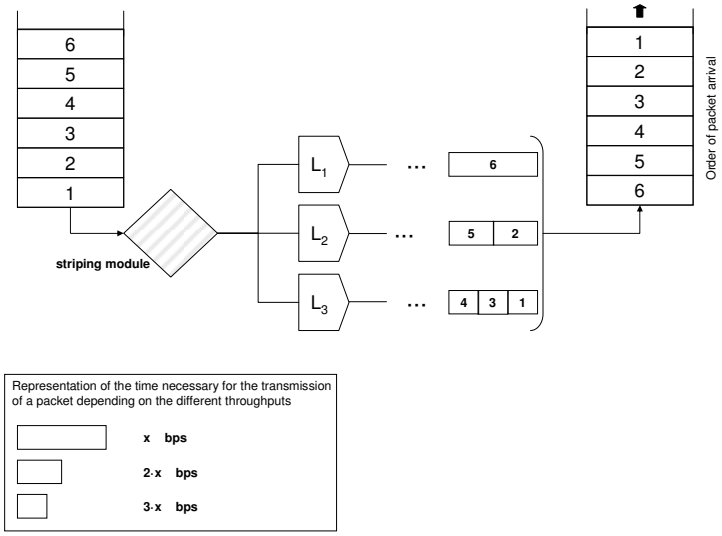


Figure 3-11: Ordered strategy for striping over the available links

The algorithm to select the outbound interface first considers the aggregated capacity of the communication system considered. We can define the normalized relative capacity of each of the n links as:

$$\chi_i = \left\lfloor \frac{T_{\max}}{t_i} \right\rfloor$$

while the normalized aggregated capacity of the system when using striping is:

$$\mathrm{X} = \sum_{i=1}^{n} \chi_i$$

This capacity represents the number of packets that can be transmitted in T_{max}.

Afterwards we define the link transmit times as the time at which successive transmissions would be done in each link.

$$t_{ij} = j_i \cdot t_i \quad \ni \quad t_{ij} \le T_{\max}$$

The selection is done so that the respective t_{ij} are the lowest one possible. It should be noted that each time a packet is assigned to one of the L_i, the respective j_i is incremented.

As is shown in Figure 3-12 in the case of colliding t_{ij}, the fastest L_i is selected.

The advantage of the ordered strategy is that packets are served to the upper layers in the correct order so that communication is transparent to them. The sequential strategy implies that upper layers (whether transport or application layer) are able to perform the reordering of the traffic flow. Nevertheless, the ordered method needs to buffer the packets to be transmitted in advance before starting the transmission so that the appropriate order can be chosen to maximize the system performance. This increases the complexity of the striping module implementation and also assumes that the source traffic generation rate is higher than the transmission one such that the striping module buffer can always be filled with the necessary number of packets.

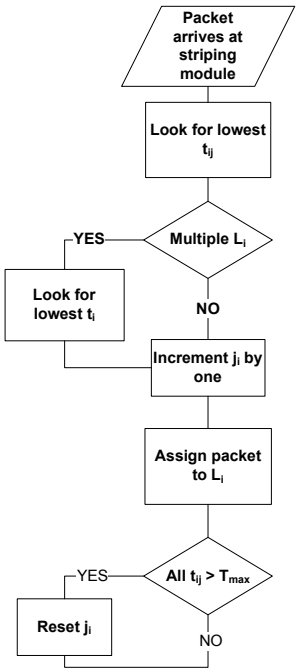


Figure 3-12: Ordered strategy flowchart

3.1.6.2 UCL best link selection approach

When the UCL path optimization module is used to select the best link possible for intra-cluster communication, it will handle the cost functions used during the decision-making process. There are two possibilities that have been considered for tackling this problem. In the first one, the decision is assumed to be a handover process meaning that the traffic flow is considered as a unique entity and link selection is done per flow. Additionally, it follows a conservative approach meaning that it only changes to another network interface upon degradation of the conditions of the link used or the moment another possible network interface become available. In the second case, each packet to be transmitted is considered independently from the previous one. In this sense, a utility function is evaluated for all the possible link-layer interfaces and the one with the best result is selected. The policy that has been foreseen is in both cases based on link status and battery level.

Figure 3-13 shows the handover process as followed in the first of the solutions considered.

As has been mentioned, in this approach the handover process is triggered only if channel conditions on the current link are degraded or upon appearance of a new possibility for communication. If it is due to degradation of channel conditions and there are other network interfaces available for continuing with the communication, the QoS characteristics of the other networks are evaluated. Mainly, the supported bandwidth and the observed SNR are considered. Then for those that fulfil the necessary requirements the power consumption is checked. It should be noted that if quality can be assured with more than one link layer technology, the one with lower power footprint would be selected. This

decision-making algorithm can be extended with more parameters by simply adding more steps.

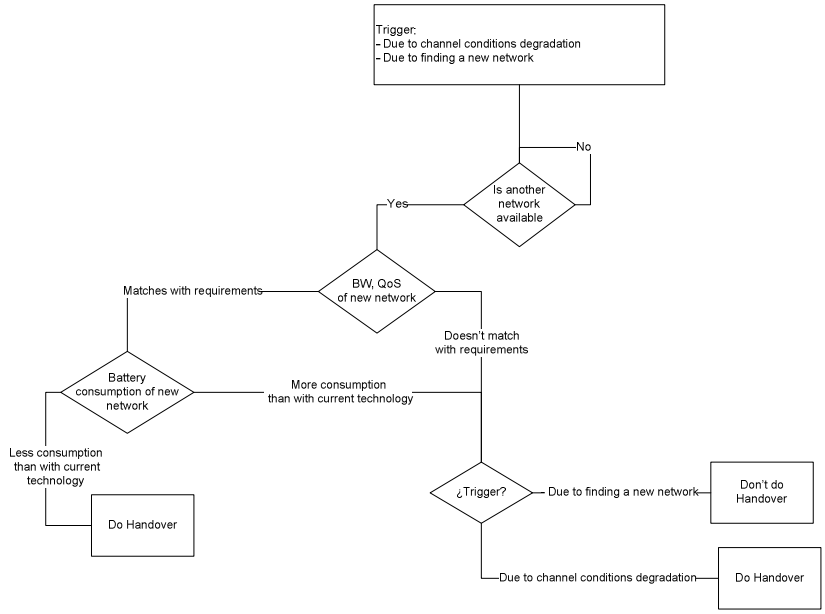


Figure 3-13: Handover decision process flow

For the second option a utility function derived as:

$$U = \sum_{n=1}^{m} w_n \cdot p_n$$

Here, w_n denotes a weighting factor and p_n represents the value of the different parameters used in the decision-making process. It is important to note that for each wireless technology different thresholds are established for the assignment of the parameter value. For example, for a SNR of 15dB an IEEE 802.11b interface is ranked higher than an IEEE 802.11a one for the same ratio. The Utility of the current and the possible new interface is evaluated and the higher one is chosen. As an optimization in this process, hysteresis lags are applied in order to prevent the ping-pong effect. This option requires more extensive computation but offers actual selection of the best link for every packet.

3.1.7 Security

From a security perspective, one of the most important design goals of UCL is to make sure that use of a legacy, radio-specific security system does not cause any additional security vulnerabilities. In order to accomplish this, the UCL uses the session keys derived and exchanged in order to provide confidentiality, integrity and origin authentication through the encryption of all the traffic exchanged between two neighbouring trusted nodes. Figure 3-14 shows how signature and encryption is applied over the payload of the MAC frame. In this way we avoid using specific radio interface features, providing a homogeneous security framework on top of the underlying heterogeneity.

The communication architecture that we are considering is based on pair-wise trust relationships. Every pair of personal nodes shares a long-term trust relationship that is

enforced when they communicate with each other. When two personal nodes meet they authenticate each other and exchange link level session keys, derived from the secret key they share, that are used to secure that particular link. These session keys are used to encrypt the IP datagram, using the AES algorithm [133], and to securely sign the packet, using SHA-256. Thus, only the counterpart neighbour is able to decrypt the information and verify the signature of the packet.

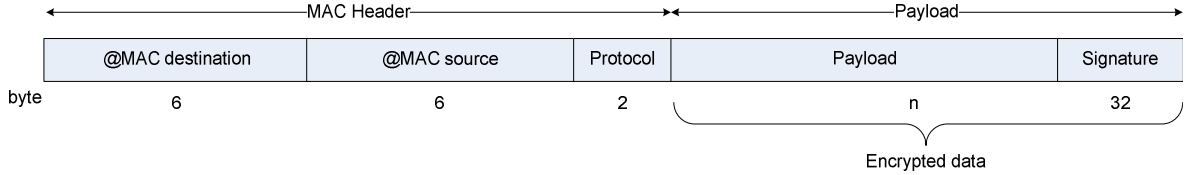


Figure 3-14: Packet encryption format

As is shown in Figure 3-14 the communication between two personal neighbours is protected through the encryption of the complete MAC frame payload (i.e. the complete IP datagram including the IP headers). The MAC header is not encrypted since the source and destination addresses would not be understood by the underlying technologies and transmission/reception would not be possible. Additionally, a cryptographic signature is added to the packet in order to assure the integrity of the packet. These extra security features can only be applied when both nodes are UCL enabled.

Block cipher algorithms like AES require their input to be an exact multiple of the block size. If the plaintext to be encrypted is not an exact multiple, you need to pad before encrypting by adding a padding string. When decrypting, the receiving party needs to know how to remove the padding in an unambiguous manner. AES uses a 16-byte block so that before encryption it has to be checked that the payload plus the signature can be set in a number of 128-bit blocks. If this is not the case, padding has to be applied. This fact limits the Maximum Transmission Unit (MTU) of the UCL since the limit of 1500 bytes for the MAC frame has to be satisfied. Hence, the MTU for the UCL is such that the following conditions are guaranteed: $MTU \leq 1500 - 32$ and $\mathrm{mod}(MTU, 16) = 0$.This leads to an MTU of 1456 bytes

In multihop scenarios at cluster level, the end-to-end security is assured by securing each of the links of the communication. By definition, all the nodes in a cluster are personal, so the packet is protected by the security of each of the links that forms the end-to-end route. If not the packet has to be encrypted and decrypted in every link of the route with the additional overhead that this implies.

As can be seen in Figure 3-15, the Session Key (SK) used for encryption always corresponds to the one associated with the next personal node in the transmission route. When the packet is to be sent to one of the neighbouring nodes, the final destination of the packet and the next node in the route are the same, but when intermediate nodes are to be used the packet has to be encrypted with the session keys corresponding to each of the links that form the whole route. Each of the intermediate nodes decrypt the packet using the SK that they share with the node from which the packet is received and encrypt it again with the SK that corresponds to the next node in the communication path. For

example in Figure 3-15 Node 2A acts as a relay for communication between Node 1A and Node 3A. When Node 1A wants to communicate with Node 3A it has to first encrypt the packet with the session key it shares with Node 2A (i.e. SK(1A,2A)), when this packet is received it is decrypted and relayed towards Node 3A after encrypting it again with the corresponding key (i.e. SK(2A,3A). This way the communication is always protected.

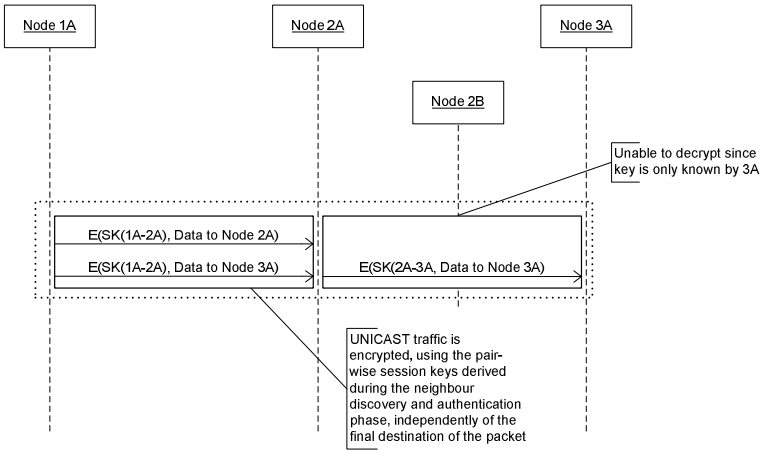


Figure 3-15: Encryption path in multihop scenarios

3.2 UCL DATA FLOW

Once the components in the UCL architecture are presented, the flow of user data across the UCL will be introduced in this section. Taking into account the information about the node's neighbourhood provided by the Neighbour Discovery module, the UCL focuses on enhancing the transmission and reception procedures by providing security and path optimization features in addition to the management of multiple interfaces.

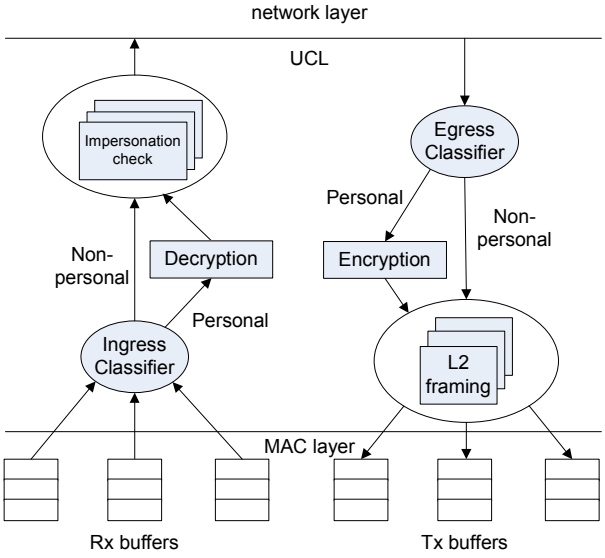


Figure 3-16: UCL high-level internal operation and admission control

UCL enables communication both with UCL-enabled devices and legacy ones, assuring backward compatibility and increasing the communication possibilities of a node. Hence, the UCL will not only deal with personal traffic but also with incoming and outgoing

packets from/to non-personal nodes. The downstream and upstream traffic management takes these possibilities into account and as will be shown, these procedures cope with all of them.

3.2.1 Downstream Data Flow – Transmission

UCL can be considered as an overlay DLC layer on top of all the different link layer interfaces the device has. In this sense, all the packets that are transmitted by the device go through the UCL. Figure 3-17 depicts the process followed by the packets on their way through the UCL. As the packet traverses the UCL, its type and destination is analyzed so that it can be redirected to the most suitable network interface, adapted to support legacy operation, or protected with the appropriate security mechanism.

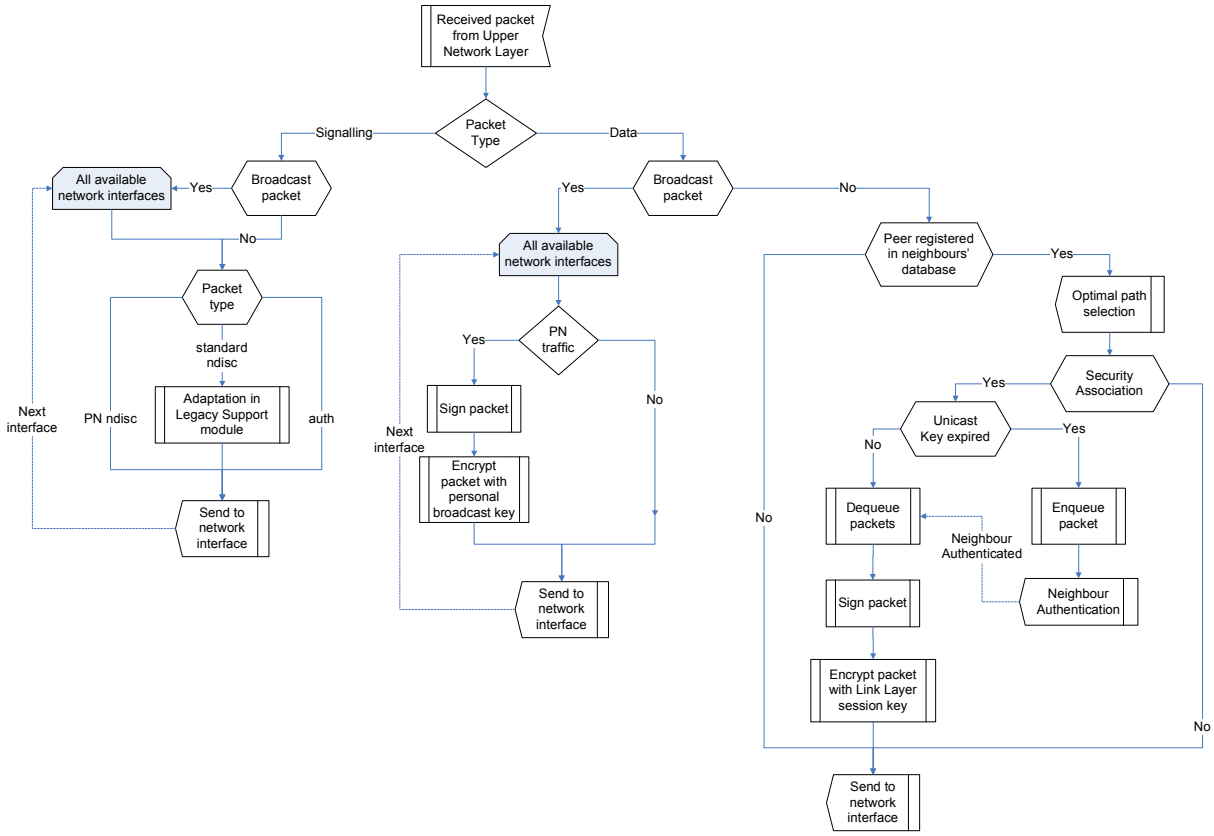


Figure 3-17: UCL downstream data flow diagram

Packets arriving at the UCL transmission function may be of two types, signalling or data packets. When a packet arrives, it firstly has to be classified since it will be treated differently depending on its nature.

When a signalling packet is to be transmitted, it is firstly analyzed to determine whether it is a broadcast or unicast packet. For broadcast ones, the packet is sent through each of the network interfaces managed by the UCL following a cyclic approach. In either of the two cases, the packet follows a similar process. There are three kind of signalling packets that are processed. Neighbour discovery (i.e. beacons, acknowledgement) and Authentication packets (i.e. session key establishment and node configuration ones) are created within the

UCL and both the security mechanisms and outgoing network interface are already included and defined within the packet. Hence, they are directly sent to the corresponding network interface (as defined by the source MAC address on the packet). These signalling packets are not encrypted nor are their transmission paths optimized. The main reason is that the session is established at link level between the network interfaces involved and any modification in the transmission path will cause an incorrect session establishment. The packets that fall into the third type are the legacy neighbour discovery ones (i.e. ARP and ICMPv6). As already explained, this kind of packets requires special handling to provide backward compatibility with non-UCL nodes. When one of these packets is to be transmitted, the packet is adapted within the Legacy Support module and then sent to the corresponding network interface.

Data packets are also classified into broadcast and unicast. Broadcast data packets are also sent through all the network interfaces available in the node following a cyclic approach. If the destination IP address indicates that this is a packet containing PN traffic, then the packet is encrypted and signed using the node's broadcast key. Thus, only other personal nodes will be able to decrypt and check the signature of the packet. If it is not PN traffic, then the packet is transmitted unencrypted. For unicast packets, the destination MAC address is first checked. If it is not registered on the neighbour's database, meaning that the destination node is non-UCL, the packet is not encrypted and is sent in a legacy manner without using enhancements in packet transmission.

If the destination MAC address corresponds to one of the registered nodes, independently of whether it is personal or not, path optimization techniques are called. The main issue is that the packet reaches its destination without suffering modification in the information it contains following the most optimal path (quickest, lowest packet loss, less power consumer, etc.). Once the outgoing network interface has been selected on the Path Optimization module and MAC header has been correspondingly modified, the relationship with the peer node is checked. If it is not a personal node, then the packet is sent through the outgoing network interface without any further modification. If it is a personal node then a valid unicast session key is fetched and used to encrypt and sign the packet before sending it through the selected network interface. The link-layer session key derived while peers mutually authenticate themselves is utilized. Depending on the output interface selected, a different key is used. If by the time the packet is to be sent, the derived session key validity time has expired, a new authentication and session key exchange procedure is triggered. In the meantime, packets to be sent to that peer are queued waiting for the authentication procedure to finish so that they can be dequeued, encrypted, signed and finally sent.

3.2.2 Upstream Data Flow – Reception

The scheme followed for dealing with incoming traffic is shown in Figure 3-18. Traffic is classified depending on the identity of the source. The packet source MAC address is used as the index for searching in the neighbours' database.

Packets that arrive at the UCL from nodes that are not registered (mainly meaning that it is a non-UCL-enabled device), are redirected to the upper layers after passing some security checks (mainly checking the IP header of data packets to assure that the destination IP address corresponds to the node's public IP address).

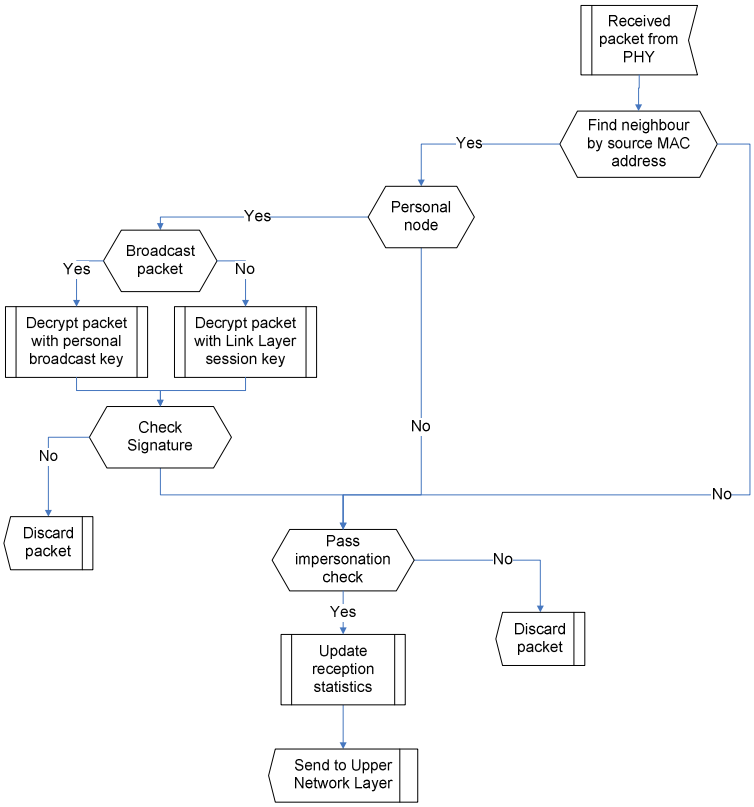


Figure 3-18: UCL upstream data flow diagram

The process that follows traffic from registered nodes is a little bit more elaborated and depends firstly on the ownership of the originator of the packet and on whether the packet is unicast or broadcast.

Packets received from non-personal nodes are also checked in order to avoid impersonation attacks before accounting them and passing them to the higher layers. Packets coming from personal nodes are catalogued depending on the dispersion. Unicast traffic is first decrypted using the corresponding link-layer session key and then the integrity of the information is checked by comparing with the signature attached to the packet. Similarly, broadcast traffic is decrypted and its signature checked using the peer's broadcast key.

Once all security checks have been performed on the packet, the packet follows the standard path in the network stack. If any security check is not successfully passed, the packet is discarded. As can be seen in Figure 3-19, the impersonation check consists of the verification of the destination IP address. For non-personal nodes the only permitted destination IP address is one of the node's public addresses. Thus, it is assured that foreign nodes can only access public services offered by this node and are not able to inject traffic in the personal cluster.

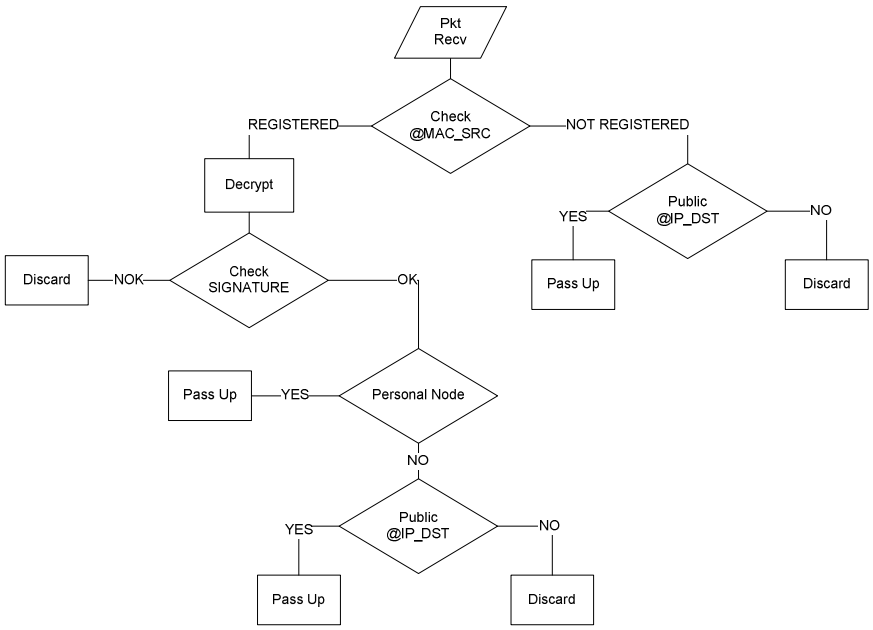


Figure 3-19: UCL impersonation attack check procedure

Before the packet leaves the UCL, link-layer context information for the source node is updated.

3.2.3 Threat Analysis

The main aim of the threat analysis is the construction of a mitigation plan that involves the selection of the appropriate countermeasures to obtain a secured system.

The first step of analyzing a system for security is determining its vulnerabilities. In order to perform this vulnerability study we used a more generic system view that allows us to define generic attacks. In general, a PN cluster can be viewed as a Mobile Ad hoc Network (MANET). In [134] several MANET attacks are identified. Basically, these attacks can be categorised into two main trends. While passive attacks attempt to discover valuable information by listening to the traffic, active ones try to disrupt the operation of the MANET protocols. In order to make informed security decisions, we need to estimate what the cost is to the attacker to exploit various vulnerabilities and what the exploited vulnerabilities impact will be on the system. We can also estimate how often the attacker will be successful when attempting a particular attack on the system. Using all this information, a risk value can be assigned to each node in the attack tree to determine where the system's weakest points are.

In order to determine the threats we also need to identify the potential threat-sources. According to [135] we can classify the threat sources in: natural, human, and environmental. A threat-source is defined as any circumstance or event with the potential to cause harm to a system. The natural sources can be: floods, earthquakes, tornadoes, landslides, avalanches, electrical storms, and other such events that can damage the devices and the infrastructures involved in the federation. The human threats are events that are either enabled by or caused by human beings, such as unintentional misuse or deliberate actions. Environmental threats can be: long-term power failure, pollution,

chemicals, liquid leakage. Besides these considerations we need to consider the specific threats that can affect intra-cluster communications.

The threats with higher risk, namely, spoofing and identity theft, can be prevented respectively with appropriate mechanisms of authorization and authentication and with a proper encryption of the identity information.

Eavesdropping and disclosure of information can be mitigated providing access to services and information only to authenticated and authorized users according to privacy regulation, and sending such information only on encrypted channels.

DoS can be mitigated with appropriate network infrastructures such as firewalls and intrusion detection systems, but also adapting the level of security in the system.

The combination of these countermeasures can be considered in this specific case the mitigation plan.

The solutions implemented at the UCL use symmetric encryption (AES-256, claimed to be secure beyond 2031, has been used for the encryption) for defending against the first kind of attacks (i.e. passive attacks). As will be commented through the experimental evaluation of the solution, the overhead of this encryption should be bearable if the subjective enhancement in security terms is considered.

Concerning the second category, the most typically referenced attacks are the ones modifying route request packets, those using spoofing techniques, and those using fabrication, the attacker intentionally floats error messages on the network falsifying existence of valid routes. However, in the literature [136], another attack is proposed, the so-called rushing attack, leading to Denial of Service situations, which exploits particular vulnerabilities of on-demand routing protocols.

The main characteristic of the UCL security implementation is that only authenticated nodes are allowed to be part of the network. In this sense, a variant of the rushing attack, namely, wormhole-like attacks [137] are prevented as the attacker will never be recognised as a valid network member and its packets will be blocked at the UCL. Similarly, spoofing attacks are also tackled since MAC addresses are simply a method of identification but the access grant is not given until pair-wise keys are used for neighbour authentication and session key exchange. Personal IP addresses (i.e. the ones used within the PN cluster) are never disclosed to non-personal nodes so IP spoofing can also be defended against. If brute force is used for IP spoofing (i.e. the attacker tries with several IP addresses), an impersonation check is also used to avoid non-personal nodes injecting traffic into the personal network.

CHAPTER 4
DYNAMIC INTERFACE SELECTION BASED ON CROSS-LAYER INFORMATION

The purpose of this chapter is to present and analyze the results from the experimental evaluation performed on the implemented solution for Dynamic Interface Selection based on Cross-Layer Information. The functionalities of the components described in Chapter 3, their correctness and their interfaces have been evaluated and debugged thoroughly on a testbed consisting of Linux laptops and PDAs. In this chapter, we will report on the results obtained from a set of measurements campaigns and analyze them thoroughly. Valuable lessons can be learnt about the current performance of the implemented software and optimal parameter settings and guidelines can be provided for future extensions that can help to improve performance. For more details on how the software works, refer to Appendix A. Using a testbed for performance evaluation has both advantages and disadvantages. The main advantage is that the software is deployed on real devices and in real-world scenarios, revealing problems that cannot be discovered or simulated with simulation tools. Of course, a testbed has some limitations, mainly regarding the size and variety of the test setups. To mitigate this limitation analytical and simulation-based studies have been also carried out so that conclusions can be derived for larger setups.

4.1 INTRODUCTION

This section presents the results obtained from the measurement campaign carried out in order to prove and validate the benefits introduced by the UCL solution in terms of selection of the most appropriate network interface for outgoing traffic. As was explained in Chapter 3, one of the key components of the UCL architecture is the Path Optimization module. This module uses Multiple Attribute Decision Making (MADM) algorithms in order to decide which of the many available outgoing interfaces to use in order to guarantee the best possible system performance. In this section the optimization in terms of enhanced throughput, reduced packet loss and energy saving that the UCL achieves by means of dynamic interface selection based on link-level information (observed SNR) will be analysed. This mechanism represents a step forward in the cooperative networks paradigm.

These analyses are based on a fully experimental measurement campaign as well as on simulation and analytical studies that will be compared so that accurate conclusions can be derived. The final objective of these analyses is to demonstrate that in a real world scenario where a mobile personal node equipped with multiple air interfaces is involved in PN communication, a considerable optimization is obtained when UCL uses cross-layer information and multiple network interfaces management for selecting the best connection to use for delivering the user traffic.

The measurement campaign follows a twofold approach. On the one hand, several tests were performed to quantify the performance degradation introduced by the UCL in user-level communications. On the other hand, a full set of tests was carried out in order to show the UCL aptitudes and communication optimization. There are two optimization approaches implemented on the UCL:

- Selection of optimal interface: Based on the channel conditions experienced by the different wireless interfaces managed by the UCL, it will decide which one to use.
- Striping of flows: Whenever it is possible and necessary for an application flow, the UCL will be able to use multiple interfaces at the same time to transmit a single flow.

4.2 MEASUREMENT CAMPAIGN SCENARIO

This section describes the environment in which the measurement campaign was carried out. Figure 4-1 shows the scenario where all the tests were performed. Four different locations were selected, each of them showing gradually worse channel behaviour. In this sense, a fixed laptop was placed at the red mark while the corresponding peer in the measurements was set at each of the different locations marked in Figure 4-1.

The scenario is an office environment but it could equally well be an in-home scenario. Basically, the purpose of selecting this environment was to test the UCL in a real-world scenario which allowed us to extract conclusions that can be directly mapped onto real situations a user experiences.

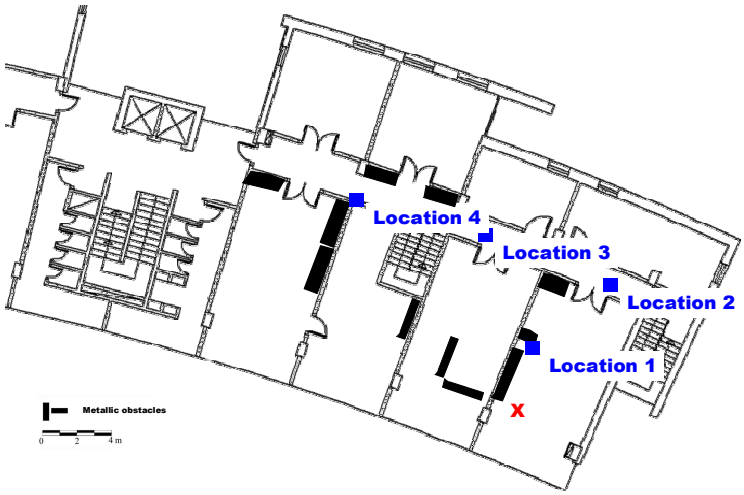


Figure 4-1: Measurement campaign environment

The first task carried out during the measurement campaign was to analyse the scenario and describe the characteristics of each of the selected locations. During the measurement campaign two wireless technologies, IEEE 802.11a and IEEE 802.11b, were used in order to recreate the heterogeneity that will be present in the future personal networking scenarios. Besides, both technologies operate in different frequency spectrum ranges and use different PHY and MAC mechanisms. Hence, the generic conclusions obtained could be extrapolated to other wireless technologies. In this sense, the laptops used during the measurement campaign had the two wireless access technologies.

Figure 4-2 shows the distribution of the SNR at each of the locations analysed. These distributions were obtained by sending a 30-second-long UDP traffic flow between the two laptops and analysing the SNR registered in the receiver of this flow. Table 4-1 summarizes the main parameters of the received SNR at each of the locations. Note that in the Location 4, the only available results are for IEEE 802.11b since the channel conditions were so bad for IEEE 802.11a that the measurements could not be used to appropriately describe the location.

As can be seen, the selected locations offers a good range of situations ranging from very good channel conditions to poor ones that will produce a deep degradation of the communications.

Table 4-1: Channel characteristics

	IEEE 802.11b		IEEE 802.11a	
	Mean SNR	**Std. deviation**	**Mean SNR**	**Std. deviation**
Location 1	43.94	2.29	28.97	2.06
Location 2	37.09	1.76	26.75	1.46
Location 3	25.83	1.14	16.78	0.99
Location 4	15.4	1.13	---	---

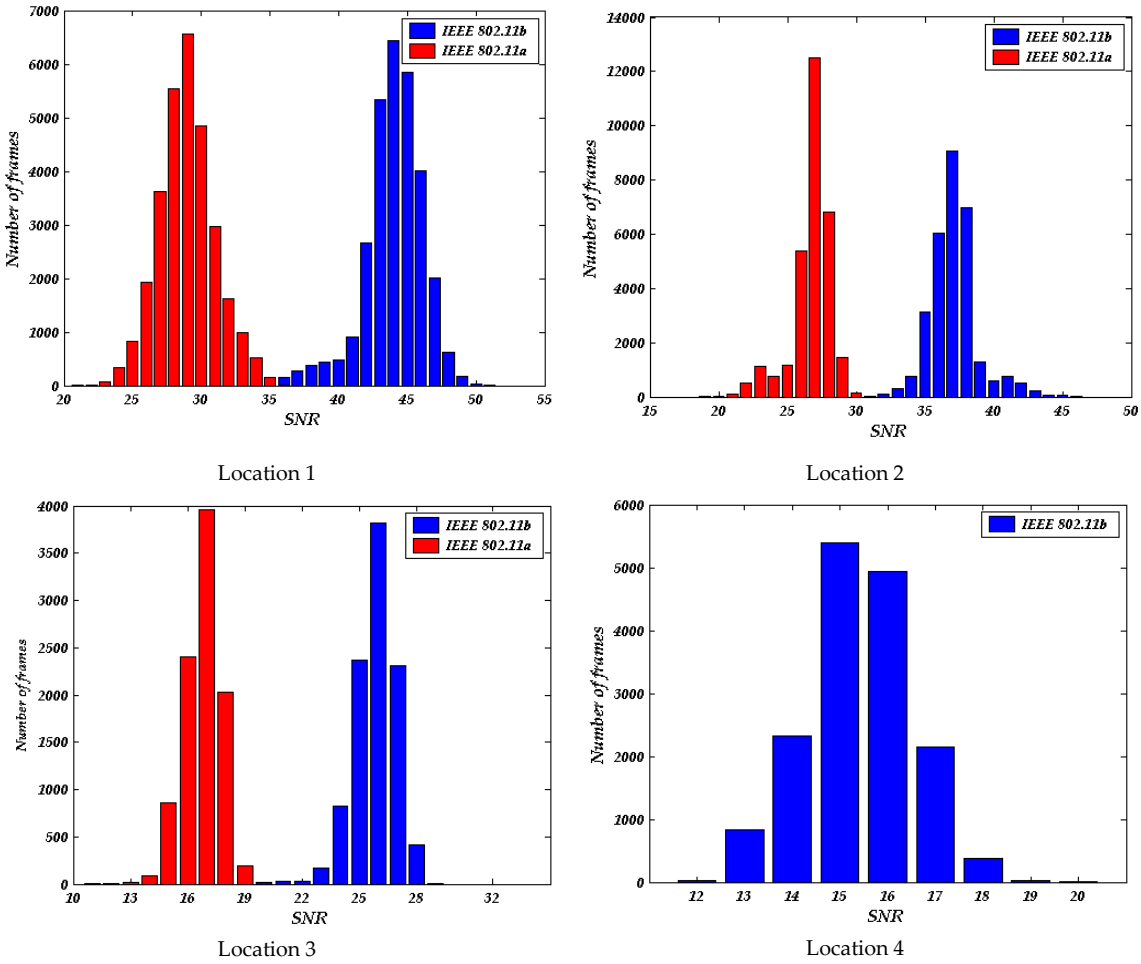


Figure 4-2: Received frame SNR distribution at the four locations

Link quality metrics such as ETX (Expected Transmission Count Metric) [138], MTM (Medium Time Metric) [139], WCETT (Weighted Cumulative Expected Transmission Time) [140], etc. have been proposed as metrics to estimate the link quality and replace the minimum hop count metric, which is widely used by current routing protocols, to select paths in order to increase network capacity. However, there are still some limitations for applying these metrics to real implementations such as the actual collection of the parameters required to calculate them. Besides, the SNR has been proven to be an adequate representation of the link quality [141][142]. Hence, our approach can be considered valid, not only because it matches the wireless channel behaviour quite faithfully but also because the implementation carried out is ready to support richer metrics with which better decisions can be made.

4.3 UCL OVERHEAD ANALYSIS

Before any further tests were done, the first step in validating the UCL implementation performed is to quantify the performance degradation introduced. Both UDP and TCP traffic will be analysed. The tests carried out to measure this degradation consist of comparing the performance achieved with and without the UCL implementation loaded

when the channel conditions are ideal. Thus, the differences between the two situations can only be attributed to the UCL and not to the channel conditions.

Table 4-2 shows the results obtained for both UDP and TCP. For UDP, the tests consisted of a 10.000 packets of 1500 bytes each sent between two nodes placed one next to the other. For TCP traffic, an FTP session exchanging a 10-Mbyte file was established between the two nodes. In both cases, the test was repeated five times in order to obtain a more accurate result. Table 4-2 shows the mean value obtained in the tests and the standard deviation.

As can be seen, the tests were performed using two different configurations for the UCL. In the first one, the security mechanisms were enabled. In this case, the differences between the two situations are appreciable but still quite low. Nevertheless, it is important to note that the UCL performs the encryption and decryption of the transmitted frames which, as will be shown in Chapter 5, represents most of the overhead imposed by the UCL. In the second case, the impact of the UCL is negligible since security mechanisms are disabled. For the rest of the tests presented in this chapter this was the configuration used in order to avoid adding entropy to the evaluated system.

Table 4-2: UCL performance degradation comparison

# Test	UDP Throughput (Mbps)	Std. Dev	TCP Throughput (Mbps)	Std. Dev
IEEE 802.11a				
1-5	32,49	0,45	28,25	0,91
UCL Security Enabled				
1-5	29,33	0,59	25,34	1,3
UCL Security Disabled				
1-5	32,14	0,52	28,01	0,85

Finally, the comparison has only been made with the IEEE 802.11a wireless interface since the UCL will always select this interface taking into account the SNR observed in the test.

4.4 UCL SELECTION OF OPTIMAL INTERFACE

As already mentioned, the laptops used during the measurement campaign were equipped with two different wireless interfaces which behave differently in the distinct locations studied. The measurements carried out firstly present the performance obtained when using each of them. This approach will allow us to show the performance of each of the wireless interfaces to be used and estimate the most appropriate SNR threshold levels at which to decide to change the output interface to send the traffic through. At the end of the section we will introduce some mobility scenarios where the output interface is changed dynamically depending on the SNR experienced at each moment.

4.4.1 UDP traffic characterisation

The tests performed in this section show the performance obtained in each of the different locations when 10.000 UDP packets of 1500 bytes each are transmitted from one laptop to the other.

UDP traffic is mostly used by real-time applications that are mainly affected by packet loss and throughput. In this sense, the parameters shown in the different cases are the mean throughput obtained throughout the test and the Frame Error Rate (FER) observed.

Location 1:

Table 4-3 present the results of the tests performed in Location 1.

Table 4-3: Location 1 UDP statistics

# Test	Throughput (Mbps)	FER (%)	SNR (dB)
IEEE 802.11a			
1	33,38	6,28%	32,31
2	33,16	7,05%	31,89
3	34,33	3,91%	34,57
4	33,30	6,51%	34,07
5	34,49	3,32%	31,14
IEEE 802.11b			
1	6,01	0,00%	45,94
2	5,99	0,01%	44,50
3	6,01	0,00%	44,70
4	6,01	0,00%	43,92
5	6,00	0,00%	44,63

Figure 4-3 shows the evolution of the throughput throughout one of the tests carried out. As it can be seen in both cases, the throughput is maintained near its mean.

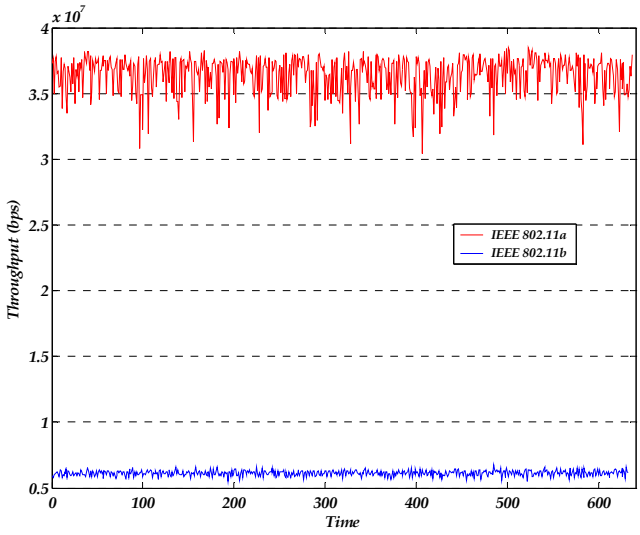


Figure 4-3: Location 1 UDP traffic immediate throughput evolution

As can be seen both wireless interfaces show a low Frame Error Rate (FER) which represents a negligible packet loss (which is null in the case of IEEE 802.11b). In such a situation, the most appropriate selection would be the IEEE 802.11a interface as it provides much greater throughput.

Location 2:

Table 4-4 shows the results of the tests performed in Location 2.

Table 4-4: Location 2 UDP statistics

# Test	Throughput (Mbps)	FER (%)	SNR (dB)
IEEE 802.11a			
1	21,11	39,31%	24,06
2	25,58	27,41%	25,83
3	27,56	22,00%	25,97
4	32,75	8,16%	28,47
5	27,91	21,04%	25,65
IEEE 802.11b			
1	6,00	0,00%	38,75
2	5,70	0,07%	35,36
3	5,88	0,02%	35,60
4	5,88	0,06%	35,22
5	5,90	0,01%	34,66

Figure 4-4 shows the evolution of the throughput all through one of the tests carried out. As it can be seen in the case of IEEE 802.11a the packet loss causes some variance in the instantaneous throughput which can result in peaks that can fall under the IEEE 802.11b line, which on the contrary remains very stable.

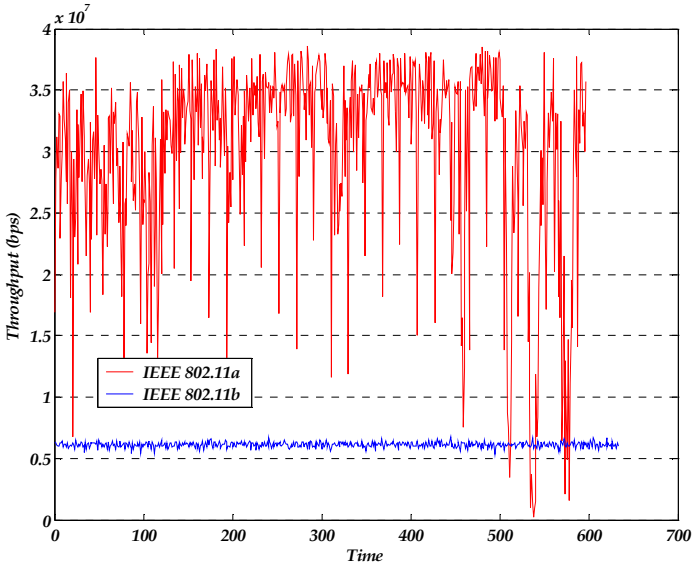


Figure 4-4: Location 2 UDP traffic immediate throughput evolution

As the distance between the two laptops increases, channel conditions start to degrade. In the IEEE 802.11b case, the reduction in the SNR barely affects the performance of the communications. In contrast, IEEE 802.11a is much more affected. Nevertheless, the reduction is not really important in terms of throughput, which remains much greaterthan in its counterpart. On the other hand, the FER is increased significantly. This increase would cause a substantial application level loss that might prevent the use of IEEE 802.11a in these circumstances. However, and taking into account the combination of both parameters, in this location we would still use IEEE 802.11a.

Location 3:

Table 4-5 presents the results of the tests performed in Location 3.

Table 4-5: Location 3 UDP statistics

# Test	Throughput (Mbps)	FER (%)	SNR (dB)
IEEE 802.11a			
1	3,51	83,73%	19,25
2	4,34	79,84%	19,21
3	5,18	84,32%	19,86
4	6,07	81,84%	18,70
5	7,76	76,88%	18,38
IEEE 802.11b			
1	5,26	3,13%	23,98
2	5,97	0,13%	26,52
3	5,73	1,09%	24,74
4	5,99	0,05%	27,10
5	5,84	0,48%	24,73

As can be derived from the results of the different tests carried out, the selection of the IEEE 802.11a interface in this location would cause an unacceptable packet loss which would severely degrade the quality of any application used over it. In contrast, the IEEE 802.11b interface performance remains almost unaltered, with a negligible packet loss and maintaining the throughput.

Figure 4-5 shows the evolution of the throughput throughout one of the tests carried out. As expected, the throughput has decreased below the IEEE 802.11b one. There are still peaks where the immediate throughput reaches nominal levels, but on average the selection of this interface would not result in throughput enhancement.

Taking into account these results, the threshold was set in the UCL to 25 dB. This means that whenever the UCL detects that the SNR observed goes below 25 dB, it will swap the output interface from IEEE 802.11a to IEEE 802.11b. The opposite will happen when the SNR climbs from low figures and goes above 25 dB.

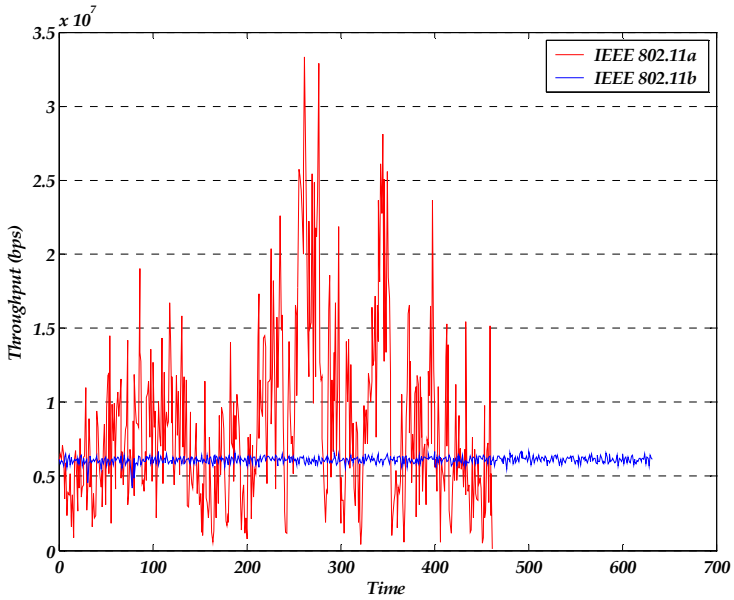


Figure 4-5: Location 3 UDP traffic immediate throughput evolution

Location 4:

Finally, and in order to offer a complete characterisation of the environment, Table 4-6 shows the results obtained from the tests carried out in Location 4.

This location gives a deficient performance in both IEEE 802.11b and IEEE 802.11a. In the latter case, the situation reaches the level where the measurements can hardly be carried out due to the poor channel conditions. Due to this, it is only possible to present the results from IEEE 802.11b since the ones from IEEE 802.11a are so biased by the high packet loss that they are not suitable to extract valid conclusions. It is important to note that even in these circumstances IEEE 802.11b does not degrade the communications performance too much.

Table 4-6: Location 4 UDP statistics

# Test	Throughput (Mbps)	FER (%)	SNR (dB)
IEEE 802.11b			
1	3,53	14,65%	16,15
2	2,92	30,10%	15,77
3	3,34	22,00%	15,36
4	2,78	18,00%	17,69
5	2,49	29,94%	17,81

Figure 4-6 shows the evolution of the throughput throughout one of the tests carried out. As it can be seen, the throughput is reduced and several troughs during the measurement fall quite low. Nevertheless, the mean performance is kept up to a reasonable level.

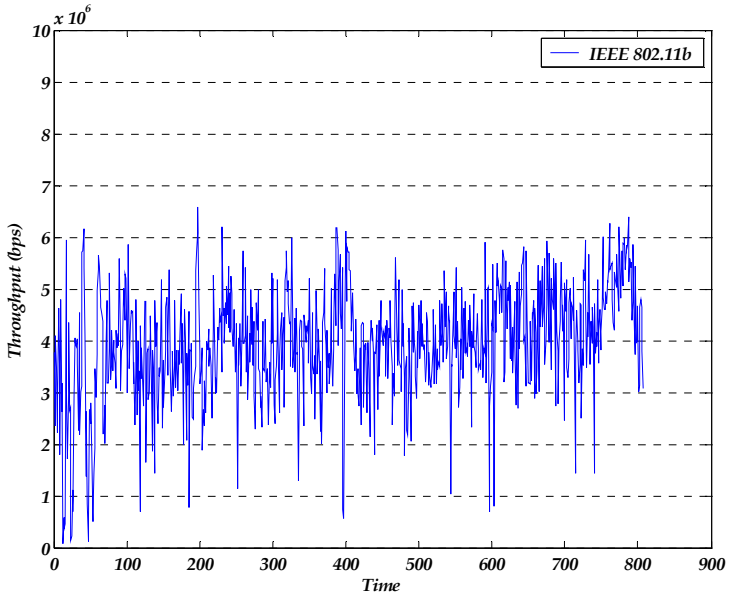


Figure 4-6: Location 4 UDP traffic immediate throughput evolution

4.4.2 TCP traffic characterisation

The tests performed in this section shows the performance obtained at each of the different locations when a File Transfer Protocol (FTP) session is established for transferring a 10-Mbyte file from one laptop to the other.

TCP traffic is highly affected by packet loss. Since TCP was designed for wired networks where packet loss is always interpreted as collisions caused by channel congestions, the congestion avoidance mechanisms of TCP force the transmitter to stop when these situations are detected. Nevertheless, in wireless channels, normally, packet loss is due to channel impairments for which stopping the transmitter is pointless.

During the TCP traffic characterization several parameters will be studied in order to obtain the most accurate picture of the communication performance. Basically, these parameters refer to the TCP retransmissions as they are mainly responsible for performance degradation.

- **Out of order pkts**: Represents the number of packets that arrive at the receiver in the wrong order. This is due to the fact that the transmission window allows transmission of a number of packets without explicit acknowledgement. If a packet is lost, the following packets can continue arriving out of order since the lost packet is the expected one.
- **Idle time max**: Is the maximum time during which the transmitter is stopped due to congestion avoidance mechanisms. Maximum time is indicative although there might have been a plethora of smaller idle times that have ruined that test.
- **Max # of retx**: Is the maximum number of consecutive retransmissions of the same packet. Congestion avoidance mechanisms are triggered and their timeouts increment as more consecutive retransmissions are done. In this sense, a big

number of consecutive retransmissions typically lead to long idle times in the transmitter.

- **Avg. retx. time**: Is the average time taken between two transmissions of the same packet. Congestion avoidance timeouts are set between retransmissions and increased with every retransmission of the same packet. The higher this parameter is, the more time transmission has been stopped due to action of congestion avoidance mechanisms.
- # **of retx**: Counts the total number of retransmissions in a test.

Location 1:

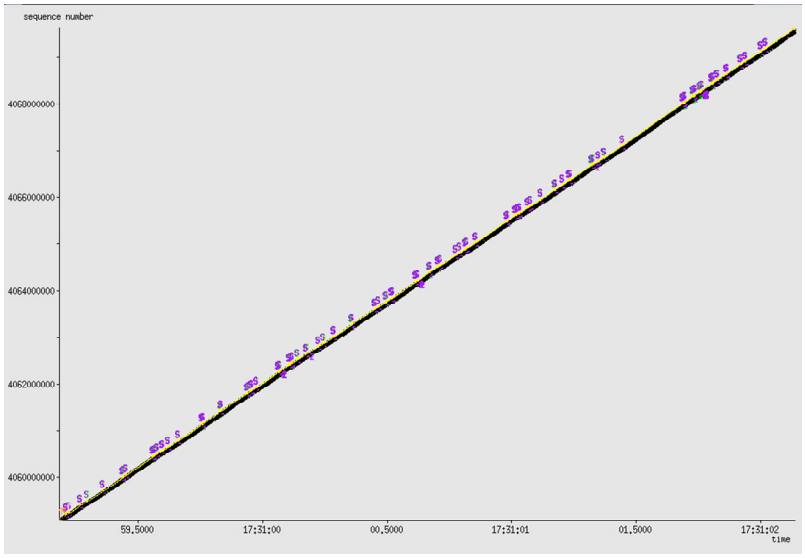
Table 4-7 presents the results obtained from the tests carried out in Location 1.

Table 4-7: Location 1 TCP statistics

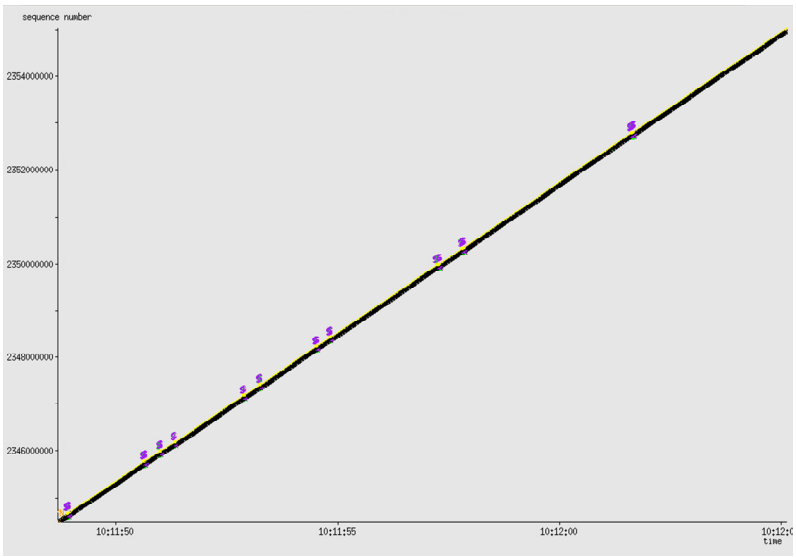
# Test	Throughput (Mbps)	Total pkts	Out of order pkts	Idle time max (ms)	Max # of retx	Avg. retx time (ms)	# retx
IEEE 802.11a							
1	27,71	7306	61	40	1	5,7	61
2	28,36	7323	76	7,5	2	6,4	78
3	27,83	7327	80	40	2	5,8	82
4	28,38	7315	70	7,7	1	6,6	70
5	28,11	7330	83	9,6	2	5,4	85
IEEE 802.11b							
1	5,11	7260	11	112,2	1	70,1	11
2	5,11	7283	5	108,7	1	102,5	5
3	5,08	7292	7	105,3	1	69,2	7
4	5,09	7293	7	103,1	1	102,4	7
5	5,10	7316	5	105,8	1	100,9	5

As it can be seen, in both cases the channel can be considered ideal maintaining the number of retransmissions at a reduced level. Besides, the errors occur in an independent fashion preventing the transmitter to misleading situations which might cause long transmitter idle times. Under these ideal conditions, the best choice would be the IEEE 802.11a interface. Note that the selection of the interface in the case of TCP traffic is more direct than in the case of UDP since the only parameter to compare is the final throughput.

Figure 4-7 shows the evolution of the received bytes with time. As can be seen, both interfaces offer perfect behaviour with a linear evolution. In the case of IEEE 802.11a the slope of the graph corresponds to 28 Mbps while only 5 Mbps are achieved using 802.11b. This situation is only possible in absence of channel impairments so there is no idle time in which the transmitter remains silent awaiting pointlessly for channel decongestion.



IEEE 802.11a



IEEE 802.11b

Figure 4-7: Location 1 TCP traffic time-sequence

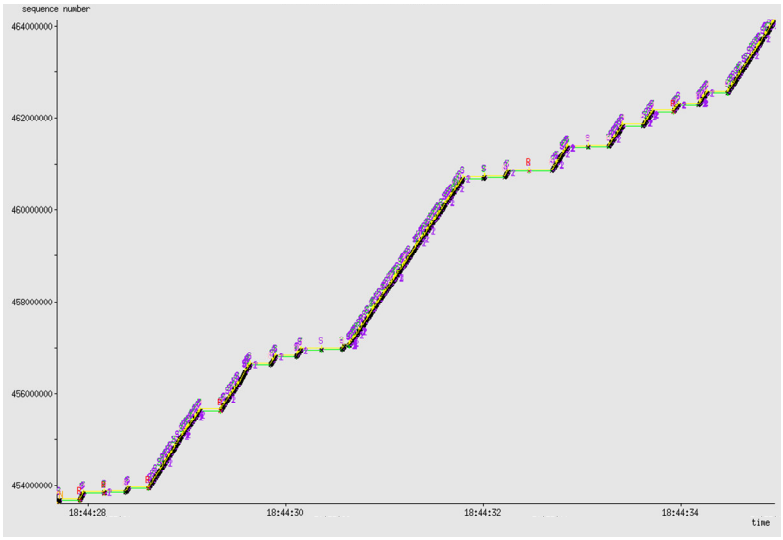
Location 2:

Table 4-8 presents the results obtained from the tests carried out in Location 2.

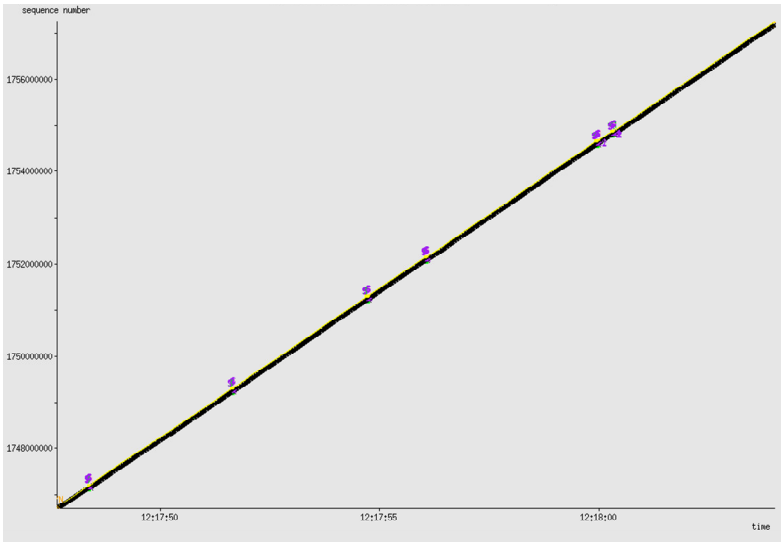
Although the performance is severely reduced due to the poorer channel conditions, in most of the cases IEEE 802.11a behaves better than its counterpart. As can be seen, the number of retransmissions is increased considerably, but they still occur in an independent manner which does not trigger the TCP congestion avoidance mechanisms. Hence, the performance is not completely broken down and remains better than IEEE 802.11b, which on the contrary remains stable in an almost error-free channel.

Figure 4-8 shows two time-sequence graphs showing IEEE 802.11a and IEEE 802.11b behaviour in Location 2 respectively. As can be seen, IEEE 802.11b presents a linear

behaviour, typical in ideal channel conditions. On the contrary, IEEE 802.11a presents a disrupted behaviour mixing periods where the communication is almost perfect with others where the transmitter is idle. Nevertheless, the channel is not so bad and IEEE 802.11a outperforms its counterpart.



IEEE 802.11a



IEEE 802.11b

Figure 4-8: Location 2 TCP traffic time-sequence

However, test #5 using the 802.11a interface highlights the vulnerability of TCP protocol to channel impairments [143]. In this test, a burst of lost packets results in a large idle time which makes the throughput decrease even below the 802.11b level. It would therefore be interesting to distinguish different thresholds for different types of traffic (e.g. TCP and UDP). The UCL enables this possibility, also allowing the definition of more complex QoS functions which would allow the decision to be made not only taking into account the transport-layer protocol but also application specific requirements and not only SNR but also other link-layer parameters.

Table 4-8: Location 2 TCP statistics

# Test	Throughput (Mbps)	Total pkts	Out of order pkts	Idle time max (ms)	Max # of retx	Avg. retx time (ms)	# retx
IEEE 802.11a							
1	9,01	7517	238	411,9	3	49,8	272
2	11,56	7546	285	231,7	2	21,9	301
3	9,72	7670	406	409,9	2	17,9	425
4	13,70	7418	159	401,9	3	34,5	173
5	3,43	7601	311	6431	7	71,7	356
IEEE 802.11b							
1	5,13	7261	8	108,9	1	90,3	8
2	5,12	7274	7	109,7	1	99,5	7
3	5,11	7318	2	108,8	1	92,6	2
4	5,11	7302	6	104,2	1	102,4	6
5	4,94	7273	8	118,1	1	96,1	8

Location 3:

Table 4-9 shows the results obtained from the tests carried out in Location 3.

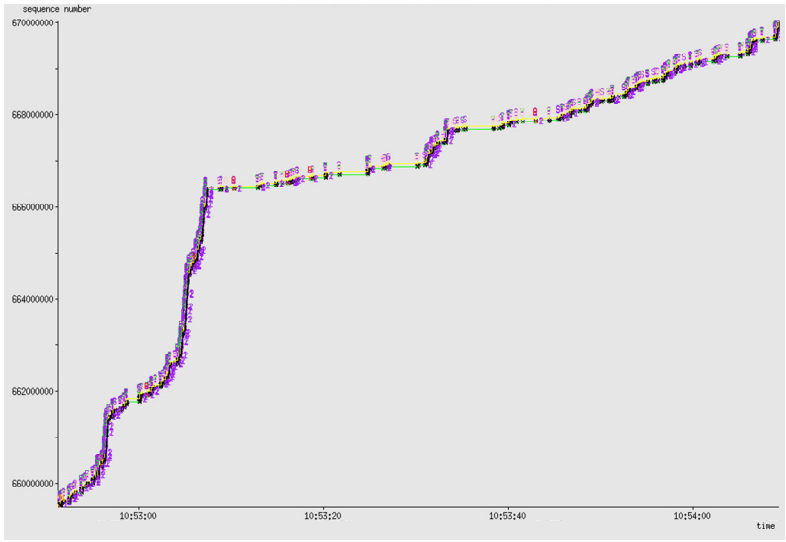
Table 4-9: Location 3 TCP statistics

# Test	Throughput (Mbps)	Total pkts	Out of order pkts	Idle time max (ms)	Max # of retx	Avg. retx time (ms)	# retx
IEEE 802.11a							
1	1,08	8173	678	1663,7	5	136,6	928
2	3,50	7790	428	815,9	3	66,6	545
3	2,20	7512	220	6463	6	241,3	267
4	4,00	7666	344	819,9	4	75,3	421
5	1,30	7647	313	7657,4	4	362	402
IEEE 802.11b							
1	5,16	7277	10	102,5	1	85,8	10
2	5,06	7320	24	112,7	1	69,9	24
3	5,10	7270	24	92,4	1	58,9	24
4	5,13	7262	12	88,1	1	81,1	12
5	5,14	7262	14	98,8	1	71,8	14

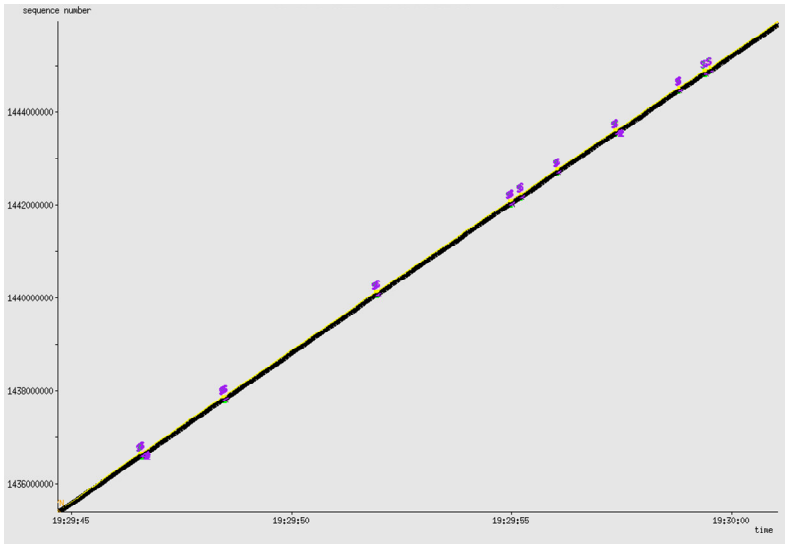
As happens with UDP traffic, Location 3 presents very poor channel conditions which results in high packet loss and a large number of retransmissions. This triggers the congestion avoidance mechanisms of TCP, leading to long periods where the transmitter

remains silent. Moreover, it is important to note the high variability that was experienced during the measurement campaign. This variability makes it really difficult to extract valid measurements since in several tests the FTP session was aborted due to expiration of maximum idle time.

A typical example of the behaviour of each interface can be seen in Figure 4-9. The transmissions over IEEE 802.11a presents a poor performance interrupted continuously with transmitter idle periods. The 802.11a graph corresponds to test #2 which does not have a really long maximum idle time (see for example test #5) but the number of retransmissions needed prevents higher throughputs being achieved. It is clear that under these circumstances the most suitable choice should be the 802.11b interface.



IEEE 802.11a



IEEE 802.11b

Figure 4-9: Location 3 TCP traffic time-sequence graphs

Location 4:

Table 4-10 presents the results obtained from the tests carried out at Location 4.

Table 4-10: Location 4 TCP statistics

# Test	Throughput (Mbps)	Total pkts	Out of order pkts	Idle time max (ms)	Max # of retx	Avg. retx time (ms)	# retx
IEEE 802.11b							
1	1,68	7790	430	3197,9	4	94,6	544
2	4,34	7392	136	415,9	2	48,4	147
3	1,83	7851	499	835,7	4	79,8	606
4	4,09	7392	130	465,9	3	63	147
5	2,19	7712	392	461,9	3	81,5	467

As can be seen, at Location 4 statistics can only be presented for 802.11b. The situation shows such poor channel conditions that it is impossible to finish any FTP session using the 802.11a interface. In contrast, 802.11b shows acceptable behaviour reducing its throughput but maintaining good performance. Note that even using 802.11b the radio channel impairments cause a large number of retransmissions.

Figure 4-10 shows the time-sequence graph of test #5. As can be seen, the evolution is not as linear as in the former locations but it remains free from long periods of transmitter inactivity.

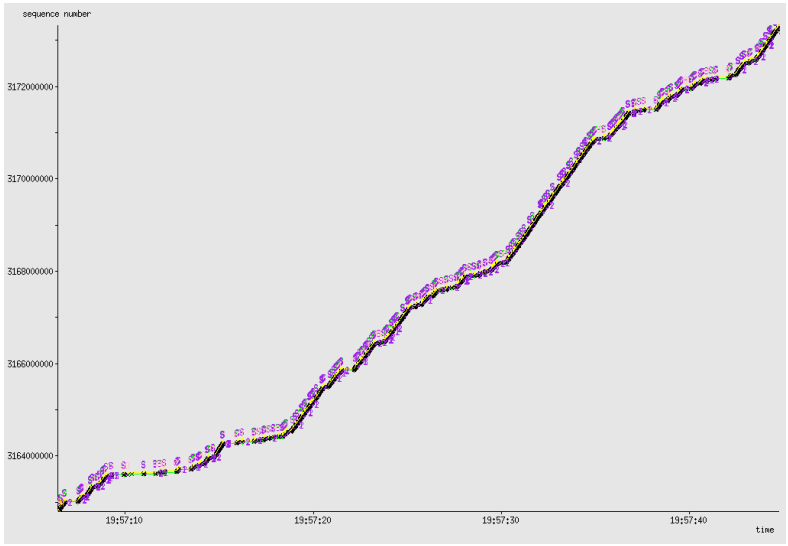


Figure 4-10: Location 4 TCP traffic time-sequence graphs for IEEE 802.11b

4.4.3 Dynamic interface selection

As has been mentioned, the decision-making process within the UCL can be as complex as desired, adapting the selection of the output interface for each individual packet to be sent

to multiple parameters (e.g. application specific QoS requirements, channel congestion, etc.).

4.4.3.1 Experimental evaluation of interface selection approach based on SNR

In our case, we have used a simpler approach to validate the optimization achieved with the UCL by deciding which interface to send the packet through, taking into account the SNR observed in each of the channels and the transport level protocol used. Observing the characterization performed in both UDP and TCP traffic, the SNR level used as a threshold for switching from one output interface to the other was set in the UCL implementation to 25 dB. As mentioned before, it might be necessary to increase the value when TCP traffic is used but we prefer not to do so as the region where TCP is vulnerable to channel impairments comprises a wide range of possible SNR values so we have chosen the mean value within this region.

As is shown in Figure 4-11, the tests performed for validating the optimal selection of the output interface to send the packets through consisted of moving one of the laptops from Location 1 to Location 3 and back again while the other remained fixed on the red mark in Figure 4-1. This test was carried out five different times, using first the UCL with its optimal interface selection option enabled, then the IEE 802.11a interface only and finally the IEEE 802.11b interface only.

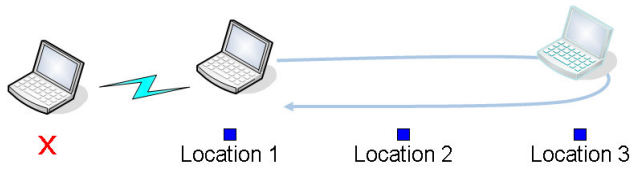


Figure 4-11: Dynamic network interface adaptation test scenario

Table 4-11 presents the results from the different tests carried out using UDP traffic. In these tests, a 1500-byte UDP packets flow lasting 40 seconds was exchanged between the two laptops.

It is important to note, that in all the tests performed we have tried to repeat exactly the same movements and speed. Nevertheless, the results might vary slightly from one to the other due to the impossibility of replicating exactly the tests. However, the different repetitions of the same experiment allow us to extract valid conclusions.

As can be seen, the UCL outperforms any of the two interfaces working in isolation. In the case of IEEE 802.11a the UCL not only achieves a higher throughput but also keeps the application loss to an almost negligible level which would assure the quality of the user experience. In the case of IEEE 802.11b the UCL greatly improves the resulting throughput, increasing it by 3 to 4 times, adapting itself to the channel conditions at all times and only selecting the slower interface if really required, establishing a trade-off between packet loss and raw performance.

Table 4-11: Moving scenario UDP statistics

# Test	Throughput (Mbps)	Application Loss (%)
UCL (optimal interface selection enabled)		
1	19,5	0,41%
2	23,4	0,06%
3	22,2	0,60%
4	19,1	0,26%
5	24,1	0,05%
IEEE 802.11a		
1	21,5	3,64%
2	17,8	6,62%
3	21,7	3,72%
4	17,9	7,50%
5	20,8	5,89%
IEEE 802.11b		
1	5,96	0,00%
2	5,96	0,02%
3	5,93	0,00%
4	5,96	0,00%
5	5,97	0,00%

Figure 4-12, Figure 4-13 and Figure 4-14 show some examples of the instantaneous and mean throughput evolving with time and also how many packets are lost at each moment.

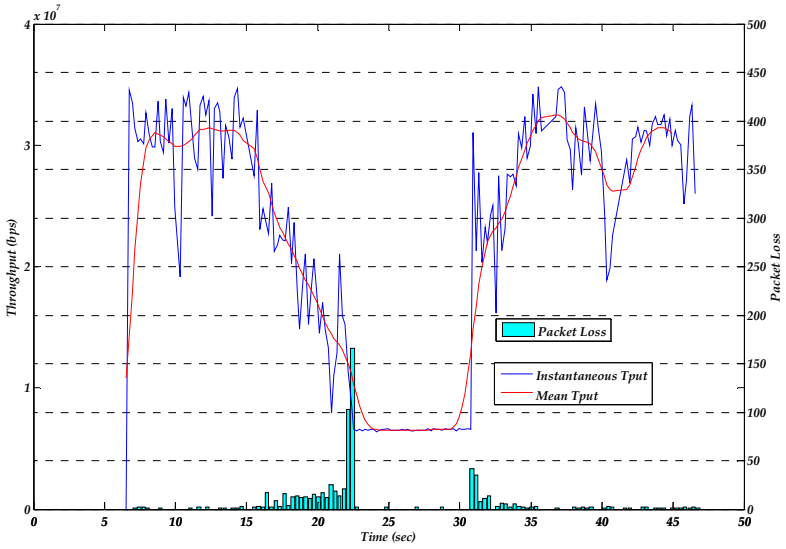


Figure 4-12: Moving scenario UDP traffic immediate throughput and packet loss evolution using the UCL

As can be seen, when the UCL is enabled, the 802.11a interface is used while the channel is good (i.e. Locations 1 and 2), but as soon as it observes that the SNR decreases (this fact can be seen in the increase of the number of lost packets) the UCL decides to change the

output interface to the 802.11b one. This event provokes a throughput decrease, but it is maintained while the mobile node roams from Location 2 to 3 and back, and the packet loss is nullified. When the mobile node re-enters the area where the SNR rises above the defined threshold, the 802.11a interface is selected again making the throughput increase without experiencing the disadvantage of high packet loss.

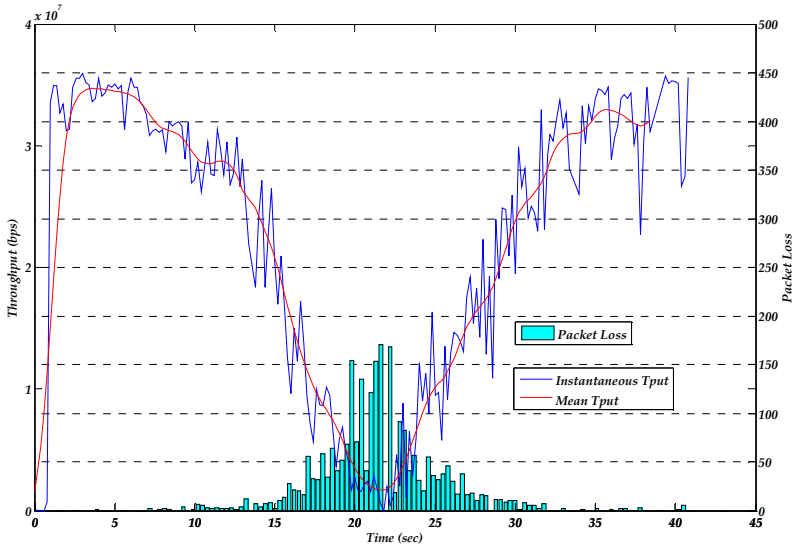


Figure 4-13: Moving scenario UDP traffic immediate throughput and packet loss evolution using the IEEE 802.11a interface only

As expected, during the period in which the laptop is moving around the locations that shows the poorest channel conditions the number of packets lost increases significantly when forcing the communication to go through the 802.11a interface only. This situation lasts for approximately 10 seconds (i.e. a quarter of the whole experiment) only, causing a severe degradation in the measurement. The UCL optimizes this situation by anticipating the dreadful conditions and selecting an output interface that shields the communication from the channel impairments at the cost of slightly reducing the throughput.

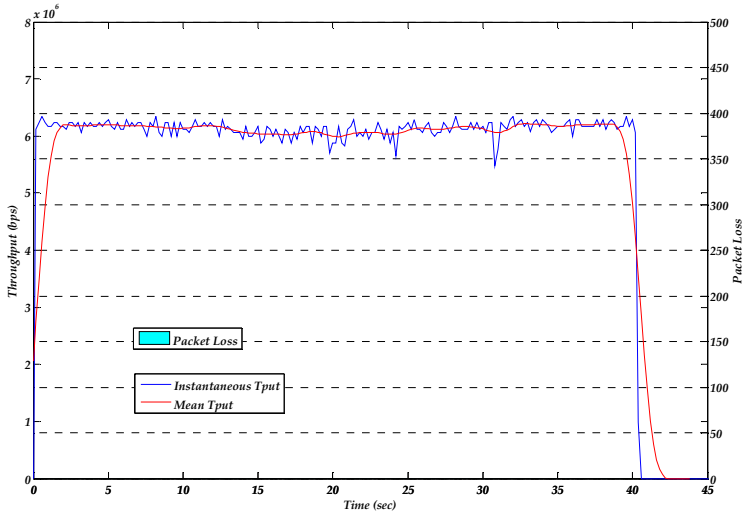


Figure 4-14: Moving scenario UDP traffic immediate throughput and packet loss evolution using the IEEE 802.11b interface only

The behaviour when using the IEEE 802.11b interface alone shows perfect stability and a null packet loss in most of the cases. Nevertheless, the throughput obtained is much worse than the other two options.

Table 4-12 presents the results from the different tests carried out using TCP traffic. In these tests, a 1500-byte TCP packets flow lasting 40 seconds was exchanged between the two laptops.

Table 4-12: Moving scenario TCP statistics

# Test	Throughput (Mbps)
UCL (optimal interface selection enabled)	
1	19,7
2	19,1
3	19,3
4	17,3
5	20,1
IEEE 802.11a	
1	12,7
2	13,5
3	11,3
4	11,9
5	12,0
IEEE 802.11b	
1	5,15
2	5,14
3	5,17
4	5,15
5	5,17

In the case of TCP, the erroneous interpretation of packet loss as channel congestion triggers the TCP congestion avoidance mechanisms which makes the throughput reduce significantly when using only the IEEE 802.11a.

Figure 4-15 shows how the UCL swaps the interface used from 802.11a to 802.11b before the conditions are so degraded that the congestion avoidance mechanisms are triggered. Thus, the throughput is maintained using the IEEE 802.11b and immediately after the laptop returns to the area where the channel conditions are good, it is able to swap again to the higher binary rate interface without having to wait for congestion avoidance timeouts to expire.

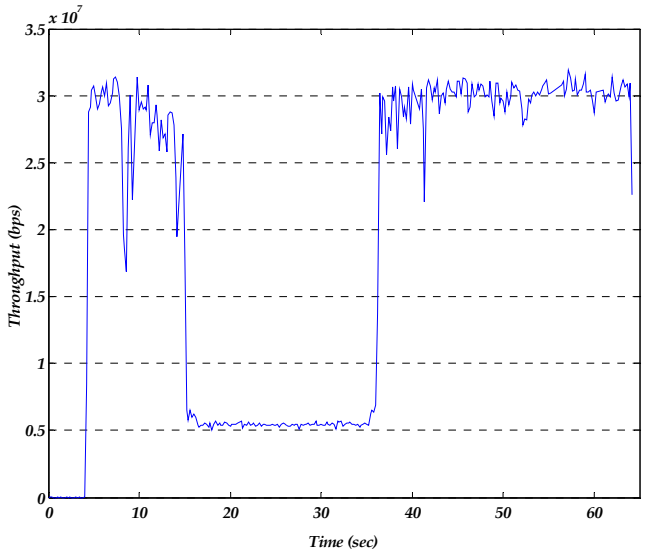


Figure 4-15: Moving scenario TCP traffic immediate throughput evolution using UCL

As can be seen in Figure 4-16 there is a period of time in which the instantaneous throughput is 0. This period corresponds to the time in which the laptop is around Location 3 plus the amount of time that the transmitter needs to start sending packets again, after its congestion avoidance mechanisms have removed the corresponding timeouts.

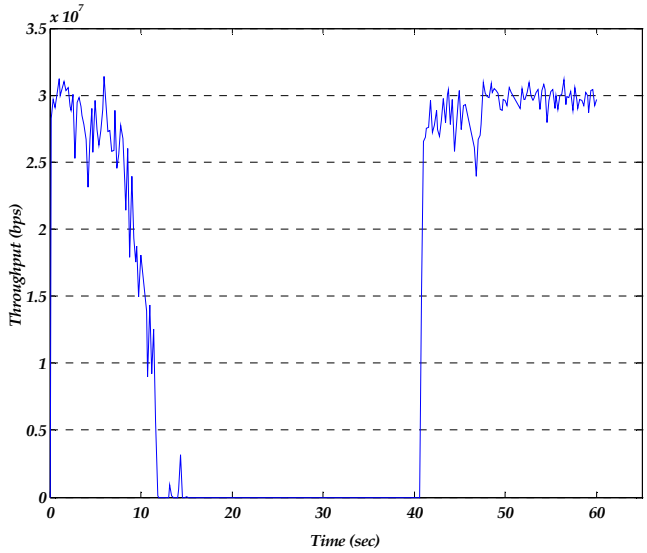


Figure 4-16: Moving scenario TCP traffic immediate throughput evolution using the IEEE 802.11a interface only

Similarly to the UDP case, in Figure 4-17 it can be seen that when the transmission is forced to go through the 802.11b the throughput is always maintained stable around 5 Mbps.

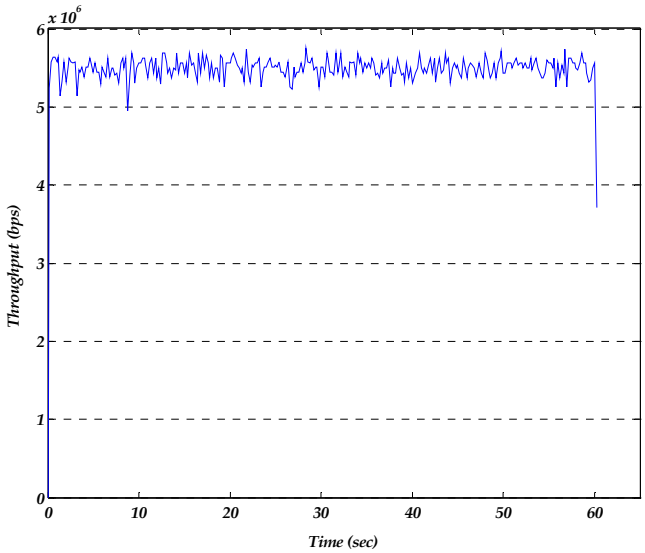


Figure 4-17: Moving scenario TCP traffic immediate throughput and packet loss evolution using the IEEE 802.11b interface only

As can be seen, the UCL always selects the most appropriate interface at each moment exploiting the advantages of the two options it has. When the channel is good, it uses the interface which offers higher bandwidth but when it detects that the channel is starting to deteriorate it switches to the interface that counteracts radio channel impairments better.

4.4.3.2 Simulation of interface selection approach based on packet loss

Although the results showed that important optimizations can be obtained using the SNR as the parameter for deciding which of the available interfaces to use, other parameters have been evaluated and compared via simulation studies.

In wireless communications the nature of the channel poses several challenges for data transmission. Wireless channels experience phenomena different from those observed in wired ones. In the wired world the transmitted signals do not experience all the degradations inherent to the wireless channel. Wireless fading channels are commonly characterized by Markov models. Models can be constructed to represent the signal degradations or their effects at different levels. It is not only important to understand the impact of the channel and the degradation on the signal itself, but also to know how this affects frames or packets as they are transmitted through the air. The importance of a channel model lies in how it can be used to simplify the analysis, design and deployment of communication systems.

In our case we have used a Gilbert-Elliot model as shown in Figure 4-18.

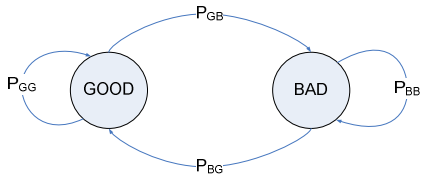


Figure 4-18: Gilbert-Elliot packet error model

We have also defined three different channels according to the observed SNR. Thus, for each of these three different channels, distinct parameters of the model were derived.

Table 4-13: Transition probability matrixes for the model of different channels

	IEEE 802.11a channels			IEEE 802.11b channels		
Frame Size	1500 bytes			1500 bytes		
Transmission Rate	54 Mbps			11 Mbps		
SNR	17 dB	24 dB	34 dB	17 dB	24 dB	34 dB
FER	--	25 %	5 %	20 %	1 %	0,01 %
P_{GG}	0,001	0,973	0,975	0,998	0,99	0,9999
P_{GB}	0,999	0,027	0,025	0,002	0,01	0,0001
P_{BB}	0,999	0,9	0,5	0,99	0,0001	0,0001
P_{BG}	0,001	0,1	0,5	0,01	0,9999	0,9999

The parameters shown in Table 4-13 have been derived from [144] and tuned with our own experimental characterization of the office scenario described in Section 4.2 so that they can be compared with our experimental measurements. They correspond to the probabilities of staying or going from one state to the other as it is depicted in Figure 4-18.

It is important to note that for the 17 dB SNR channel no model has been derived for the IEEE 802.11a interface. This is due to the fact that from the experimental characterization of this channel we can conclude that this is not an available channel for transmission with this technology.

Taking into account these parameters we simulated in Matlab® the scenario shown in Figure 4-19. We simulated a random sequence of states which modelled the reception status of the transmitted frames. A total of 60000 UDP frames were transmitted per simulation. In order to be able to compare the results from the simulations with the ones already presented from the experimental measurements campaigns, the number of frames transmitted over each of the modelled channels was selected to mimic the movements taken during the measurements campaigns. In this sense, in the scenario simulated two thirds of the frames where transmitted over the best channel while the other third was equally transmitted over the other two channels. It is important to note that IEEE 802.11 MAC implements an ARQ scheme by which each frame is retransmitted a given number of times if an error occurs. Typically this number is fixed to 4. Thus, Frame Error Rate (FER) is not equivalent to Packet Error Rate (PER). This fact was also taken into account in the simulations. A Monte Carlo approach was taken so simulations were run 1000 times.

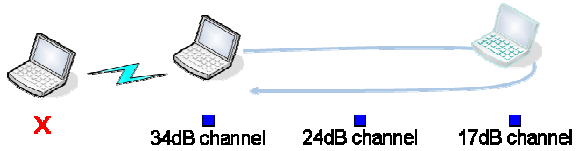


Figure 4-19: Simulation scenario for dynamic interface adaptation based on packet loss

The adaptation strategy based on packet loss taken at the UCL consisted of assessing the channel conditions using the number of lost packets. More specifically the UCL evaluated the bursts of correct and lost packets in order to know the actual status of the channel. These channels typically behave in a burst manner due to the fading processes they suffer. Hence, even when the SNR is relatively low we can correctly receive packets. This strategy attempts to take advantage of this behaviour and react quickly to these slots. The SNR strategy is slower since typical mobility in personal networks clusters obeys to a walking speed pattern and SNR variations are progressive. Thus, the SNR strategy can be seen as a little bit more conservative whereas the one based on packet loss is more aggressive leading to better throughput while the loss at application level can be increased too. Nevertheless, the number of packets lost on which the threshold is set for swapping to use a technology with a better channel can be adapted accordingly in order to obtain a trade-off between the throughput and the application loss. Similarly, the number of correct packets received through the most robust technology with which we decide to use a faster technology with the risk of having poorer channel quality can also be tuned.

All these possibilities have been simulated and the results are presented in Figure 4-20.

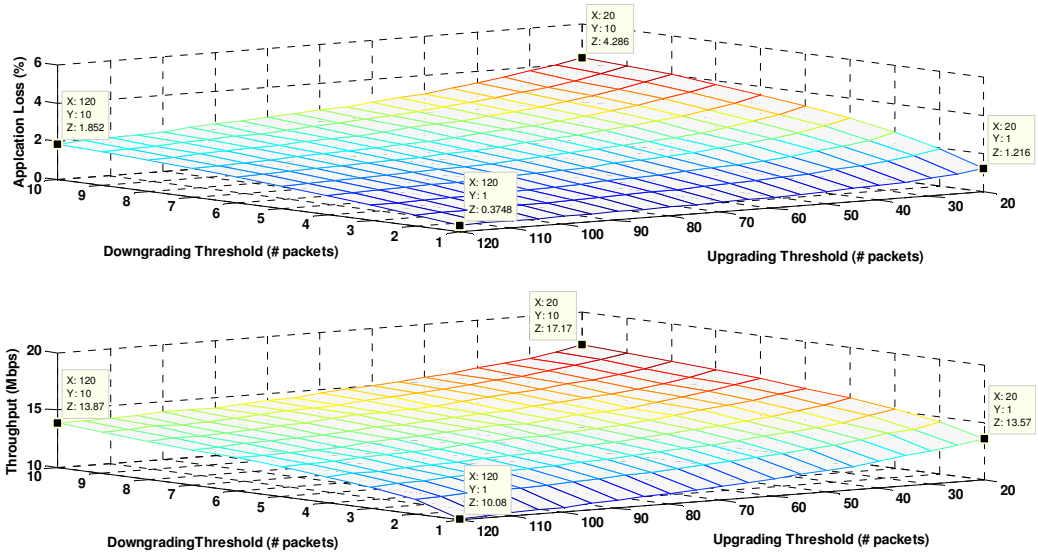


Figure 4-20: Results from simulations of packet loss based outbound interface adaptation

Figure 4-20 shows both the throughput achieved and the packet loss observed at application level for different configurations of the packet loss based adaptation strategy. As described above, the parameters that were modified during the simulations were the number of consecutive erroneous packets that lead to changing from 802.11a interface to 802.11b (downgrading threshold) and the number of consecutive correct packets that lead to swapping from the 802.11b interface to the 802.11a one (upgrading threshold).

When the downgrading threshold is increased, we accept the risk of losing more packets caused by maintaining the 802.11a interface (hoping that channel situation will not be too bad and the erroneous packet burst will not last). The lower the downgrading threshold is the lower the application loss. However, this also leads to reduced throughput. Conversely, when decreasing the upgrading threshold we assume that the channel gets

better sooner and thus change to the interface that allows better throughput. The price that is paid in this case is that this optimism might not fit totally with the reality and the channel is still not optimal thus degrading the system in terms of application level loss.

Figure 4-21 shows a comparison of the system behaviour when using the two proposed adaptation mechanisms.

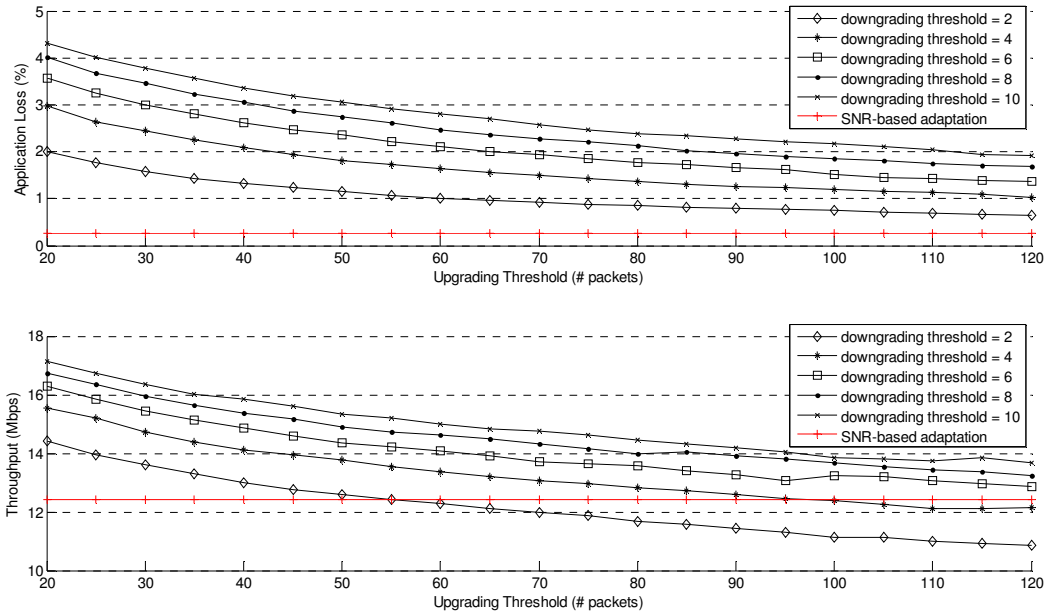


Figure 4-21: Comparison of different adaptation approaches

The red line in the figure corresponds to the system behaviour when the SNR observed in the channel is used for selecting the technology to be used. On the X-axis we have modified the upgrading threshold value while we show different results for different values of the downgrading threshold.

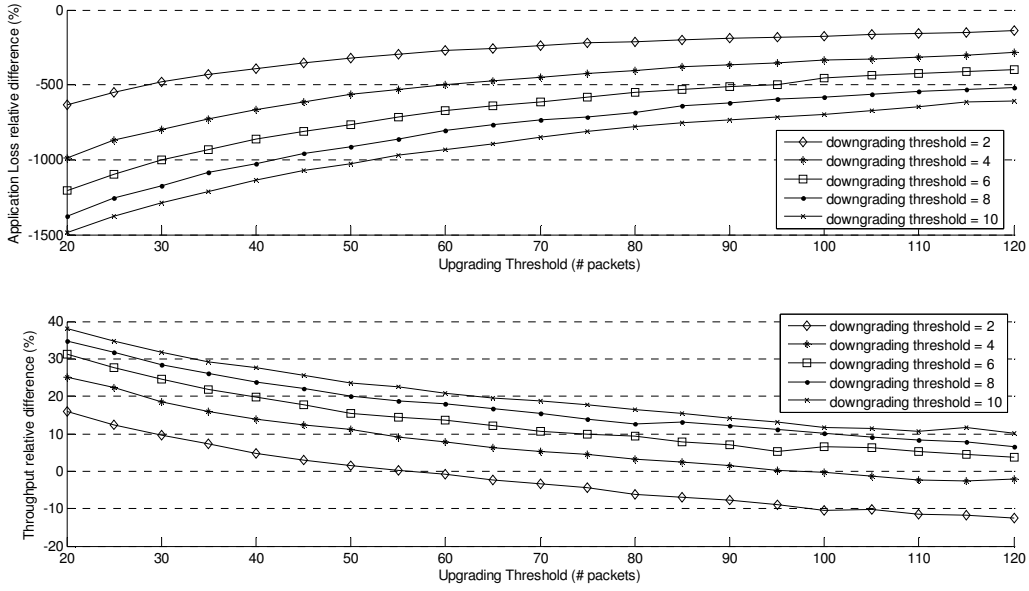


Figure 4-22: Relative difference in system performance using distinct adaptation strategies

As can be seen in Figure 4-21 and Figure 4-22 depending on the configuration selected for the packet loss based scheme we faced situations in which the simpler SNR based approach resulted in better behaviour when assessing individual performance parameters. The configurations studied for the packet loss based approach lead to higher packet loss as a consequence of the aggressive strategy for deciding when the channel might be good enough for using the 802.11a interface. The benefit is an increased overall throughput. When we choose a conservative strategy we obtain the opposite behaviour where application loss is highly reduced at the cost of lowering the throughput. However, it is interesting to observe that different adaptation strategies lead to optimization on different aspects. This allows us to conclude that a cognitive utilization of all of them might result on overall system optimization. Additionally, application level quality of service parameters might also be applied during the decision process so that the strategy that best suits can be applied. For example, if application loss is a critical issue, then a more conservative approach (i.e. SNR-based adaptation) should be taken whereas if a certain level of application loss is possible we can go for more aggressive strategy. UCL enables this flexibility which is actually one of the most important features of the solution proposed in this thesis.

Finally, we compared the UCL adaptive selection schemes with the non-UCL situation on which only 802.11a or 802.11b were used under the same moving scenario described in Figure 4-19.

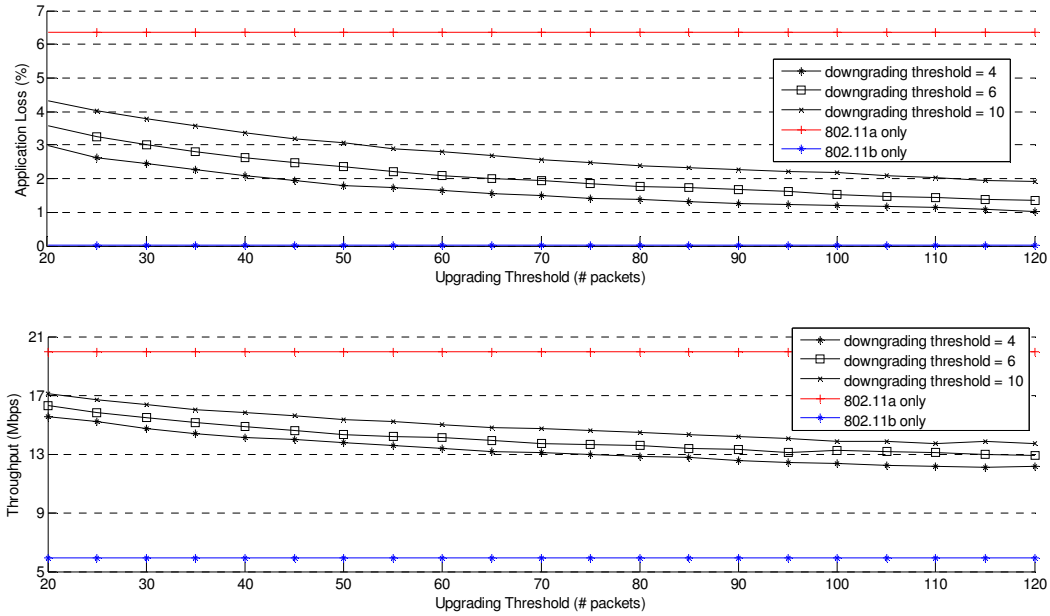


Figure 4-23: Comparison with non-UCL situation

Although the throughput achieved using only the 802.11a is around 30% higher than the one reached when using the UCL, it is at the cost of an unbearable application level loss. For any application this level of degradation would make it impossible to actually provide a decent quality of experience to the final user. This loss is reduced two to three times when adequate adaptation strategy is used. On the contrary, when only 802.11b is used the

packet loss is kept to the minimum[1] but a lot of the capacity of the system is unexploited. As can be seen, UCL dynamic interface adaptation allows optimized exploitation of the available resources.

4.4.4 Power-aware optimization based on dynamic interface selection

In the previous section only traffic-related parameters have been evaluated to assess the goodness of the dynamic interface selection strategies implemented at the UCL. Nevertheless, there is another important aspect that is exploited by intelligently selecting the most appropriate output interface, the power efficiency.

Taking the results from the experimental and simulation-based analyses done, we have derived the optimization in terms of power efficiency reached through the UCL dynamic interface selection mechanisms.

The WLAN card can be in any one of these five physical states: off, sleep, listen, receive and transmit. The majority of the power consumed is in receiving and transmitting states. Although it has been stated that 802.11a technology is more efficient than 802.11b [145] [146], a blind selection of the 802.11a interface might result in more efficient use of the power, but quality of experience would be severely degraded due to the packet loss that affects the communication when the channel is not good enough.

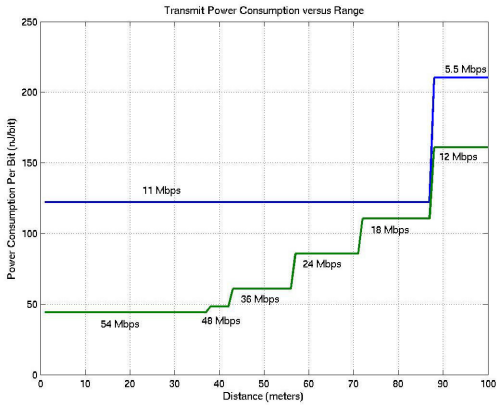


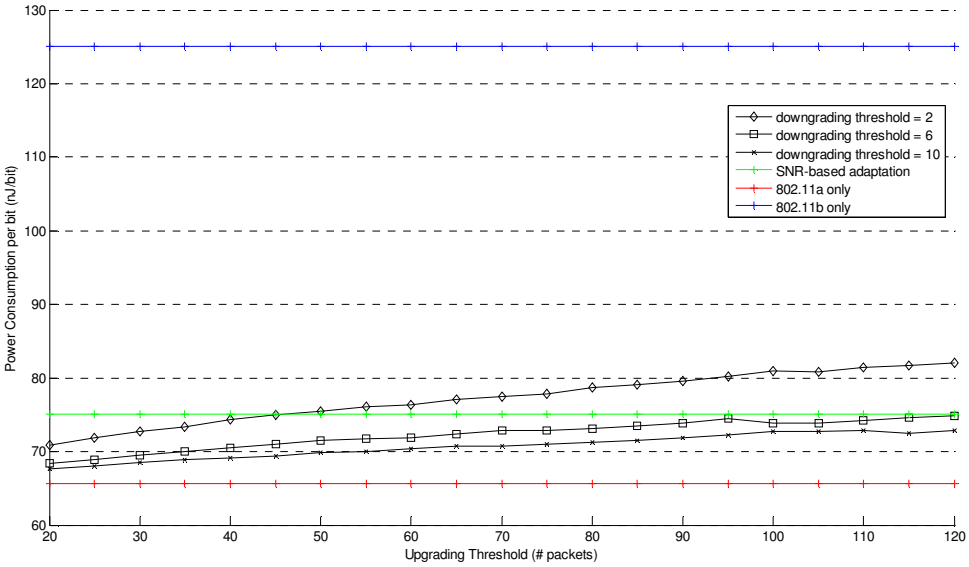
Figure 4-24: Energy consumed per transmitted bit (802.11b vs. 802.11a) (source [145])

Based on power consumption per bit values that can be seen in Figure 4-24 we have simulated the moving scenario as described in the previous section measuring the efficiency of the different strategies in terms of power consumption.

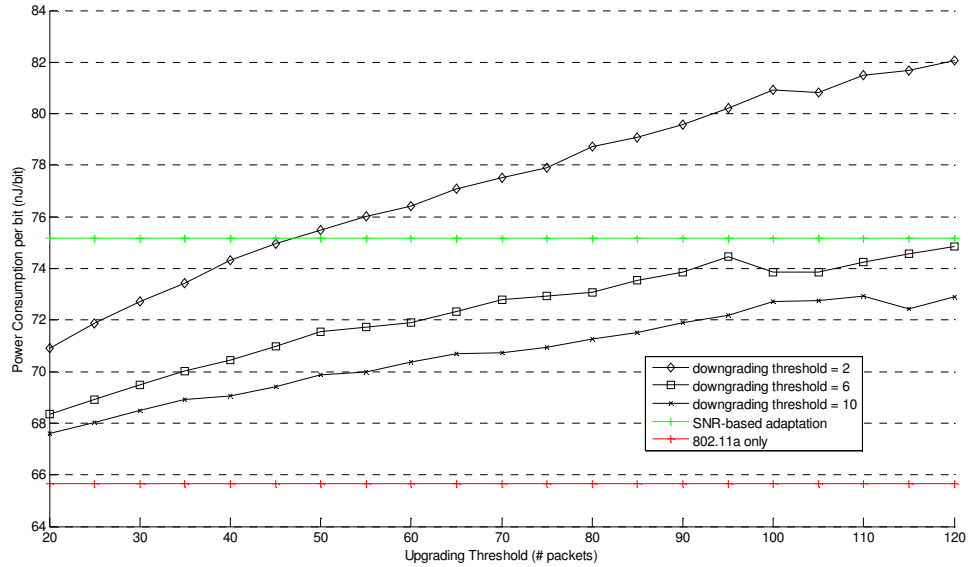
As is shown in Figure 4-25 a blind selection of the 802.11a interface would lead to better use of power since the consumed battery per unit of information is a bit lower than in any of the other strategies. Its better inherent efficiency compensates the amount of errors and necessary retransmissions that ruin the transmission when the channel is not good enough. In contrast, when only using 802.11b the efficiency is much lower even though the

[1] Note that the scenario implies passing through a channel that even for IEEE 802.11b radio and MAC leads to packet loss.

low frame error rate experienced is much better than on any of the channels that compose the moving scenario.



(a) including 802.11b



(b) excluding 802.11b

Figure 4-25: Power consumption efficiency of UCL vs non-UCL approaches

However, it can be seen that an intelligent use of the available interface leads to a sub-optimal power consumption, but does not jeopardize the system performance as would be the case with an 802.11b only approach, where throughput is reduced or with an 802.11a only approach, where application level loss is much higher.

It can be seen in Figure 4-26 that the left Y-axis show the power consumption per bit and the right Y-axis the throughput obtained in each of the cases.

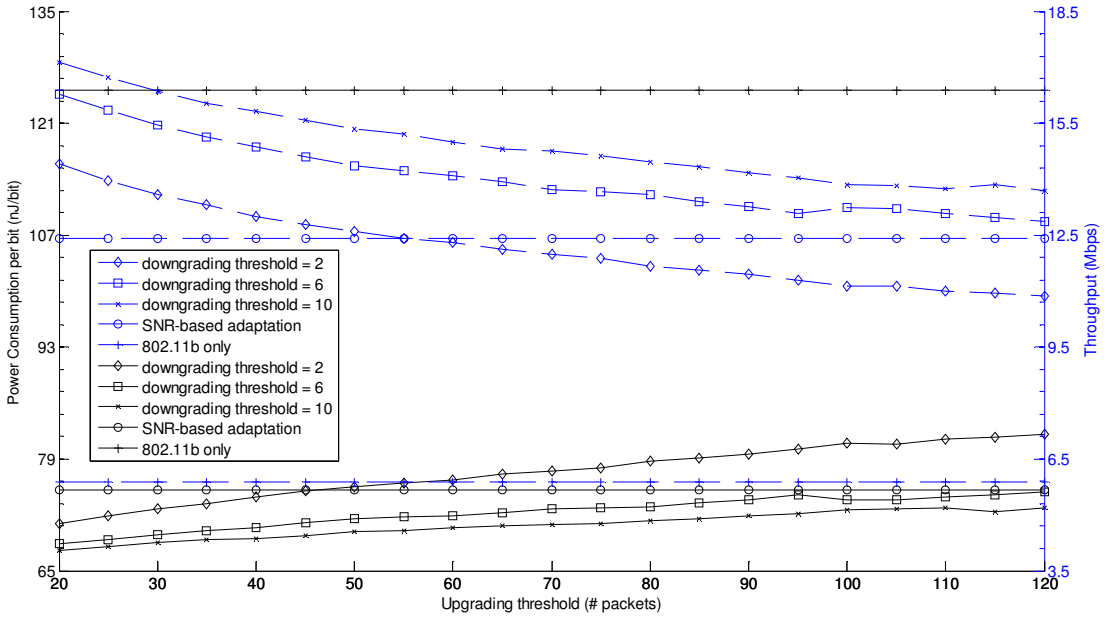


Figure 4-26: Throughput versus power consumption efficiency of UCL and 802.11b only approaches

As can be seen, when selecting the appropriate parameters for the packet loss based adaptation strategy the UCL achieved better power consumption per bit figures while highly increasing the system throughput. For example, using 6 packets as the downgrading threshold and 65 packets as the upgrading one we can have almost 2.5 times more throughput while still reducing the overall power consumption per bit by 40%, which means almost halving the total power consumed during the simulation. When the SNR-based approach is used we also double the throughput while the power consumption is still 40% less than when using the 802.11b interface only.

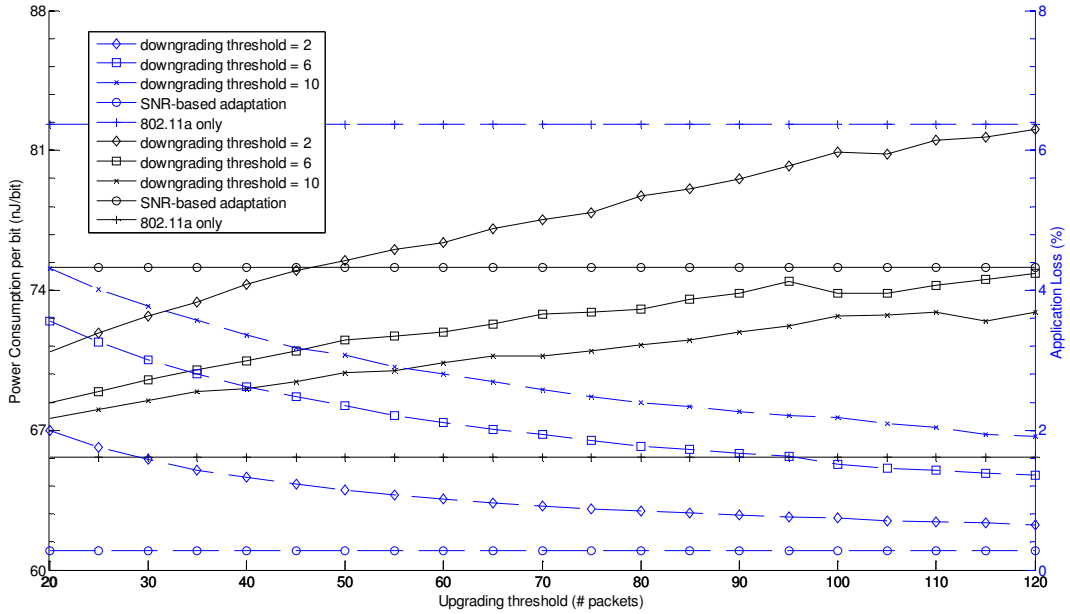


Figure 4-27: Packet loss versus power consumption efficiency of UCL and 802.11a only approaches

It can be seen in Figure 4-27 the left Y-axis represents the power consumption per bit and the right Y-axis the packet loss at application level obtained in each of the cases.

Similarly, when selecting the appropriate parameters for the packet loss based adaptation strategy the UCL achieved relatively smaller power consumption per bit figures, but highly reduced the packet loss at application level. For example, using 6 packets as the downgrading threshold and 65 packets as the upgrading one we can obtain three times less packets lost at the application level, while still keeping the overall power consumption per bit only 10% higher. When the SNR-based approach is used the packet loss is reduced drastically (around 20 times less), while keeping the power consumption only 15% more than that when using the 802.11a interface all the time.

Table 4-14: Moving scenario experimental power consumption efficiency statistics

# Test	Total Packets	Application Loss (%)	Total Power Consumption (J)	Power Consumption per bit (nJ/bit)
UCL (SNR-based interface selection enabled)				
1	70071	0,41%	52,72	62,69
2	83355	0,06%	61,94	61,93
3	80130	0,60%	60,68	63,11
4	68386	0,26%	51,18	62,36
5	86129	0,05%	63,98	61,91
IEEE 802.11a				
1	80003	3,64%	68,43	71,28
2	68040	6,62%	63,06	77,24
3	80525	3,72%	69,03	71,44
4	69385	7,50%	65,78	79
5	79168	5,89%	71,99	75,78
IEEE 802.11b				
1	21390	0,00%	32,15	125,27
2	21407	0,02%	32,2	125,36
3	21295	0,00%	32,01	125,27
4	21371	0,00%	32,13	125,27
5	21400	0,00%	32,17	125,27

Using the same power consumption parameters used for the simulations we measured the efficiency of the SNR-based adaptation strategy implemented in the UCL prototype. Table 4-14 presents the results from the measurement campaign.

It is important to note that not only the total power consumed is important, but the power consumed per bit transmitted and correctly received is also a critical parameter. In this sense, although the overall power consumed in 802.11a tests were quite similar to the case when enabling the UCL SNR-based adaptation strategy, the efficiency is around 20% better in the latter case.

Figure 4-28 shows the result from the experimental comparison in a graphical manner.

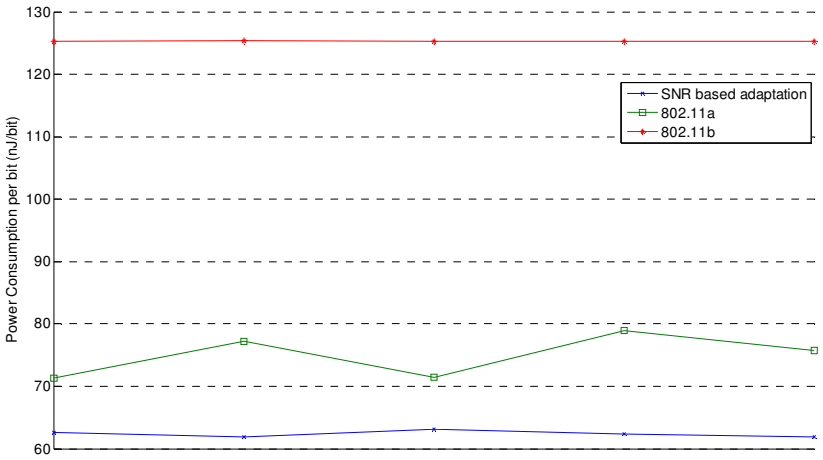


Figure 4-28: Moving scenario experimental power consumption efficiency

4.4.5 Striping of user-data flows at UCL level

Grouping all the air interfaces offers the possibility, among others, of using many of them at the same time for communicating with the same peer whenever both peers share these air interfaces. This situation could be exploited when the characteristics of any of the wireless technologies do not cope with the traffic flow requirements, but multiplexing over several fulfils the flow expectations.

4.4.5.1 Analysis of the UCL striping strategies

As already presented in paragraph 3.1.6.1 we defined the normalized relative capacity of a link as:

$$\chi_i = \left\lfloor \frac{T_{\max}}{t_i} \right\rfloor$$

while the normalized aggregated capacity of the system when using striping is:

$$\mathrm{X} = \sum_{i=1}^{n} \chi_i$$

This capacity represents the number of packets that can be transmitted following any of the two striping strategies developed for the UCL in T_{max}.

The transit times are inversely proportional to the transmission speed. Hence, the throughput for each of the n links available can be expressed as $BW_i = \frac{1}{t_i}$, expressed in bits per second if we assume t_i to be the time to transmit one bit. Similarly $BW_{\min} = \frac{1}{T_{\max}}$ is the binary rate of the slowest of the available links used for striping.

The achievable throughput using UCL striping techniques is calculated as:

$$Tput_{UCL} = \mathrm{X} \cdot BW_{\min}$$

Compared with a strategy that would consist in a uniform distribution of the packets in any link without considering the availability of the links, where the achievable throughput would be reduced to $Tput_{naive} = n \cdot BW_{\min}$, our strategy leads to fast links being well exploited. In contrast, if the resources are optimally used we could reach $Tput_{opt} = \sum_{i=1}^{n} BW_i$.

For a normalized BW_{min}=1 we can conclude that the difference between the maximum achievable throughput and the one that is obtained using the UCL striping techniques corresponds to the error of the floor function. In this sense, the degradation of the UCL striping methods compared with the optimal one obeys the following rule:

$$err = \sum_{i=1}^{N} \left(\frac{1}{t_i} - \left\lfloor \frac{1}{t_i} \right\rfloor \right)$$

Hence, we can establish a linear limit for this degradation such that $err \leq N \quad \forall \quad t_i$

Figure 4-29 shows the respective achievable throughputs using the different striping methods in a system where one IEEE 802.11b, one IEEE 802.11a and one Bluetooth interface are available.

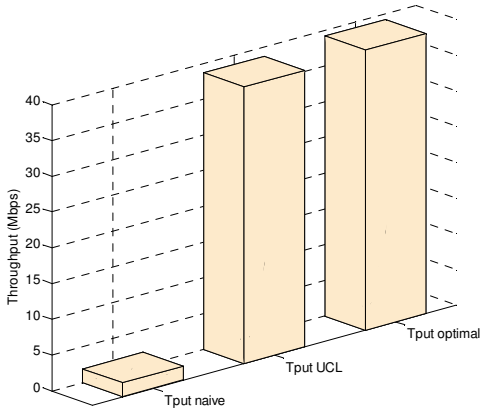


Figure 4-29: Achievable throughput in an 802.11b-802.11a-Bluetooth system

As can be seen the UCL approach almost reaches the optimal behaviour, only 1% is lost with respect to the optimal solution.

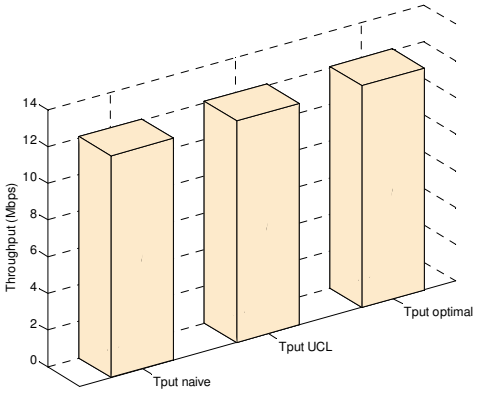


Figure 4-30: Achievable throughput in an 802.11b-802.11b system

Figure 4-30 shows the respective achievable throughputs using the different striping methods in a system where two IEEE 802.11b interfaces are available.

In this case, since all the available interfaces are identical there is no difference between any of the strategies studied.

These analyses will be used together with the one in the following section to derive conclusions on the appropriacy of the implementation of the sequential striping strategy performed in the UCL framework.

4.4.5.2 Experimental evaluation of UCL sequential striping strategy

Figure 4-31 shows the set-up of the testbed over which the measurement campaign was performed to validate the striping features in the UCL. Two laptops equipped with two IEEE 802.11b interfaces working in the ad hoc mode and at a maximum bit rate of 11 Mbps were used. They were running Linux OS and the UCL was loaded on both of them.

Figure 4-31: Striping experiment set-up

Table 4-15 shows the throughput obtained when sending 10,000 UDP packets of 1500 bytes from one of the laptops in Figure 4-31 to the other.

Table 4-15: Throughput using UCL multiplexing capacity

Avg. Throughput (Mbps)	Variance
10.84	0.05

Figure 4-32 shows the comparison of the instantaneous throughput evolution during one of the tests when using the UCL (blue line) and when using a native configuration (red line).

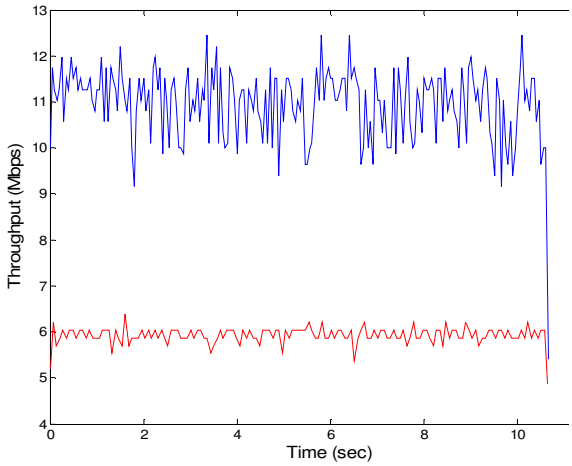


Figure 4-32: Instantaneous throughput comparison between striped and non-striped scheme

Since the UCL is loaded and detects the possibility of striping the transmission among the two 802.11b interfaces, packets are sent over both of them. The experiment was repeated ten times showing that the performance achieved is almost double the one that is achieved by a single 802.11b interface [147]. The difference compared with the theoretical value resulting from adding the throughput obtained with each of the individual interfaces is due to the fact that, due to hardware constraints, the behaviour is not really parallel, but there is a pipelined situation.

4.4.6 Conclusions about UCL validation results

In the above sections we have presented a complete set of results obtained from different measurement campaigns that have proven firstly the feasibility of the implemented UCL and secondly the optimization achieved through its use.

First of all, being able to carry out experimental measurement campaigns demonstrates one of the main achievements in this thesis, which is the actual implementation of the UCL framework in a real prototype. The tests do not only validate the functionalities at the UCL, but also show the benefits of the cognitive cross-layer optimization approaches implemented. Besides, additional analytical and simulation studies have been carried out in order to obtain elements for comparison with the results in the real-world tests and to be able to study, in scenarios that cannot be easily set up in a real testbed, the behaviour of the solutions proposed in this thesis.

The corresponding tests have been made in order to verify that the inclusion of the UCL in the communication protocol stack does not introduce any major overhead in terms of performance. Additionally, it is important to note that the tests performed to measure the degradation introduced by the UCL is due to the security sub-module operation, which performed the encryption and decryption of all the traffic sent.

The optimizations introduced by the UCL, through the selection of the most appropriate output interface and the possibility of striping traffic destined to a node through different wireless interfaces, have been presented. The results presented in this thesis demonstrate the suitability of the mechanisms implemented in the UCL in order to optimize the communications within PN clusters and proves the appropriacy of the cognitive cross-layer approach for improving the intra-cluster communications over heterogeneous scenarios.

Finally, during the different validation tests, possible enhancements have been identified that can be supported by the UCL although they have not been implemented yet. The results presented in this section show that the solutions developed in this thesis represent a step forward in the cross-layer optimization paradigm. It paves the way for more sophisticated strategies that fully develop the cognitive network concept. In this sense, the solutions implemented exploit the learning-assessing-adapting approach precluded on the cognitive networking. The UCL is the framework over which advanced techniques for assessing the system state can be developed. Taking advantage of the global view that the convergence layer provides (access to information from multiple layers and access technologies), a complete map of the system can be inferred. After matching it with the

user requirements, the UCL can take the appropriate decisions in order to best serve the final users' wishes, optimizing their quality of experience not only by providing enhanced bandwidth but also improving the power consumption efficiency or enhancing the service provision by lowering the packet loss due to wireless channel impairments.

CHAPTER 5
SECURE COMMUNICATIONS OVER HETEROGENENOUS WIRELESS ENVIRONMENTS

A fundamental property of the PN is that personal devices form private multi-hop clusters in an ad hoc manner whenever they come across each other. In this concept strong emphasis is put on security and privacy of communications.

The purpose of this chapter is to present and analyze the protection mechanisms that allow efficient communication while safeguarding the user communications privacy. In order to assess the performance, different evaluation models and tests have been performed. The evaluation and assessment procedure followed consisted of both the development of analytical models and the performance of a set of experimental tests. The functionalities of the components presented in Chapter 3, their correctness and their efficiency have been evaluated thoroughly through comparison of the analytical and experimental results. Additionally, we will report on the results obtained from a set of measurement campaigns that compared the solution implemented within the UCL for securing the communications with other alternative approach. As such, valuable lessons can be learnt about the current performance of the implemented software and optimal parameter settings and guidelines can be provided for future extensions that can help to improve performance. For more details on how the software works, please refer to Appendix A.

5.1 INTRODUCTION

Much of the work on ad hoc networking has been concentrated on the unicast routing algorithms and several ad hoc routing protocols have been specified in the Internet Engineering Task Force (IETF). Nevertheless, current MANET routing protocols [148][149] inherently trust all participants of the network. This naive trust model allows malicious nodes to paralyze an ad hoc network by inserting erroneous routing updates, replaying old messages, changing routing updates or advertising incorrect routing information.

The main set of requirements this work was based on were produced under the conditions and assumptions of a personal networking scenario. Nevertheless, they fit in many future wireless communications paradigms where private mobile ad hoc networks are to be used. The main requirements are:

- Only trusted nodes can be part of the network. More specifically, a security association must exist between network members. The security associations are managed and monitored by the user.
- The privacy of the communications has to be assured. Sensitive information must be exchanged confidentially between the nodes in the network.
- The network has to support the heterogeneity. Multiplicity of wireless technologies must not be an obstacle and the flaws in terms of security of any of these technologies must not jeopardize the whole system security.
- Unicast and broadcast traffic must be supported. Secure communications must not be limited to point-to-point interactions.

This section presents the results obtained from derived analytical models and the measurement campaigns carried out in order to assess the performance of the solutions implemented at the UCL to enable secure communications over heterogeneous wireless environments.

5.2 SECURE LINK ESTABLISHMENT

Figure 5-1 shows the secure link layer establishment procedure. As was described in Chapter 3, after the beacon is received a four-way handshake, based on the Extensible Authentication Protocol (EAP), is initiated by acknowledging this beacon.

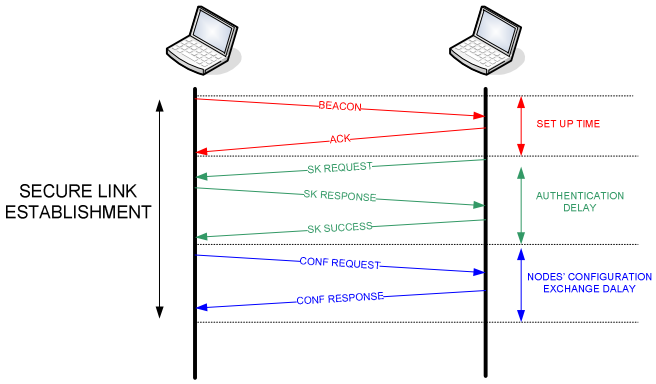


Figure 5-1: Secure link establishment procedure

In order to assess the performance, different evaluation models and tests have been performed. The evaluation and assessment procedure followed consisted of both the development of analytical models and the performance of a set of experimental tests. The following sections present the results of these studies.

5.2.1 Analytical assessment of secure link establishment time

As is shown in Figure 5-1, the secure link establishment time can be divided into three different steps, namely Discovery time, Authentication delay and Configuration exchange time.

– Discovery time

This value measures the amount of time it takes from the moment one node is started up till it is discovered by another peer (already up and running).

$$T_{disc} = T_{tx-BEAC} + T_{tx-ACK} + 2 \cdot T_{proc} \quad (5.1)$$

The discovery time, T_{disc}, involves the transmission of the beacon and its corresponding acknowledgement. As described in (5.1) T_{disc} will consist on the time it takes to transmit the beacon packet, $T_{tx\text{-}BEAC}$, plus the time it takes to transmit the acknowledgement, $T_{tx\text{-}ACK}$, plus twice the time it takes the node to process each of the frames, T_{proc}.

$$T_{tx} = T_{acc} + T_{prop} + t_{tx}$$
$$t_{tx} = \frac{number\ of\ bytes}{Binary\ rate} \quad (5.2)$$

Basically, the transmission time of a frame can be divided into the time the node needs to access the channel, T_{acc}, the propagation time, T_{prop}, and the time it takes for transmitting the frame itself, t_{tx}.

For the analytical modelling of the Discovery time, we are assuming T_{prop} and T_{proc} to be negligible. Hence from (5.1) and (5.2) we have:

$$T_{disc} = 2 \cdot T_{acc} + t_{tx-BEAC} + t_{tx-ACK} \quad (5.3)$$

All the different times in (5.3) are dependent on the link layer technology used. In our experiments we have used three different ones, namely IEEE 802.11a, IEEE 802.11b and Bluetooth.

Table 5-1: Discovery time over different wireless technologies

	T_{acc}	$Rate_{BEAC}$	# $bytes_{BEAC}$	$Rate_{ACK}$	# $bytes_{ACK}$	T_{disc}
IEEE 802.11a	0,17 ms(*)	6 Mbps	94	54 Mbps	94	0,51 ms
IEEE 802.11b	0,87 ms(*)	1 Mbps	94	11 Mbps	94	2,59 ms
Bluetooth	---	DH5	94	DH1	94	12,5 ms

(*) time obtained by analyzing IEEE 802.11 access protocol

Table 5-1 presents the different values needed to calculate the Discovery time.

Results are presented using the maximum binary rate for each of the analyzed wireless access technologies.

– Authentication delay

This value measures the amount of time it takes to complete the derivation and exchange of the link-level session keys between two personal nodes.

$$T_{auth} = 3 \cdot T_{tx-SK\_*} + 3 \cdot T_{proc}$$
$$T_{proc} = \frac{number\ of\ bytes}{encryption / decryption\ rate} \quad (5.4)$$

In this case, the Authentication delay, Tauth, does not only comprise the pure transmission of the frames over the air, Ttx-SK_*, but the processing time, Tproc, must also be considered since the encryption and decryption of the EAP messages is comparable to the rest of periods.

As depicted in (5.2) the transmission time is fully dependent on the wireless technology used. Table 5-2 presents the parameters with which Tauth can be calculated.

Table 5-2: Authentication delay over different wireless technologies

	T_{acc}	Rate	# $bytes_{EAP\_*}$	Encryption/Decryption (μs/byte)	T_{auth}
IEEE 802.11a	0,17 ms(*)	54 Mbps	106	0,02(**)	0,55 ms
IEEE 802.11b	0,87 ms(*)	11 Mbps	106	0,02(**)	2,59 ms
Bluetooth	---	DH5	106	0,02(**)	18,76 ms

(*) time obtained by analyzing IEEE 802.11 access protocol

(**) time obtained from a benchmark carried out on a Pentium IV processor

– Configuration exchange time

This value measures the amount of time it takes to exchange each node's IP addresses once they are authenticated.

As shown in Figure 5-1, this exchange consists of a couple of messages sent between the pair of nodes whose transmission time is defined as $T_{tx\text{-}CONF\_*}$. It is important to take into account that these messages are also encrypted so the resulting T_{conf} expression is:

$$T_{conf} = 2 \cdot T_{tx-CONF\_*} + 2 \cdot T_{proc} \quad (5.5)$$

where T_{proc} is calculated as in (4).

Table 5-3 shows the analytical results for the Configuration exchange time.

Table 5-3: Configuration exchange time over different wireless technologies

	T_{acc}	Rate	# $bytes_{CONF\_*}$	Encryption/Decryption (μs/byte)	T_{conf}
IEEE 802.11a	0,17 ms(*)	54 Mbps	100	0,02(**)	0,37 ms
IEEE 802.11b	0,87 ms(*)	11 Mbps	100	0,02(**)	1,72 ms
Bluetooth	---	DH5	100	0,02(**)	12,5 ms

(*) time obtained by analyzing IEEE 802.11 access protocol

(**) time obtained from a benchmark carried out on Pentium IV processors

In summary, Table 5-4 presents the aggregated time for the secure link establishment. The results of the analytical model will be compared with the results obtained from the measurement campaign carried out to assess the impact of the negotiation of the link-layer session keys in the system. As can be seen, the total time for the establishment of the secure link is negligible.

Table 5-4: Secure link establishment aggregated time over different wireless technologies

	T_{disc}	T_{auth}	T_{conf}	**T_{est}**
IEEE 802.11a	0,51 ms	0,55 ms	0,37 ms	**1,43 ms**
IEEE 802.11b	2,59 ms	2,59 ms	1,72 ms	**7,90 ms**
Bluetooth	12,5 ms	18,76 ms	12,5 ms	**43,76 ms**

The analysis done assumes that the first beacon is sent at time t=0. This is correct if one of the nodes is turned on while the other is running in the same radio domain since the first action a node takes when it is started up, is to start sending beacons. In a more general case, where both nodes are previously running and they enter in the same radio domain, T_{est} would depend on the inter-beacon interval.

Basically, the time for receiving a beacon from a node that is in the same radio domain can be modelled with a uniform random variable between 0 and T_{BEAC}:

$$f(t) = \begin{cases} \dfrac{1}{T_{BEAC}} & t < T_{BEAC} \\ 0 & t > T_{BEAC} \end{cases} \tag{5.6}$$

In this case, since both nodes are sending beacons independently at the same rate and it is not relevant which of them receives the beacon first (reacting with an acknowledgement), the probability of receiving a beacon is given by the addition of both beaconing random variables:

$$f(t)=\begin{cases}\dfrac{4\cdot t}{T_{BEAC}} & 0\le t\le \dfrac{T_{BEAC}}{2}\\ \dfrac{4\cdot\left(T_{BEAC}-t\right)}{T_{BEAC}} & \dfrac{T_{BEAC}}{2}\le t\le T_{BEAC}\\ 0 & t>T_{BEAC}\end{cases} \tag{5.7}$$

On average, we would then need to add $E[T]=\frac{T_{BEAC}}{2}$ to the already calculated T_{est} which takes into account that the order of magnitude of the respective times can be assumed to depend exclusively on the time it takes to receive the first beacon.

5.2.2 Experimental assessment of secure link establishment time

In order to test the real performance of the system implemented, several experimental measurement campaigns have been carried out. These experiments have a twofold approach; on the one hand they will be used to assess the functionality and performance of the system under real world conditions, thus, we have configured different test conditions that can be associated to a real user scenario; on the other hand, we will be able to compare the experimental results with the analytical studies to assess the validity of both analyses.

Figure 5-2 shows the different set-ups for the measurements.

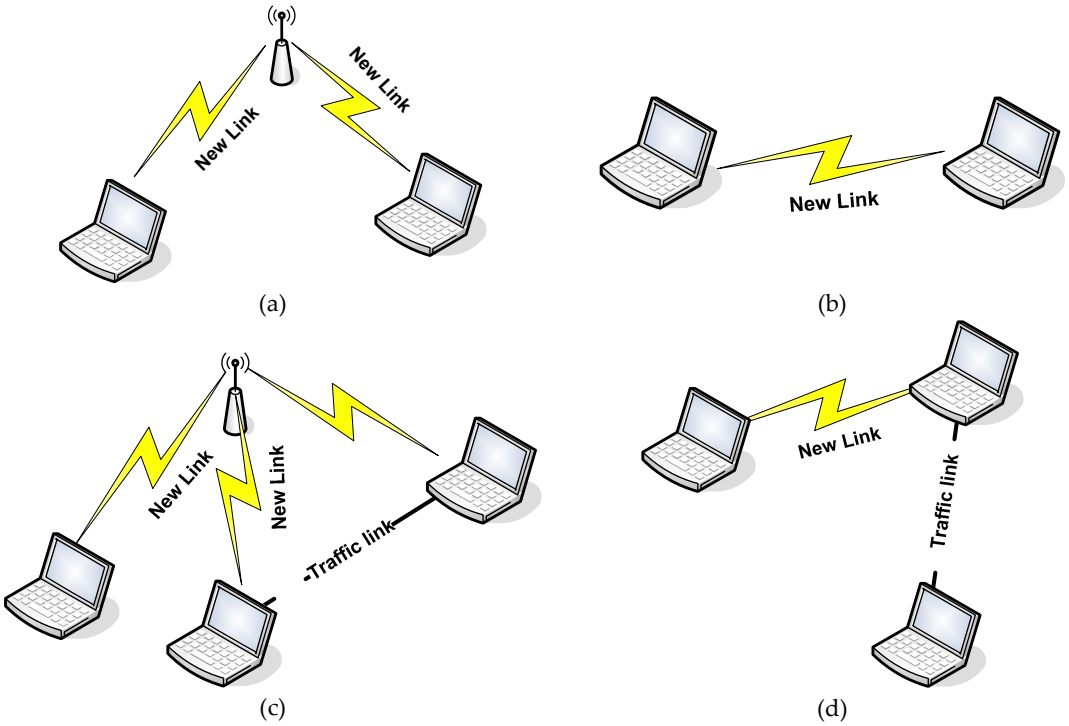


Figure 5-2: Secure link establishment measurement setup

The tests were performed using three of the most relevant wireless technologies that are available nowadays, namely IEEE 802.11a, IEEE 802.11b and Bluetooth. For the 802.11a

and 802.11b technologies the two possible operation modes, ad-hoc (case b in Figure 5-2) and infrastructure (case a in Figure 5-2) were used during the measurements.

Additionally, the effect of the background traffic on the secure link establishment was also evaluated. A third node was used during the tests in order to emulate the fact that one of the nodes involved in the secure link establishment was maintaining a traffic transaction when the establishment is triggered (cases c and d in Figure 5-2). We used both TCP and UDP bulk transactions to evaluate the different possibilities of background traffic. For the Bluetooth case, the characteristics of the technology itself (i.e. peer-to-peer) prevent us from analyzing any cases other than the one presented in (b) in Figure 5-2.

The tests were performed with Pentium IV based laptops. As discussed in the previous section, for the general case of two nodes meeting and starting the secure link establishment, the test is dependent on a triangular random variable that is given by the addition of the uniform beaconing process. In this sense, we have focused on the actual time it takes as shown in Figure 5-1. To this end, we set one of the laptops to be continuously running and then boot up the other one. Since the first action taken by the node is to start sending beacons we can measure the time for establishing the secure link. We repeated each test 8 times in order to correctly assess the performance of the system implemented.

For our tests we have considered three different scenarios for each topology and wireless technology:

- No traffic in the network. In this case the only packets exchanged were those related with the secure link establishment. This evaluates the ideal conditions where the T_{est} is minimum.
- UDP traffic in the network. A bulk UDP traffic transfer is introduced between a third node and the one involved in the secure link establishment that is always up. The *iperf* tool [150] is used to generate and send these UDP packets.
- TCP traffic in the network. Using the same topology as in the previous case, this time the third node created an FTP session to download a file while completing the secure link establishment.

It should be noted that all the tests were done under ideal channel conditions so no packet loss appears.

5.2.2.1 Secure link establishment time results

In Table 5-5 the results from the experiments are presented for the different wireless access technologies analyzed.

The same results are graphically presented in Figure 5-3, Figure 5-4 and Figure 5-5.

Table 5-5: Experimental results of measurement campaign

	T_{disc}	T_{auth}	T_{conf}	**T_{est}**
IEEE 802.11a	0.67	0.78	0.45	**1.90**
	0.71	0.71	0.42	**1.84**
	0.58	0.82	0.49	**1.89**
	0.61	0.72	0.41	**1.74**
	0.59	0.91	0.45	**1.95**
	0.67	0.76	0.49	**1.92**
	0.72	0.77	0.47	**1.96**
	0.77	0.80	0.45	**2.02**
IEEE 802.11b (Ad-hoc mode)	2.96	2.96	1.66	**7.59**
	2.79	2.39	2.41	**7.58**
	3.81	2.44	2.11	**8.36**
	3.45	3.63	1.88	**8.96**
	2.78	2.78	1.64	**7.20**
	2.78	2.42	1.64	**6.84**
	2.79	2.41	2.44	**7.64**
	2.78	2.42	1.64	**6.84**
IEEE 802.11b (Infrastructure mode)	37.42	3.235	7.86	**48.52**
	23.78	3.905	3.832	**31.52**
	40.98	3.475	3.041	**47.50**
	32.59	5.675	2.69	**40.96**
	30.07	4.225	4.51	**38.81**
	32.59	5.675	2.69	**40.96**
	33.92	6.875	6.97	**47.77**
	31.23	4.134	4.67	**40.03**
Bluetooth (DH5)	11.91	22.34	14.89	**49.14**
	11.11	20.84	13.89	**45.84**
	10.47	19.63	13.09	**43.18**
	13.61	25.53	17.02	**56.16**
	10.80	20.27	13.51	**44.58**
	11.26	21.12	14.08	**46.45**
	11.13	20.87	13.91	**45.92**
	11.47	21.51	14.34	**47.32**

As can be seen in the figures the results from the experimental evaluation are quite close to the analytical value derived in the theoretical study. Typically the experimental values are a little bit over the expected result but this may be due to the fact that in the analytical study we assumed that the propagation and computation time could be neglected. Looking at the results obtained, it could be expected these times might have been taken into account in the theoretical study. Nevertheless, the modelling of the computational

time would be beyond the scope of this study, and the accuracy of the analyses done has been proven in light of the results obtained in the experimental measurements.

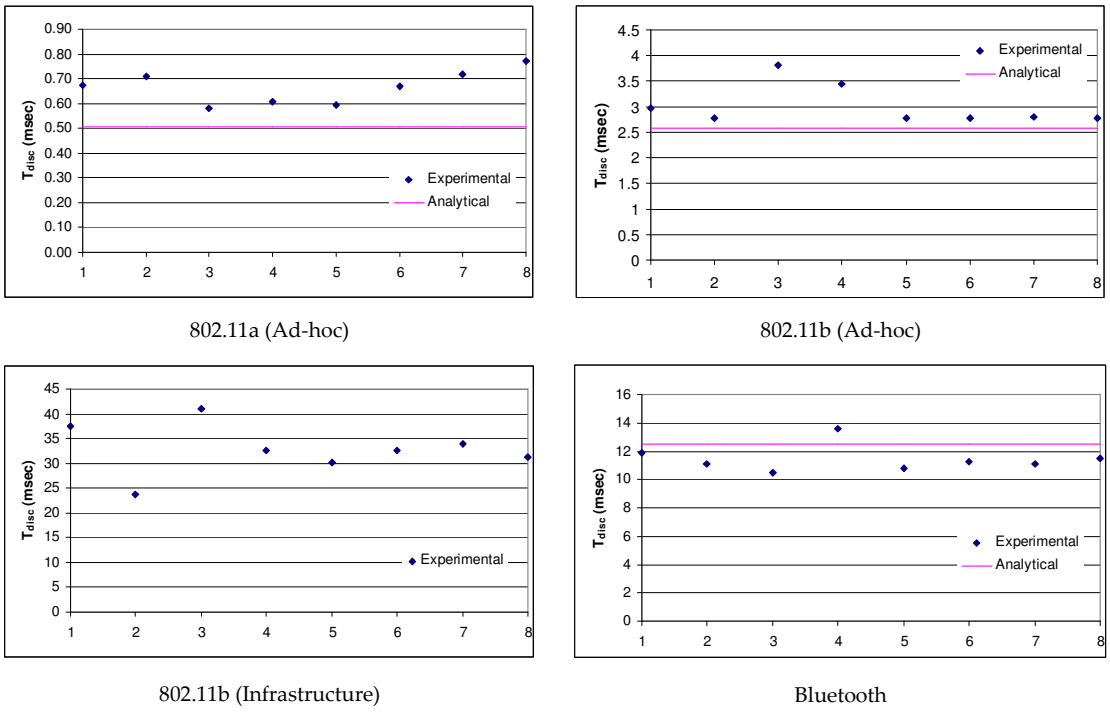


Figure 5-3: Discovery time results in the experimental platform.

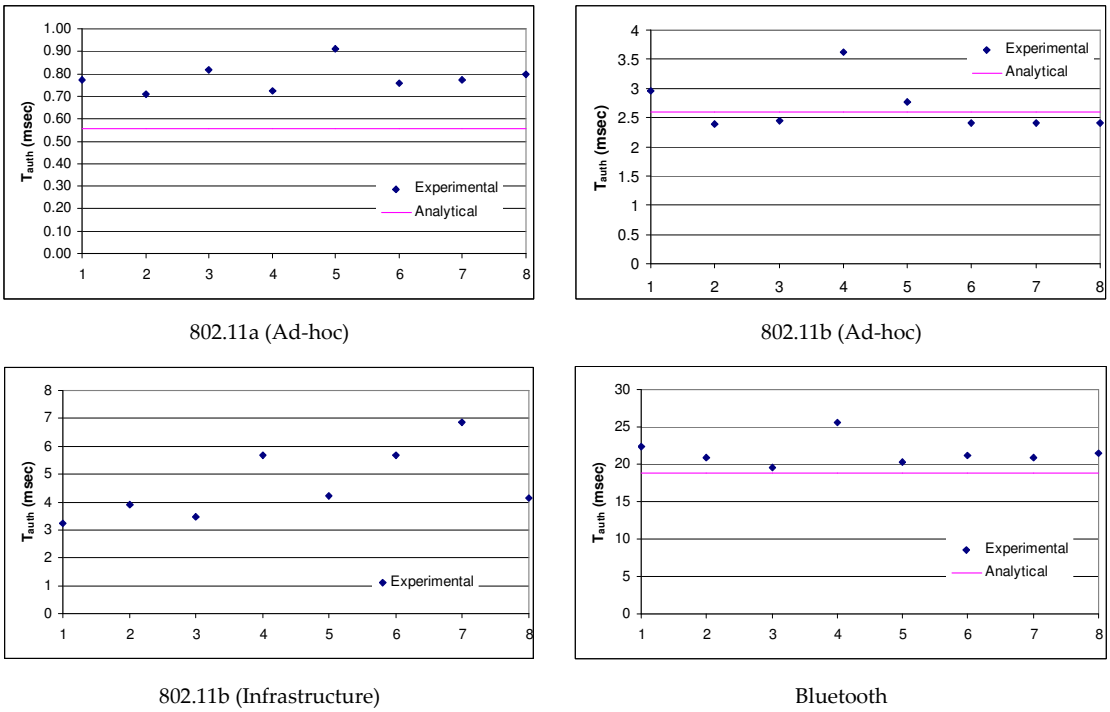


Figure 5-4: Authentication delay results in the experimental platform.

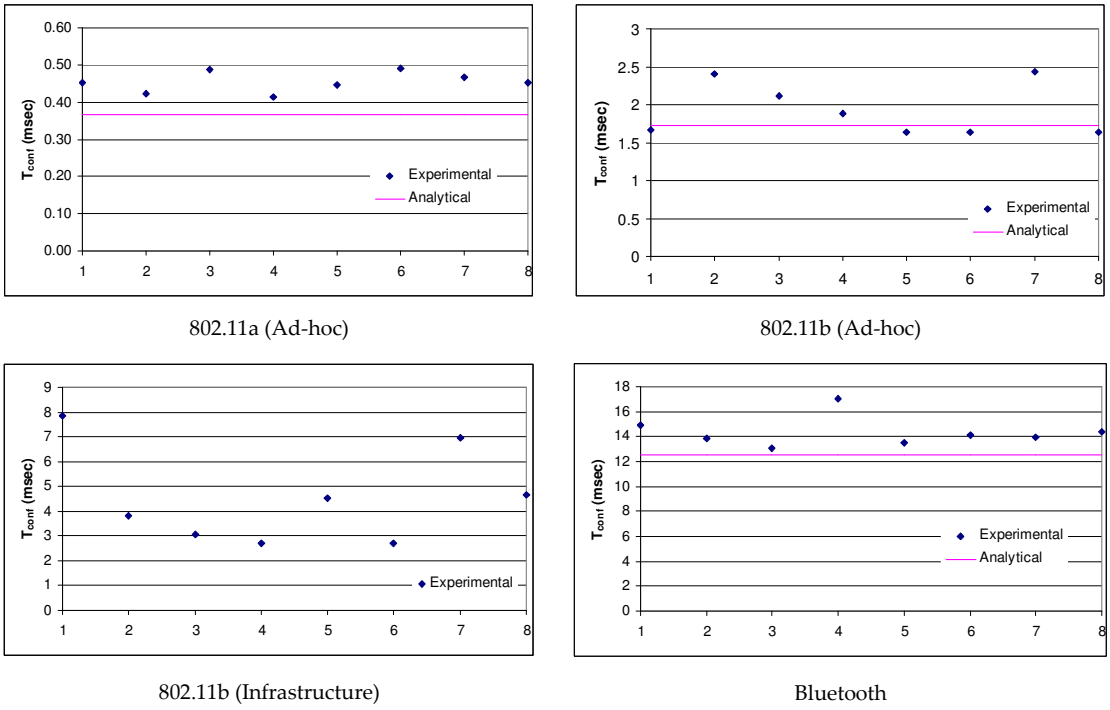


Figure 5-5: Configuration exchange time results in the experimental platform.

5.2.3 Comparison of results. Analytical vs Experimental

Figure 5-6, Figure 5-7 and Figure 5-8 show the comparison of the results obtained from the theoretical analyses and the ones obtained from the experimental measurement campaign. The comparison has been carried out for the three different wireless access technologies studied.

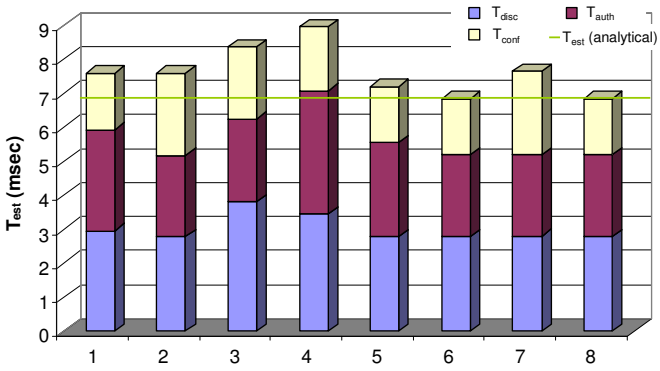


Figure 5-6: Secure link establishment time comparison (analytical vs experimental) – 802.11b (Ad-hoc)

As can be seen, the results obtained through the test platform slightly differ from the analytically obtained. Nevertheless, this difference, as already explained, is due to some of the assumptions made during the analytical modelling of the secure link establishment delay.

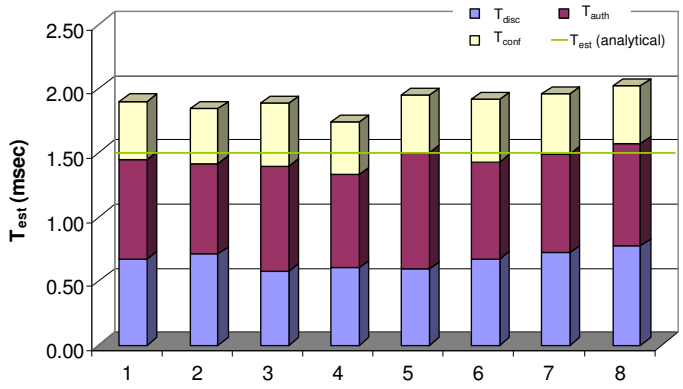


Figure 5-7: Secure link establishment time comparison (analytical vs experimental) – 802.11a

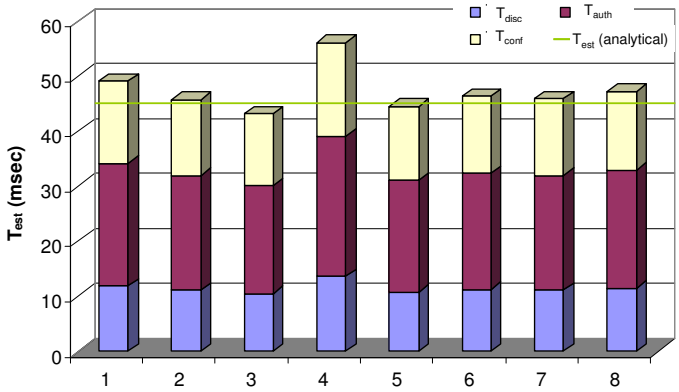


Figure 5-8: Secure link establishment time comparison (analytical vs experimental) – Bluetooth

Each of the bars on Figure 5-6, Figure 5-7 and Figure 5-8 showthe experimental result from a different test. As can be seen the delay introduced is more or less equally due to the link establishment steps although T_{auth} is typically higher than the other two. In any case it is clear that the delay introduced by the secure link establishment procedure is negligible, above all if we consider the critical system feature provided through this process.

5.2.4 Effect of the background traffic

The results presented in the previous sections were obtained under no background traffic conditions (i.e. both peers were idle when they met each other). Nevertheless, it is important to characterize the behaviour of the system when one of the peers is involved in another communication or when the radio domain is being used.

As shown in Figure 5-2, a third node was introduced and the communication was established with the first of the peers involved in the secure link establishment process. This way we can measure how the procedure is affected by this situation. Table 5-6 and Table 5-7 show the mean results obtained from the different conditions measured.

Table 5-6: Experimental results of UDP background traffic effect

	T_{disc} (msec)	T_{auth} (msec)	T_{conf} (msec)	**T_{est} (msec)**
IEEE 802.11a	21,10	24.12	15,07	**60,29**
IEEE 802.11b (ad-hoc)	192,61	220,12	137,58	**550,31**
IEEE 802.11b (infrastructure)	522,70	597,38	373,36	**1493,44**

Each test was repeated eight times in order to assess the behaviour of the secure link establishment process under background traffic conditions. As can be seen, the impact of background traffic sharply increases the secure link establishment time. This effect is due to the fact that in these tests, the packets exchanged for the link establishment process had to compete for the channel with the UDP and TCP connections. In this sense, the effect of the parallel TCP session was smaller than its UDP counterpart. This is probably due to the way packets are managed internally in the nodes. While TCP congestion control procedures do not overload the nodes' transmission internal buffers so much, in the UDP case the source's buffer is loaded without any flow control so secure link establishment related packets find it more difficult to get served.

Table 5-7: Experimental results of TCP background traffic effect

	T_{disc} (msec)	T_{auth} (msec)	T_{conf} (msec)	**T_{est} (msec)**
IEEE 802.11a	7,43	8,49	5,31	**21,23**
IEEE 802.11b (ad-hoc)	55,22	63,11	39,44	**157,77**
IEEE 802.11b (infrastructure)	423,67	484,20	302,62	**1210,49**

Moreover, it is interesting to note that the infrastructure mode is even more affected than the ad-hoc case. In any case, even under these difficult conditions, the secure link establishment time does not represent a cumbersome step and would not have deep impact in the overall system performance under real-world conditions.

5.2.5 Conclusions

A thorough measurement campaign has been carried out in order to assess the performance of the implemented system in terms of the secure link establishment delay. The results obtained have been used to support the analytical study developed and the comparison made has shown that the implementation behaves as would be expected.

As a general conclusion from the results shown, it can be stated that the complete secure link establishment process does not represent a task that would affect the overall system performance. Additionally, the measurement campaign carried out proves the functionalities of the components implemented. In this sense, different wireless access technologies have been used in order to assess the support of heterogeneity and the impact of the access technology in the process behaviour. Besides, the effect of background traffic

has also been studied so that conclusions can be derived for real-world scenarios where cluster formation takes place in conjunction with existing communications.

From the results obtained we can conclude that the secure link establishment process designed and implemented does not only fulfil perfectly the requirements imposed on Personal Network cluster formation, but it has shown that its small footprint on the overall system performance also makes it viable for the formation of private ad-hoc networks of any kind, from sensor networks to more advanced devices.

Another important feature of the implemented secure link establishment procedure is that it can feed this information to the routing protocol that determines the routes in the cluster. The required modifications to any routing protocol that is placed on top mainly affect the way the routing protocol gets information about neighbouring nodes. For example, while in the original AODV [151] this is handled by unauthenticated Hello messages, which can be replaced by the result of the neighbour discovery and authentication procedure so that only trusted neighbours are fed into the routing protocol. Further modifications might be possible in which the same routing protocol could be used to route both secure and unsecure communications. In this case, it would be necessary to differentiate the routing tables from both networks. The routing protocol would not need to take care of security considerations since all the security checks (integrity assurance, impersonation defence, etc.) are done below, at UCL level.

5.3 LINK BREAK DETECTION

Figure 5-9 shows the procedure by which a peer discovers that one of its neighbours has disappeared (e.g. switch off, exits radio domain, etc.).

As can be seen, the breakage of a link is detected upon the loss of three consecutive beacons. The wireless medium is a highly variant environment where overload or fading effects can lead to the loss of some packets even in favourable channel circumstances. Hence, during the specification a trade-off between quick detection and avoidance of overhead was agreed in order to set the number of consecutive beacons that must be lost before the link is said to be broken.

It has to be noted that any packet coming (e.g. actual user data traffic, etc.) from the neighbour resets the counter to zero. Hence, the link break detection time ranges between four times the inter-beacon interval and three times the inter-beacon interval.

Functionality tests were carried out to assess the correct behaviour of the implemented solution. Tests were performed using different wireless technologies (i.e. IEEE 802.11a/b/g and Bluetooth) and in varying channel conditions (i.e. from high observed SNR to lower ones). In all the cases the behaviour of the system was as expected with no abnormal situations occurring. No further performance tests were performed since this is a system feature that does not exhibit performance results.

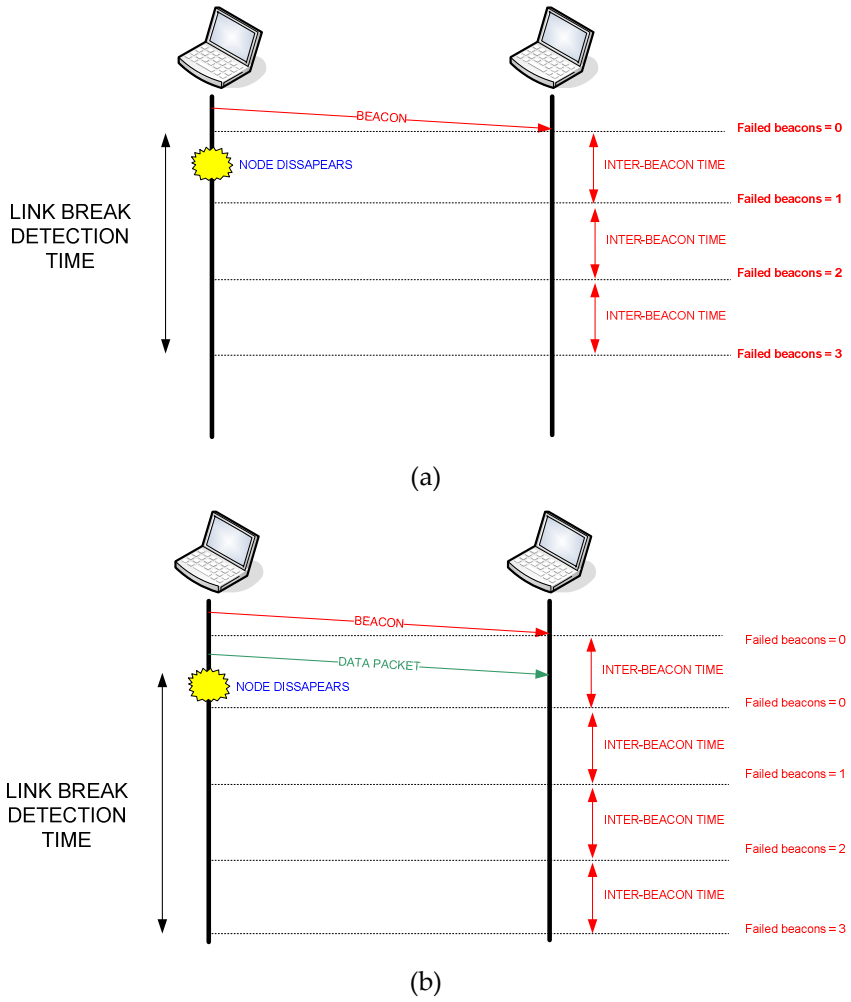


Figure 5-9: Link break detection procedure. Without background traffic (a) and with background traffic (b)

5.4 SECURE LINK USAGE

From a performance point of view, this is a link-level system feature that requires a more careful study. When traffic is exchanged between personal nodes at cluster level, additional mechanisms are enforced at the Universal Convergence Layer [152]. These mechanisms, basically cryptographic processing, protect the packet so that privacy, origin authentication and integrity, among other issues, are assured.

As is shown in Figure 5-10 the communication between two personal neighbours is protected through the encryption of the complete MAC frame payload (i.e. the complete IP datagram including the IP headers). Additionally a cryptographic signature is added to the packet in order to assure the integrity of the packet.

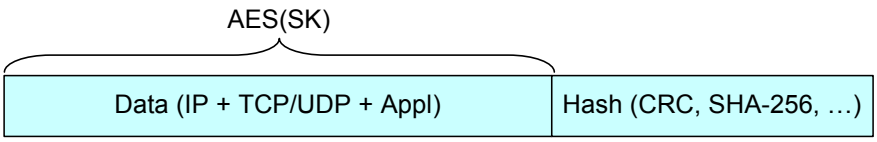


Figure 5-10: Data PDU format

Our communication architecture is based on pair-wise trust relationships. Every pair of personal nodes shares a long-term trust relationship that is enforced when they communicate with each other. As has been described in the Secure Link Establishment sections, when two personal nodes meet they authenticate each other and exchange link-level session keys that are used to secure that particular link. These session keys are used to encrypt the IP datagram, using AES algorithm [127], and to securely sign the packet, using SHA-256. This way, only the counterpart neighbour is able to decrypt the information and verify the signature of the packet.

On multihop scenarios at cluster level, the end-to-end security is assured by securing each of the links of the communication. By definition, all the nodes in a cluster are personal, so the packet is protected by the security of each of the links that forms the end-to-end route. The counterpart is that the packet has to be encrypted and decrypted in every link of the route with the additional overhead that this implies.

In order to assess the performance of this system feature, different evaluation models and tests have been performed. The evaluation and assessment procedure followed consisted of both the development of analytical models and the performance of a set of experimental tests. Throughput assessment analyses and measurements as well as comparison with competing solutions have focused the performance studies. The following sections describe the results of these studies.

5.4.1 Analytical assessment of the secure link usage

First of all let's compare the theoretical throughput that can be achieved in the case that the link is not secured (traditional transmission) with the one that is obtained when security mechanisms are applied. For the first analysis a single link with no packet loss is assumed. Different wireless technologies are analyzed and both TCP and UDP traffic are studied.

$$Throughput = \frac{number\ of\ bytes}{T_{trans}} \tag{5.8}$$

The factor in (5.8) that is dependent on the approach taken is T_{trans}. A thorough analysis has been carried out in order to assess the time it takes to transmit a certain number of bytes and correspondingly the throughput achieved.

First of all, the basic transmission time, $T_{trans\text{-}*}$, that is, the one that is obtained when standard transmission is done was derived. Taking into account the specification of the different wireless technologies analyzed [153][154][155] the results obtained are shown in Table 5-8.

In (5.9) the modelled transmission time for the UDP case over 802.11 links is shown:

$$\begin{aligned} &T_{trans-802.11}(\#bytes, bitrate) = T_{acc-Data} + T_{TX-Data}(\#bytes, bitrate) + T_{ACK-802.11} \\ &T_{acc-Data} = DIFS + BO + PLCP \\ &T_{ACK-802.11} = SIFS + PLCP + ACK_{802.11}(bitrate) \end{aligned} \tag{5.9}$$

where the following notation is used:

T_X(# bytes, bitrate): Time as a function of the number of bytes and the bitrate.

DIFS: DCF interframe space.

BO: the back-off time that the 802.11 stations wait before transmitting.

PLCP: MAC preamble.

$T_{TX\text{-}Data}$: Is the time for transmitting the 802.11 frame. It is dependent on the MAC payload size and on the selected bitrate.

SIFS: Short interframe space.

$ACK_{802.11}$: the time to transmit an 802.11 level acknowledgement. It is dependent on the selected bitrate.

The different parameters in (5.9) are correspondingly substituted for their values. These values are different in 802.11a and 802.11b cases, but the expression remains the same.

In (5.10) the same expression is derived for the Bluetooth case:

$$T_{trans-BT}(\#bytes, bitrate) = T_{TX-Data}(\#bytes, bitrate) \tag{5.10}$$

In (5.11) and (5.12) the T_{trans} is derived for the TCP case.

$$\begin{aligned} T_{trans-802.11} = N \cdot \left(T_{acc-Data} + T_{TX-Data}\right) + \left(T_{acc-Data} + T_{TX-ACK-TCP}(bitrate)\right) \\ + \left(N+1\right) \cdot T_{ACK-802.11} \end{aligned} \tag{5.11}$$

where the following notation is used:

$T_{TX\text{-}ACK\text{-}TCP}$(bitrate): Is the time for transmitting the TCP level acknowledgement. It is dependent on the selected bitrate.

$$T_{trans-BT} = N \cdot T_{TX-Data} + T_{TX-ACK-TCP}(bitrate) \tag{5.12}$$

In these cases, we are considering a complete cycle of the TCP congestion window where N data packets are transmitted per TCP acknowledgement.

Table 5-8: Maximum achievable throughput over different wireless technologies

	TCP T_{trans} (μs) (N·Data + ACK)	**Tput (Mbps)**	UDP T_{trans} (μs)	**Tput (Mbps)**
IEEE 802.11a (54 Mbps)	2145	**27**	393,5	**29,93**
IEEE 802.11b (11 Mbps)	10245,09	**5,65**	1884	**6,25**
Bluetooth (DH5)	101250	**0,572**	18750	**0,628**

The access technologies under study can work at several bitrates but only the maximum rate has been analyzed for the sake of space. Nevertheless, similar analyses for the rest of available bitrates can easily be done. The study has been done assuming that 1500-byte MAC frames are sent. In this sense, although the MAC level throughput is higher than the one in Table 5-8, at user level, the exposed throughput is the presented one. Besides, when analyzing the TCP case, congestion window behaviour was emulated by allowing N data

packets per TCP acknowledgement. The bigger the number of packets allowed in the window the lower the overhead introduced by the TCP protocol itself. In our analysis, N was fixed at 5. The T_{trans} on the TCP case in (4) represents a complete cycle; that is, 5 TCP data packets and 1 TCP ACK. The throughput is calculated accordingly as 5 times the TCP payload (i.e. the total MAC frame length minus the IP and TCP headers' length) divided by the total T_{trans}. For the UDP case, the calculation is easier since the protocol is simpler and the throughput is calculated as the UDP payload (i.e. the total MAC frame length minus the IP and UDP headers' length) divided by the total T_{trans}.

It is interesting to note that the results obtained for the Bluetooth maximum achievable throughput have been derived taking into account a payload offered to the MAC layer of 1500 bytes. Thus, the results vary from the ones usually managed [156] as they count on optimally filled Bluetooth MAC frames (i.e. no need for fragmentation of the packet). For the TCP case, the assumption made when modelling the Bluetooth performance was that both peers were using asymmetric connection operation mode where the server was occupying five slots (for the downstream traffic) while the client has allocated one (for the TCP acknowledgements). The TCP congestion window behaviour was modelled as described above.

Once the base throughput has been identified, let's analyze the overhead introduced by the different security additions that the system introduces for secure link usage. First of all, let's analyze the effect of the secure signature introduced.

In (5.13) and (5.14) the modelled T_{trans} for the UDP case is shown for both 802.11 and Bluetooth.

$$T_{trans-802.11} = T_{acc-Data} + T_{TX-Data} + T_{SHA-256}(\#bytes, hash\ rate) + T_{ACK-802.11} \tag{5.13}$$

$$T_{trans-BT} = T_{TX-Data} + T_{SHA-256}(\#bytes, hash\ rate) \tag{5.14}$$

where the following notation is used:

> **$T_{SHA-256}$(#bytes, hash rate)**: Is the time for generating the signature. It is dependent on the size of the block from where the signature is derived and on the rate at which the hash is generated.

In (5.15) and (5.16) the modelled T_{trans} for TCP case is presented for both 802.11 and Bluetooth.

$$\begin{aligned} T_{trans-802.11} = {} & N \cdot \left(T_{acc-Data} + T_{TX-Data} + T_{SHA-256-Data}\right) + \\ & + T_{acc-Data} + T'_{TX-ACK-TCP} + T_{SHA-256-ACK-TCP} + \\ & + \left(N+1\right) \cdot \left(T_{ACK-802.11}\right) \end{aligned} \tag{5.15}$$

$$T_{trans-BT} = N \cdot \left(T_{TX-Data} + T_{SHA-256-Data}\right) + T'_{TX-ACK-TCP} + T_{SHA-256-ACK-TCP} \tag{5.16}$$

where the transmission time for the beacon has to take into account the addition of the 32-byte signature overhead.

The overhead introduced in all these cases is due to both the 32 bytes attached to the IP datagram before sending it to the MAC level and to the need to calculate the actual signature. Although the SHA-256 function is quite fast, the impact of the hash calculations has to be taken into account. Table 5-9 shows the results for these analyses when fixing the maximum binary rate in the different access technologies.

As can be seen, in general the reduction in terms of throughput is between 2.6 and 4.1%. The main difference between the UDP and TCP cases is that in the TCP case, the TCP acknowledgement is also signed so that more overhead is included. In principle, the impact of the signature does not seem relevant. The different values of overhead are due to different factors. TCP is in general more affected due to the fact that actual data size is less than in the UDP case (TCP header is bigger and MAC payload is fixed to 1500 bytes) and because TCP acknowledgements are also influenced. When comparing the results between the different wireless technologies studied, the overhead is bigger the faster the technology is. This is due to the fact that the time it takes to do the signature process become comparable with the actual transmission time, hence the impact is bigger.

Table 5-9: Throughput over different wireless technologies considering UCL signature overhead

	TCP T_{trans} (μs) (N·Data + ACK)	**Tput (Mbps)**	**Overhead (%)**	UDP T_{trans} (μs)	**Tput (Mbps)**	**Overhead (%)**
IEEE 802.11a (54 Mbps)	2185,55	**25,92**	**4,02 %**	400,74	**28,75**	**3,94 %**
IEEE 802.11b (11 Mbps)	10304,91	**5,5**	**2,78 %**	1891,24	**6,09**	**2,55 %**
Bluetooth (DH5)	101287,26	**0,572**	**0,04 %**	18757,24	**0,614**	**2,21 %**

The benefits obtained through the adoption of the secure signature are well worth the small throughput reduction introduced. Nevertheless, assuring the origin authentication and integrity of the communication is not enough, but it is necessary to completely protect user communication and offer full privacy. It is then necessary to encrypt the communications using the link-layer session keys derived during the secure link establishment procedure.

In (5.17) and (5.18) the modelled T_{trans} for the UDP case is presented for both 802.11 and Bluetooth.

$$T_{trans-802.11} = T_{acc-Data} + T_{TX-Data} + T_{SHA-256} + T_{AES}(\#bytes, hash\ rate) + T_{ACK-802.11} \quad (5.17)$$

where the following notation is used:

> **T_{AES}**: Is the time for encrypting the IP datagram. It is dependent on the size of the block to be encrypted and on the rate at which the hash is generated. It is important to note that only the IP datagram is encrypted not the full MAC frame.

$$T_{trans-BT} = T_{TX-Data} + T_{SHA-256} + T_{AES}(\#bytes, hash\ rate) \quad (5.18)$$

In (5.19) and (5.20) the modelled T_{trans} for the TCP case is presented for both 802.11 and Bluetooth.

$$\begin{aligned} T_{trans-802.11} = {} & N \cdot \left(T_{acc-Data} + T_{TX-Data} + T_{SHA-256-Data} + T_{AES-Data}\right) + \\ & + T_{acc-Data} + T'_{TX-ACK-TCP} + T_{SHA-256-ACK-TCP} + T_{AES-ACK-TCP} + \\ & + \left(N+1\right) \cdot \left(T_{ACK-802.11}\right) \end{aligned} \quad (5.19)$$

$$\begin{aligned} T_{trans-BT} = {} & N \cdot \left(T_{TX-Data} + T_{SHA-256-Data} + T_{AES-Data}\right) + \\ & + T'_{TX-ACK-TCP} + T_{SHA-256-ACK-TCP} + T_{AES-ACK-TCP} \end{aligned} \quad (5.20)$$

As summarized in Table 5-10 the effect of the traffic encryption is appreciable and might result in around 10 % overhead when working with very high data rate technologies. On the contrary, when the binary rate of the wireless technology decreases, the overhead is greatly reduced. This is basically due to the fact that although the security mechanisms implemented perform at very high rates, the transmission time becomes comparable with the encryption time for faster access technologies. In any case, the security benefits that the encryption brings clearly compensate this reduction in the overall throughput.

Table 5-10: Throughput over different wireless technologies considering UCL encryption overhead

	TCP T_{trans} (μs) (N·Data + ACK)	**Tput (Mbps)**	**Overhead (%)**	UDP T_{trans} (μs)	**Tput (Mbps)**	**Overhead (%)**
IEEE 802.11a (54 Mbps)	2331,75	**24,29**	**10,04 %**	429,7	**26,81**	**10,42 %**
IEEE 802.11b (11 Mbps)	10451,11	**5,42**	**4,14 %**	1920,2	**5,99**	**4,02 %**
Bluetooth (DH5)	101436,3	**0,558**	**2,39 %**	18786,2	**0,613**	**2,36 %**

Looking at the results of the analyses done, it might be thought that the system performance is severely degraded, but it is important to take into account that the subjective benefits are difficult to measure in order to make objective comparisons. Nevertheless, we can compare the solution adopted with the most adopted solution for securing Internet communication, namely IPSec tunnelling [97].

In [157] an experimental study of the IPSec Linux OS implementation is described. Based on the results from this study we derived the encryption and decryption time for a set-up similar to the one we used for deriving the UCL delay for encrypting and decrypting each byte in a 1400-byte block. The result highlighted that it takes approximately 0.07 μs per

byte to perform AES-128 and SHA-1 processes in IPSec Linux implementation for a Pentium IV based laptop.

Table 5-11 summarizes the analytical results of the maximum achievable throughput using IPSec over a link of different wireless technologies.

Table 5-11: Throughput over different wireless technologies considering IPSec tunnelling overhead

	TCP T_{trans} (μs) (N·Data + ACK)	**Tput (Mbps)**	**Overhead (%)**	UDP T_{trans} (μs)	**Tput (Mbps)**	**Overhead (%)**
IEEE 802.11a (54 Mbps)	2660,5	**20,84**	**22,83 %**	494,02	**22,89**	**23,49 %**
IEEE 802.11b (11 Mbps)	10797,68	**5,13**	**9,18 %**	1984,52	**5,7**	**8,81 %**
Bluetooth (DH5)	101350,5	**0,55**	**1,76 %**	18850,52	**0,6**	**2,14 %**

As can be seen the results obtained with our approach that benefits from a faster cryptography implementation and smaller header overhead outperforms the IPSec one. IPSec can suffer a reduction of a quarter of the maximum achievable throughput in the worst case doubling the impact on the performance of all the access technologies studied. It is important to note that the results shown in Table 5-11 are derived without considering the Internet Key Exchange (IKE) protocol [100] that precedes every IPSec session. This would have a negative impact on the figures presented. Nevertheless, we preferred not to include this in the modelling for the sake of simplicity. The model would then be perfectly valid for the case where long data transmissions follow the key exchange so that the key exchange overhead can be neglected without incurring in big error. When the analyses of the approach taken in the project (Table 5-10) were carried out, the secure link establishment delay was not considered either in the light of the aforementioned. However, it is important to clarify that the secure link establishment process takes place the moment two nodes meet (i.e. they hear each other's beacon) and is valid for a random period of time of the order of several minutes while the IKE process has to take place just before any IPSec session can be started. This difference between proactive versus reactive approach would increase the overhead in the case of IPSec more than in the approach presented in this thesis if these two phases were considered.

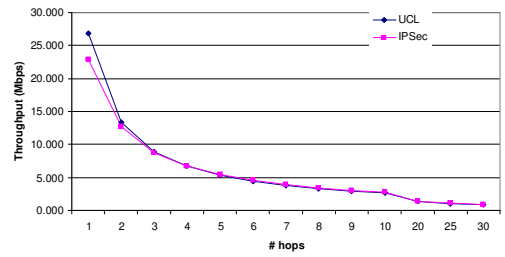
However, the IPSec approach has yet another difference as well as the one described. While in the UCL the security is carried out in a pair-wise basis, IPSec does this on an end-to-end basis. When it is a one hop link, there is no difference but in multihop scenarios IPSec might perform better since encryption and decryption is done only at the communication ends while the UCL way of encryption mandates a hop-by-hop encryption and decryption.

The results of the comparison in multihop scenarios are summarized in Table 5-12

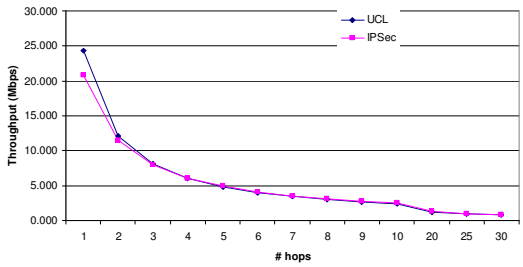
Table 5-12: Analytical results of comparison of secure communication approaches in multihop scenarios

	# of hops	TCP		UDP	
		UCL Throughput (Mbps)	IPSec Throughput Mbps)	UCL Throughput (Mbps)	IPSec Throughput (Mbps)
IEEE 802.11a (54 Mbps)	2	**12,36**	**11,52**	**13,41**	**12,75**
	4	**6,18**	**6,08**	**6,7**	**6,76**
	8	**3,09**	**3,13**	**3,35**	**3,48**
	10	**2,47**	**2,52**	**2,68**	**2,8**
	20	**1,24**	**1,27**	**1,34**	**1,42**
IEEE 802.11b (11 Mbps)	2	**2,73**	**2,63**	**3**	**2,92**
	4	**1,36**	**1,33**	**1,5**	**1,48**
	8	**0,68**	**0,67**	**0,75**	**0,746**
	10	**0,55**	**0,54**	**0,6**	**0,597**
	20	**0,27**	**0,27**	**0,3**	**0,299**
Bluetooth (DH5)	2	**0,279**	**0,274**	**0,307**	**0,301**
	4	**0,140**	**0,137**	**0,153**	**0,151**
	8	**0,07**	**0,069**	**0,077**	**0,075**
	10	**0,056**	**0,055**	**0,061**	**0,06**
	20	**0,028**	**0,027**	**0,031**	**0,03**

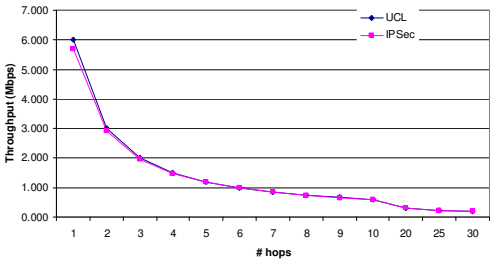
Figure 5-11 shows how the two approaches converge the more hops are considered. In the 802.11a case IPSec finally performs better than the UCL encryption when considering clusters greater than 4 hops wide. Nevertheless, it is important to note that this secure communication mechanism is targeted at cluster level so we should take into account the feasibility of a 4-hop wide cluster (i.e. 4 hops between two personal nodes only using other personal nodes as relays) and that the difference even when we are modelling large clusters is not that high.



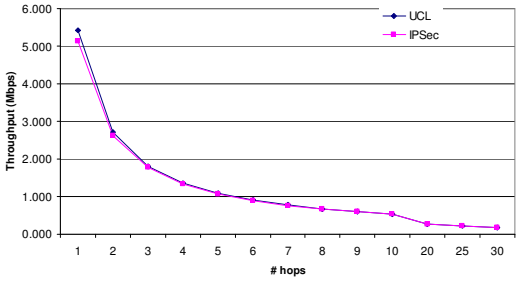
802.11a UDP



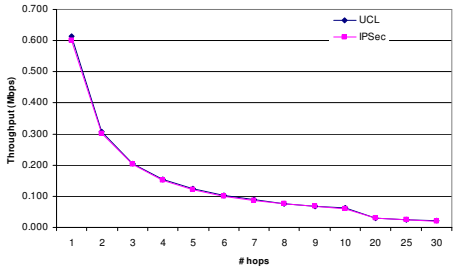
802.11a TCP



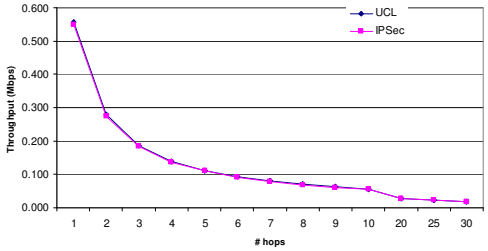
802.11b UDP



802.11b TCP



Bluetooth UDP



Bluetooth TCP

Figure 5-11: Throughput comparison in multihop scenarios

The relative difference in terms of throughput between the two approaches studied can be better seen in Figure 5-12. This difference has been calculated using the following expression:

$$RD = \left(1 - \frac{Tput_{IPSEC}}{Tput_{UCL}}\right) \times 100 \quad (\%) \tag{5.21}$$

When the relative difference values are above zero then the UCL performance is better than the IPSec one and vice versa.

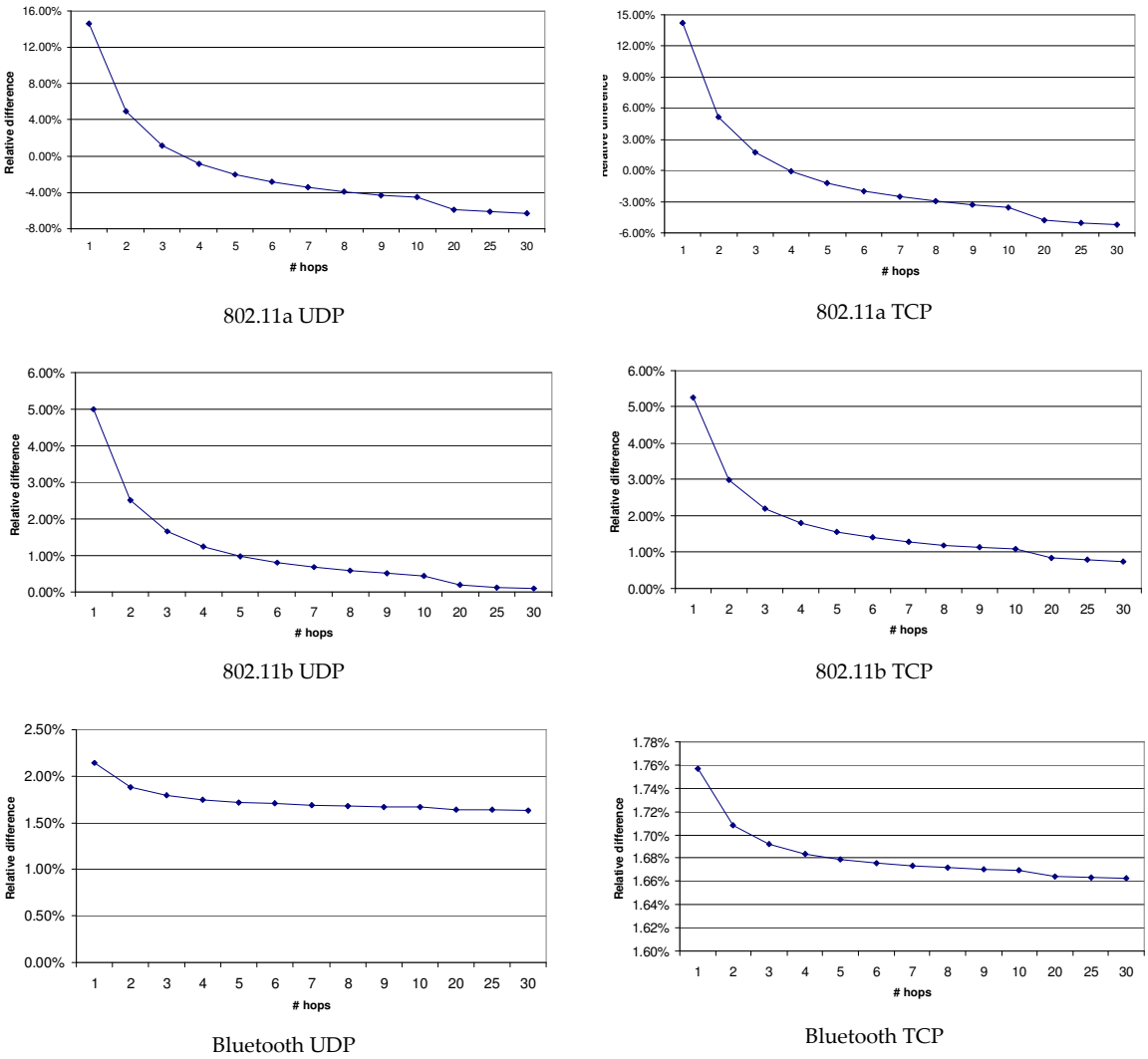


Figure 5-12: Relative difference between UCL encryption and IPSec in multihop scenarios

In the 802.11a case, when the number of hops is low the secure link usage approach taken performs better than the IPSec one. For large numbers of hops the overhead of encrypting and decrypting on a hop-by-hop basis makes UCL encryption behaviour worse. UDP traffic is more affected by the hop-by-hop approach, while in TCP the UCL-based secure link usage performs better even in a 5-hop wide cluster situation. When the binary rate is smaller, in IEEE 802.11b and Bluetooth, the impact of a slower encryption approach and a greater header overhead like the one imposed by the IPSec solution, means the UCL-based encryption is better even for large clusters (more than 20-hop wide).

5.4.2 Experimental assessment of secure link usage

Once the theoretical performance of the system has been evaluated, the real behaviour of the implemented prototype will be presented. Using the same testbed as described for the secure link establishment time, a measurement campaign was carried out to assess the performance of the system implemented. Both UDP and TCP traffic was characterized. For

the UDP case a bulk transfer of 10000 UDP packets was performed. For the TCP case, an FTP session was established to exchange a 10-Mbyte file. In both cases, the tests were carried out under ideal channel conditions and repeated 8 times.

The main objectives of this measurement campaign were to compare the results with the estimated performance derived in the previous section and to assess the behaviour of the implemented system components that are involved in this scenario.

5.4.2.1 One-hop UDP performance results

First of all we consider the system behaviour when UDP traffic is exchanged. Table 5-13 shows the mean value calculated over the different repetitions performed.

Table 5-13: Experimental UDP traffic throughput

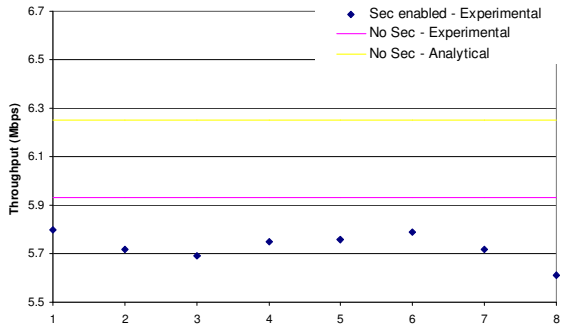
Technology	Throughput (Mbps)		Overhead
	No Sec	Sec enabled	
IEEE 802.11a	32,49	29,33	**9,73 %**
IEEE 802.11b (ad-hoc)	5,93	5,74	**3,30 %**
IEEE 802.11b (infra)	2,96	2,71	**8,57 %**
Bluetooth	0,634	0,533	**15,98 %**

In Table 5-13 the *No Sec* column corresponds to the value obtained when the UCL was not loaded into the system so it represents the raw maximum achievable throughput. The results obtained when the UCL is loaded and security techniques are applied are presented in the *Sec enabled* column.

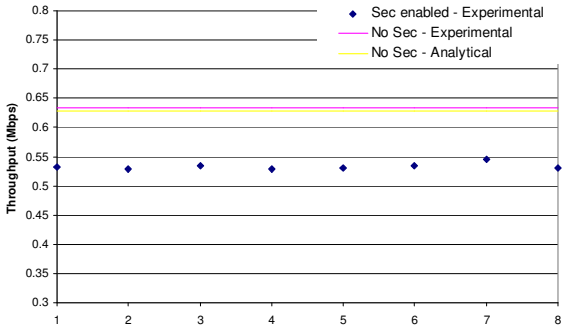
Taking into account these results and comparing them with the overhead calculated analytically, we can conclude that the implementation we are validating behaves as expected, at least in relative terms. The performance that we measured when applying the UCL security mechanisms is lowered by a similar percentage as was derived theoretically in the previous sections.

In absolute terms the values obtained in the experimental measurements depend on the particular components that we are using in the measurement campaign set-up. It is important to note that these components are off-the-shelf hardware whose behaviour might not completely fulfil the standard specification. Only in the Bluetooth case do the experimental results not follow the analytical study. These results oblige a thorough analysis of how Bluetooth connections are handled within the UCL implementation since additional computational overhead might be introduced unnecessarily.

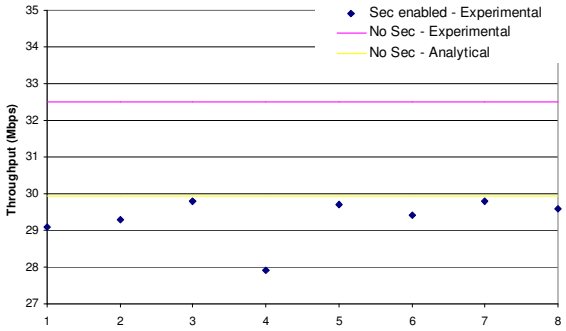
Figure 5-13 shows the results of the different tests and compares them with the analytical value and with the mean value derived from the experimental tests when no security was applied.



IEEE 802.11b (ad-hoc)



Bluetooth



IEEE 802.11a (ad-hoc)

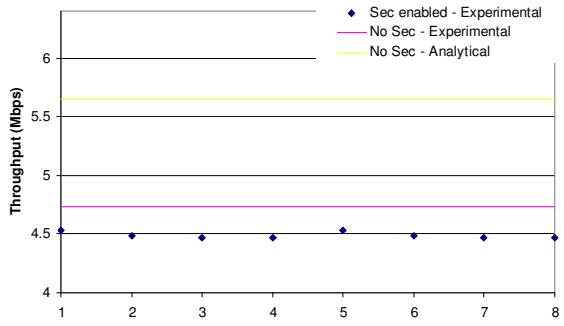
Figure 5-13: Experimental results for UDP traffic and comparison with analytical values

5.4.2.2 One-hop TCP performance results

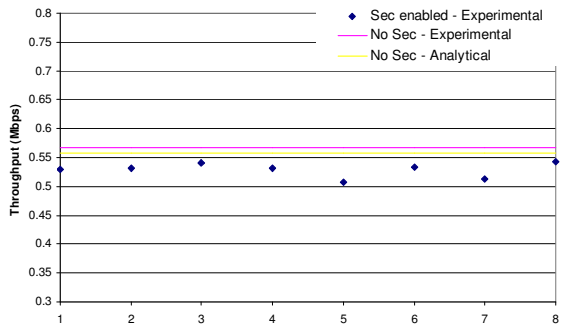
In the TCP case, as shown in Table 5-14, the values of overhead obtained through the experimental measurements are slightly higher than in the UDP case. This might be due to the way we have modelled the TCP window in the analytical case and the processing delays that we have neglected in the analytical study. It affects the TCP communications more due to its symmetric nature. The misalignment when measuring the throughput over Bluetooth links remains, as in the UDP case.

Table 5-14: Experimental TCP traffic throughput

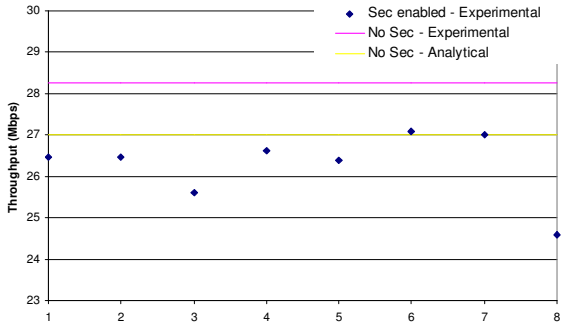
Technology	Throughput (Mbps)		Overhead
	No Sec	Sec enabled	
IEEE 802.11a	28,25	25,34	**10,30 %**
IEEE 802.11b (ad-hoc)	4,73	4,49	**5,19 %**
IEEE 802.11b (infra)	2,16	1,92	**10,89 %**
Bluetooth	0,57	0,53	**6,66 %**



IEEE 802.11b (ad-hoc)



Bluetooth



IEEE 802.11a (ad-hoc)

Figure 5-14: Experimental results for TCP traffic and comparison with analytical values

Figure 5-14 shows the results of the different tests and compares them with the analytical value and with the mean value derived from the experimental tests when no security was applied.

In the TCP case, the values of overhead obtained through the experimental measurements are slightly higher than in the UDP case. This might be due to the way we have modelled the TCP window in the analytical case and the processing delays that we have neglected in the analytical study. This affects the TCP communications more due to its symmetric nature.

5.4.2.3 Multihop scenarios performance results

As was done in the analytical study section, multihop scenarios have also been studied in the measurement campaign. PN clusters tend to be multihop wireless networks and it is interesting to evaluate the behaviour of the implemented solution in this kind of scenarios.

A number of laptops were placed following a string layout as is shown in Figure 5-15.

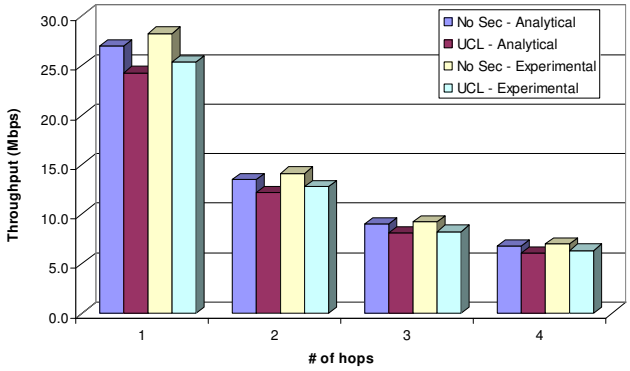
Figure 5-15: Experimental set-up for multihop scenario measurement campaign

All the laptops were located close to each other in order to guarantee perfect link conditions avoiding the effect of channel impairments. To force the multihop scenario, a MAC filtering mechanism was implemented so that the nodes that were supposed not to be neighbours did not see each other although they were in the same radio domain. Multihop routes were statically configured. Similarly to the single hop case, bulk 10000 UDP packets transfer was used to test this kind of traffic while an FTP session was used to generate the TCP traffic.

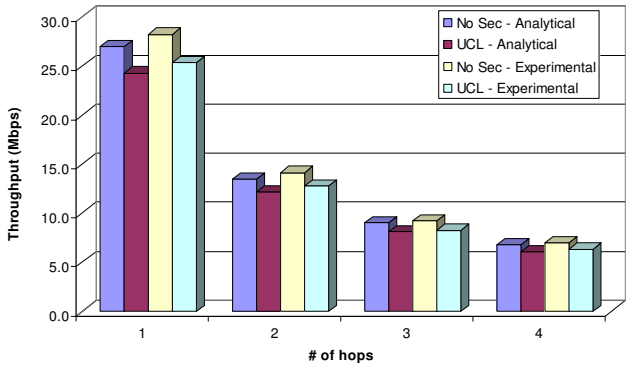
Table 5-15 and Table 5-16 summarizes the results (mean values are shown) obtained in the measurement campaign (mean values are shown) for multihop scenarios.

Table 5-15: Experimental UDP traffic throughput in multihop situations

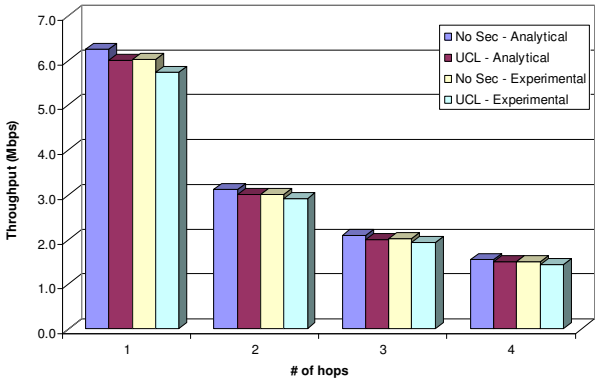
# of hops	Technology	Throughput (Mbps)		Overhead
		No Sec	Sec enabled	
2 hops	IEEE 802.11a	15,95	14,45	**9,43 %**
	IEEE 802.11b	3,01	2,92	**3,19 %**
3 hops	IEEE 802.11a	10,78	9,89	**8,29 %**
	IEEE 802.11b	2,02	1,94	**4,07 %**
4 hops	IEEE 802.11a	8,08	7,37	**8,69 %**
	IEEE 802.11b	1,52	1,45	**4,41 %**



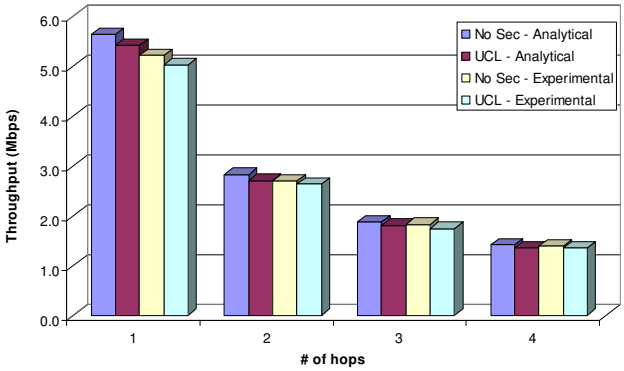
802.11a UDP



802.11a TCP



802.11b UDP



802.11b TCP

Figure 5-16: Throughput comparison for multihop scenarios. Analytical and experimental results

As can be derived from the overhead figures obtained, the behaviour of the experimental measurements corresponds with that expected. When comparing the overhead calculated analytically with the experimentally obtained one, we can observe that the performance measured when applying the UCL security mechanisms is lowered by a similar percentage as was derived theoretically. We see that it is around 4% for the 802.11b case (for both TCP and UDP traffic) and around 10% for the 802.11a case (for both TCP and UDP traffic). These are the values that were derived using the analytical model.

Figure 5-16 presents the comparison of the throughput obtained (mean results are represented) for different multihop configurations. The different approaches used (analytical and experimental) are presented together.

As can be seen, for each of the multihop configuration the analytical and experimental values obtained show similar values and what is more important, experimental and analytical evaluations show that the overhead introduced when applying the UCL security mechanisms is approximately the same. The coincidence in the relative difference supports the results obtained on both analyses.

Table 5-16: Experimental TCP traffic throughput in multihop situations

# of hops	Technology	Throughput (Mbps)		Overhead
		No Sec	Sec enabled	
2 hops	IEEE 802.11a	14,14	12,73	**10,02 %**
	IEEE 802.11b	2,7	2,64	**2,26 %**
3 hops	IEEE 802.11a	9,17	8,26	**9,97 %**
	IEEE 802.11b	1,83	1,75	**4,54 %**
4 hops	IEEE 802.11a	7,02	6,33	**9,83 %**
	IEEE 802.11b	1,39	1,36	**2,87 %**

5.4.2.4 Comparison with IPSec solution

As was done analytically, a measurement campaign of the IPSec solution has been carried out using the same testbed used to assess the UCL solution. The aim of this test is to be able to compare experimentally the performance of the UCL and IPSec solutions.

For these tests, we configured the different laptops following the same string layout as in the previous sections and established IPSec tunnels (using ESP in tunnel mode) between the communicating laptops. For multihop configurations, intermediate nodes were simply used to route the packets while the cryptographic tasks were only performed at the communication endpoints.

Table 5-17: Experimental UDP traffic throughput in multihop situations using IPSec

# hops	Technology	Throughput (Mbps)		Overhead
		No Sec	Sec enabled	
1 hop	IEEE 802.11a	32,49	28,84	**11,25 %**
	IEEE 802.11b	6,03	5,52	**8,39 %**
2 hops	IEEE 802.11a	15,95	14,38	**9,85 %**
	IEEE 802.11b	3,01	2,87	**4,63 %**
3 hops	IEEE 802.11a	10,78	10,04	**6,89 %**
	IEEE 802.11b	2,02	1,93	**4,43 %**
4 hops	IEEE 802.11a	8,08	7,51	**7 %**
	IEEE 802.11b	1,52	1,49	**1,56 %**

Table 5-17 and Table 5-18 present the mean values obtained through the experimental measurement campaign of the IPSec solution. When comparing with the analytically derived results we see that the tendency is maintained and the more hops are considered the less the impact the IPSec solution has in the communication. This is the expected behaviour since the encryption and decryption process is only done at the communication endpoints and the longer the communication is, the greater the transmission time and the smaller the influence of the encryption and decryption process.

Table 5-18: Experimental TCP traffic throughput in multihop situations using IPSec

# hops	Technology	Throughput (Mbps)		Overhead
		No Sec	Sec enabled	
1 hop	IEEE 802.11a	28,25	24,84	**12,06 %**
	IEEE 802.11b	5,23	4,95	**5,28 %**
2 hops	IEEE 802.11a	14,14	12,57	**11,11 %**
	IEEE 802.11b	2,7	2,58	**4,52 %**
3 hops	IEEE 802.11a	9,17	8,42	**8,21 %**
	IEEE 802.11b	1,83	1,76	**3,77 %**
4 hops	IEEE 802.11a	7,02	6,47	**7,74 %**
	IEEE 802.11b	1,39	1,38	**1,43 %**

In absolute terms, the values differ since the analytical model assumed some parameters that might not exactly correspond with the ones in the actual Linux implementation used for the tests. Even in the testbed there were different laptops that might work slightly

differently. Nevertheless, this issue does not reduce the validity of either the analytical or the experimental assessment of the system since both maintains the same trend when the number of hops increases.

Once the correctness of the measurements for the IPSec solution has been assessed, we compared them with the results obtained from the UCL experimental validation.

Figure 5-17 shows the relationship between the throughput obtained with the IPSec and UCL solutions depending on the number of hops. On the left axis the UCL (blue-diamond line) and the IPSec (red-asterisk line) throughput are shown. As can be appreciated the difference between them is hardly noticeable. To shed some light on the actual relationship between the two solutions, on the right axis the overhead introduced by the IPSec with respect to the UCL solution is represented (dashed-square line). As can be seen, as long as the number of hops increases the overhead introduced by the IPSec reduces. This means that the Linux IPSec implementation actually outperforms the UCL implementation. Nevertheless, as already pointed out, even for quite large clusters the difference between the two solutions is small.

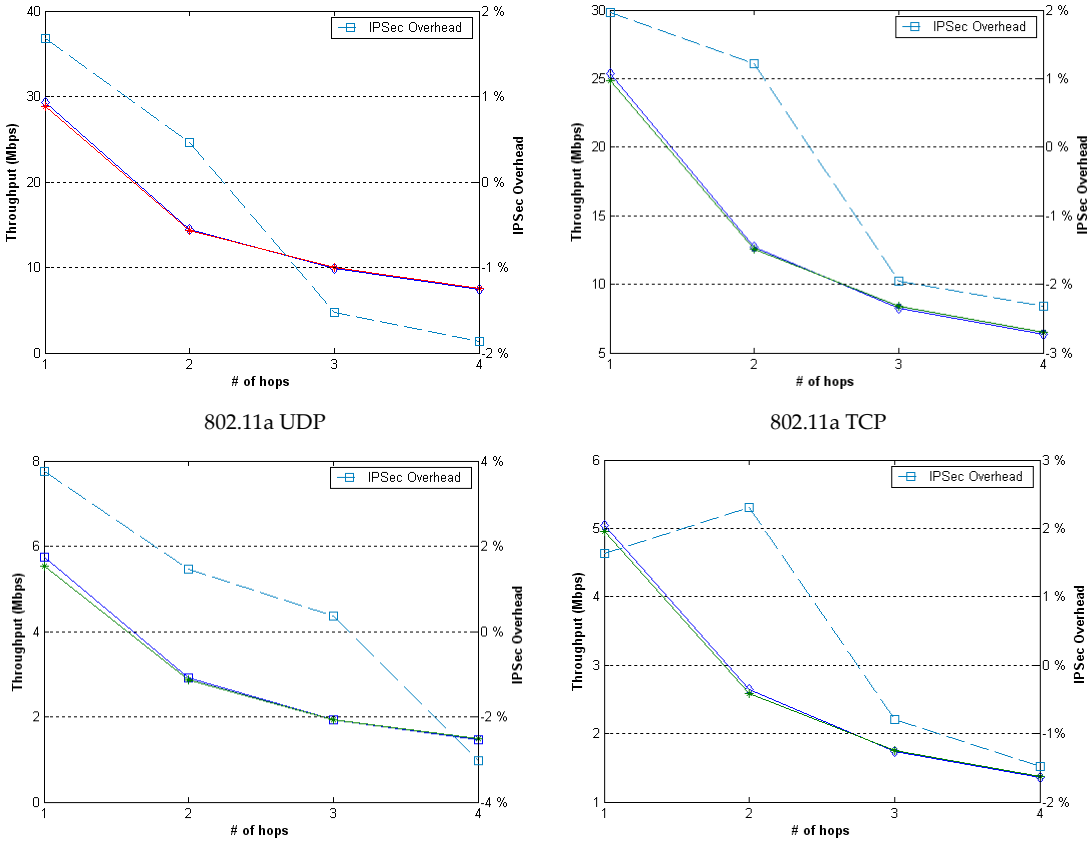


Figure 5-17: UCL vs IPSec throughput comparison for multihop scenarios

5.4.3 Conclusions

The evaluation that has been presented supports the approach taken for securing the communications in multihop wireless clusters of personal devices proving at the same

time that the implementation done fulfils the requirements imposed. By comparing the approach taken with IPSec, we have proven that it not only presents comparable performance in large clusters but it outperforms IPSec when typical small and dynamic clusters are considered.

As has been shown the results obtained with our approach, which benefits from faster cryptography implementation and smaller header overhead, are better than the IPSec solution when limited number of hops are considered, while they are worse when the number of hops increases. Nevertheless, there are other important features that cannot be measured so directly and that imply a key security improvement for the ad hoc networking scenarios.

First of all, it is important to note that the results presented are obtained without considering the Internet Key Exchange (IKE) protocol [100] that precedes every IPSec session. In our case, the keys used for the tunnel establishment were statically fed to the IPSec implementation. This would have a negative impact on the figures presented. The results would be valid for the case where long data transmissions follow the key exchange so that the key exchange overhead can be neglected without incurring a large error. Nevertheless, it is important to clarify that the secure link establishment process used within the UCL takes place the moment two nodes meet (i.e. they hear each other's beacons) while the IKE process has to take place just before any IPSec session can be started. This difference between proactive versus reactive approach would increase the overhead in the case of IPSec more than in the UCL approach case if these two phases were considered.

More importantly, most of the attacks performed against ad hoc networks [158] would be prevented if the IP addresses of the nodes within the network were hidden to those other nodes not allowed to be part of the network. This way, malicious nodes would not be able to inject traffic or spurious control information in the network so that not only the communication integrity is assured but also the network stability is protected. While the UCL solution provides this feature thanks to its hop-by-hop encryption behavior, when IPSec end-to-end security is used, the IP header must be visible for each of the intermediate nodes so that they can route the datagrams without having to decrypt the actual payload. Hence, although IPSec has shown better performance for certain multihop configurations, the characteristic that allows this advantage (i.e. end-to-end secure association) represents a security disadvantage that leads to some security vulnerabilities when ad-hoc networks are considered. These vulnerabilities should be addressed using other security mechanisms that would increase the overhead, thus leading to worse performance than exhibited by the UCL implementation.

Moreover, the solution adopted by the UCL does not require the nodes to have any other IP address than the private personal network addresses [159] while when using IPSec in tunnel mode, the nodes must have been served by the access network or network manager with additional IP addresses. Of course, the IPSec transport mode might be used, but in this case, the exposed IP address (as described in the previous paragraph) would not be the one from the access network but the private personal network one with all the security

threats that this would lead to. This not only has security implications but from a network configuration point of view implies better performance.

Finally, it is important to note that the proposed approach for securing intra-cluster communications works transparently to the upper layers so that any routing protocol can be placed on top of the UCL. This is not the case when IPSec is used since the routing protocol itself would need to establish and maintain the tunnels.

CHAPTER 6
CONCLUSIONS AND FUTURE WORK

In this chapter, a short summary of the key point of each chapter is given and a conclusion is provided with respect to the problem statement. Furthermore, the chapter provides an outlook of the future steps in the work presented.

6.1 CONCLUSIONS

Telecommunications technology has achieved tremendous success over the past several decades and the networking research and development community continues to thrive as there are still a wide range of topics to be tackled on the path towards ubiquitous networks. In particular, the next generation networks will be seen as a new initiative to bring together all heterogeneous systems under the same framework, where the service shortcomings of one system can be complemented by other systems. Network convergence is therefore regarded as the next major challenge in the evolution of telecommunications technologies and the integration of computers and communications.

Personal Networks arose as a challenging concept that potentially includes all of a person's devices capable of network connection whether in his or her vicinity or remotely located. The work towards enabling this vision transparently for users results in major extensions of the present Personal Area Networking (PAN) and Ambient Intelligence (AI) paradigms. PNs are configured in an ad-hoc fashion, as the opportunity and the demand arise to support personal applications. PNs consist of communicating clusters of personal and foreign digital devices connected through various suitable communications means.

The achievement of this paradigm led to the identification of a set of optimizations in intra-cluster communications that were needed to fully support it. Firstly, heterogeneity will be a fundamental characteristic of next generation wireless communications since more and more personal devices are equipped with multiple network access technologies so that the user can have access to the different services that the different operational environments provide. However, Next Generation Networks (NGN) will comprise such a diverse number of possibilities that the users cannot be expected to take technical decisions on their own. It is necessary to provide mechanisms that intelligently select the optimal available access network based on context information such as user preferences, power consumption, link quality, etc. Finally, users need to trust the system that supports their personal communications. Within a personal network the most confidential information might be exchanged and the user need to be sure that this will never be disclosed. If the system fails in these features, NGN in general and PNs in particular will never happen.

This thesis has contributed with the development of the mechanisms that tackle the abovementioned challenges. The design and specification of a convergence framework, the so-called Universal Convergence Layer (UCL), has been the first topic addressed. This framework aims to manage all the network access interfaces with which a device is equipped so that they can be transparently used by upper layers as if the node were equipped with a single access technology.

On the other hand, the UCL enables the cross-layer optimization paradigm. Its privileged location within the protocol stack gives the UCL the possibility to support both bottom-up and top-down information flow. During the design, we have specified the intra-cluster communication process within the UCL and how we support backward compatibility. We have specified the complete protocol architecture as well as defined all the components

that are necessary to fulfil the required functionalities. The design followed a modular approach on which each of the required functionalities was provided by a different component. This design characteristic enables functionality extendibility so that new features can be added to the UCL in order to further optimize the system performance.

As identified from the requirements the work in this thesis has focused on two main areas, dynamic intelligent interface selection and security.

In the first area, the objective was the production of adaptive and flexible strategies to provide the best available connectivity for each operating environment and scenario. We designed and developed two mechanisms based on cross-layer information that optimize the system performance by exploiting the heterogeinity in terms of access technologies. Both strategies provide a resource management plane for handling, selecting and optimally using the multiple available radio technologies. While the first one concentrates on selecting the most appropriate outbound interface taking into account the conditions faced in each of the underlying wireless channels, the latter takes advantage of the multiplicity of access technologies by using all available interfaces in an intelligent way to improve the overall system performance. All the above aspects have been integrated into a fully working system, whose performance has been evaluated in all its aspects and for a wide range of setups and parameter values. The results have been analyzed in detail, offering valuable lessons about the current performance of the implemented software, which is promising, and optimal parameter settings. The results presented have shown that the solutions developed in this thesis represent a step forward on the cross-layer optimization paradigm. It paves the way for more sophisticated strategies that fully develop the cognitive network concept. In this sense, the solutions implemented exploit the learning-assessing-adapting approach precluded on the cognitive networking.

The UCL also plays a key role in security issues as an enabler for providing link-layer security mechanisms that ensure data confidentiality and integrity, authenticity and non-repudiation. The techniques implemented for node authentication combined with traffic encryption in ad-hoc networks are motivated by a vision of a PN where the user's devices organise themselves in clusters. Nevertheless, the mechanisms implemented are generic for the formation of private multi-hop clusters in an ad hoc manner, thus, they can be applied to many other fields just as well, for instance military or emergency networks. The analyses performed have proven the validity of the mechanisms developed. By comparing them with state-of-the-art approaches we have shown that not only subjective benefits can be obtained but performance can also be optimized. Within the first kind of benefits we highlight the fact that most of the attacks performed against ad hoc networks would be prevented if the IP addresses of the nodes within the network were hidden to those other nodes not allowed to be part of the network. While the UCL solution provides this feature thanks to its hop-by-hop encryption strategy, current solutions for IP security makes use of an end-to-end approach that obliges transmitting the IP header to be transmitted in clear. This vulnerability could be addressed using other security mechanisms that would increase the overhead, thus resulting on worse performance than the one exhibited by the UCL implementation. Finally, it is important to note that the proposed approach for securing intra-cluster communications works transparently to the upper layers so that any

routing protocol can be placed on top of the UCL. This is not the case for the IP security suite since the routing protocol itself would need to establish and maintain the tunnels. Pertaining to the actual performance optimization, the analytical and experimental studies carried out have presented the gain achieved for different scenarios and wireless access technologies.

The biggest advance in the state-of-the-art comes from enabling the user to have easy, affordable and seamless control of their devices over heterogeneous communications networks. They are empowered to communicate efficiently and securely with their selected interaction groups, no matter what kind of access is available for them to use.

6.2 FUTURE WORK

Current data networking technology limits a network's ability to adapt, often resulting in sub-optimal performance. Limited in state, scope, and response mechanisms, the network elements (consisting of nodes, protocol layers, policies, and behaviours) are unable to make intelligent adaptations. Communication of network state information is stifled by the layered protocol architecture, making individual elements unaware of the network status experienced by other elements. Any response that an element may make to network stimuli can only be made in the context of its limited scope. The adaptations that are performed are typically reactive, taking place after a problem has occurred. We have advanced the idea of cognitive networks, which enable the removal of these limitations by allowing networks to observe, act, and learn in order to optimize their performance.

Cognitive networks are motivated by complexity. Particularly in wireless networks, there has been a trend towards increasingly complex, heterogeneous, and dynamic environments. While wired networks can also take on any of these characteristics (and are not excluded from potential cognitive network applications) wireless networks are a natural target because of their inter-node interactions and the size of their system state space. The research carried out in this thesis has addressed some of these issues but has shortcomings from the network perspective. Cognitive networks represent a new scope and approach in dealing with this complexity.

Despite similarities, cognitive networks reach far beyond the scope of cross-layer designs. Cognitive networks can support trade-offs between multiple goals and in order to do so perform Multiple Objective Optimization (MOO), whereas cross-layer designs typically perform single objective optimizations. Cross-layer designs perform independent optimizations that do not take into account the network-wide performance goals. Trying to achieve each goal independently is likely to be sub-optimal, and as the number of cross-layer designs within a node grows, conflicts between the independent adaptations may lead to adaptation loops [160]. This pitfall is avoided in a cognitive network by jointly considering all goals in the optimization process.

In this thesis we have proposed techniques exploiting the cross-layer optimization approach. Nevertheless, the UCL is the framework over which advanced techniques for assessing the system state can be developed. Taking advantage of the global view that the convergence layer provides (access to information from multiple layers and access

technologies) a complete map of the system can be inferred. After matching it with the user requirements, the UCL can take the appropriate decisions in order to best serve the final users' wishes optimizing their quality of experience not only by providing enhanced bandwidth but also improving the power consumption efficiency or enhancing the service provision by lowering the packet loss due to wireless channel impairments. Last but not least, learning capabilities can be introduced so that optimization strategies can be dynamically tuned taking into account previous experiences and behaviours.

Some, but not all, of the research paradigms that should be studied in order to turn UCL into a cognitive networking framework are:

- Sensor networks. This kind of networks are rapidly becoming one of the most addressed topics in the research community due to their inherent capacity of easy deployment and gathering of useful information from the environment. It is precisely this capacity of extracting information from the environment that makes them particularly interesting for the cognitive networking concept. Extracting the information from the sensor networks and feeding it to the cognitive engines might lead to advantageous situations on which the network can adapt to the parameters acquired from the sensors.

- Context awareness. Being aware of the situation of the network or the requirements imposed by a particular service, is of critical importance in order to correctly react to any possible system misbehaviour. Nevertheless, context is such a heterogeneous concept and might come from such a plethora of different sources that it is mandatory to first treat the raw information so that it can become useful. Besides, it is necessary to define the characteristics of the context such as freshness, reliability, etc. in order not to make a decision based on a piece of context that might not be completely trustworthy. Finally, it is necessary to make the context information available to the components that implement the network adaptation mechanisms so that the can act accordingly.

- Channel model feeding into cognitive loop. Wireless channel impairments are an important challenge to be overcome when dealing with wireless networks. Nevertheless, as has been demonstrated in this thesis, there are ways to tackle this issue. A detailed and robust channel model might be of extraordinary help if it is intelligently used within the cognitive loop. Thus, if by analyzing the current channel parameters (by means of sensing the SNR, frame loss bursts, etc.) we can infer its behaviour it would be possible to take the best adaptation decisions. It would be particularly important to develop an adaptive channel model that could exploit the learning capacity so that its internal parameters could also be tuned to better map the real channel behaviour.

- Network synchronization. There is an implicit assumption that the cognitive network implements configuration changes synchronously. The details of actually making this happen with high reliability are likely to be complex. The implications of nodes' switching configuration at different times may be worse than if no adaptation had been performed at all. Also, the varying topology of the network means that not all nodes will receive notification of configuration changes at the same time.

- Green wireless. Power efficiency is one of the key challenges for future networks. Already today, about 3% of the world-wide electrical energy is consumed by the ICT infrastructure. Lowering the energy consumption of future wireless systems demands greater attention than ever, ensuring that ICT-enabled solutions are available and fully deployed in an ecologically and economically sustainable way. The goal should be to improve the energy efficiency of existing and future wireless systems without compromising the user's perceived quality. UCL-enabled solutions might work on this direction.

Additionally, this thesis has mainly focused on the PN concept while still envisaging the limitation of this concept as it is restricted to the formation of a network belonging only to one person. It would be necessary to further develop the concept of PN Federation (PN-F) as the internetworking of multiple PNs belonging to different persons that have a common objective. In this sense, not only the secure self-establishment of the PN-F network needs to be tackled but also the necessary service enablers such as context and service provision management frameworks have to be carefully addressed so that such a challenging concept can actually be developed.

APPENDIX A: IMPLEMENTATION DETAILS

LINUX KERNEL AND LOADABLE KERNEL MODULES

The implementation of the Universal Convergence Layer and its building blocks has been done taking as a basis the Linux operating system. The main reason for this choice is the fact that most of the operations to be performed by the UCL are done before any of the networking mechanisms. In this sense, it is necessary to have complete access to the network stack and this is only possible with an open source Operating System (OS). Additionally, the aim of the projects that have served as the framework for the development of the work carried out in this thesis has been to demonstrate the feasibility of the PN concepts through the development of pilot services. The need for including devices with a reduced footprint (such as PDAs) limited the choices among the open source operating systems to Linux.

There are frameworks that enable tweaking the normal network stack operation that do not require kernel-level programming but experience tell us that they imply additional computational overhead that leads to reduced system performance. Hence, the UCL has been developed as an add-on made to the Linux kernel.

If you want to add code to a Linux kernel, the most basic way to do that is to add some source files to the kernel source tree and recompile the kernel. In fact, the kernel

configuration process consists mainly of choosing which files to include in the kernel to be compiled.

However, you can also add code to the Linux kernel while it is running. A chunk of code that you add in this way is called a loadable kernel module. These modules can do lots of things, but there are four main things Loadable Kernel Modules (LKMs) are used for:

- Device drivers. A device driver is designed for a specific piece of hardware. The kernel uses it to communicate with that piece of hardware without having to know any details about how the hardware works.
- Filesystem drivers. A filesystem driver interprets the contents of a filesystem (which is typically the contents of a disk drive) such as files and directories. There are lots of different ways of storing files and directories on disk drives, on network servers, and in other ways.
- System calls. User space programs use system calls to get services from the kernel. For example, there are system calls to read a file, to create a new process, and to shut down the system. Most system calls are integral to the system and very standard, so are always built into the base kernel (no LKM option). However, you can invent a system call of your own and install it as an LKM. Or you can decide you do not like the way Linux does something and override an existing system call with an LKM of your own.
- Network drivers. A network driver interprets a network protocol. It feeds and consumes data streams at various layers of the kernel's networking function.

The UCL is a mixture between the first and last of the categories. As such, the UCL development follows a twofold approach within Linux OS, both user and kernel space, for handling with the wireless interfaces and the information packets that flow through them.

The following section will explain in more depth the Linux features that will be used in the UCL development.

Linux Ethernet virtual device

Traditionally, network interfaces are considered by the OS as devices, most of them associated to a physical component that is in charge of transmitting and receiving data packets (normally an Ethernet card). A network interface is a software object that can process incoming and outgoing packets, while the actual reception and transmission mechanisms remain hidden inside the interface driver, which interacts with the underlying hardware. Even though most interfaces are associated to physical devices (or, for the loopback interface, to a software-only data loop), it is possible to design network interface drivers that rely on other interfaces to perform actual packet transmission.

The idea of a virtual interface can be useful to implement special-purpose processing in data packets while avoiding hacking the network subsystem of the kernel. In this sense, the virtual interface could be considered as a tool for customizing network behaviour. To

control the operation mode of the virtual device (add new ports, enable security issues, etc.) user space programs interact with kernel modules.

From the abovementioned, one or multiple interfaces can appear as one large interface to the participating hosts by binding them to the same virtual network interface. In our case the virtual network interface will manage all the wireless network interfaces. In this sense the UCL resembles a network bridge which could be considered the baseline for UCL development. However, to achieve the functionalities in terms of heterogeneity, security and data transfer management it is necessary to extend these functionalities and adapt their behaviour to the new requirements.

Security libraries

OpenSSL [161] is a full-strength general purpose cryptography library that provides cryptographic functionality to user space applications. In the latest Linux kernel versions some cryptographic features have been included that allows the application of encryption and decryption to packets without having to queue them to user space. As a result of this, an improvement in performance has been obtained. However, not all cryptographic algorithms included in OpenSSL has been exported to the Linux kernel, so the most complex operations have still to be run in user space (i.e. Diffie-Hellman protocols).

As defined in the MAGNET project, security is going to be tackled in several layers, taking into account the trust relationship between peers. The UCL will cope with security at a link level, before the packet is actually sent to its destination. To handle packets at this level it is necessary to develop some kernel code and limitedly complex cryptographic algorithms have to be used, in order to achieve a lower performance decrease. Cryptographic operations will only be used for deriving a session key, based, for instance, on a mutual authentication mechanism. Taking into account these premises, Linux kernel cryptographic modules offer the suitable algorithms (DES, AES, HMAC, SHA, MD5, etc.).

NEIGHBOUR DISCOVERY AND AUTHENTICATION MODULE

As has been already mentioned, a PN is a collection of one's most private devices, referred to as personal devices/nodes, which forms a virtual network where collocated personal devices organize themselves in clusters which are in turn interconnected over the Internet. Physically neighbouring PN nodes must authenticate each other and establish short-term link-level security associations, based on the long-term pair-wise keys exchanged during the imprinting, as the first step towards the cluster self-organization. Direct secure communication is possible in an ad-hoc manner after the two nodes have correctly authenticated each other, exchanged the link-level session key and exchanged each other’s personal network address.

The neighbour discovery and authentication module is in charge of the aforementioned tasks. In this section we will present the module implementation details and the main insights needed to understand the module operation and its role in the overall PN/PN-F system.

SW Architecture and implementation details

The neighbour discovery module is implemented as a Linux kernel module. It is mainly divided in three components, as shown in Figure A-1, that interplay to accomplish the tasks the module is responsible for.

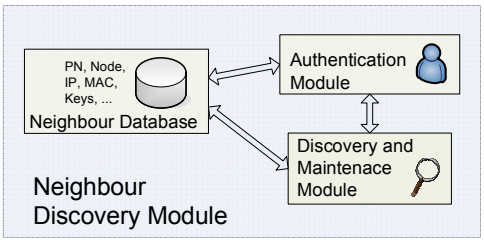


Figure A-1: Neighbour Discovery module high-level architecture diagram

Neighbour database

The information that is retrieved by the Neighbour Discovery module is stored in an internal database and made available to the rest of the system through several interfaces.

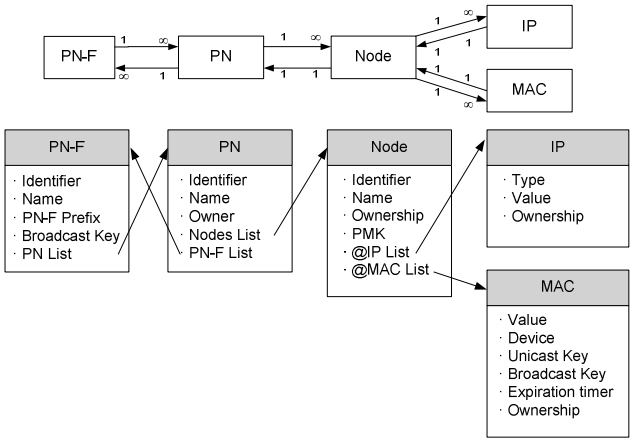


Figure A-2: Neighbour database organization

As is shown in Figure A-2 the structure of the database starts from the PN to which the neighbouring nodes belong. This first table contains a list of nodes. Each entry on this list contains the information about a node, basically its identifier and the Primary Master Key (PMK – the one exchanged during the imprinting procedure). Besides, a list of the IP and MAC addresses of this node is also included on each node record. For each IP address, the type of address (i.e. IPv4 or IPv6), its value and its category (PERSONAL or PUBLIC) is stored. Finally, for the MAC addresses, not only its value and the network interface to which this entry corresponds are stored but the session and broadcast keys are stored in each MAC record. It is important to note that, although a node is univocally identified by a node identifier provided within the beacon payload, a different authentication process is performed for each of the air interfaces, through which it is possible to communicate with the neighbour. Thus, there will be different keys with the same node if they belong to more than one radio domain.

Additionally, the neighbour discovery module also holds the information pertaining to PN-Fs to which the PN belongs and those that are publicly advertised in the node radio

domain. In this sense, the information stored refers to the PN-F's basic parameters such as name, identifier and list of members.

Figure A-2 also presents the model of the possible relationships between entities in the Neighbour Database. Hence, a node can only belong to one PN, that is, a node record can only be present in the node list of one PN. In contrast, a PN can have any number of node records in its nodes list. A similar situation appears in the relationship between IP addresses and Nodes. While a node can have many IP addresses, an IP address can only belong to one node. For the MAC addresses, the relationship is the same. A MAC address record can only be associated with one node while this node can be reached through multiple network interfaces, thus the MAC addresses list for that node will contain multiple MAC address records. Regarding PN-Fs the relationship is mutual since a PN-F has multiple members (i.e. multiple PNs associated) and a PN can be member of several federations.

Each table is organized as an independent hash list that contains the necessary pointers to link each entry to the associated entries on the other lists. The use of hash lists helps in fetching information when necessary.

```
struct ndisc_fdb_info {
	struct hlist_head pnf[NDISC_PNF_HASH_SIZE]; //ndisc_fdb_pnf_entry
	struct hlist_head pn[NDISC_PN_HASH_SIZE]; //ndisc_fdb_pn_entry
	struct hlist_head node[NDISC_NODE_HASH_SIZE]; //ndisc_fdb_node_entry
	struct hlist_head ip[NDISC_IP_HASH_SIZE]; //ndisc_fdb_ip_entry
	struct hlist_head mac[NDISC_MAC_HASH_SIZE]; //ndisc_fdb_mac_entry
};
```

From these *heads* the lists are expanded with the different *entries*.

```
struct ndisc_fdb_pnf_entry{
	struct hlist_node pnf_hlist; // for handling global PN-F list (base on hash)
	// General information
	u8 id[PNF_ID_LENGTH]; // Unique PN-F identifier
	u8 name[PNF_NAME_LENGTH]; // User friendly descriptive name

	union {
		u32 ipv4;
		struct in6_addr ipv6;
		unsigned char ip[0];
	} u;

	u8 bcast_key[PNF_BKEY_LENGTH];

u8 membership;
};

struct ndisc_fdb_pn_entry {
	struct hlist_node pn_hlist; // for handling global pn list (base on hash)
	struct hlist_head pnf_hlist; // for handling pnf_hlist from pn
	struct hlist_head node_hlist; // for handling node_hlist from pn
	// General information
	u8 id[PN_ID_LENGTH]; // Unique PN identifier
	u8 name[PN_NAME_LENGTH]; // User friendly descriptive name
	// Imprinting and security
```

```
        u8 ownership;
        u8 owner[PN_NAME_LENGTH];
        u8 imprinted;
        u8 pmk[PMK_LENGTH];
};

struct ndisc_fdb_node_entry {
        struct hlist_node node_hlist; // for handling global node list (base on hash)
        struct hlist_head ip_hlist; // for handling ip_hlist from node
        struct hlist_head mac_hlist; // for handling mac_hlist from node
        struct hlist_node hlist;     // for handling node_hlist from pn
        struct ndisc_fdb_pn_entry *pn; // pointer to pn
        // General information
        u8 id[NODE_ID_LENGTH]; // Unique node identifier
        u8 name[NODE_NAME_LENGTH]; // User friendly descriptive name
        // Imprinting and security
        u8 ownership;
        u8 imprinted;
        u8 pmk[PMK_LENGTH];
};

struct ndisc_fdb_mac_entry {
        struct hlist_node mac_hlist; // for handling global mac list (base on hash)
        struct hlist_node hlist; // for handling mac_hlist from node
        struct ndisc_fdb_node_entry *node; // pointer to node
        // General information
        struct net_device *dev; // In device
        unsigned char addr[ETH_ALEN]; // device MAC (kept for virtual deleting)
        // Imprinting and security
        u8 ownership; // Personal, foreign
        struct ucl_key ukey; // Unicast key
        struct ucl_key bkey; // Bcast key
        unsigned long key_exp_time; // Session key expiration time
};

struct ndisc_fdb_ip_entry {
        struct hlist_node ip_hlist; // for handling global ip list (base on hash)
        struct hlist_node hlist; // for handling ip_hlist from node
        struct ndisc_fdb_node_entry *node; // pointer to node
        // General information
        int type; // IPv4 or IPv6
        union {
                u32 ipv4;
                struct in6_addr ipv6;
                unsigned char ip[0];
        } u;
        // Imprinting and security
        u8 ownership;
};
```

Each time a new neighbour is detected the corresponding entries are created and indexed such that all the information can be retrieved at any moment. The idea behind indexing all the information pertaining to one node is that in this way it is possible to check all the reports from one node given a known parameter. For example, given the reception of a

packet we can check from the MAC address if this is a personal node and which key to use for decrypting the frame.

Note that although the work in this thesis is focused on the PN scenario, the implementation done also covers PN-F scenarios where more than one PN interplays.

Discovery and maintenance

The discovery and maintenance mechanisms are based on periodic beaconing. Specific beacon packets are defined by using a pre-defined Ethernet Type field values in the MAC header. The same Ethernet Type value is used for the beacon packets and for the configuration exchange ones. Hence, the reception function (*ndisc_packet_handler*) has to split the handling of the two kinds of packets.

```
static struct packet_type ndisc_packet_type = {
        .type = ETH_P_NDISC,
        .func = ndisc_packet_handler,
};

int ndisc_packet_handler(struct sk_buff *skb, struct net_device *dev,
                         struct packet_type *pt, struct net_device *orig_dev)
{
        struct ndisc_hdr *hdr;
        hdr = (struct ndisc_hdr *)skb->data;
        switch (hdr->type) {
                case NDISC_BEACON:
                        ndisc_beacon_handler(skb, dev, pt);
                        break;
                case NDISC_CONF:
                        ndisc_conf_handler(skb, dev, pt);
                        break;
                default:
                        printk("Unknown packet");
        }
        kfree_skb(skb);
        return 0;
}
```

The beacon handler function parses the information contained in the beacon packet and triggers the corresponding actions. It creates the node and linked MAC and IP entries on the database if this is the first beacon received and initiates the authentication procedure. If the entries were already created by a previous beacon, the timeouts detecting the node disappearance are re-started. For beacons received from the same node but through a different network device the corresponding entries are created.

Authentication

When a beacon is received from a personal node the authentication mechanism is triggered in order to certify this status.

As already pointed out the authentication procedure is embedded in EAP protocol.

Figure A-3 shows the EAP message format.

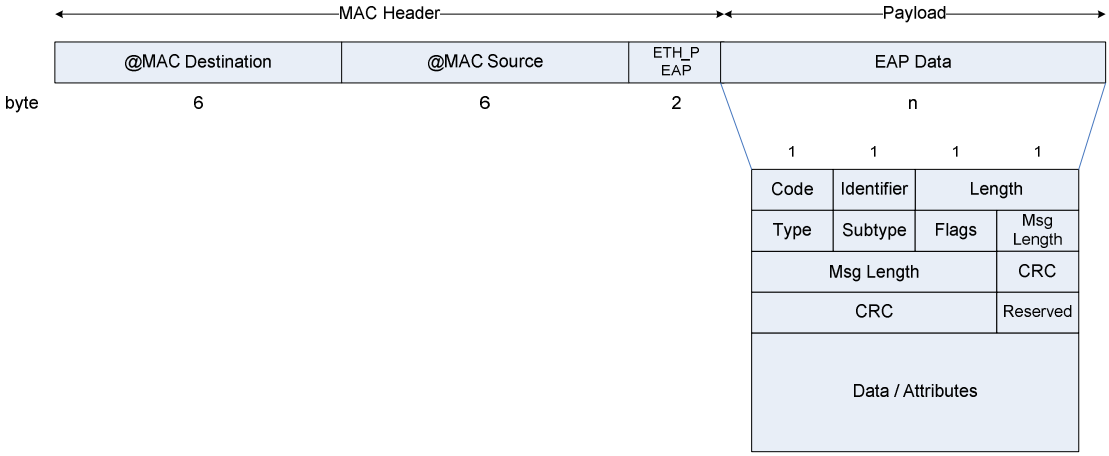
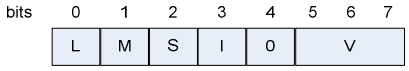


Figure A-3: EAP Message Format

EAP MAGNET Type has been defined as 0x08. Subtype defined as MAGNET SK is 0x01.

Flags have the following format and semantics:



L = Length included

M = More fragments

S = Start – Identity first fragment

I = Integrity Checksum included

V = 100 for EAP-MAGNET

Integrity checksum is a cryptographic hash computed over the message and must be present if the messages are fragmented.

Data consists of a sequence of attributes that follow the 4-byte aligned header.

Figure A-4 shows the structure of an attribute.

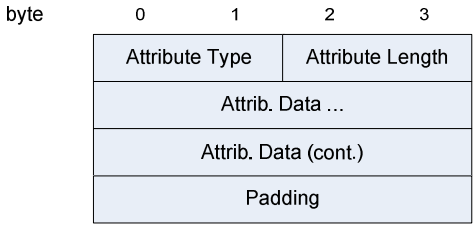


Figure A-4: EAP-MAGNET attribute format

Attribute Length is the actual length of the attribute data, including the attribute type and attribute length fields. Each attribute must be padded to preserve the 4-byte alignment.

Attribute type is one of the following:

Name	Notation	Code
Padding	AT_PADDING	6
MT Nonce	AT_NONCE_MT	7
MT Expiration Time	AT_EXP_TIME_MT	8
MT Broadcast Key	AT_BCAST_KEY_MT	9
Node Identity	AT_IDENTITY	14
Supported Versions	AT_VERSION_LIST	15
Version Selected	AT_SELECTED_VERSION	16
Server Nonce	AT_NONCE_S	21
Server Expiration Time	AT_EXP_TIME_S	22
Server Broadcast Key	AT_BCAST_KEY_S	23
Initialization vector for encryption	AT_IV	129
Encrypted field	AT_ENCR_DATA	130

- **AT_PADDING**

Attribute used when padding is needed for encryption or any other attribute that encapsulates information in it.

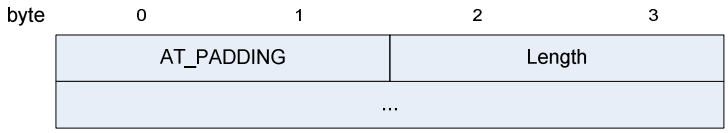


Figure A-5: AT_PADDING attribute format

- **AT_NONCE_MT - AT_NONCE_S**

Random sequence of bytes used to perform authentication. Attribute length is a variable.

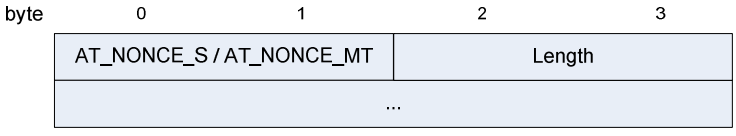


Figure A-6: AT_NONCE attribute format

- **AT_BCAST_MT - AT_BCAST_S**

Broadcast key of the nodes. It should be encapsulated within AT_ENCR_DATA.

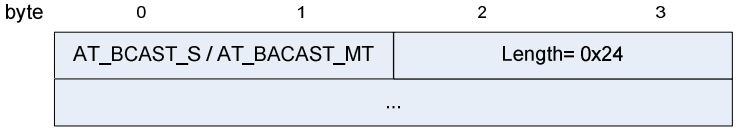


Figure A-7: AT_BCAST attribute format

– **AT_EXP_TIME_MT - AT_EXP_TIME_S**

Short indication of the session key maximum expiration time. Time is expressed in seconds. Currently values are randomly chosen in the range from 1 to 65535 seconds.

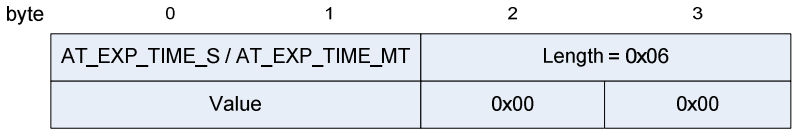


Figure A-8: AT_EXP_TIME attribute format

– **AT_ENCR_DATA**

Attribute encapsulating encrypted information. Data is encrypted using a symmetric cryptographic algorithm (AES). Any other attribute or sequence of attributes can be held in an encrypted form inside.

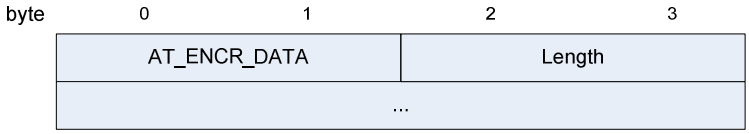


Figure A-9: AT_ENCR_DATA attribute format

The following is an example of the authentication mechanism. It is important to note that the AT_ENCR_DATA attribute is not encrypted as it is an example for showing the procedure.

EAP request / MAGNET_SK

```
01                                          EAP Code - Request
0F                                          Identifier
00 5C                                       Message Length
08                                          EAP Type - MAGNET
   01                                       EAP-MAGNET Subtype - SK
   04                                       Flags
   00                                       Reserved
   00 82                                    AT_ENCR_DATA
      00 54                                 Length
      00 15                                 AT_NONCE_S
         00 15                              Length
         4D 7C 21 B7 0F C3 67 12 F9 67      Value
         8E 6D 34 81 48 96 46 00 00 00
      00 17                                 AT_BCAST_KEY_S
         00 24                              Length
         64 78 35 93 B6 51 C0 7D 8B C5      Value
         FA 33 D5 70 E1 4E 47 DC A9 B6
         D4 B8 69 A4 00 72 9B 13 CD 83
         8C 1E
      00 16                                 AT_EXP_TIME_S
         00 06                              Length
         01 2C 00 00                        Value = 300 sec.
      00 06                                 AT_PADDING
         00 0C                              Length
         00 00 00 00 00 00 00 00
```

EAP response / MAGNET_SK

```
02                                                  EAP Code – Response
0F                                                  Identifier
00 8C                                               Message Length
08                                                  EAP Type – MAGNET
   01                                               EAP-MAGNET Subtype – SK
   04                                               Flags
   00                                               Reserved
   00 82                                            AT_ENCR_DATA
      00 84                                         Length
      00 15                                         AT_NONC
E_S
         00 15                                      Length
         4D 7C 21 B7 0F C3 67 12 F9 67              Value
         8E 6D 34 81 48 96 46 00 00 00
      00 17                                         AT_BCAST_KEY_S
         00 24                                      Length
         64 78 35 93 B6 51 C0 7D 8B C5              Value
         FA 33 D5 70 E1 4E 47 DC A9 B6
         D4 B8 69 A4 00 72 9B 13 CD 83
         8C 1E
      00 07                                         AT_NONCE_MT
         00 09                                      Length
         3A 4A 8F F5 DD 00 00 00                    Value
      00 09                                         AT_BCAST_KEY_MT
         00 24                                      Length
         46 CB 0E 2F CF 3E D9 83 E0 4F 47 EC        Value
         29 8E A0 4C E5 9F 67 5B D1 D9 26 91
         19 F0 1D C1 72 49 1C 7B
      00 08                                         AT_EXP_TIME_S
         00 06                                      Length
         00 F0 00 00                                Value = 240 sec.
      00 06                                         AT_PADDING
         00 0C                                      Length
         00 00 00 00 00 00 00 00
```

EAP success / MAGNET_SK

```
03                                                  EAP Code – Success
0F                                                  Identifier
00 3C                                               Message Length
08                                                  EAP Type – MAGNET
   01                                               EAP-MAGNET Subtype – SK
   04                                               Flags
   00                                               Reserved
   00 82                                            AT_ENCR_DATA
      00 34                                         Length
      00 07                                         AT_NONCE_MT
         00 09                                      Length
         3A 4A 8F F5 DD 00 00 00                    Value
      00 09                                         AT_BCAST_KEY_MT
         00 24                                      Length
         46 CB 0E 2F CF 3E D9 83 E0 4F 47 EC        Value
         29 8E A0 4C E5 9F 67 5B D1 D9 26 91
         19 F0 1D C1 72 49 1C 7B
```

Similarly to the handler defined for the beacon packets, a function handling all the packets exchanged during the authentication is declared.

```
static struct packet_type eap_packet_type = {
        .type = __constant_htons(ETH_P_EAP),
        .func = eap_recv,
};
```

It is important to note that when the node is switched on, it already has some information pertaining to other personal nodes as a result of the imprinting procedure. This information is loaded into the neighbouring database prior to the initialization of the node interfaces. The information loaded consists of the node identifier and the shared secret. When the authentication procedure is launched, the PMK is fetched from the database and the LMSK is derived from it. The EAP exchange can then start.

Interfaces

As was already said and as will be explained in Appendix C, the implementation of the neighbour discovery and authentication is part of a vertically integrated system. In this sense, several interfaces are implemented for the module to interplay with the other system blocks.

- Signalling/control interfaces
    - newLink
        - **Type:** multidirectional out
        - **Protocol:** Netlink socket
        - **Description:** An event in the form of a netlink message is broadcast to any module listening to the socket when a new personal node is detected.
        - **Corresponding components:** PN Manager, PN Routing Protocol module
    - linkBreak
        - **Type:** multidirectional out
        - **Protocol:** Netlink socket
        - **Description:** An event in the form of a netlink message is broadcast to any module listening to the socket when a personal node leaves the close vicinity.
        - **Corresponding components:** PN Manager, PN Routing Protocol module
    - setNodeVisibility
        - **Type:** unidirectional in
        - **Protocol:** /proc filesystem
        - **Description:** Set in order to hide node's presence by stopping sending beacons
        - **Corresponding components:** PN Manager

- setNodeID
  - **Type:** unidirectional in
  - **Protocol:** ioctl
  - **Description:** This information goes in the NDISC beacons and identifies the node from its neighbours
  - **Corresponding components:** PN Manager
- setPNID
  - **Type:** unidirectional in
  - **Protocol:** ioctl or /proc filesystem
  - **Description:** This information goes in the NDISC beacons and identifies the node from its neighbours
  - **Corresponding components:** PN Manager
- loadPMK
  - **Type:** unidirectional in
  - **Protocol:** ioctl passing list of structures with per node information (Node ID, Node Name, PMK, PN ID, PN name, type of node – PERSONAL, FRIEND, FEDERATED)
  - **Description:** Used to inform the NDISC about all the nodes with which a security association is established. It is used once just after the NDISC module is loaded
  - **Corresponding components:** Trust establishment module

- Data interfaces:
  - Display neighbours table
    - **Type:** unidirectional out
    - **Protocol:** /proc filesystem
    - **Description:** User may want to display the current status of the neighbours table (all entries, devices, etc.).
    - **Corresponding components:** PN Manager, SCMF

UNIVERSAL CONVERGENCE LAYER

The first main objective of the UCL is to hide the complexity of the available air interfaces and to offer a unique interface to the upper layers. This module will handle this task by discovering and managing the different network resources (set them up, acquire statistics for feeding cross-layer optimization techniques, etc…).

UCL aims to masquerade multihoming by aggregating the different network interfaces (one per access technology the node is equipped with) on a single interface. Management of the heterogeneity of wireless interfaces is recommended to be implemented in a kernel space level. Although management programs run in user space in order to provide easy access to the list of devices and their main features, it is necessary to enable some Linux kernel modules which will be the suppliers of the information.

In this section we will present the module implementation details and the main insights needed to understand the module operation and its role in the overall PN/PN-F system. It will be focused on the implementation of the UCL framework without developing further the insights of each of the UCL building blocks.

SW Architecture and implementation details

As can be seen in Figure A-10 the UCL is divided in two main parts. One of them is sited in the user space while the other is implemented in the kernel space. While the former is used during the node start-up for creating the framework and discovering the network interfaces to be managed through the UCL, the kernel space one is in charge of all the operations that deal with inbound and outbound traffic flows.

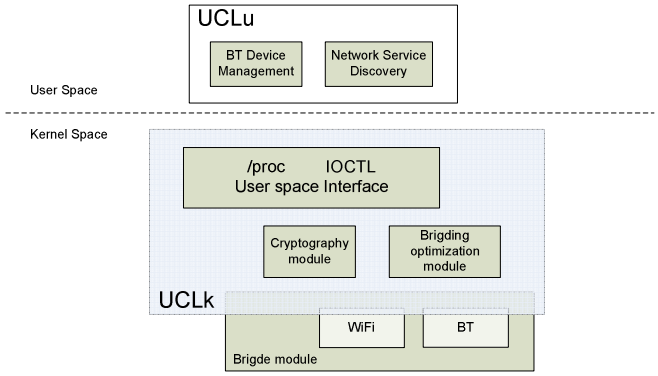


Figure A-10: UCL low-level architecture specification

UCLk

```
static struct net_device *new_ucl_dev(const char *name)
{
      dev = alloc_netdev(sizeof(struct net_bridge), name, ucl_dev_setup);
      ...
      return dev;
}

void ucl_dev_setup(struct net_device *dev)
{
      memset(dev->dev_addr, 0, ETH_ALEN);

      ether_setup(dev);
      dev->do_ioctl = ucl_dev_ioctl;
      dev->get_stats = ucl_dev_get_stats;
      dev->hard_start_xmit = ucl_dev_xmit;
      dev->open = ucl_dev_open;
      dev->set_multicast_list = ucl_dev_set_multicast_list;
      dev->change_mtu = ucl_change_mtu;
      dev->mtu = 1496;
      dev->destructor = free_netdev;
      SET_MODULE_OWNER(dev);
      dev->stop = ucl_dev_stop;
      dev->tx_queue_len = 0;
      dev->set_mac_address = NULL;
      dev->priv_flags = IFF_EBRIDGE;
}
```

The UCL kernel module provides the framework over which the other components operate. It creates the network device and sets it up generating a virtual interface that provides the necessary functions to operate in a pseudo-bridge manner.

Upon introduction of the network interfaces with which the node is equipped, new ports are added to the UCL structure.

When a packet is received by any of the controlled network interfaces, it is passed to the UCL framework created by this kernel module. Similarly, when a packet is to be transmitted, the UCL transmission function is the anchor point for the network layer.

Additionally, it defines other functions that will be used for interfacing with the UCL such as *ioctl* framework.

Bridge module and UCLu

By definition, a bridge is a device that separates two or more network segments within one logical network (e.g. a single IP-subnet). The UCL implementation is built around the Linux bridge implementation. In its roots, this implementation enables the management of several interfaces on the same machine as a unique virtual interface with a single network identifier. This is theoretically all that is required from the UCL, but the addressing and other control functionalities of Linux bridging are mainly founded on the IEEE 802.1d protocol which does not fulfil the required behaviour for the UCL. Thus, we only take the gathering capacity of the Linux bridging framework but trim all the other control characteristics.

Additionally, the UCL is meant to automatically load all the interfaces of a multi-homed device into its virtual domain so that they are all managed from it. To fulfil this requirement the user space part of the UCL (UCLu) was implemented. In our implementation it looks for all the wired, WiFi and Bluetooth network interfaces and adds them to the UCL bridging framework.

For the Bluetooth case a special mechanism is implemented due to the implicit characteristics of this technology. BNEP profile is used for the Bluetooth devices. This profile creates a point-to-point link between the piconet coordinator and the slave. Thus, when a node is switched on, it might the first Bluetooth device on the radio domain and as such it will have to act as a piconet coordinator. Other nodes with Bluetooth devices entering in the area will be slaves in this piconet and they will establish the link with the coordinator. The node with the coordinator role will have as many Bluetooth interfaces as slaves in the piconet. Moreover, when the coordinator switches off or leaves the coverage area, a new device has to take the responsibility and reorganize the Bluetooth piconet. All these tasks are also performed in the UCLu.

```
int UCL::serviceStart()
{
	PRINTF("\nUCL Initializing ...\n");

	if (initBridge() != 0) {
		return -1;
	}
	if (m_bRunWired) {
		PRINTF("Adding Wired interface ...\n");
		addNetworkInterface(m_pcWiredInterfaceName);
	}
	if (m_bRunWiFi) {
		PRINTF("Seaching WiFi interfaces ...\n");
		m_bExistsWiFi = (getWiFiInterfaces(&m_mapWiFiIf) == 0) ? true : false;
		if (m_bExistsWiFi) {
			mapWiFiLLCT_t::iterator mapIter;
			for (mapIter = m_mapWiFiIf.begin();
			    mapIter != m_mapWiFiIf.end();
			    mapIter++) {

			      WiFiLLCT *wifi = mapIter->second;
			      addNetworkInterface(mapIter->first);
			      wifi->serviceStart();

			}
		} else {
			PRINTF("Not found ...\n");
		}
	}
	if (m_bRunBT) {
		PRINTF("Seaching Bluetooth interfaces ...\n");
		std::set<int> setBT;
		m_bExistsBT = (getBTDevices(&setBT) == 0) ? true : false;
		if (m_bExistsBT) {
			m_BTLLCT = new BTLLCT(this);
			m_BTLLCT->serviceStart();
		} else {
			PRINTF("Not found ...\n");
		}
	}
	PRINTF("UCL service start ended\n");
	PRINTF("UCL Initiliazing ...OK\n");
	return 0;
}
```

Bridging optimization module and Cryptography module

At the UCL the packet transmission procedure follows the path described in Figure 3-17. Basically these two modules implement the functionalities specified in Sections 3.1.6 and 3.1.7.

The bridging optimization module operation is the first to be called in the downstream flow. As a result of this module the outbound hardware interface is selected. This information is used for both applying the corresponding security keys on the cryptography module and to appropriately compose the MAC layer header.

```
static void ucl_deliver(const struct net_bridge_port *to, struct sk_buff *skb)
{
      struct ndisc_fdb_mac_entry *fdb_mac;
      …
      skb->dev = to->dev;

      // UCL Path optimization
      fdb_mac = ucl_path_opt(&skb);
      ...
      // UCL sign and encryption
      ucl_apply_security(&skb, fdb_mac);
      ..
      // Fill MAC source address field with proper address
      memcpy(eth_hdr(skb)->h_source, skb->dev->dev_addr, ETH_ALEN);
     dev_queue_xmit(skb);
}
```

User space interface

In order to interact with the rest of the system[2] special purpose interfaces have to be deployed. It has to be taken into account that in contrast to the rest of the system components, the UCL is implemented in the kernel space.

The two ways implemented for interplaying with the UCL are based on *ioctl* commands and on the */proc* filesystem.

Most devices can perform operations beyond simple data transfers, so user space must often be able to make special requests or inform the kernel module handling the device about certain parameters. These operations are usually supported via the *ioctl* method, which implements the system call by the same name.

The */proc* filesystem is a virtual filesystem that permits a novel approach for communication between the Linux kernel and user space. In the */proc* filesystem, virtual files can be read from or written to as a means of communicating with entities in the kernel, but unlike regular files, the content of these virtual files is dynamically created.

It is important to note that although the neighbour discovery module and the UCL are independent components, they are tightly intertwined and some of the interfaces of the former, described in the previous section, are actually implemented in the UCL module. These are mainly the ones implemented through *ioctl* commands.

Interfaces

- Signalling/control interfaces
    - addInterface
        - **Type**: multidirectional in
        - **Protocol**: ioctl

[2] As has already been mentioned, the implementation done during the development of this thesis was integrated in a full-scale Personal Networking prototype.

- **Description**: UCL can control all PAN and WAN interfaces in a node. As they may vary, it is necessary to adapt to context modifications in these terms. This way, new air interfaces can be added by the user or automatically without having to switch off the device.
- **Corresponding components**: PN Manager

- removeInterface
  - **Type**: multidirectional in
  - **Protocol**: ioctl
  - **Description**: Similarly to the previous one, new air interfaces can be removed by the user or automatically without having to switch off the device.
  - **Corresponding components**: PN Manager
- configureWANInterface
  - **Type**: unidirectional out
  - **Protocol**: writing on /proc the details of newly available WAN interfaces.
  - **Description**: To instruct the PN Routing about new available connection to Internet.
  - **Corresponding components**: PN Routing Protocol module
- getPossibleDevs
  - **Type**: unidirectional in
  - **Protocol**: Shared memory.
  - **Description**: UCL module will use this interface in order to ask the Pn Neighbouring module for the list of devices that allow communication with the node represented with the MAC address passed as an argument.
  - **Corresponding components**: Neighbour Discovery module
- getKey
  - **Type**: unidirectional in
  - **Protocol**: Shared memory.
  - **Description**: UCL module will use this interface in order to ask the PN Neighbouring module for the key to use for encrypting or decrypting the IP datagram.
  - **Corresponding components**: Neighbour Discovery module

APPENDIX B: MAGNET VERTICALLY INTEGRATED PROTOTYPE

For the PN and PN-F paradigm, conceptual solutions have been proposed and evaluated and the most promising solutions have been selected and implemented in x86 and ARM architectures for a Linux-based platform [162][163][164]. In the following subsections we will first briefly summarize the implementation of the PN concept. Figure B-1 presents the Birds Eye View of the different components that have been implemented and integrated into the PN and PN-F platform.

The basic approach taken to develop the PN concept was to implement the PN as a secure and self-organising overlay network consisting of all the nodes that belong to the PN. This overlay network has its own private IP addressing space, creating a confined and private network in which personal nodes (PN nodes) can freely communicate with each other and on top of which a service discovery platform and PN applications can be deployed.

A basic requirement to develop this overlay network, is the ability to discriminate between personal nodes and non-personal nodes (i.e. foreign nodes). This discrimination is stored as a property of the corresponding trust relationship. This bilaterally secure association between the PN nodes is negotiated using the Certified PN Formation Protocol (CPFP).The CPFP protocol is based on asymmetric cryptography and uses the novel Elliptic Curve

Cryptography algorithms to generate the secret keys. The concept behind is that each new node must be introduced to the PN by the user during a procedure called imprinting. After a successful imprinting, the new personal device receives a valid PN certificate and is ready to establish secure associations with any other personal node based on each other's PN certificates. While the first step (i.e. the introduction to the PN through the imprinting) has to be monitored by the user, the subsequent secure associations that the node establishes with each of the other personal nodes are done automatically and transparently to the user.

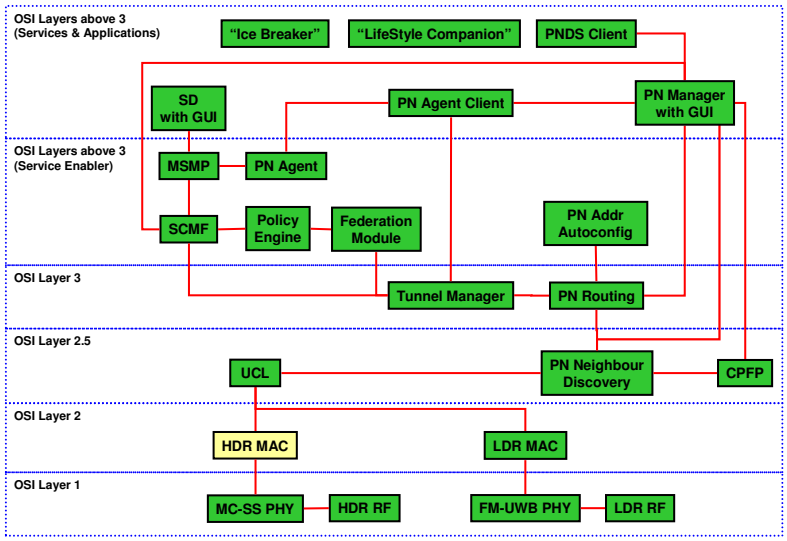


Figure B-1: PN and PN-F system birds eye view

Table B-1: Integrated components in the PN/PN-F system overview

Component Name	Main Functionalities
Trust Establishment module	This component is responsible for the Certified PAN Formation Protocol (CPFP); the so-called imprinting procedure. Every node that wants to be included in the PN needs to establish a long-term trust relationship with the rest of the nodes.
Neighbour Discovery and Authentication module	This module is in charge of the discovery and authentication of the surrounding personal nodes. The first step in the cluster formation is to find out which nodes are in the surroundings. Additionally, it is necessary to authenticate them based on the pair-wise long-term trust relationship they share after the imprinting
Universal Convergence Layer	This component has a twofold objective. On the one hand, it hides the possible heterogeneity in terms of multiple air-interfaces on the same node from the upper layers, so that it supports nodes with multiple interfaces. On the other hand, it provides encryption and filtering of the traffic, so it assures a secure and private connectivity level.

PN/PN-F Routing module	This module is responsible for routing all the PN internal traffic. Irrespective of whether it is a route within the cluster or it is a packet destined to a node in a remote cluster, they are routed by this module.
Dynamic Tunnel Establishment module	This module is in charge of the tunnels negotiation, establishment and adaptation. Each of the PN clusters is interconnected across the Internet via secure tunnels that are established between the different clusters' Gateways. These tunnels are set up automatically and adapted if the cluster mobility requires it.
PN Agent framework	The PN Agent Framework is responsible for registering and maintaining up-to-date information concerning major PN components, mainly in terms of availability, location and contact points. The key components that are registered within this PN Agent are the Cluster attachment points (i.e. Cluster gateways and Edge nodes if any are used), the MAGNET Service Management Protocol (MSMP) SMNs (Service Management Nodes) and the CMNs (Context Management Nodes) on the Secure Context Management Framework (SCMF).
PN/PN-F Directory Service	The PN Directory enables PN Federations. It acts as a trusted 3rd party by providing x509 certificates to MAGNET users. In addition the PN Directory can be used to store and publish PN federation profiles which contain information about PN federations and members of these federations; who created them and who maintains them (i.e. has the ability to add or remove members or edit member attributes). It is also designed to make it possible for people to use aliases instead of their real names.
Federation Module	Distributed functionality (mostly contained in the Federation Manager, FM), which manages the participation of the PN in PN Federations and the resulting PN-F profiles and participation profiles.
MAGNET Service Management Platform	MSMP contains components for the fulfilment of the service discovery and service provision processes.
Secure Context Management Framework	It constitutes a distributed agent framework which is dedicated to gathering, processing and distributing various types of information, commonly known as context information. The framework carries out the required functionality and provides context sensitive applications, services and other networking components, and provides easy access to context information.
MAGNET Air-Interfaces driver	This module provides the interface between the software based platform and the hardware network interfaces developed

Next, physically neighbouring PN nodes can authenticate each other and establish short-term link-level security associations based on the long-term pair-wise keys exchanged

during the imprinting. Direct secure communication is then possible at the link level. In order to be able to obtain IP communication, an address configuration protocol with duplicate address detection allows PN nodes to automatically generate a unique PN IP address from the private IP addressing space assigned to the PN. After the establishment of a secure link, ad hoc routing information is exchanged. The result of the above procedures is a secure and self-organising cluster in which PN nodes can communicate over one or multiple hops.

In order to get full PN connectivity, clusters at different geographical locations need to be interconnected through PN Gateway Nodes that have access to the Internet. To secure inter-cluster connectivity GW nodes will use CPFP over the insecure channel to derive a secure key which will be used to set up an IPsec tunnel between the clusters. A new PN entity called the PN Agent was designed and implemented for maintaining up-to-date information about all the PN cluster attachment points. This PN Agent provides name registration/deregistration/discovery, publish, subscribe and name resolution functions at PN and PN Federation level. During the PN formation process, the PN Gateway Nodes register themselves with the PN Agent (mainly in terms of attachment point to the Internet - public/private IP addresses and ports) and get, as a registration response, the location information of the Cluster Gateway Nodes of all the remote PN Clusters. This remote PN Gateway information will be maintained up to date by the PN Agent through binding updates. The PN Gateway information in the PN Agent is used to dynamically establish and maintain tunnels between the PN Gateway Nodes. Finally, after the exchange of routing information over these tunnels, full inter-cluster connectivity within the PN IP addressing space is possible, allowing secure communication between every pair of PN nodes.

Additional mechanisms have been implemented to improve communication. A universal convergence layer manages all network interfaces and hides the heterogeneity of the underlying interfaces from the routing layer and PN IP addressing space. Extensions have been implemented to be able to take into account NAT boxes. As well as unicast functionality, cluster-wide and PN-wide broadcasting functionality is also supported. Moreover, the combination of mechanisms to deal with dynamics (such as cluster splits and merges) and private PN addressing allows applications to maintain connectivity despite mobility. Finally, a PN Manager GUI presents the user with an interface to use, manage and monitor the implemented software. Users interact with the system; trigger service discovery, etc through this GUI.

On top of this network overlay, a service discovery and management platform (called MAGNET Service Management Platform, MSMP) and a Secure Context Management Framework (SCMF) are implemented. The MSMP offers the user viewing, managing and secure access to all PN resources and services. Its structure follows a twofold approach. A Service Management Node (SMN) is selected for each PN cluster. The SMN discovers and manages services within its cluster and interacts with other clusters' SMNs in a peer-to-peer fashion via a service overlay. This SMN is also responsible for discovering and

advertising remote services within the cluster. The Secure Context Management Framework provides access to all context and user profile information within a PN. The SCMF consists of context agents running on all PN nodes. Applications can access all information through the context agent running on their local node. Context agents access local context information through retrievers that provide a uniform interface to context sources. Examples of context sources are sensors, the networking stack, and the operating system. User profile information is stored in a storage component within a context agent. For accessing context information, the Context Access Language (CALA) is used. CALA provides a synchronous query/response as well as an asynchronous subscribe/notify interaction style.

In order to develop the PN-F concept, a mechanism is needed to define new PN-Fs and to add PNs to this PN-F, resulting in a PN-F creation and participation protocol. The PN-F Creator generates a PN-F Profile containing the main details of the PN-F (i.e. identification, means to proceed with the participation protocol and policies that rule the federation) and store it in the SCMF. The PN-F Profile is made public and candidates (i.e. other PNs) dialogue with the creator to see whether they are allowed to enter on the PN-F or not.

In order to proceed with the next step in the PN-F participation phase, the PN-F Creator and potential PN-F members (i.e. other PNs) need to be able to authenticate each other and to establish a security association that can be used to secure all ensuing communication. A new PN component, called Personal Network Directory Service [165], is also introduced as the identity provider (i.e. trusted third party entity). The PNDS, operated by a service provider, acts as a Certificate Authority (CA) providing X.509 certificates which associate public key with a particular user. The PNDS certificates are used by CPFP to establish bilateral trust relationships between the PNs that are later enforced each time the two PNs communicate under the auspices of any federation.

After this authentication and security association step, the PN-F member can actually join the PN-F. A PN-F participation profile that lists the services that the new PN-F member will make available within the PN-F is created and stored in the SCMF. At this stage, each member knows in which PN-Fs she/he participates, which other PNs are currently member of the PN-F and, optionally, what services are made available by these members. This information can in any case be retrieved through a PN-F wide service discovery mechanism since the MSMP implementation has been extended to support this feature too.

The concept of a network overlay selected to provide secure PN-F communication enables all PN nodes of the PN-F members to become part of the PN-F overlay. In order to separate the internal PN communication from any PN-F communication, every PN-F will also have its own PN-F addressing space (defined in the PN-F profile) and every node involved will obtain a unique PN-F IP address within this addressing space. In a similar way to the PN, the PN-F overlay will be established. Neighbouring clusters of different PNs are discovered through the use of beacons. When establishing secure associations with a PN, a pair-wise (1 key for each pair of PNs) primary master key is exchanged (using

the PNDS certificates through CPFP). This key is then used for deriving link-level session keys used to secure the link between nodes of different PNs. Using this secure link, PN-F routing information can be exchanged, forming a PN-F cluster. For the interconnection of clusters of PNs at different locations, the PN Agents of the respective PNs are used. All cluster location information can be retrieved by contacting the PN Agents of the other PN-F members. Tunnels are then established using the primary master key as the basis and routing information is exchanged, creating full end-to-end secure PN-F connectivity.

The service discovery framework is extended to allow PN-F service discovery and use. Higher-level SMNs, called PN-F Agents, are introduced. The PN-F Agent implements all the PN SMN functionality but is exclusively dedicated to storing and discovering PN-F resources and services at PN level. One PN-F Agent per federation is activated within a PN. The PN-F Agents of each participant interact in a peer-to-peer manner via a PN-F service overlay to provide PN-F wide service discovery according to PN-F participation profiles. The service related functions provided by the GUI are extended to the PN-F case.

In the PN-F case, the SCMF also provides access to context information from the members of the PN-F [166]. The SCMF of each PN has a dedicated Context Management Gateway (CMG). The CMGs interact with each other, exchanging context information, while enforcing the privacy policies of the user.

TESTBED DESCRIPTION

Testing the functionality of even a single PN needs several clusters with a reasonable number of devices in each one of them. Therefore, a minimum set of hardware devices was defined as a requisite to set up and test a PN and PN-F platform. Indeed, a fully operational system has now been built through a process of conformance verification, integration, and interoperability testing of the different baseline components described before.

Different partners are hosting different parts of the PN/PN-F platform in their premises, as shown in Figure B-2, with all individual parts interconnected via the Internet to form a distributed testbed, as a prerequisite to validate how the developed PN/PN-F solutions and pilot applications function over real-life network conditions.

In order to promote and ease the usage of the system among the partners and developers of pilot services in particular, a set of system installation and usage guidelines have been developed. The distributed testbed is set up in four different laboratories across Europe and has been the cornerstone for the integration process. The testbed is composed of both laptops and PDAs in order to showcase the feasibility of the system to be run on real user equipment. All the integrated components are implemented to run over the Linux Operating System. For the high capable devices like laptops, the Ubuntu distribution was selected while for the PDA-like devices, the project decided to use Nokia Internet Tablets. Accordingly, easy to install SW packages have been created for the two selected kinds of MAGNET nodes, while the respective installation guides are planned to set all software

from scratch. Thereby, the PN/PN-F Platform for pilot services is now a reality, and the necessary means have been set up to promote its use among application developers and potential end users.

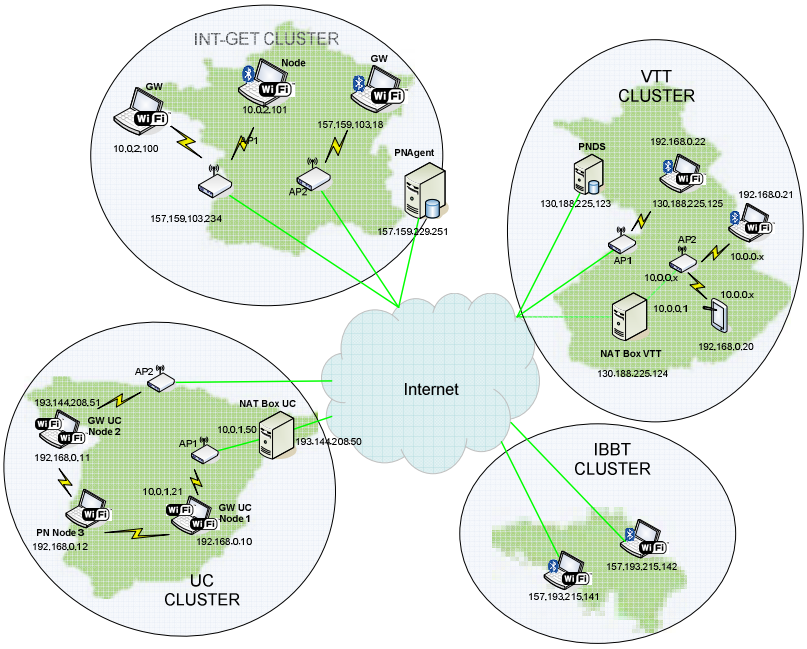


Figure B-2: Physical location of the remote testbed

TESTBED SCENARIOS AND OBJECTIVES

The PN/PN-F platform is currently being used for testing the developed PN and PN-F functionality, as described in the previous sections.

The availability of this dedicated "always on" testbed was the only viable way to guarantee that all participating partners can assess the real usability of the pilot applications and the performance of the underlying platform. Since the objective of the platform is not only to prove the feasibility of a PN system but to support the pilot services on top of it, it is possible to assess the usability of the PN concept from a user-centric viewpoint. Indeed, the same platform will evolve via further performance testing to serve also as the platform to support the pilot services.

For the actual component interoperability testing to be executed with the overall PN system, it is necessary to prepare and perform an interoperability testing plan. It includes end-to-end performance, robustness and reliability testing. In addition, it must be assured that the integrated components fulfil requirements about the system specification as well as supporting all the selected Pilot Services. A complete set of test cases (see Table B-2) were set up and carried out to guarantee the suitability of the solutions implemented and integrated.

Table B-2: MAGNET system prototype test scenarios

Test case	Components involved	Scope	Functionalities tested
Node personalization	Trust establishment module	Cluster	Imprinting of new personal node and transitive generation of long-term trust relationship with former personal devices.
Secure intra-cluster communication	Neighbour discovery module and UCL	Cluster	Secure link-layer establishment procedure and packet protection so that privacy, origin authentication and integrity are assured.
Cluster formation	Neighbour discovery module and PN/PN-F routing module	Cluster	Establishment of links to neighbouring personal nodes and routing information exchange so that all other nodes in the cluster have a route to the new personal node.
PN formation	PN/PN-F routing module, Dynamic tunnel establishment module and PN Agent framework	PN	PN Gateway node successfully registers within PN Agent. Establishment of tunnels between remote clusters and inter-cluster routing protocol is able to send messages over tunnel(s) so that a communication routes to all other nodes within the PN are created.
PN dynamics	PN/PN-F routing module, Dynamic tunnel establishment module and PN Agent framework	PN	Re-configuration of the overlay on presence of events caused by cluster mobility (cluster splitting and merging, etc.)
Network specifics	Dynamic tunnel establishment module and PN Agent framework	PN	Maintenance of the network overlay under special circumstances of the access network such as NAT traversal or firewalls.
Service discovery and usage	MSMP	Cluster / PN	Service/Application is registered within the service platform so that it is discoverable through its cluster's SMN. All the descriptions of PN services/applications that match specified service attributes are returned from the Cluster SMN in an SD response.

PN-F establishment (ad-hoc case)	Federation Manager, CPFP, PN/PN-F routing module, MSMP, SCMF	Cluster	PN-F Member discovers published federation and registers in it. Upon acceptance from the PN-F Creator the federation overlay is automatically configured on member's nodes (nodes from all PN-F members become part of the same federation cluster). If no trust relationship exists in advance common Certificate Authority certificates are used to set this relationship. Federation parameters are exchanged through corresponding profiles. Service and context is enabled if such resources are shared in the federation.
PN-F establishment (infrastructure case)	PNDS, Federation Manager, CPFP, Neighbour discovery, UCL, PN/PN-F routing module and Dynamic tunnel establishment	PN-F	PN-F Member fetches federation information from PNDS and starts federation establishment through the Internet. Upon acceptance from the PN-F Creator the federation overlay is automatically configured in member's nodes (Members' clusters are interconnected through tunnels that are dynamically established). If no trust relationship exists in advance common Certificate Authority certificates are used to set this relationship. Federation parameters are exchanged through corresponding profiles. Service and context is enabled if such resources are shared in the federation.
Secure context management	SCMF	Cluster / PN / PN-F	Access to context information (both synchronously and asynchronously) at all levels of the SCMF architecture.

For the development of the PN-F connectivity and networking, the same concept of a network overlay was used. As such, in order to support secure PN-F communication, all PN nodes of the PN-F members become part of a PN-F overlay. This means that essentially the same components, solutions and protocols are being used for both the PN and PN-F overlays. Therefore, similar test scenarios could be defined for evaluating the behaviour of the PN-F solutions at connectivity and network level.

PILOT SERVICES

The transparent and seamless federation of networks and services comprises a highly novel feature in the MAGNET project. By federating networks, the number of services potentially available to users increases dramatically as does the radius of action for

individual users. Services can now be based on collaborative behaviour among individual users and/or nodes in a networked infrastructure. The impact of physical distances diminishes as nodes become interconnected using MAGNET technology. MAGNET's selected Pilot Services centre on two main topics: 'The Lifestyle Companion' and 'The Ice-breaker'.

These two MAGNET pilot services intended to demonstrate the MAGNET project in its entirety are initially analyzed and described. This analysis focuses on their general importance for MAGNET as seen from a user's perspective. Through this analysis, three key issues are identified as being of special importance for the two pilot services; federation, trust and initiative.

The Lifestyle Companion service targets primarily diabetic users who have a desire to keep in good shape by using a fitness centre as well as monitoring and adjusting their blood glucose level. This service allows these users to automatically monitor and register their blood glucose concentrations and receive recommendations for insulin injections based on this input. The service also offers a "personal trainer" functionality by which the service acts as a fitness trainer guiding the user through fitness programs in a fitness centre, keeping track of repetitions, load settings, etc. This service comprises the following core MAGNET functionalities;

- Proximity-based PN formation (enabling the user to easily interconnect an number of MAGNET-enabled nodes within a PN).
- Location/context-aware service-discovery (providing the user with service-related information based on the current physical location of the user)
- LDR transmission (wireless transmission of low-rate data between MAGNET-enabled nodes)
- Automated proximity-based PN federation (enabling quick and easy inter-node communication when required)

The Icebreaker scenario is situated in a conference where data-handling workers e.g. journalists meet physically and exchange information. This scenario allows a mobile worker to cooperate with one or more colleagues using shared digital material and to share computational power of individual devices. The service also provides means for exchanging "digital business cards" easing the establishment of new business contacts and acting as a social 'icebreaker'. Whenever it were convenient for the user to view data on a larger screen than currently available (e.g. a small handheld device), the service offers in addition the possibility of 'beaming' information from one device to another. Large quantities of data on a PDA can in this way be projected to a larger (MAGNET-enabled) public screen for closer inspection. The scenario addresses:

- context awareness applied as physical access control: mobile device as ticket to the conference, and as settings change of the mobile devices according to the activities of the users

- Location/context-aware service-discovery in the shape of the 'Icebreaker' services, offered to the guests at the fair, and notification of the users of 'Icebreaker' according to the context.

- PN discovery based on filtered search in local area allowing the user to explore the presence of available PNs applied as social software embedded in a physical setting

- seamless high data rate transfer from device to device (a mobile unit to a screen) to demonstrate the 'Public Screen' concept and device discovery / device management and in general demonstrating the concepts: 'digital handshaking' and 'digital hospitality'

- exchange of a digital business card as demonstration of a temporary PN-F formation

- forming of a long term ad hoc PN-F for the purpose of collaborative work

APPENDIX C: PUBLICATIONS

BOOK CHAPTERS

Ernö Kovacs, **Luis Sánchez** (eds.) et al, "Chapter 3: Personal Networks and Personal Network Federations", My Personal Adaptive Global Net, Springer, To appear in 2009.

Juha Zidbeck, **Luis Sánchez** (eds.) et al, "Chapter 7 PN and PN-F Testbed Realisation", My Personal Adaptive Global Net, Springer, To appear in 2009.

PUBLICATIONS ON INTERNATIONAL JOURNALS

Luis Sánchez, Jorge Lanza, Luis Muñoz, Kimmo Ahola, Mikko Alutoin, "Securing the communication in Private Heterogeneous Mobile Ad-hoc Networks", To appear on Wireless Personal Communications Journal, Springer, To appear on 2009.

Luis Sánchez, Jorge Lanza and Luis Muñoz, "Extending Private Personal Area Networks to Personal Network Federations in Heterogeneous Ad-hoc Scenarios", Special issue on Personal Networks of Telektronikk journal, vol. 103, no. 1-2007, pp. 34–44, February 2007.

Luis Sánchez, Jorge Lanza and Luis Muñoz, "Experimental Assessment of a Cross-Layer Solution for TCP/IP Traffic Optimization on Heterogeneous Personal Networking Environments", Lecture Notes in Computer Science (Vol 4217), pp. 284–296, September 2006.

Dimitris M. Kyriazanos, Michael Argyropoulos, **Luis Sánchez**, Jorge Lanza, Mikko Alutoin, Jeroen Hoebeke and Charalampos Z. Patrikakis, "Overview of a Personal Network Prototype", IEC Annual Review of Communications (volume 59), pp. 521–534, 2006.

Luis Muñoz, Ramón Agüero, Johnny Choque, José Ángel Irastorza, **Luis Sánchez**, Marina Petrova and Petri Mähönen, "Empowering Next-Generation Wireless Personal Communication Networks", IEEE Communications Magazine, vol. 42, no. 5, pp 64–70, May 2004.

Ramón Agüero, **Luis Sánchez**, Johnny Choque, Luis Muñoz and José Ángel Irastorza, "WPANs Heading towards 4G", UPGRADE Vol. V, issue no. 1 WireIess Networks - Telecommunications' New Age, pp 63–68, February 2004.

PATENTS

Jordi Jaen-Pallares, **Luis Sánchez**, Jorge Lanza, Luis Muñoz, "Method for distributing a Certificate Revocation List in a network", request for a grant of a European patent, nº 08167197.6, October 2008.

PUBLICATIONS ON NATIONAL JOURNALS

Ramón Agüero, **Luis Sánchez**, Johnny Choque, Luis Muñoz and José Ángel Irastorza, "Las WPAN en el trayecto hacia la 4G", NOVATICA Nº 167, January-February 2004.

PUBLICATIONS ON INTERNATIONAL CONFERENCES

Jorge Lanza, **Luis Sánchez**, Luis Muñoz, "Experimental Comparison of Two Solutions for Securing Heterogeneous Ad-Hoc Network Communications", 11th International Symposium on Wireless Personal Multimedia Communications, Lapland, September 2008.

Luis Sánchez, Jorge Lanza, Jeroen Hoebeke, Ingrid Moerman, Mikko Alutoin, Kimmo Ahola, Jordi Jaen Pallares, Marc Girod Genet, Martin Bauer, Joachim Zeiss, "Assessing Personal Networks on a pan-European Testbed", ICT Mobile and Wireless Communications Summit 2008, Stockholm, June 2008.

Rasmus Olsen, Martin Bauer, **Luis Sánchez** and Jorge Lanza, "Self Organisation of Context Agents in Personal Networks and Federations", 10th International Symposium on Wireless Personal Multimedia Communications, Jaipur, September 2007.

Luis Sánchez, Jorge Lanza and Luis Muñoz, "Federating Personal Networks over Heterogeneous Ad hoc Scenarios", 12th IFIP International Conference on Personal Wireless Communications, Springer IFIP Series, vol. 245, pp. 38–53, September 2007.

Joachim Zeiss, **Luis Sánchez** and Sandford Bessler, "Policy driven formation of federations between personal networks", 16th IST Mobile & Wireless Summit Communications Summit, Budapest, July 2007.

Luis Sánchez, Jorge Lanza, Juan Rico and Luis Muñoz, "Implementation of a Vertical Handover Solution for Coexisting Ad-hoc Multihop and Infrastructure Access Networks based on IEEE 802.11", 7th International Workshop on Applications and Services in wireless Networks, pp. 15–21, Santander, May 2007.

Luis Sánchez, Jorge Lanza and Luis Muñoz, "Cluster Head Selection and Maintenance over Heterogeneous Mobile Wireless Personal Area Networks. An experimental approach", 9th International Symposium on Wireless Personal Multimedia Communications, San Diego, September 2006.

Luis Sánchez, Jorge Lanza, Martin Bauer, Rasmus. Olsen and Marc Girod-Genet, "A Generic Context Management Framework for Personal Networking Environments", 3rd Annual International Conference on Mobile and Ubiquitous Systems – Workshops, pp. 1–8, San Jose, July 2006.

Martin Bauer, Rasmus L.Olsen, Martin Jacobsson, **Luis Sánchez**, Jorge Lanza, Mohamed Imine and Neeli Prasad, "Context Management Framework for MAGNET Beyond", 15th IST Mobile & Wireless Summit Communications Summit, Mykonos, June 2006.

Jeroen Hoebeke, Gerry Holderbeke, Ingrid Moerman, Wajdi Louati, Wassef Louati, Marc Girod Genet, Djamal Zeghlache, **Luis Sánchez**, Jorge Lanza, Mikko Alutoin, Kimmo Ahola, Sami Lehtonen and Jordi Jaen Pallares, "Personal Networks: From concept to a demonstrator", 15th IST Mobile & Wireless Summit Communications Summit, Mykonos, June 2006.

Majid Ghader, Rasmus L. Olsen, Venkatesha Prasad, Martin Jacobsson, **Luis Sánchez**, Jorge Lanza, Wassef Louati, Marc Girod Genet, Djamal Zeghlache and Rahim Tafazolli "Service Discovery in Personal Networks; design, implementation and analysis", 15th IST Mobile & Wireless Summit Communications Summit, Mykonos, June 2006.

Jorge Lanza, **Luis Sánchez** and Luis Muñoz, "Performance Evaluation of a Cross-layer based Wireless Interface Dynamic Selection on WPAN/WLAN Heterogeneous Environments: An Experimental Approach", 6th International Workshop on Applications and Services in Wireless Networks, pp. 139–146, Berlin, May 2006.

Luis Sánchez, Jorge Lanza and Luis Muñoz, "Self-Configuring Private Personal Area Networks: The first stage towards Personal Networking", 8th International Symposium on Wireless Personal Multimedia Communications, Aalborg, September 2005.

Luis Sánchez, Jorge Lanza, Luis Muñoz and Julian Pérez, "Enabling Secure Communications over Heterogeneous Air Interfaces: Building Private Personal Area Networks", 8th International Symposium on Wireless Personal Multimedia Communications, Aalborg, September 2005.

Mikko Alutoin, Kimmo Ahola, Sami Lehtonen, **Luis Sánchez**, Jorge Lanza, Jeroen Hoebeke, Gerry Holderbeke, Ingrid Moerman, Marc Girod Genet, Wassef Louati, "Self-organisation and mobility in Personal Networks", 8th International Symposium on Wireless Personal Multimedia Communications, Aalborg, September 2005.

Mikko Alutoin, Sami Lehtonen, Jeroen Hoebeke, Gerry Holderbeke, Ingrid Moerman, **Luis Sánchez**, Jorge Lanza, Djamal Zeghlache, Wajdi Louati, "Towards Self-organising Personal Networks", 1st ACM Workshop on Dynamic Interconnection of Networks, Collogne, September 2005.

Ramón Agüero, Johnny Choque, **Luis Sánchez**, Luis Muñoz, Fabrice Lucas and Sylvie Colin, "Empowering Cross-Layer Optimization: Dynamic Context Awareness in a Heterogeneous Wireless Environment", 7th International Symposium on Wireless Personal Multimedia Communications, September 2004

Mirko Presser, R. Tafazolli, István Z. Kovács, D. Dahlhaus, J. Farserotu, F. Platbrood, **Luis Sánchez** and Karsten Schoo, "MAGNET 4G Personal Area Network Air-Interfaces for Personal Networks", 13th IST Mobile Communications Summit, Lyon, June 2004

Ramón Agüero, **Luis Sánchez**, Johnny Choque, Roberto Sanz, Luis Muñoz and José Ángel Irastorza, "On the Implementation and Experimental Characterization of the Dynamic Source Routing Protocol for Mobile Ad Hoc Networks", 6th International Symposium on Wireless Personal Multimedia Communications, Yokosuka, October 2003.

Marta García, Johnny Choque, **Luis Sánchez** and Luis Muñoz, "An Experimental Study of Snoop TCP Performance over the IEEE 802.11b WLAN", 5th International Symposium on Wireless Personal Multimedia Communications, Honolulu, October 2002.

PUBLICATIONS ON NATIONAL CONFERENCES

Luis Sánchez, Jorge Lanza and Luis Muñoz, "Redes Personales: la implementación del concepto", XVII JORNADAS TELECOM I+D, Valencia, September 2007

Ramón Agüero, Luis Muñoz, Johnny Choque, **Luis Sánchez** and Jorge Lanza, "Estudio experimental de los protocolos IP en redes inalámbricas multi-salto basadas en el protocolo DSR", V Jornadas de Ingeniería Telemática, Vigo, September 2005.

Roberto Sanz, Ramón Agüero, **Luis Sánchez**, Johnny Choque and Luis Muñoz, "Mejora de las prestaciones de la pila TCP/IP en entornos inalámbricos multisalto", IV Jornadas de Ingeniería Telemática, Gran Canaria, September 2003.

Ramón Agüero, Marta García, **Luis Sánchez**, Johnny Choque and Luis Muñoz, "Efecto Combinado de Técnicas de Nivel de Enlace Independientes en el Comportamiento de los Protocolos de Transporte de Internet sobre Redes de Área Local Inalámbricas", IV Jornadas de Ingeniería Telemática, Gran Canaria, September 2003.

Ramón Agüero, **Luis Sánchez**, Marta García, Johnny Choque and Luis Muñoz, "Análisis Experimental del Comportamiento de TCP sobre IEEE 802.11b y del protocolo Snoop como Mecanismo de Mejora", IV Jornadas de Ingeniería Telemática, Gran Canaria, September 2003.

Luis Sánchez, Verónica Gutiérrez, Ramón Agüero, Luis Muñoz, "Propuestas para el despliegue de redes de acceso inalámbricas de bajo coste basadas en tecnología WLAN", IV Jornadas de Ingeniería Telemática, Gran Canaria, September 2003.

OTHER CONTRIBUTIONS

Luis Muñoz, **Luis Sánchez**, Jorge Lanza, Mikko Alutoin, Sami Lehtonen, Djamal Zeghlache, Marc Girot Genet, Wajdi Louati, Ingrid Moerman, Jeroen Hoebeke, Gerry Holderbeke, Majid Ghader and Martin Jacobsson, "A proposal for Self-Organizing Personal Netowrks", World Wireless Research Forum meeting 15, Paris, December 2005.

Invited Conference Lecture MAGNET 2nd Training Event, **Luis Sánchez**, "Service and System aspects of Personal Networks beyond 3G: Networks and Interworking", Rome, November 2005.

Invited Conference Lecture MAGNET 1st Training Event, **Luis Sánchez**, "Evolution of Personal Networks towards 4G: Technology, Services and Markets: Personal Networks: Networking Technologies towards Pervasive Communications", Copenhagen, January 2005.

Invited Conference Lecture SympoTIC 03 - Joint WS NEXWAY & PRODEMIS - **Luis Sánchez**, Luis Muñoz, "Paving the way for the Next Generation: A new family of Wireless PANs. The PACWOMAN Case", Bratislava, October 2003.