HELSINKI UNIVERSITY OF TECHNOLOGY
Department of Electrical and Communications Engineering
Networking Laboratory

UNIVERSIDAD POLITÉCNICA DE MADRID
E.T.S.I. Telecomunicaciones

Juan Francisco Redondo Antón

Ad Hoc Networks – design and performance issues

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Telecommunications Engineering

Espoo, May 2002

Supervisor: Professor Jorma Virtamo

Abstract of Master's Thesis

Author:	Juan Francisco Redondo Antón
Thesis Title:	Ad hoc networks – design and performance issues
Date:	May the 28th, 2002 **Number of pages:** 121
Faculty:	Helsinki University of Technology
Department:	Department of Electrical and Communications Engineering
Professorship:	S.38 – Networking Laboratory
Supervisor:	Professor Jorma Virtamo

The fast development wireless networks have been experiencing recently offers a set of different possibilities for mobile users, that are bringing us closer to voice and data communications "anytime and anywhere". Some outstanding solutions in this field are Wireless Local Area Networks, that offer high-speed data rate in small areas, and Wireless Wide Area Networks, that allow a greater mobility for users.

In some situations, like in military environment and emergency and rescue operations, the necessity of establishing dynamic communications with no reliance on any kind of infrastructure is essential. Then, the ease of quick deployment ad hoc networks provide becomes of great usefulness. Ad hoc networks are formed by mobile hosts that cooperate with each other in a distributed way for the transmissions of packets over wireless links, their routing, and to manage the network itself. Their features condition their design in several network layers, so that parameters like bandwidth or energy consumption, that appear critical in a multi-layer design, must be carefully taken into account.

This work, with the aim of identifying open questions and research problems, is a literature survey on ad hoc networks that addresses some of their critical design and performance issues. As the throughput per node is identified as the limiting factor in their performance, the bounds that available network capacity can achieve are considered, as well as the different factors that impact on this capacity and, consequently, that suggest possibilities to increase it. Furthermore, the problems routing algorithms for ad hoc networks have to solve are investigated, in addition to the medium access control mechanisms. There exists a significant dependence between the medium access control layer and the network layer, and, indeed, a complete integration of the diverse design solutions available in the different network layers of the protocol stack is still not achieved. This thesis also includes a review about some possibilities for power-aware design and quality of service mechanisms in ad hoc networks.

Keywords: ad hoc networks, capacity, medium access control, routing protocols, power-aware design

Resumen de Proyecto Fin de Carrera

Autor:	Juan Francisco Redondo Antón		
Título del proyecto:	Redes Ad hoc – diseño y eficiencia		
Fecha:	28 de Mayo de 2002	**Número de páginas:**	121
Universidad:	Helsinki University of Technology		
Departmento:	Departmento de Ingeniería Eléctrica y de Comunicaciones		
Laboratorio:	Networking Laboratory		
Supervisor:	Jorma Virtamo		
Coordinador:	Ángel Álvarez Rodríguez		

El rápido desarrollo que las redes inalámbricas han experimentado en los últimos años ofrece a los usuarios diferentes soluciones que nos aproximan hacia las comunicaciones posibles en cualquier momento y desde cualquier lugar. Algunas de las opciones más destacadas en este campo son las redes inalámbricas de área local (WLANs), que aportan alta velocidad en la transmisión de datos, y las redes inalámbricas de área extendida (WWANs), que permiten una mayor movilidad para los usuarios.

En algunas situaciones, como en entornos de aplicación militar o en operaciones de emergencia, la necesidad de establecer comunicaciones dinámicas sin contrar con ningún tipo de infraestructura se convierte en imprescindible. Entonces, la facilidad de rápido despliegue que las redes "ad hoc", herederas naturales de las redes inalámbricas de conmutación de paquetes, proporcionan, resulta de gran utilidad. De hecho, hoy en día, como en los 70 —cuando comenzó el interés por las redes inalámbricas de conmutación de paquetes—, son los proyectos militares los que lideran la actividad investigadora en este campo, dado que la aplicaciones que pueden obtener mayor provecho de la flexibilidad y dinamismo de las redes ad hoc son militares por excelencia.

Las redes ad hoc están formadas por dispositivos móviles que cooperan los unos con los otros de manera distribuida para llevar a cabo la transmisión de los paquetes por los enlaces inalámbricos que forman la red, el encamiento de dichos paquetes y la gestión y el mantenimiento de la red misma. Sus peculiares características y limitaciones condicionan sobremanera el diseño en varios de los niveles OSI de red, de forma que parámetros como el ancho de banda o el consumo energético, que tienden a ser críticos en un diseño multinivel como el que parece apropiado para las redes ad hoc, deben ser tenidos en cuenta de manera especialmente cuidadosa.

Este Proyecto Fin de Carrera, con el objetivo de identificar cuestiones abiertas y problemas de investigación en este campo, es un estudio acerca de las redes ad hoc que se centra en temas críticos relativos a su eficiencia y diseño.

Palabras clave: redes ad hoc, capacidad, control de acceso al medio, protocolos de encaminamiento, limitaciones energéticas

El hecho de que cada nodo no sólo sea responsable de mandar sus propios datos por la red, sino que tenga además que reenviar el tráfico procedente de otros usuarios (naturaleza "multihop" de la red ad hoc), hace que el "throughput" o capacidad disponible por nodo se reduzca considerablemente. Se ha demostrado que los límites teóricos de la capacidad por nodo para una red ad hoc bidimiensional, en las mejores condiciones posibles —cuando los nodos están colocados óptimamente, el rango de transmisión es elegido igualmente de manera óptima y los patrones de tráfico son conocidos—, decrecen con el número de nodos n como $1/\sqrt(n)$. Este límite no es alentador en absoluto en términos de escalabilidad de la red. De hecho, ha sido observado en los análisis de eficiencia de los protocolos de encaminamiento ad hoc que es precisamente la capacidad el mayor factor coercitivo. Consecuentemente, el estudio de los factores que afectan a dicho "throughput" está totalmente justificado, y el conocimiento de los parámetros involucrados y su correcta elección permiten sugerir posibilidades para el incremento, en lo posible, de esta capacidad.

En este trabajo hemos identificado tres de estos factores: los patrones de tráfico, la localización y movilidad de los nodos y el rango de transmisión. Patrones de tráfico predominantemente locales constituyen el mejor caso en términos de capacidad, mientras que los patrones de tráfico aleatorios tienden a ser la peor situación posible para redes ad hoc. Así pues, esto sugiere que hay posibilidades de formar redes con un número de usuarios considerable y, al mismo tiempo, una capacidad por nodo razonable, puesto que los patrones de tráfico de carácter local son bastante comunes. Se ha demostrado que las estrategias que contemplan la inclusión de nodos en una red estática con la misión exclusiva de reenviar paquetes de otros no merecen la pena, porque el número de nodos adicionales por introducir en la red para obtener mejoras significativas en la capacidad es enorme. Sin embargo, los efectos del reenvío de paquetes no propios parecen ser interesante en el caso de redes móviles, y mecanismos como la diversidad de usuarios en el reparto de los paquetes, cuando gestionen de forma adecuada los altos retardos que involucran, pueden ser, en un futuro, una fuente de incremento potencial de la capacidad.

Otras técnicas como la recepción multipaquete (MPR), auspiciada por los continuos avances en el tratamiento de señal y por las tecnologías de múltiples antenas, han cambiado algunos de las consideraciones iniciales hechas por los protocolos de acceso al medio y afectarán pronto a su diseño. En términos de capacidad, la recepción multipaquete no cambia la ley asintótica del "throughput", aunque sí mejora su coeficiente, sobre todo en el caso de redes con alta densidad de usuarios.

El rango de transmisión es uno de los parámetros críticos en el diseño de las redes ad hoc. Por un lado, afecta al consumo energético de los dispositivos móviles, y por otro, al "throughput" y a la conectividad, de forma regula estas dos propiedades de manera inversa (cuando hace que aumente la primera, disminuye la segunda, y viceversa). En este sentido, es posible estimar, de forma teórica y también práctica, el rango de transmisión crítico, que es aquél toma el mínimo valor posible (para maximar el "throughput", puesto que as disminuye las interferencias) que garantiza la conectividad.

Con una perspectiva más global, hemos de decir que uno de los retos en el diseño de las redes ad hoc es la integración adecuada de los diferentes niveles de red y la asignación de responsabilidades para cada de ellos. En nuestro campo de estudio, hay varios detalles que impiden la independencia total en el diseño de la red, como propone la arquitectura OSI. No está aún claro si ciertas funcionalidades como el control de potencia, la seguridad o la calidad de servcio deberían ser gestionadas por una única capa de red o si se debería considerar una visión multicapa. Nuestra opinión es que, hoy en día, no es posible diseñar una red ad hoc sin tener en cuenta las relaciones multicapa que inducen ciertos factores como el ancho de banda restringido, los frecuentes cambios en la topología de la red y las limitaciones energéticas.

De todos modos, como es lógico, en la literatura sobre el tema aparecen diversas soluciones para los distintos niveles de red. En cuanto al nivel de control de acceso al medio (MAC), crítico en comunicaciones inalámbricas, la mejor posibilidad disponible para una implementación real parece ser, hasta la fecha, IEEE 802.11, aunque no fue específicamente diseñado ni para comunicaciones "multihop" ni para soportar eficientemente redes ad hoc. Esa es la razón por la que muchas otras soluciones han aparecido, siendo especialmente interesantes aquellas que prestan atención, además de a la eficiencia y a la gestión de la capacidad, al control de potencia y a la calidad de servicio.

Por otro lado, en el nivel de red, también hemos estudiado los problemas que los algoritmos de encaminamiento tienen que resolver en el caso de redes ad hoc. Un gran número de propuestas ha sido presentado, y las continuas mejoras en los protocolos llevan a soluciones cada vez más sofisticadas. Sin embargo, no está claro si es posible que un protocolo de encamienamiento cumpla al mismo tiempo las diversas necesidades de todos los posibles contextos de aplicación de las redes ad hoc. Por supuesto, la escalabilidad de los protocolos también ha de ser considerada para evaluar sus posiblidades de aplicación práctica. De cualquier manera, se está llevando a cabo una gran labor a nivel de algoritmos de encaminamiento, aunque quizás el mayor reto que tengan por delante sea la correcta integración en ellos de seguridad, gestión de la potencia o calidad de servicio.

La gestión eficiente del control de potencia en redes inalámbricas ad hoc puede reportar enomes beneficios, principalmente por su impacto sobre la vida de las baterías de los dispositivos móviles (otro factor limitante en las redes ad hoc) y sobre la capacidad de la red. Así pues, los ahorros energéticos y el uso eficiente de los recursos de la red son razones suficientemente importantes que hacen de la gestión de la potencia uno de los mayores retos de las comunicaciones inalámbricas. Varios mecanismos han sido propuestos en distintos niveles de red, dado que el control de potencia influye, a través del rango de comunicación empleado, tanto al "throughput" (puesto que afecta a los niveles de interferencia) como al enrutamiento (a través de la conectividad).

Nuestro estudio se centró básicamente en los protocolos de encaminamiento y de control de acceso al medio que tienen como criterio principal de diseño las restricciones energéticas a las que la red está sometida. La complejidad de su implemtación puede ser bastante alta e incluso puede conllevar carga adicional para los protocolos, pero se ha observado que estos mecanismos pueden mejorar las prestaciones de la red, sobre todo si se tienen en cuenta algunas métricas que han sido definidas para evaluar la eficiencia energética de los protocols.

Igualmente, se hacen necesarias soluciones que lleven a soportar calidad de servicio dentro del contexto de las redes ad hoc, ya que muchas de las posibles aplicaciones tienes requerimientos de retardo o ancho de banda. El comportamiento dinámico de las redes ad hoc, manifestado en variaciones constantes del estado de los enlaces y en cambios en la topología de la red, dificulta de gran manera garantizar calidad de servicio. Para hacerlo, la coordinación entre los niveles de red y de acceso al medio se hace fundamental, así como la posible incorporación de un sistema de señalización encargado de la reserva de los recursos.

Todos estas cuestiones y características de las redes ac hoc, la mayoría de las veces relacionadass las unas con las otras en el diseño de la red, suponen un gran incentivo para la comunidad investigadora, que habrá de tener un papel decisivo para proporcionar una arquitectura flexible, eficiente y estable para las redes ad hoc, que, sin duda alguna, revolucionarán las comunicaciones inalámbricas en un futuro cercano.

Acknowledgements

This Master's Thesis has been written at the Networking Laboratory of the Helsinki University of Technology as a part of the AHRAS project funded by the Finnish Army. The Thesis has also been financially supported by a grant from the European Union within the Erasmus mobility program.

First of all I would like to thank professor Jorma Virtamo for giving me the opportunity of researching in this laboratory and for his supervision and contribution to this Thesis. Working here was an interesting experience and I want to remember also all the personnel of the lab, especially Aleksi, with whom I worked in the same research project and who helped me with the language corrections, and Eemeli, my office mate, with whom I have shared laughter and many hours of work. I hope that my stay here has been valuable for other persons, as the presence of my Spanish predecessors in the lab —Juan, Moisés and Nacho, who is brave enough as to settle in Finland— was for me.

Likewise, I want to thank Ángel Álvarez and all the people in charge of the Erasmus issues, for a task that is absolutely recognized by all the students who were given the chance of having such a great experience within this exchange program.

I have to say that I consider myself fortunate to have met so great people in Finland: Javi, Iván, Albert, Christian, Simone, Carlos, Alfredo ... Also Ales, my flatmate and good friend during these months. The atmosphere in the group was great and we have spent wonderful moments together. Thanks all of you, guarros!

I would also like to remember all the unforgettable moments I have spent during my years studying in Madrid. Now that I am finishing, I want to dedicate it to all my friends there (Peña de Trago forever!), for being like they are. Also to all the chiquis that made every weekend funny in places like Chatarra, Milenium, Black Jack and, recently, Helsinki Club.

Finally, I would like to thank my parents and my sister, María, for everything they have done during these last 23 years, for their love, their guidance, their understanding, and their constant and unconditional support. Thanks for so many things ...

May the 28th, 2002

Espoo, Finland

Juan Francisco Redondo Antón

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This master's thesis deals with some important issues concerning ad hoc networks, natural inheritors of packet radio networks. This literature survey analyzes relevant topics in different network layers, such as the capacity they can work with and the parameters that affect it. As well, other design issues such as wireless problems (specially Medium Access Control protocols), routing characteristics, power-aware operating and quality of service are explained.

1.1 Purposes and Scope of Study

The great increase of the popularity of mobile wireless devices in recent years has resulted in a continuous improvement of their power, memory and communication features, nowadays common in business and personal life. Simultaneously, the possibilities of communication of these devices have grown, so that, once assumed their capability to access public or private networks such as Internet —by using wireless networks adaptors—, the expected desire consists on achieving the faculty of interconnection between these mobile terminals in order to share information.

The circumstances of these communications can vary tremendously depending on the context, and the traditional infrastructure networks are not functional in some occasions, for instance, when the network has to be established quickly or when the terrain becomes inhospitable.

Therefore, the demand of easily deployable infrastructureless networks is utterly justified. Such networks are called ad hoc networks (AHNs) and they possess the important characteristic of having decentralized topology,

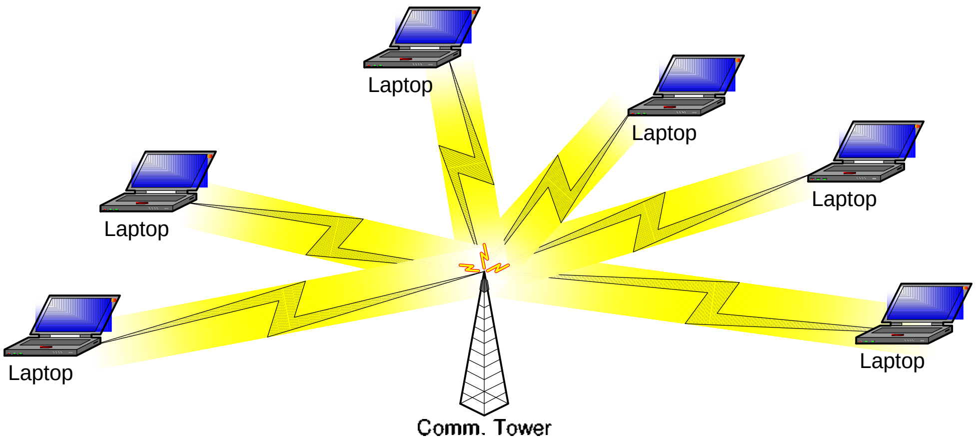

Figure 1.1: A fixed network

so that the devices don't need the use of any kind of fixed network element for the communication between them, but they administrate the network resources in a distributed way. With this decentralized architecture, they should be able to adapt themselves to changes in the number of the nodes, their location and the traffic pattern requirements. Furthermore, the mobility of the nodes is another relevant point to allow for, and two communicating endpoints, which may be travelling too, should be unaware of any other mobility in the network.

The wireless nature of the links used in ad hoc networks and the fact that the nodes employ a part of their resources (bandwidth, power, etc.) to send other nodes' packets, limits the throughput available for each node. When the number of nodes becomes high, scalability problems appear and this issue turns out to be critical. Therefore, capacity should be managed properly and the parameters affecting it should be designed carefully.

All these attributes that define the behavior of this kind of networks have important implications in their engineering, affecting to several network layers. Obviously, the form of routing, for instance, is especially specific, since two nodes that may establish an information exchange might not be able to communicate directly, so the routing scheme becomes multihop and the cooperation between stations turns out to be fundamental in order to manage mobility and to deliver packets. Therefore, in this environment, the hosts need to behave also as routers, because they should forward the packets they receive from their neighbors by following a particular routing algorithm.

Other important aspects may appear as requirements in the design, such as power constraints of the devices, security of the transmissions and quality

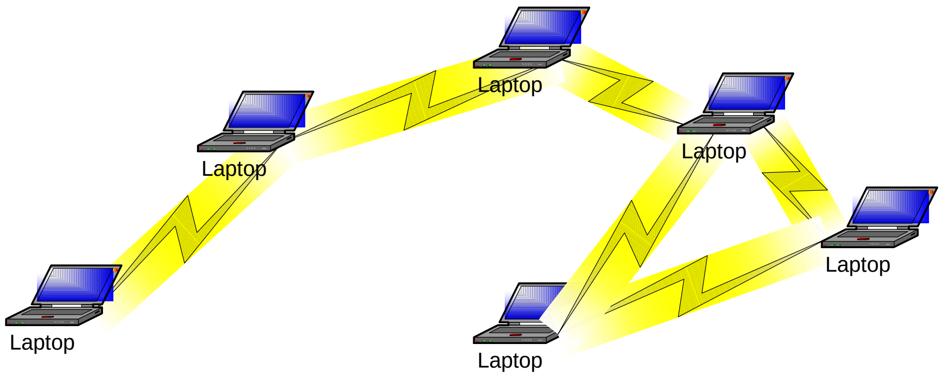

Figure 1.2: An example of ad hoc network

of service in the delivery of the packets, so they have to be taken into account too. Consequently, they should be integrated in the architecture of the network to perform satisfactorily for the framework they were planned for, even when the size of the networks happens to be large.

Managing efficiently power control in wireless ad hoc networks may carry important benefits, mainly because of its impact on battery life of the devices and on the carrying capacity of the network. Thus, energy savings and efficient use of the network resources are important reasons that make power management one of the most challenging problems is wireless communications.

This work, with the aim of identifying open questions and research problems, is a literature survey about ad hoc networks that addresses some of their critical design and performance issues. As the throughput per node is identified as the limiting factor in their performance, the bounds that available network capacity can achieve are dealt with, as well as the different factors that impact on this capacity and, consequently, that suggest possibilities to increase it. Furthermore, the problems routing algorithms for ad hoc networks have to solve are investigated, in addition to the medium access control mechanisms. There exists a significant dependence between the medium access control layer and the network layer, and, indeed, a complete integration of the diverse design solutions in the different network layers of the protocol stack has not been achieved yet. This thesis also includes a review about some possibilities for power-aware design and quality of service mechanisms in ad hoc networks.

1.2 Structure of the thesis

In chapter 2, a general overview about wireless communication, medium access techniques and the difficulties they inherently have, is given, with the aim of understanding the implications they have for ad hoc networks and the criteria we can take into account when selecting a medium access control protocol for them.

Chapter 3 deals with the evolution of ad hoc networks from the packet radio networks, as well as their main features, advantages and drawbacks and the applications they may be useful for.

Chapter 4 describes the issue of capacity in wireless ad hoc networking, as well as the theoretical limits presented in the literature. Furthermore, we also analyze the diverse factors that have a bearing on it, focusing on the important issue of the selection of the transmission range.

In chapter 5 we introduce the basic concepts concerning routing in ad hoc networks and, after a classification of them, we describe and compare the way of functioning of some of them. Furthermore, we consider mobility as an critical parameter for the routing protocols, and explain the mobility models that can be taken into account in simulations and then describe some protocols intended for high mobility.

Chapter 6 examines the power control mechanisms available in the context of ad hoc networks. It focuses on the protocols and algorithms that appear in the literature in the Medium Access Control and Network layers, as well as the possible power-aware metrics for them.

Chapter 7 summarizes the possible strategies for Quality of Service (QoS) support in ad hoc networks.

In chapter 8 we show the results obtained by simulation of the probability of having a fully connected network as a function of the transmission range.

Finally, in chapter 9 we present the main conclusions of our work, as well as possible alternatives for future research about problems that are still open in the design of ad hoc networks and that, if they are properly solved, can improve the performance of ad hoc networks and make them a suitable solution for certain communication contexts.

Chapter 2

Wireless Networks

In this chapter, a general overview about wireless communication, medium access techniques and the difficulties they inherently have, is given, with the aim of understanding the implications they have for ad hoc networks and the criteria we can take into account when selecting a medium access control protocol for them.

2.1 Introduction to Wireless Networks

The rapid growth of the Internet and the development of a great amount and variety of applications that are open to be used by plenty of final users worldwide, have made the communications and networking scenario undergo a change. The generalized access to the information highways involves not only new opportunities for the investors, but a challenge for the research community in order to provide the suitable means for the communications without restrictions.

In this sense, wireless networks are the best candidates to offer, with all the challenges it implies, mobility and location-independent multimedia communications. Even so, the great increase in its popularity is recent —we should remember that the foundations of the technology have existed commercially for more than ten years— and it mainly arises from two causes [1]: the recent availability of affordable, portable, low-power, wireless communications devices due to hardware improvements and, on the other hand, the ubiquitous and affordable access to telephone and data networks caused by the quick deployment of communications infrastructure.

Thus, various emerging wireless products, with a great standardization effort, launch out into the capture of the market with different solutions and proposals in several situations, from high-speed km-range wireless access (such as LMDS) to low-speed very-short-range device interconnection (IrDA —Infrared Data Association—). Here we present briefly some of these products, which utilize different wireless medium access techniques, some of those will be considered in the following section:

- **Broadband Wireless** (BW) allows simultaneous delivery of voice data and video. It is considered a competing technology with Digital Subscriber Line (DSL). It is normally used in metropolitan areas and call for a direct sight between the transmitter and the receiver. BW includes two different technologies operating in licensed frequency bands:
  - LMDS (Local multi-point distribution service) is a high bandwidth wireless networking service in the 28-31 GHz range with a coverage average distance of one mile from the LMDS transmitters.
  - MMDS (Multi-point distribution service) works in the 2 GHz licensed frequency bands and it has wider coverage than LMDS (up to 35 miles), but its throughput rates are lower.
- **Wireless Wide Area Networks** (WWAN) are computer data networks that may extend over a large geographical region. As they were designed, like traditional analog networks, for voice rather than data transfer, they have some inherent problems, although some 2G (second generation) and new 3G (third generation) digital cellular networks are fully integrated for data/voice transmission. With the advent of 3G networks, transfer speeds should also increase greatly.

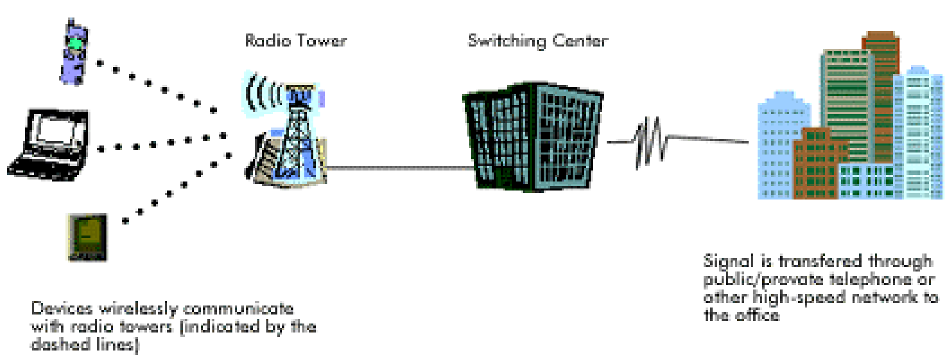

Figure 2.1: An example of Wireless WAN

- **Wireless Personal Area Networks** (WPAN) [2] are becoming essential when personal devices like mobile phones, computers and PDAs need to share data, have access to the Internet, share peripherals and network in other ways. Technologies such as Bluetooth [3] introduce the opportunity of creating networks with all kind of computing and communication devices in order to form ad hoc Wireless Personal Area Networks.

- **Wireless LANs** (WLAN) [4] are implemented as an extension to wired LANs and they offer high-speed data communication in small areas such as a building, an office or a campus. They let users move around in a confined area within the coverage of the network without experiencing disruption in connectivity. Therefore, compared to wired LANs, Wireless Local Area Networks provide installation and configuration flexibility and the freedom inherent in the network mobility.

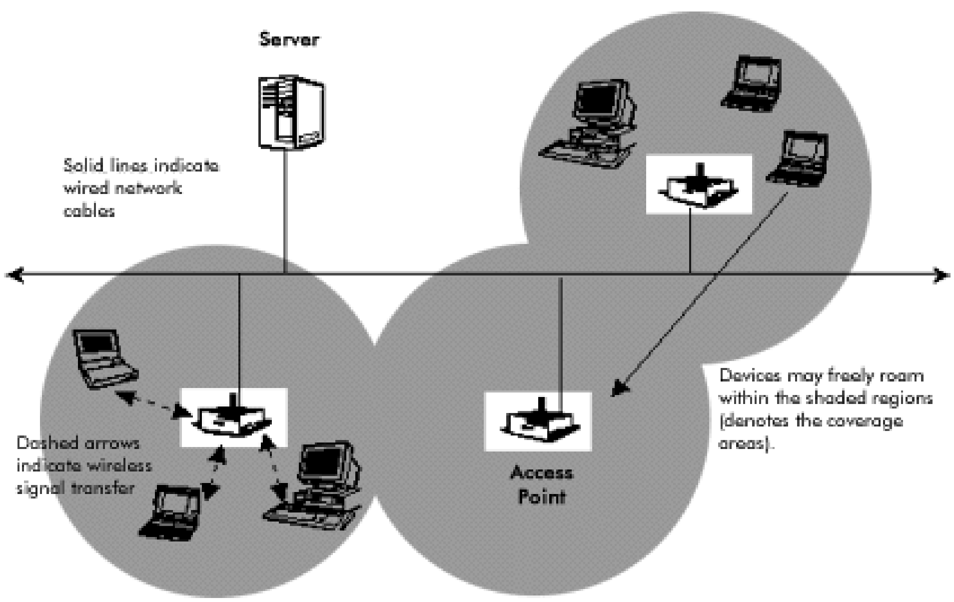

Figure 2.2: An example of Wireless LAN

All of them have to cope with the constraints of the wireless medium, such as electromagnetic interference, physical obstacles and the management of such a scarce resource as bandwidth is. Besides, topics such as standardization, data security and power consumption and ease of use of the devices are to be considered too.

2.2 Wireless Medium

The FCC allocated in 1985 three different frequency bands, without license requirements, for Industrial, Scientific and Medical (ISM bands)purpose operating respectively in 900, 2.4 and 5.7 GHz. These bands have been typically utilized for wireless LAN applications. Nevertheless, until 1997 the first internationally interoperability standard, IEEE 802.11 [5], was not approved. It offers not one standard, but three: two for spread spectrum (direct sequence and frequency hopping) and one for infrared transmission technology. Afterwards, others standards such as HiperLAN 1 [6] and 2 [7] have been developed and plethora of new techniques and modification proposals have been brought out. As for the demanded increase of capacity, the introduction of smart or adaptive antennas for mobile communication systems is becoming very attractive, not only due to the great rise of the capacity (3 times for TDMA and 5 times for CDMA), but also for the radio coverage enlargement, the better security and the possibility of introducing new services based on user location. As Lenhe and Pettersen present in their detailed survey [8], smart antenna technology is founded on the dynamic adaptation of the generated antenna beam to the current radio conditions, by making the radiation pattern be maximized in the desired direction.

2.2.1 Transmission Technologies

The diverse transmission methods that can be used are the following ones:

- **INFRARED (IR)**
  Infrared systems fall back on very high frequency (VHF) electromagnetic waves just below those for visible light to convey data. Although the common way of transmission is through a direct line of sight (LOS), the utilization of omnidirectional transmitters by reflecting signals off walls with the use of diffuse technology allows the formation of multihop networks. Any type of transmission can be sent regardless of the protocol —ATM, Fast Ethernet, Token Ring— because infrared operates at the physical layer of the OSI model.

  This technology is simple and inexpensive and can achieve 10 Mbps data rates, however, the spectrum sharing with the sun and fluorescent lights is a source of heavy interference.

- **MICROWAVE (MW)**
  Microwave transmission is between infrared and spread spectrum in terms of range and throughput. It is capable of generating data rates of 10 Mbps and sending signals several miles. It is most commonly used

in building-to-building applications. Limitations such as environmental restrictions (need of a clear LOS), high cost, license requirement (in the 18 GHz band) and safety issues have subdued the growth of microwave technology. Future challenges for microwaves are related to the development of products in the 5 GHz ISM band (dedicated exclusively to high date rate communication systems), where RadioLAN [9], by using narrowband technology, takes advantage of factors such as greater available bandwidth and reduced sources of interference.

- **RADIO**
  Spread Spectrum based systems use Radio Frequency (RF) transmission as the physical medium. Two major techniques exist to spread the required bandwidth over a large range of frequency than would be necessary to simply transmit the data [10]:

  - Direct Sequencing Spread Spectrum (DSSS) uses a wide frequency band in order to transmit every bit modulated onto a pseudo-random spreading code (chipping code), what involves that the system is less sensible to bit errors and the risk of eavesdropping is low. If it is utilized together with Code Division Multiple Access (CDMA), several LANs can share the same band —a study on distributed assignment of codes for multihop packet-radio networks was done in [11].
  - Frequency Hopping Spread Spectrum (FHSS) modulates the data on a carrier whose frequency changes to different channels (1 MHz each one), if they are free, in a pattern known to both transmitter and receiver. With this scheme the effects of interference are belittled and in terms of security the communication is difficult to be intercepted.

2.2.2 Multipath Interference

In radio transmissions, it is very important to take into account the called multipath interference (see Figure 2.3), which occurs when a signal bounces off of walls or other barriers and the interferences that are generated reach destination at different times. This kind of interference affects IR, RF and MW systems. FHSS intuitively resolves this problem by hopping to other frequencies. Other systems, such as array antennas, apply algorithms to avoid this. A subset of multipath interference is Rayleigh fading [12]. This happens when signals arrive from different directions and the difference in path length is a multiple of half the wavelength. This can utterly cancel the signal, so it is a source of potential danger. Rayleigh fading does not affect IR because the wavelengths used in IR are very small.

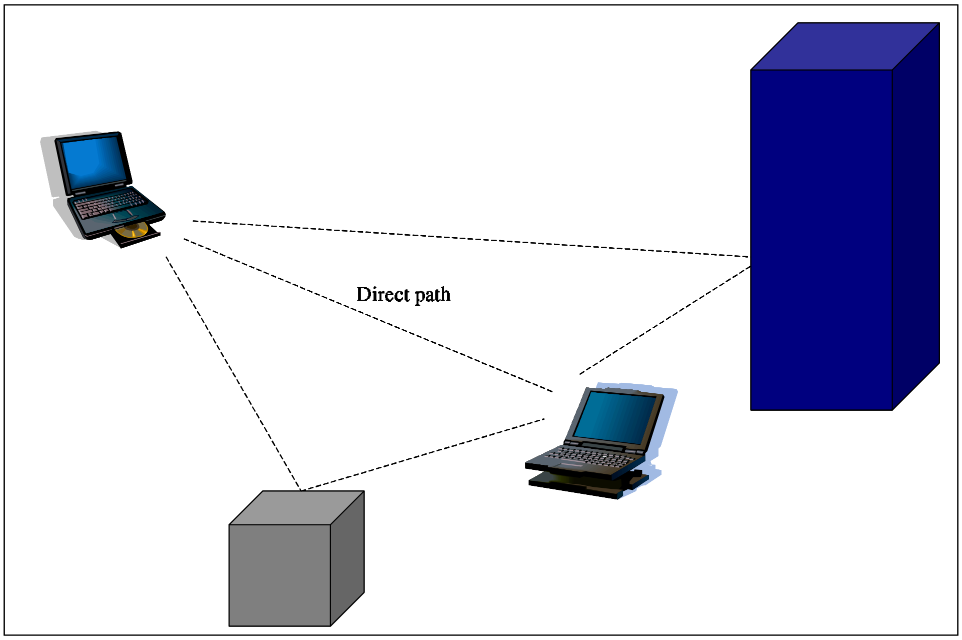

Figure 2.3: Multipath interference

2.3 Medium Access

In wireless networks, spectrum is a scarce and precious resource, so that the capacity is to be managed properly in order to offer as much bandwidth as possible for the applications running on the systems. Hence, the selection of the wireless medium access control protocol is an important issue to take into account. Therefore, first of all, the requirements a wireless network MAC Protocol should fulfill (see [13] and [14]) must be taken into account:

- *Delay* characteristics are important in every application, but especially because ad hoc networks should serve not only the mandatory asynchronous data service but also time-bounded multimedia applications such as voice and video. Delay can also cause problems for all data services where the preservation of the sequence of packets is extremely important.

- *Throughput* is definitely one of the most critical considerations in the design. We should consider not only theoretical throughput, but also operating throughput (which, practically, is more important).

- *Transparency to different PHY layers*, that can be achieved by a Physical Dependent Layer, a Physical Convergence Layer, and an appropriate MAC-PHY interface in each station, like IEEE 802.11 does.

- *Fairness of access.* Situations in which some stations have a preference to access the channel should be avoided.
- *The maximum number of nodes and highest coverage area* should not be limited by the MAC protocol, which should be able to handle the geographical size and number of nodes network without revealing appalling degradation of service. In this sense, scalability is considered a desired property for the network.
- *Robustness against co-channel access and interference* is a big challenge that, if properly faced, may allow several networks operate in the same region.
- *Power consumption* is to be allowed for, because most wireless devices, such as sensors in an ad hoc network, have limited battery power.
- *Support for multimedia* becomes fundamental for many applications nowadays, owing to the high level of convergence of data, voice and video.
- *Security* of the transmissions should be provided.

Of course, in the ad hoc network case we want to face, the medium access control protocol should manage appropriately *multihop* transmission, as well as the ability of supporting *multicast*.

Likewise, the lack of centralized control (as cellular networks have) hampers the use of TDMA or dynamic assignment of frequency bands; FDMA is inefficient in dense networks; and the node mobility and the subsequent time-varying neighborhood makes CDMA difficult to implement. Therefore, random access turns out to be the best option for the medium access control protocol in AHNs.

Classical Problems

In most wired LANs, like IEEE 802.3 (Ethernet), Carrier Sense Medium Access with Collision Detection (CSMA/CD) [15] is the common protocol used in the MAC level. The stations listen to the medium before transmitting until the medium is spare. While they are putting information on the air, they can detect packets collisions and if this happens, they abort the transmission, wait a random period of time and follow the same procedure once and again. The collision detection technique cannot be applied when the medium is wireless, since one node's transmitting signal masks any others arriving at the node. Moreover, carrier sense mechanisms turn out to be inadequate, due to the hidden terminal and the exposed terminal problems, which are shown in Figure 2.4.

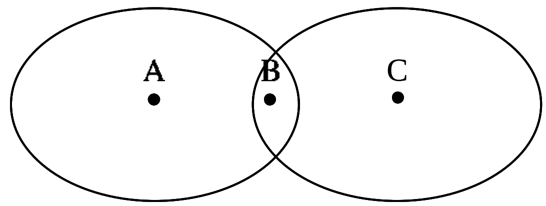

Figure 2.4: The hidden terminal problem and the exposed terminal problem

Stations C and A cannot hear each other because they are out of their communication range. Hence, they are unconscious potential competitors to sent data to B and the information may be lost because the carrier sense becomes ineffective. This is the *hidden terminal problem*.

An *exposed node* is one that belongs to the neighborhood of the sender but out of the range of the receiver. Let's suppose that node B is transmitting or trying to transmit to A, so that, when node C listen to the channel because it also wants to send information, it can hear the transmission made by B and thinks that the channel is busy. This fact denies any other parallel exchange of data between C and another neighbor station (like D) out of the range of B. Consequently, the bandwidth is underused.

After CSMA [16] —the first protocol to be used in multihop packet-radio networks—, plenty of protocols have been proposed in order to solve the hidden terminal and exposed terminal problems. Now we present some of the options.

2.3.1 MACA Protocol

MACA (Multiple Access Collision Avoidance) [17] was used as the starting point for the IEEE 802.11 Wireless LAN standard. It is based upon a control messages exchanges between the source and the sink (Figure 2.5). When a station A wants to transmit DATA, it sends to the receiver B a REQUEST-TO-SEND (RTS), which is answered with a CLEAR-TO-SEND (CTS) if B is prepared for the communication. If it is not, A will not hear the CTS and, after a time out, will consider that a collision occur and use a binary exponential backoff (BEB) algorithm in order to know when to send a new RTS. When finally possible, A dispatches the data while every B neighbor (that received the CTS) is silent. Meanwhile, sender neighbors can transmit to other possible destinations if they were not in the range of B (they did not receive the CTS).

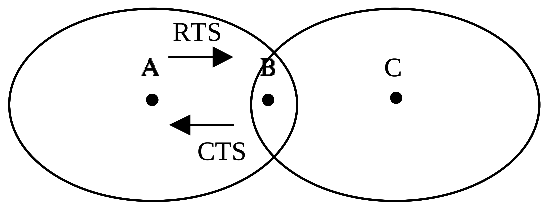

Figure 2.5: RTS-CTS dialogue

2.3.2 MACAW Protocol

MACAW [18] was designed to look up the efficiency of MACA in several manners. First, by introducing a more complex packet exchange that includes an ACK frame after each successful reception in order to quickly detect the lost frames (MACA waits for the transport layer to become aware of it). MACAW also adds carrier sensing for stations that are trying to send a RTS. Moreover, with the aim of improving the system performance, some modifications of the backoff algorithm and additional information exchange for congestion situations were included.

2.3.3 IEEE 802.11

IEEE 802.11 [5] standard sets up specifications on the parameters of both the physical (PHY) and medium access control (MAC) layers for two different type of networks: ad hoc (in the sense of distributed) and client/server networks. The PHY layer, which actually handles the transmission of data between nodes, can use either Direct Sequence Spread Spectrum (DSSS), Frequency Hopping Spread Spectrum, or infrared (IR) pulse position modulation. IEEE 802.11 supplies data rates of either 1 Mbps or 2 Mbps.

The MAC layer is a set of protocols that is responsible for maintaining order in the use of a shared medium. The 802.11 standard specifies a carrier sense multiple access with collision avoidance (CSMA/CA) protocol. In this protocol, when a node has a packet to be transmitted, it first listens to ensure no other node is transmitting. If the channel is clear, it then transmits the packet. Otherwise, it chooses a random *backoff factor* which determines the amount of time the node must wait until it is allowed to transmit its packet. During periods in which the channel is clear, the transmitting node decrements its backoff counter. When the backoff counter reaches zero, the node transmits the packet. Since the probability that two nodes will choose the same backoff factor is small, collisions between packets are minimized.

The handshaking procedure applied in order to prevent the disruptions caused by the hidden terminal problem is similar to the MACA packet exchange, but including acknowledgements when the reception is successful.

Security requirements are addressed in the standard as an optional attribute. The data security is accomplished by a complex encryption technique called Wired Equivalent Privacy Algorithm (WEP), that, as it is used at the two lowest layers of the OSI model, it does not offer end-to-end security.

Power management is supported at the MAC level for those applications requiring mobility under battery operation. Provisions are made in the protocol for the portable stations to go to low power *sleep* mode during a period of time previously defined.

There are different specifications in the 802.11 family:

- **802.11a** is an extension to 802.11 that applies to Wireless LANs and provides up to 54 Mbps in the 5.2 GHz band. 802.11a uses an OFDM (Orthogonal Frequency Division Multiplexing) encoding scheme rather than FHSS or DSSS.
- **802.11b** (also referred to as 802.11 High Rate) becomes an extension to 802.11 that applies to wireless LANS and provides 11 Mbps transmission (with a fallback to 5.5, 2 and 1 Mbps) in the 2.4 GHz band. 802.11b uses only DSSS. 802.11b was 1999 ratification to the original 802.11 standard, allowing wireless functionality comparable to Ethernet.
- **802.11g** is a contentious standard (not approved yet and with design criteria differences among vendors) that will also applicable to wireless LANs and provides 54 or more Mbps in the 2.4 GHz band. It is expected to be commercially available during 2002.

Unfortunately, due to different radio frequency and modulation, the 802.11 standard offers no provisions for interoperability between the different physical layers, so that end users equipped with radio cards of one type are not able to connect with the other types access points. Therefore, the decision of using the suitable standard for every particular case should be analyzed. A rough approximation lead us to conclude that 802.11a is better when the performance is a key feature (i.e. when there is a need to support end applications such as video, voice and the transmission of large images and files) and RF interference in the 2.4 GHz band is probable (wireless phones, Bluetooth [3], etc.). On the other hand, higher range capabilities and the lack of necessity to migrate to IEEE 802.11a (when investments have been done in 802.11b, first in the market) may suggest the use of 802.11b. Anyway, some vendors are already proposing dual 802.11a/b chipsets that can detect the solution used by the other communication point and act by employing the proper standard.

To make this standardization issue even more complex, IEEE 802.11i sub-committee was formed to address security concerns within the standards, as well as 802.11e draft specification adds quality of service (QoS) features and multimedia support to the existing 802.11b and 802.11a standards.

2.3.4 HiperLAN

HiperLAN (High Performance Radio Local Area Network) is a set of wireless local area network (WLAN) communication standards mainly used in European countries. There are two specifications: HiperLAN/1 [6] and HiperLAN/2 [7]. The European Telecommunications Standards Institute (ETSI) has adopted both of them. We will explain basically HiperLAN type 1, which offers 24 Mbps, because it was intended to be from the beginning a multihop MAC protocol, so it may be useful for ad hoc networks.

The HiperLAN/1 MAC protocol is based on a carrier-sensing mechanism, but is quite different in its details from that used in the IEEE 802.3 standard (Ethernet) or the IEEE 802.11 standard discussed earlier. If the medium has been sensed as free for a sufficient length of time immediate transmission is allowed. If not, a channel access cycle is formed to reduce the probability of collisions, so that only the higher priority traffic can be transmitted in the first phase (prioritization phase) and the rest of the traffic contends within the elimination phase to transmit in the last phase (yielding phase), which ends with the reception of an acknowledgement.

The HiperLAN technical committee explicitly wanted to support Quality of Service (QoS) for packet delivery. QoS support is provided via two mechanisms: the priority of a packet (high or normal) and the packet lifetime, whose initial values are assigned by the application using the HiperLAN services. The residual lifetime of a packet together with its priority are used to determine its channel access priority. Channel access priority can fall into one of five categories, and is used for the prioritization phase. No other explicit mechanism is used to support the desired QoS, unlike the time-bounded services of the IEEE 802.11 standard.

There is support for packet encryption in the HiperLAN packet transmission mechanism. The standard stays away from defining the particular encryption method used, but defines methods to inform the receiver which of a particular set of encryption keys has been used to encrypt the packet.

2.3.5 BLUETOOTH

Bluetooth [3] is a new short-range radio technology for connecting voice or data capable devices. Originally Bluetooth was designed just as a cable-replacement technology for devices within a small area (in the same room), forming between them only point-to-point radio connections. But Bluetooth is currently evolving to become also an important networking technology, enabling the formation of short-range wireless networks.

This has lead to a new networking category: The Personal Area Network (PAN). This concept is becoming an essential networking term, when personal devices like mobile phones, computers and PDAs need to share data, have access to the Internet, share peripherals and network in other ways. Bluetooth brings the possibility of creating ad hoc wireless Personal Area Networks [19], formed by all kind of computing and communication devices, which will incorporate Bluetooth technology.

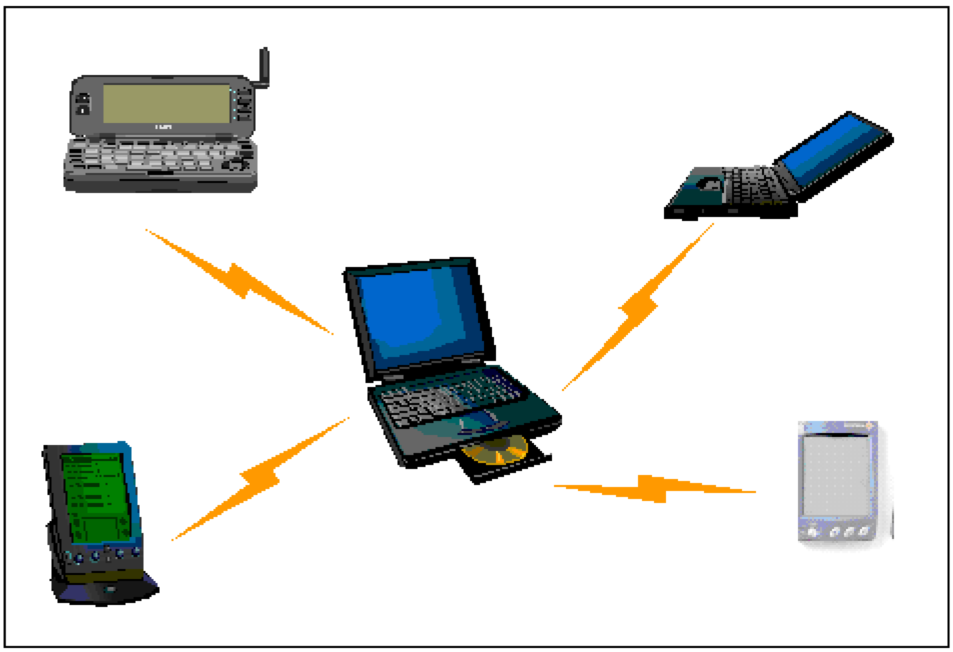

Figure 2.6: Example of Bluetooth ad hoc network

About the radio interface, Bluetooth radio links operate in the unlicensed 2.4 GHz band, going from 2.402 to 2.480 GHz, and Gaussian Frequency Shift Modulation (GFSK) is the selected modulation scheme. Bluetooth uses frequency hopping by dividing the available bandwidth into 79 1MHz frequency intervals, using only one of the intervals at a time, but hopping very quickly from one to another continuously. This is done following a pseudo-random interval sequence, which both the transmitter and the receiver know. The hopping rate is 1600 hops/s, and every radio channel uses a different hopping sequence. When there is more than one channel there is a probability of instant collisions, but normally this probability is small.

2.3.6 Other solutions

In order to improve the performance of the medium access network layer or to solve some of the problems the medium access control protocols may have, several proposals for AHNs appear in the literature.

FAMA (presented in [20] and whose performance was evaluated by the authors in [21]) utilizes carrier sensing and collision avoidance in order to assign the channel control for a particular station, so that guarantees collision-free packet transmission. The MACA-BI protocol shown in [22] removes the Rts part of the handshake in order to be more robust at controlling packet collisions and turn-around time problems.

Some of the protocols make use of contention access, what has higher efficiency at low loads, and others utilize allocation-based access in order to achieve stability. Other solutions try to combine the two ways of access. HRMA [23], CHMA [24] and MACA-CT [25] use reservations methods with frequency hoping spread spectrum, while ADAPT [26], ABROAD [27], CATA [28] and FPRP [29] employ contention for or within TDMA slots. In order to look up the data channel utilization, DBTMA [30] makes use of a channel control with transmit and receive busy tones for Rts/Cts handshake.

SEEDEX

A recent study by Rozovsky and Kumar proposes a random schedule called SEEDEX [31] as an alternative to making explicit reservations for each packet. Each node knows the transmission schedule of its 2-hop neighbors —this could be used in a multicast environment— by mean of a procedure of broadcasting of the current state of the seeds that generate the pseudo-random transmission schedule, as shown in Figure 2.7.

With these seeds, every node can estimate in which of two possible states (S —“Silent”— and PT —“Possibly Transmit”—) every of its two-hop neighbor is. If a node T wants to transmit to a node R, it will do it when T is in PT state and the receiver R is is S state. Moreover, when this event happens, it will only transmit with certain probability, depending on the number of neighbors of R that are —by T’s estimation— in PT state in the current time slot.

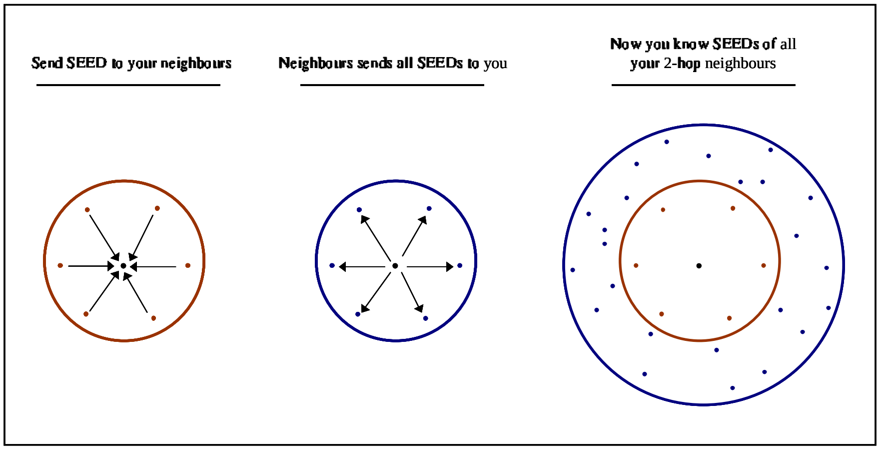

Figure 2.7: Procedure for exchange seeds in a 2-hop neighborhood in SEEDEX

This mechanism can be enhanced if it is employed in a hybrid way, i.e., only on the RTS packets —in the context of the RTS, CTS, DATA, ACK handshake used in IEEE.802.11 [5]—, which are used to make reservations. Therefore, the DATA packets (normally longer) will be free of collisions, because they will be in "reserved" slots. The authors call it SEEDEX-R and their results of the performance evaluation show that:

- The maximum throughput that can be offered is about 10% better than the reachable with IEEE 802.11.
- The mean delay remains fairly constant and 40% lower than in IEEE 802.11, while the factor of reduction of the delay jitter with the former reference is about five.

It is interesting to point out that, by modifying the probability p for every node to be in PT state, SEEDEX can provide different levels of throughput for different flows.

Other ideas

A different prospective based on the possibility of improving the performance of the MAC protocols is studied by Tzamaloukas in [32], who proposes receiver-initiated collision avoidance MAC protocols, which outperform classical sender-initiated ones under certain conditions.

The use of multiple directional antennas in the terminals is a another way of improving the performance with the MAC scheme proposed in [33] and applied to on-demand routing protocols in [34].

In the MAC level, some protocols have also been designed taking into account power saving considerations as a priority. We will deal with the implications of this kind of strategies and their description in chapter 6.

2.4 Conclusion: election of the MAC Protocol

Nowadays, it seems that the best choice for the implementation of the physical and medium access control networks layers for wireless ad hoc networks is the use of IEEE 802.11 (particularly IEEE 802.11b), due to the existence of a standard that guarantees interoperability and the level of flexibility it offers. Currently it is the major standard for Wireless LANs and it is widely supported by vendors such as Cisco, Lucent, Apple, etc. It is the most widespread alternative used in testbeds and simulations for AHNs as well.

Nevertheless, Xu and Saadawi concluded in the June 2001 IEEE Communications Magazine [35] that the existing version of the IEEE 802.11 does not operate properly in multihop AHNs, given that it was not designed to support it. The problems they address are:

- When configured as multihop network, 802.11 still has the hidden node problem, provided that the CTS the receiver sends may not be heard by all receiver's neighbours and, for that reason, they may transmit and intercept the packet intended to the receiver.
- 802.11 MAC does not give heed to the exposed node problem. As carrier sense range is typically larger than communication range to offer some better immunity from hidden terminal, spatial reuse is obviously reduced and the exposed node problem is intensified.
- The binary exponential back off always favours the latest successful node by immediately resetting its contention windows to the minimum value, leading to severe unfairness.

Moreover, one additional problem is the inability of this protocol to support multipoint communication efficiently, what may be needed, for instance, for routing.

In this sense, it is worthy to point out the enormous dependence of the routing protocol in wireless multihop AHNs on the MAC and physical layers, although the network layer model tries to make it transparent. That is the one of the motivations of the great research effort in order to improve medium access techniques and to fit them appropriately for wireless multihop AHNs, which is a still not attained challenge.

Chapter 3

Mobile Ad Hoc Networks

In this chapter we deal with the evolution of ad hoc networks from the packet radio networks, as well as their main features, advantages and drawbacks and the applications they may be useful for.

3.1 History and Evolution

The starting point for the study of the packet radio networks [36] matched with the interest of military organizations such as Defense Advanced Research Projects (DARPA) in the mid-1960's and the consequent establishment of ARPANET in 1969. Initiated in 1970, the ALOHANET, based at the University of Hawaii, was the first large-scale packet radio project.

The PRNET, or Experimental Packet Radio Network, project was proposed by SRI International of Menlo Park and funded by the DARPA starting in 1979 and running for four and a half years (the technology description of the system can be found in [37] and [38]). This was followed by the Vancouver Amateur Digital Communication Group (VADCG) development of a Terminal Node Controller (TNC) in 1980, which automatically divides the messages into packets, keys the transmitter, and then sends the packets, as well as the analogous operations as a receiver. Subsequently, the Amateur Packet Radio Network (AMPRNET) was developed and is used by shortwave radio amateurs, called "hams". By 1985, about 30,000 hams around the world had equipment capable of transmitting data reliably.

The main limitation of these networks consisted in the fact of that each node could not contact several other stations simultaneously. It was overcome with the construction of the first Metropolitan Area Network (MAN), which had a star topology, where every station was able to contact the hub node, but they could not detect each other. The next step, in 1991, came with the utilization of Carrier Sense Multiple Access (CSMA) performing through a MAN repeater with two different radio channels, devoted respectively to transmission and reception.

Other packet radio systems, such as Cellular Data Packet Data, have been used, but the limitations imposed by the wireless medium (multipath propagation, high data error rate, constrained bandwidth) have not permitted the existence of high data rates. In recent times, the emergence of new license-free products for Wireless LANs —based upon spread spectrum techniques— has considerably increased the data rates.

These results are just those that are being used by ad hoc networking community in order to implement the appropriate solution for their purposes, owing to the conceptual identity between packet radio networks that have evolved towards Wireless LANs and ad hoc networks. This ad hoc networking community is leaded by the MANET working group [39], created in 1999 by IETF in order to provide with a framework for the study of Mobile Ad hoc NETworks.

3.2 Main features

An Ad Hoc Network consists of a group of mobile hosts forming a temporary network on wireless links without the aid of any centralized administration or standard support services regularly available on the wide-area network to which the host may normally be connected.

The RFC 2501 by MANET working group in IETF [40] points out the following ones as some of the relevant characteristics of AHNs:

1. Dynamic topologies: nodes can move freely in arbitrary directions and with capricious speed. Therefore, the network must adapt itself to unpredictable changes in its topology, which is typically multihop.

2. Bandwidth constrained: the restrictions imposed by the wireless channel, such as multiple access, multipath interference, noise, fading and limited availability of spectrum, together with the inherent problems the medium access control protocol has to deal with them, make the throughput for each node be much less than the radio's maximum data transmission rate.

3. Energy-constrained operation: the devices that form part of the AHN may be power limited due to the circumstances of their functioning (as in sensors networks, for instance, where maximizing the average network life is a design criteria), hence, the routing algorithms must

manage properly this issue, which can be complicated if *dozing mode* is accepted for the terminals.

4. Limited physical security: mobile wireless networks are susceptible to be have security lacks and can be attacked quite easily. Existing security techniques are applied in the date link layer in order to reduce the risk, but some mechanisms can also be introduced in the network layer. On the other hand, the fact of being a decentralized network provides additional robustness against single point of failures.

In addition, some networks (e.g. mobile military networks or highway networks) may be relatively large (e.g. tens or hundreds of nodes per routing area), although the need for scalability is not unique to AHNs. However, owing to the preceding characteristics, the mechanisms required to achieve scalability likely are more complicated. Actually, most of the existing protocols break down for large networks.

These characteristics create a set of underlying assumptions and performance concerns for protocol design that extend beyond those guiding the design of routing within the higher-speed, semi-static topology of other packet networks such as the fixed Internet.

3.3 Advantages and drawbacks

Ad hoc networks are really an alternative to fixed networks in some operational situations, but an analysis of their advantages and drawbacks, that was summarized in [41] from [39] and [42], will help us know the applications and the contexts they may be useful for. Below, we present some of those.

Among the advantages of AHNs are:

- Fast installation: the level of flexibility for setting up AHNs is high, since they do not require any previous installation or infrastructure and, thus, they can be brought up and torn down in very short time.
- Dynamic topologies: nodes can arbitrarily move around the network and can disappear temporally from the AHN, so the network topology graph can be continuously changing at undetermined speed.
- Fault tolerance: owing to the limitations of the radio interfaces and the dynamic topology, AHNs support connection failures, because routing and transmission control protocols are designed to manage these situations.

- Connectivity: the use of centralized points or gateways is not necessary for the communication within the AHN, due to the collaboration between nodes in the task of delivering packets.

- Mobility: the wireless mobile nodes can move at the same time in different directions. Although the routing algorithms deal with this issue, the performance simulations show that there is a threshold level of node mobility such that protocol operation begins to fail.

- Cost: AHNs could be more economical in some cases as they eliminate fixed infrastructure costs and reduce power consumption at mobile nodes.

- Spectrum reuse possibility: owing to short communication links (node-to-node instead of node to a central base station), radio emission levels could be kept at low level. This increases spectrum reuse possibility or possibility of using unlicensed bands.

Some of the problems AHNs have are:

- Bandwidth constraints: as commented above, the capacity of the wireless links is always much lower than in the wired counterparts. Indeed, several Gbps are available for wired LAN, while, nowadays, the commercial applications for wireless LANs work typically around 2 Mbps.

- Processing capability: most of the nodes of the AND are devices without a powerful CPU. Furthermore, the network tasks such as routing and data transmission cannot consume the power resources of the devices, intended to play any other role, such as sensing functions.

- Energy constraints: the power of the batteries is limited in all the devices, which does not allow infinitive operation time for the nodes. Therefore, energy should not be wasted and that is why some energy conserving algorithms have been implemented (COMPOW [43], PARO [44] and MBCR [45] are some examples). A study of the energy saving and capacity improvement potential of power control in multihop wireless networks is done in [46].

- High latency: when an energy conserving design has been applied it means that the nodes are sleeping or idle when they do not have to transmit any data. When the data exchange between two nodes goes through nodes that are sleeping, the delay may be higher if the routing algorithm decides that these nodes have to wake up.

- Transmission errors: attenuation and interferences are other effects of the wireless links that increase the error rate.

- Security: [47] analyses some of the vulnerabilities and attacks AHNs can suffer. The authors divide the possible attacks in passive ones, when the attacker only attempts to discover valuable information by listening to the routing traffic; and active attacks, which occur when the attacker injects arbitrary packets into the network with some proposal like disabling the network. Other security issues such as availability, authenticity, integrity, confidentiality and privacy are discussed in [48].

- Location: the addressing is the another problem for the network layer in AHNs, since the information about the location the IP addressing used in fixed networks offers some facilities for routing that cannot be applied in AHNs. The way of addressing in AHNs, of course, has nothing to do with the position of the node. In [49] a recent proposal on an IPv6-based addressing scheme for AHNs can be seen.

- Roaming: the continuous changes in the network connectivity graph involve that the roaming algorithms of the fixed network are not applicable in AHNs, because they are based on the existence of guaranteed paths to some destinations.

- Commercially unavailable: AHNs are yet far from being deployed on large-scale commercial basis.

3.4 Applications

Once known the features of ad hoc networks, the examples of their potential practical use are only limited by imagination. So, the set of applications for AHNs is diverse, from small, static networks that are constrained by power sources, to large-scale, mobile, highly dynamic networks. Typical applications are those in which survivable efficient and dynamic communications must be established. Some examples are:

- Conferences and meetings for a group of people with laptop computers that may wish to exchange files and data without mediation of any additional infrastructure between them.

- Home environment for communication between smart household appliances can be hold by an AHN between the different devices, which may share control information for their correct functioning. We can think in an AHN formed by our electrical household appliances in the kitchen, the laptop computer, which can be moved arbitrarily, the television set, the windows and doors, the air conditioning system, all together communicated to automatically perform imaginative functions related with their control and their relation with the user.

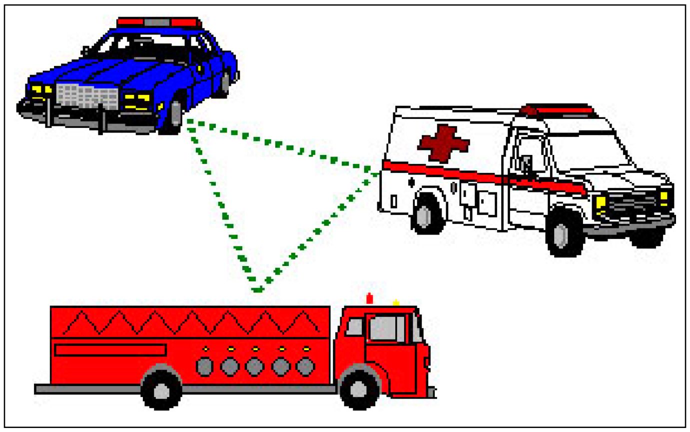

Figure 3.1: Ad hoc networks can be very useful in emergency and rescue operations

- Emergency search and rescue operations require quick and dynamic communications, and with the assistance of AHNs they could be developed in inhospitable areas.

- AHNs are suitable to be used in areas where earthquake or other natural disasters have destroyed communication infrastructures. In this case, the rapid deployment is a fundamental parameter, as well as the lack of necessity of any fixed infrastructure.

- AHNs perfectly satisfy military needs like battlefield survivability, operation without pre-placed infrastructure and connectivity beyond the LOS. Indeed, the research about packet radio network started in a military context and nowadays, the concept of Digital Battlefield [50] is a burning topic.

- For monitoring and measuring purposes a large number of small computing devices could be spread over a geographical area (sensors scattering) to form a self-sustained AHN. In this case one of the most important design criteria should be obtaining the highest average life of the network, based on power consumption requirements, as well as the need for scalability, provided that the number of sensors could be very high. Some examples of this sort of applications could be:
  - Military sensor networks to detect enemy movements, the presence of dangerous material (such as poison gases or radiation), explosions, etc.
  - Environmental sensor networks to detect and monitor environmental changes.

- Wireless traffic sensor networks to monitor vehicle traffic on a highway or in a congested part of a city.
- Wireless surveillance sensor networks for providing security in a shopping mall, parking garage, or other facility.
- Wireless parking sensor networks to determine which spots are occupied and which spots are free.

It is worthy to mention the application suggested in [51] by S. H. Humbad. He remarks the potential usability and applicability of wireless AHNs in a peer-to-peer highway vehicle network. This network could be useful, in addition to provide mobile Internet access to allow consumer and business services, to mitigate urban traffic congestion through an area-wide road pricing system.

Chapter 4

Capacity of Ad Hoc Networks

In this chapter we describe the issue of capacity in wireless ad hoc networking, as well as the theoretical limits presented in the literature. Furthermore, we also analyze the diverse factors that have a bearing on it, focusing on the important issue of the selection of the transmission range.

4.1 Introduction

The possibility of use of certain bandwidth by an ad hoc network doesn't imply, of course, the availability of that capacity for the final applications. It is assumed that different network layers include overhead and signalling information exchange that diminish the throughput achieved at the end by the network architecture. But, in the case of wireless networks, where the spectrum has to be watched over so carefully and the interferences may be utterly determinant, coping with the issue of the capacity as one of the main problems is crucial.

Particularly, as AHNs are multihop networks, which involves that every node is not only responsible for sending its own data, but also has to forward other nodes' packets, the per node throughput decreases, so that this fact becomes critical with a high number of nodes and scalability problems are always present in our context. Actually, it is commonly observed in performance analysis of ad hoc routing protocols that the capacity is the restraining cause, so that routing algorithms should take this into account in their scheme of queries, updating messages and additional information exchange. In consequence of that, as it is observed in [52], very low data rates are predisposed to be used in simulations of ad hoc protocols in order to keep away from running out of capacity.

Likewise, the increase in the number of nodes in the AHN has a great influence on the throughput the network can provide for every node, so the scalability appears also as an important issue in terms of capacity.

4.2 Bounds on capacity

In order to determine theoretical limits on the capacity of wireless ad hoc networks, Gutpa and Kumar [53] employ two different models of non-interference and successful reception of the packets. The first one, the Protocol model, is based on the existence of a guard zone —determined by Δ— for transmission of different senders over the same frequency channel at the same time, so a transmission from node X_t to node X_r is successful if, for every other node X_i,

$$|X_i - X_r| \geq (1+\Delta)|X_t - X_r|$$

The other one, the Physical model, considers the necessity of a certain level of Signal to Interference Ratio (SIR), and the condition of successful reception, when X_i selects a power level P_i, is:

$$\frac{\frac{P_t}{|X_i - X_r|^\alpha}}{N + \sum_{i\in\tau,\, i\neq t} \frac{P_i}{|X_i - X_t|^\alpha}} \geq \beta$$

The results they obtain for these models are quite similar to each other.

First, under the Protocol model and the hypothesis of n randomly located fixed nodes within a planar disk (also extended to the case of the surface of a sphere with the same results) and randomly chosen traffic pattern (i.e., the destination for every node is independently chosen as the node nearest to a randomly located point, uniformly and independently distributed), the authors report that the per node capacity C_n of a wireless network scales like

$$C_n < k_1 \cdot \frac{1}{\sqrt{n \cdot \log n}} \tag{4.1}$$

in the limit of a high enough number of nodes.

If all the nodes were optimally placed, the maximum throughput under the Protocol model would be like

$$C_n < k_2 \cdot \frac{1}{\sqrt{n}} \tag{4.2}$$

bits/sec to each node. This upper bound is supposedly possible to reach when the range of each transmission is optimally chosen and the best possible traffic pattern of transmissions that knows all the nodes locations, movements and traffic demands is used (under these hypothese, the network is called arbitrary by the authors), what is unfeasible in practice, but a challenge for the scheduling of the algorithms.

Now, we outline the proof for equation (4.2). Every bit b, with $1 \leq b \leq \lambda nT$, has to move from its source to its destination by using $h(b)$ hops, where the h-th hop covers a distance of r_b^h. Then, if the average distance between source and destination is $\overline{L}$,

$$\sum_{b=1}^{\lambda nT} \sum_{h=1}^{h(b)} r_b^h \geq \lambda nT\overline{L} \tag{4.3}$$

Furthermore, for any sub-channel m with bandwidth W_m and any slot s of duration τ, there will be, in the best case, $n/2$ transmitter-receiver pairs, so the following inequality holds:

$$\sum_{b=1}^{\lambda nT} \sum_{h=1}^{h(b)} 1(\text{the } h\text{-th hop of bit } b \text{ is over sub-channel } m \text{ in slot } s) \leq \frac{W_m \tau n}{2} \tag{4.4}$$

This equation can be summed over the sub-channels and slots, so that H is defined as

$$H \equiv \sum_{b=1}^{\lambda nT} h(b) \leq \frac{WTn}{2} \tag{4.5}$$

By taking into account geometric considerations and basing on the spatial reuse principle, it can be shown (we will do it in subsection 4.4, where we will deal with the range of transmission issue) that the maximum number of simultaneous receptions per unit area is $\frac{16}{\pi\Delta^2 r^2}$, where r is the communication range, which can be substituted in this demonstration by r_b^h. Thus, if we introduce this expression instead of the $n/2$ transmissions considered before and sum over all the sub-channels, we obtain

$$\sum_{b=1}^{\lambda nT} \sum_{h=1}^{h(b)} \frac{\pi\Delta^2}{16}(r_b^h)^2 \leq WT \tag{4.6}$$

Finally, in order to apply that

$$(\sum x)^2 \leq \sum x^2 \qquad \text{for all } x \leq 1,$$

we divide both sides of equation 4.6 by H (so that $r_b^h/H \leq 1$), and, after substituting equations (4.3) and (4.5) in the result of applying this convex property of the quadratic function, it results that

$$\lambda n T\overline{L} \leq \sqrt{\frac{16WTH}{\pi\Delta^2}} \leq \sqrt{\frac{8}{\pi}}\frac{1}{\Delta}TW \cdot \sqrt{n} \tag{4.7}$$

Therefore, the capacity per node follows the bound given by the equation 4.2, with

$$k_2 = \sqrt{\frac{8}{\pi}} \cdot \frac{W}{\Delta\overline{L}}$$

Also with a non-practical prospective, Toumpis and Goldsmith [54], by extending the previous work [53] to a 3-dimensional topology and slotting in the Shannon capacity into the link model, conclude that the capacity per node of an ad hoc network (without aggregating the capacity of all the stations), when the number of nodes is large enough, is limited by

$$k_3 \cdot \frac{1}{n^{2/3} \cdot \log n} < C_n \leq k_2 \cdot \frac{\log n}{\sqrt{n}} \tag{4.8}$$

what involves that the per node transmission rate drops off although the number of users increases the overall capacity of the network. Hence, the size of a network that has to guarantee a certain per user rate is limited.

Actually, Gupta et al. show in an experimental study [55] that the scalability law for AHNs is even worse than the theoretically presented in [53], as the per node throughput decays like $c/n^{1.68}$ bits/sec. in their experiments. They suggest that the hardware and implemented protocols should be improved, above all the efficiency of the medium access ones.

All these studies may hint that AHNs are not scalable under the current conditions. Nevertheless, Li et al. point out in [52] that all these previous studies consider a random communication pattern, where every node is equally likely to transmit to any other station all over the network, independently of location. This can be presupposed for small networks, but not for large ones, where the vast majority of a certain station's communications may occur with nearby nodes.

This traffic pattern is a parameter that has a great importance, because, as [52] —also k_2 in equation (4.2)— shows by focusing on the capacity limits inflicted by multihop traffic patterns, per node available bandwidth decreases with the expected path length. This, and other parameters that should be taken into account, will be discussed in the following section.

4.3 Parameters that can modify capacity

The assumptions considered to obtain Equations (4.1) and (4.2), which represent the scaling laws (when the number of nodes n tends to infinite) for the per node capacity of AHNs, lead us to think in a set of parameters that affect the behavior of the network in terms of capacity. Those parameters are:

- Traffic pattern
- Location and mobility of the nodes
- Range of transmission

In order to see the relationship among these parameters and how they can modify the amount of useful bandwidth that each node can have for its own traffic, it is necessary to observe two facts:

1. The area is a valuable resource in AHNs. In fact, it can be shown that the capacity is determined by the amount of spatial reuse possible in the network and this is proportional to the physical area A of the network, so that, with a uniform node density δ, $C = k \cdot A = k \cdot \frac{n}{\delta}$ for some constant k.
2. The multihop nature of the AHNs. If we consider:
   - every node as a source of packets at a rate of $\lambda(n)$.
   - the expected physical path length for the traffic pattern in the network is $\overline{L}$ from the source to the destination.
   - a fixed communication range $r(n)$.

   we have that the mean minimum numbers of hops to deliver a packet is $\frac{\overline{L}}{r(n)}$ and, as a result, the total one-hop capacity required to send and forward packets has to fulfill $C > n \cdot \lambda(n) \cdot \frac{\overline{L}}{r(n)}$.

Combining the results of both considerations, we find that the per node capacity, $\lambda(n)$, is limited by

$$\lambda(n) < \frac{kr(n)}{\delta} \cdot \frac{1}{\overline{L}} = \frac{C/n}{\overline{L}/r(n)} \tag{4.9}$$

The inequality (4.9) shows us that, as expected, the throughput available per each node is lower than the total capacity divided by the number of nodes (because the delivery of the packets is multihop) and that the decreasing factor is determined by the quotient between the expected path length (depending on the traffic pattern) and the range of communication, i.e., the expected number of hops for every packet. Anyway, in this analysis, made by Li et al. [52], the range of transmission as a factor that increased the interference and, hence, makes the throughput go down, is not considered, so, logically, we cannot assume that the higher $r(n)$ is, the higher the capacity per node will be.

In section 4.4, that will deal with the important issue of the election of the range of transmission in AHNs, the interferences will be also taken into account and another relationship will be shown.

4.3.1 Traffic pattern: locality, effects of relaying and multi-packet reception

First of all, an important fact we should take into account is that the medium access control protocols, indeed, involve a particular consideration of how to deal with the communication pattern within the network. Thus, the information obtained about the influence, in several ways, of the traffic pattern on the capacity of the AHN, could be used as a valuable resource in the design of those MAC protocols (see section 2.3), which, in fact, is one of the challenges of the field of study. In this sense, the research efforts tend to look for traffic patterns that scale well and maximize the capacity as much as possible.

Locality of traffic pattern

If each node sends only to nodes within a fixed radius, independent of the network size, then the expected path length remains obviously constant as the network size grows, and the per node capacity also stays constant.

Unfortunately, this assumption is difficult to be accomplished and the expected path length may vary to a great extent depending on the distribution of the nodes around the network.

In [56], the probability distribution of the link distances within a wireless network is found for two possible scenarios:

- The mobile stations are uniformly distributed over a rectangular area.
- The orthogonal components of the locations of the nodes have Gaussian distributions.

The shapes of both link distance distributions are quite similar to each other, so, as a conclusion, the model used for the distribution of the mobile locations can be chosen for convenience in simulation or analysis of mobile systems.

In [52], a power law distribution of the distances to the destinations is used to fix, through the parameter α of that distribution, the expected path length. This distribution is like:

$$p(x) = \frac{x^{\alpha}}{\int_{\epsilon}^{\sqrt{A}} t^{\alpha} dt}$$

Therefore, by assuming different levels of prevalence of the local traffic (measured by the average path length, dependent of the parameter α of the power law considered for the distances distribution), the authors conclude that:

- The random traffic pattern is the worst case of traffic pattern one may choose for AHNs.
- Per node capacity will scale like $\Theta(\frac{1}{\log n})$ for a certain value of $\alpha\,(-2)$. This is exactly the value chosen for the location update traffic pattern by the Grid Location Service (GLS) [57], a technique which, combined with geographic forwarding, can help the scalability of ad hoc mobile networks.
- Per node capacity may remain approximately constant if the size of the network is large and the distance distribution decays more quickly than a certain value of $\alpha\,(-2)$.

With that, the authors show that "as the power law distribution moves from a very local to a very distant destination selection, the capacity scaling moves from constant per node capacity to a $\Theta(1/\sqrt{n})$ degradation of capacity with the network size". The idea behind this result arises from the fact of that if successful transmissions over long distances occur, then there are many possible short-distance transmissions that could be done within that range, so that the resource of the spatial reuse is not being consumed optimally.

The conclusion suggests that there are some possibilities to construct large networks with reasonable per node capacities. Indeed, they give some examples of networks where the local traffic patterns are common:

- LAN users
- Telephone system
- Caching systems in the Internet at large

Effects of relaying

One of the possible alternatives we can think with the aim of increasing the capacity of AHNs, is the inclusion in the network of a certain amount of nodes acting as relay nodes, so that they can collaborate in the delivery of the packets without generating new traffic that would overload the network. On the other hand, increasing the density of nodes in the network also intensifies the interference and decreases the spatial reuse, a valuable parameter in wireless networks.

With the same assumptions they use to find the theoretical limits on capacity of AHNs, Gupta and Kumar [53] consider that adding additional homogeneous nodes in random positions around the networks, with no independent traffic needs of their own, is not worthy in terms of capacity. For a fixed random network, the scaling law is determined by equation (4.1), while if we include m additional nodes the scaling law is like

$$C_n < k_4 \cdot \frac{n+m}{n \cdot \sqrt{(n+m) \cdot \log{(n+m)}}} \tag{4.10}$$

As an example, if $n = 100$ active nodes, in order to increase the factor $\frac{n+m}{n \cdot \sqrt{(n+m) \cdot \log{(n+m)}}}$ 5 times from its value at $m = 0$, the value of m has to be 4476. To increase $k+1$ times the value of that factor (i.e. the capacity) we would have to add kn nodes.

In Grossglauser and Tse's work [58], other limits of the capacity per node are shown, by taking into account the possibility of having relaying nodes and adding mobility to the Gupta and Kumar's model. These results will be explained in section 4.3.2.

Multipacket reception

Continuous advances in signal processing and multiple antenna technologies may change the traditional concept of collisions in wireless communications, by letting the terminals receive multiple packets at the same time. This multipacket reception (MPR) can be obtained in spread spectrum techniques by assigning several codes to a single receiver and in antenna systems by proper space-time coding or beamforming [59].

The effects of these techniques on networking layers other than the physical layer have changed some of the underlying assumptions made by conventional MAC techniques, and, even, are not still fully understood, as well as their possible impact on the performance and design of MAC protocols.

In terms of capacity, Mergen and Tong [60] study the effects of multipacket reception capability on the capacity of wireless networks with regular structures. Their results indicate that having MPR capability at the physical layer given a fixed connectivity does not change the asymptotic capacity law, but only its coefficient, so that the final capacity can be improved by a factor. Moreover, they suggest that the contribution of MPR to the capacity is higher when the network connectivity is higher and that this contribution depends significantly on the type of traffic, so that it becomes very effective in nodes that receive much traffic (for instance, nodes that could act as a gateway).

Some MAC protocols that use this feature have been proposed in the literature. For instance, RCT (Receiver Controlled Transmissions) is presented in [61] as a multiple access protocol that is a hybrid of scheduled and random access protocol and that scales to networks of arbitrary sizes. Also, the Multi-Queue Service Room (MQSR) [62] protocol is designed explicitly for general MPR channels, so that it tries to accommodate groups of users with different delay requirements.

The idea of MPR could be complemented by the dual concept of multipacket transmission (MPT) —useful for source type of nodes, such as the down-link of a base station—, that can also improve capacity by itself [60]. However, when having both MPR and MPT, the capacity increases linearly with the minimum increase of the two. Therefore, the choice for one of these capabilities to be implemented in some network nodes should be made by taking into account the application and whether there is a considerable amount of nodes that act as sources (then it would be better to implement MPT in these nodes) or sinks (then MPR would be the suitable technology).

4.3.2 Location and mobility of the nodes

In principle, the location of the nodes is not a parameter the ad hoc network designer can modify at will, but having a knowledge about how the network behaves in terms of connectivity and throughput with different distributions of the devices in the space may let us obtain conclusions about the traffic patterns we should look for in order to maximize as much as possible the capacity of the whole network.

In this sense, relationships between connectivity and communication range (for example, probabilities of fully —or partially— connected network when we assume a particular distribution of the nodes depending on the transmission range) for diverse cases of such distributions of the nodes, become important, as they could let us fix one of those parameters depending of the other, although, on the other hand, we lose a grade of freedom.

Mobility

Another interesting idea to take into account is the mobility as a factor that can increase the capacity of AHNs, because, by using it, the traffic pattern can be modified in a clever way. By introducing mobility in the network model Gupta and Kumar employed in [53], Grossglauser and Tse [58] show that the average long-term throughput per source-destination pair can be kept constant even as the number of nodes per unit of area increases.

The design the authors put forward is based on the employment of mobility as a mechanism to offer multiuser diversity for the relaying of packets (see figure 4.1). They try to apply the information theoretic result of Knopp and Humblet [63] to AHNs. The result shows that the optimal way of maximizing the total information theoretic capacity is to assign the transmission at any time to the user with the best channel to transmit to the base station. In order to attain this, their strategy follows the scheme given by the Infostation architecture [64] and, thus, the transmissions are carried into effect only when the sender and the destination nodes are near. Nonetheless, this situation is quite improbable (the fraction of the time two nodes are close together is too small, of the order of $1/n$) and they show that, without relaying (only direct transmission and, so, trusting in mobility as the unique mechanism to provide diversity), the achievable throughput per source-destination pair decreases at least as fast as $n^{-\frac{1}{1+\alpha/2}}$, so that:

$$C_n < k_4 n^{-\frac{1}{1+\alpha/2}} \tag{4.11}$$

Therefore, in order to find a way to communicate only locally, they propose that the owner of the packet transmits it to as many different mobile nodes as possible (see figure 4.1), which will act as one-hop relay nodes and will transmit the packet to the destination if the assumption of the proximity attained by several independent movements is accomplished.

With this strategy, they show that there is a scheduling policy π that, by only choosing the sender-receiver pairs in a random way depending on the node locations, can achieve, with the scheme proposed, that the long-term throughput between every two nodes keeps constant. For this policy, the probability that two arbitrary nodes are scheduled as a sender-receiver pair is $\Theta(1/n)$. This probability is said to be equal to the long-term throughput between any two nodes (by taking into account the hypothesis of node locations being i.i.d., stationary and ergodic). Therefore, by adding over the throughputs of the $n-1$ routes, it is seen that the total average throughput per source-destination pair (and, consequently, also the throughput per node) is $\Theta(1)$ and, consequently, can remain constant even when the number of nodes increases:

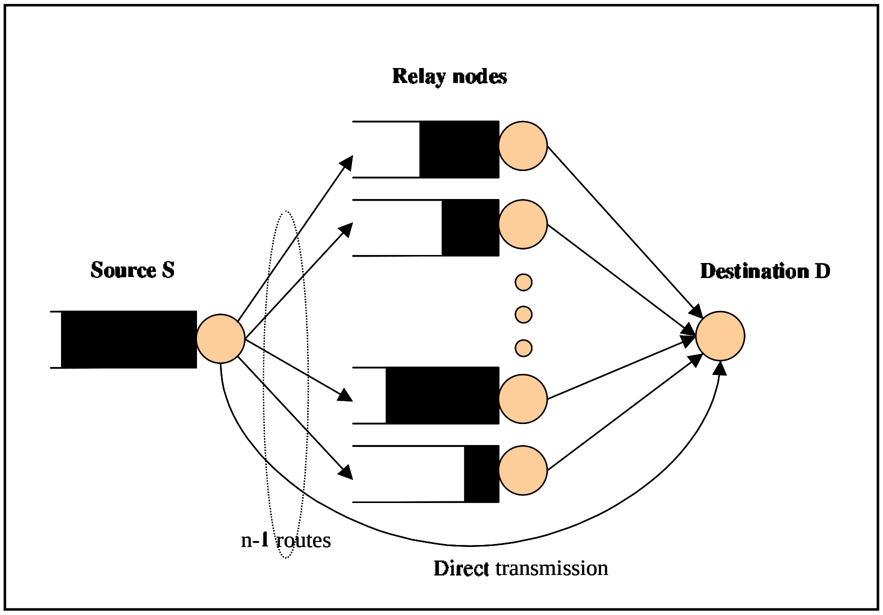

Figure 4.1: How relaying can create multiuser diversity and get benefits from nodes' mobility

$$\lim_{n\to\infty} Pr\{C_n = k_5 W \quad \text{is feasible}\} = 1 \tag{4.12}$$

In order to attain this, the packets of every source node should be delivered across all the other nodes in the network, so that every node in the network will have, in the long run, packets buffered destined to every other node.

Evidently, this approach is only suitable by assuming asynchronous applications, such as email and database synchronization, because the delays involved may become too high, but under this supposition, node mobility can notably increase the throughput capacity of the network. Likewise, the delay tolerance also allows the stations to use the transmit power in the most effective way, i.e., when the destination is closer; so it becomes a grade of freedom for the network to deliver packets.

4.4 Range of transmission

4.4.1 A common range of communication is needed

Unlike what happens in wired networks, links in a wireless network cannot always be considered bidirectional, so the fact that one node A can hear

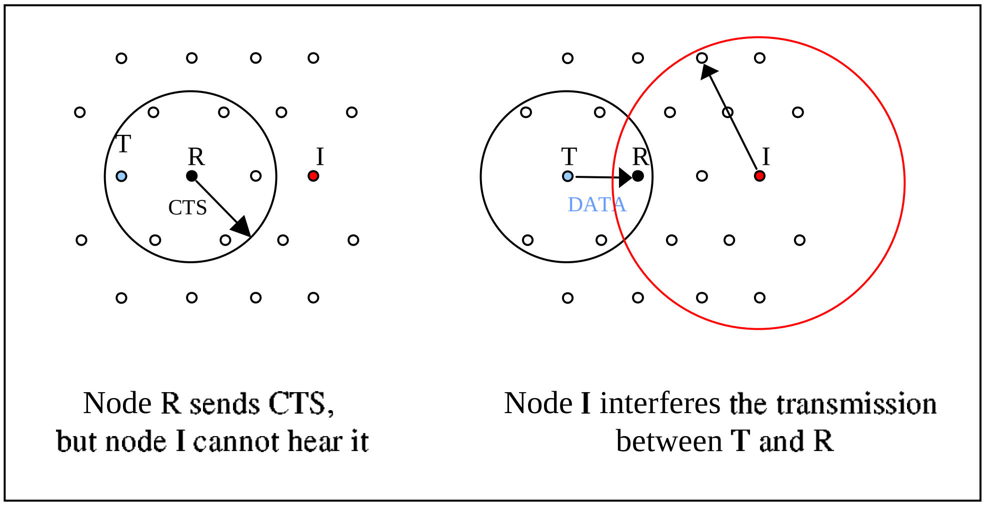

Figure 4.2: Possible interference due to different ranges of communication

another node B does not guarantee that B can hear A, due to possible obstacles, shadowing, multipath effects (see section 2.2.2), fading, different transmission powers available ...

Bidirectionality of the links is important both in the MAC protocols and in the network layer, and the simplest way to ensure it consists on using a common transmission power for all the nodes.

At the Medium Access Control layer, protocols such as IEEE 802.11 [5] employs a handshake (see section 2.3.3) where different power levels would carry the following problems:

1. When a transmitter node T sends a RTS packet, the receiver R answers with a CTS. The problem arises when an interfering node outside the range of communication of R (so it does not receive the CTS) has a communication range higher than the communication range R has, so that I would not keep silent and could interfere the transmission from T to R (see figure 4.2).

2. Acknowledgements in the MAC layer are needed owing to the unreliable nature of the wireless medium. Nevertheless, if the receiver R of a packet sends the ACK packet at a lower level than the level used by T to transmit the data to R, then the ACK may not be heard by T.

At the network layer, the problem comes, for instance, when the used protocol is intended to work with end-to-end acknowledgements that follow the forward path taken by the data.

A completely different approach to the problem of the communication range can be based on developing strategies that enable the hosts within the AHN to change their transmission power. Therefore, the problems cited above should be solved by considering, for instance, other communication schemes in the Medium Access Control layer.

This kind of strategies has been proposed by Adler and Scheideler [65]. In their work, they define an abstract model about the scheduling of the transmission of packets and the selection of the routes in power-controlled AHNs. This framework makes use of a new class of MAC protocols they propose for AHNs, the local probabilistic control protocols (LPR), that, theoretically, allow each node to decide independently of the other nodes when to send a packet, i.e, the coordination between nodes is not a requirement. Equally, these protocols ensure that each attempted transmission is successful with a constant probability.

Furthermore, they apply these ideas to the network layer, where they find a nearly optimal way for routing arbitrary permutations — paths of N hops crossing different nodes — in any static AHN. They demonstrate that for any transmission graph, there is a distributed protocol that can reach a constant factor of the optimal permutation routing time in the average case. Consequently, this can be applied for every route within an AHN, because this route will be included in, at least, one of the possible permutations.

4.4.2 Constraints on range: connectivity and throughput

The decision of selecting the range of communication for AHNs affects to two essential parameters in the behavior of any kind of network: the connectivity level and the throughput achieved for each node in the network.

1. **Connectivity**. Each node should be able to communicate with any other in the network, but this fact is influenced by the capacity of the nodes of communicating to each other in a local environment, which is determined by the communication range. The higher the communication range is, the easier close nodes can reach each other and the higher the level of connectivity is. Therefore, in the network there will be more possible routes (what can be useful for path redundancy, for example) and the probability of finding disconnected sets of nodes within the AHN will decrease.

   In order to measure the level of connectivity of the network, several metrics can be used:

- Ratio of connected pairs:

$$\begin{aligned} \text{Conn} &= \frac{\text{number of connected pairs in the network}}{\text{maximum number of possible connected pairs}} \\ &= \frac{\sum_{i=1}^{k} n_i(n_i-1)}{n(n-1)} \end{aligned}$$

  where k is the number of clusters (disconnected parts of the network) in a given instant and n_i is the number of nodes within the cluster i.

- Number of disconnected parts of the networks. Of course, when this number is 1, there is full connectivity.

- Multihop connectivity. For instance, a metric could be the average number of hops needed to reach every other node in the network.

In chapter 8 we show the results obtained by simulation of the probability of having a fully connected network as a function of the transmission range.

2. **Throughput**. The influence the communication range has on the throughput of an AHN arises from the importance of the concept of spatial reuse. When a node is successfully receiving data from another station, there is a circle around the first where no other node except for the transmitter of these data can be transmitting. Hence, as we mentioned in the beginning of this section, the area is a valuable resource in wireless networks. Now, by using this concept, we will show the dependence between the maximum number of simultaneous sender-receiver pairs in an area and the communication range r.

   Under the protocol model of interference used in [53], and nodes X_{T_1} and X_{T_2} being transmitters and X_{R_1} and X_{R_2} their corresponding receivers, the model itself considers:

$$\left.\begin{array}{rcl} |X_{T_1} - X_{R_1}| & \leq & r(n) \\ |X_{T_2} - X_{R_1}| & \geq & (1+\Delta)\cdot r(n) \quad \forall k \end{array}\right\} \text{Procotol Model}$$

   in order to assure communication and avoid interference, as can be seen in Figure 4.3.

   By using the triangle inequality, we have that

$$\begin{aligned} |X_{R_1} - X_{R_2}| &\geq |X_{R_1} - X_{T_2}| - |X_{R_2} - X_{T_2}| \\ &\geq (1+\Delta)r - |X_{R_2} - X_{T_2}| \\ &\geq (1+\Delta)r - r \end{aligned} \tag{4.13}$$

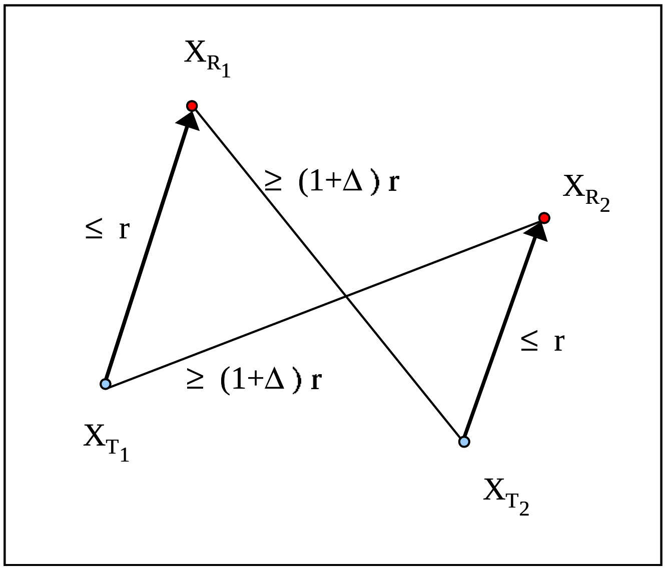

Figure 4.3: Protocol model of interference

Therefore, we conclude that the minimum distance between two receivers is

$$|X_{R_2} - X_{R_1}| \geq \Delta \cdot r, \tag{4.14}$$

which implies that disks of radius $\frac{\Delta r}{2}$ around X_{R_1} and X_{R_2} are disjoint when simultaneous transmissions are carried into effect, as shown in Figure 4.4.

Hence, the maximum number of simultaneous transmission-reception pairs within a disk of area A are:

$$\begin{aligned}
\text{Max. number of tx} &= \frac{A}{\text{minimum area of receivers disks in } A} \\
&= \frac{A}{\frac{1}{4} \cdot \pi \frac{\Delta^2 r(n)^2}{4}} \\
&= \frac{16A}{\pi \Delta^2 r(n)^2}
\end{aligned} \tag{4.15}$$

By following the same considerations taken into account to obtain equation (4.9), we have that

$$n\lambda(n)\frac{\overline{L}}{r(n)} \leq \frac{16A}{\pi\Delta^2 r(n)^2} \cdot W \tag{4.16}$$

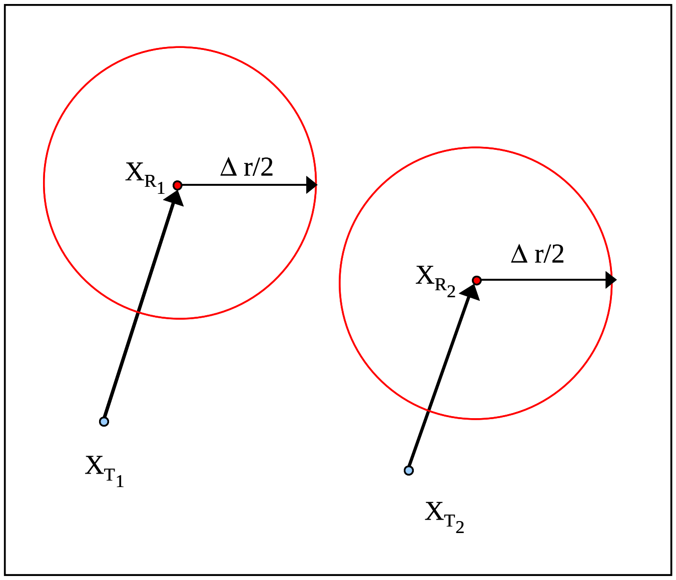

Figure 4.4: Simultaneous receivers employ disjoint disks of $\Delta r/2\ m^2$

Hence,

$$\lambda \leq \frac{16AW}{\pi\Delta^2\overline{L}} \cdot \frac{1}{n \cdot r(n)} \tag{4.17}$$

The intuitive explanation about how the communication range is a trade-off between the connectivity and the throughput is based on the fact that a too small communication range reduces connectivity because nodes cannot reach each other, but if r is too high, interference is introduced in the network (due to the decrease of the possible spatial reuse) and the capacity is reduced as shown in equation (4.17). Furthermore, a low communication range decreases the overall energy consumption in the network, so reducing it as much as possible is often a desired goal.

4.4.3 Theoretical critical power

Equation (4.17) suggests that the common range of communication r should be made as small as possible in order to reduce the interference and, consequently, maximize the average throughput per node. Nonetheless, a certain level of connectivity should be assured as well, to guarantee that a node that wants to transmit a packet at a given instant has a physical path (route) that ensures the delivery of this packet.

Therefore, after knowing this tradeoff between the throughput and the connectivity, it is important to examine what the better communication range is. This optimal communication range is such that it assures the level of connectivity we define. In principle, to our knowledge, theoretical studies have tried to find this critical value of r for the situation in which the network is fully connected. We will call to this value of r the *critical transmission range* of communication, and we will denote it by $r_{crit}(n)$.

In this case, Gupta and Kumar find in [66] the solution for a network graph $\mathcal{G}(n, r(n))$ with n randomly (uniformly distributed) located nodes in a disk $\mathcal{D}$ of unit area under a Protocol model of interference (see section 4.2).

Defining $P_c(n, r(n))$ as the probability that $\mathcal{G}(n, r(n))$ is connected, they prove the following theorem :

Theorem 4.1

$$\textit{If } \pi r^2(n) = \frac{\log n + k(n)}{n} \text{ , } \textit{then}$$

$$P_c(n, r(n)) \to 1 \iff k(n) \to +\infty$$

Therefore, if we extend this to a disk of area A, the minimum value of r to assure connectivity, i.e, the critical transmission range $r_{crit}(n)$, is:

$$r_{crit}(n) = \sqrt{\frac{A \log n + k_2(n)}{\pi n}} \text{ where } k_2(n) \underset{n\to\infty}{\longrightarrow} +\infty \tag{4.18}$$

As an extension of this result, the authors generalize the problem by considering that a node A, even being within the range of another node B, communicates with it with probability $p(n)$, $0 \leq p(n) \leq 1$, which, under these circumstances, can be considered as a measure of the reliability of the link between A and B. The authors believe that Theorem 4.1 also holds when substituting $p(n)\pi r^2(n)$ instead of $\pi r^2(n)$.

Combining equations (4.17) and (4.18) we obtain that the maximum throughput that can be achieved while assuring full network connectivity (i.e., by employing the critical communication range r_{crit}) is like

$$\lambda \leq \frac{16\sqrt{A}W}{\sqrt{\pi}\Delta^2\overline{L}} \cdot \frac{1}{\sqrt{n \ \cdot \log n}} \tag{4.19}$$

under the Protocol model of interference.

A possible future work could focus on trying to extend these studies so that the best value of the communication range r could be chosen as a function of the level of connectivity desired for the network at a precise moment. In this case, it could be interesting to use the number of disconnected parts in the network as the metric for the connectivity, in order to evaluate the benefits (in terms of throughput) of having different partitions in the network at some particular moment.

A probabilistic analysis for the range assignment problem

The problem of the assignment of the best —minimum— transmitted range to maintain the AHN connected is dealt with in [67] from a probabilistic point of view, in the context of static AHNs with homogeneous transmission range (the case we are interested in, as discussed in subsection 4.4.1).

Definition 4.1 A *graph* $\mathcal{G}$ is said to be *strongly connected* if, for every two nodes $X_a, X_b \in \mathcal{G}$, there exist directed paths from X_a to X_b and from X_b to X_a.

The kind of analysis investigates the tradeoffs between the communication range r, the number of nodes n and the dimensions of the region (given by l, that is the side of a square in the two dimensions problem or of a cube in the 3-D scenario), when r is such that it ensures strong connectivity with high probability. Another work with this prospective is [68], where the authors consider the stronger requirement of bi-connectivity, but propose a solution where centralized control is necessary. Moreover, they present two heuristics to deal with mobility for this problem.

Let's consider the bounds obtained in [67] and discuss them. For the general case of a communication graph $\mathcal{G}$ of d dimensions, the probabilistic event that is studied in depth in this study is $P\{\text{node } i \text{ is isolated}\}$ when a value of r is selected as the communication range. As "Node i is isolated" $\subseteq \mathcal{G}$ is disconnected", and $Pr\{\mathcal{G} \text{ is connected}\} = 1 - Pr\{\mathcal{G} \text{ is disconnected}\}$, at the end we can get a necessary condition for the event "$\mathcal{G}$ is connected".

Theorem 4.2 $\forall l > 0, \forall i \in \{1, .., n\}$,

$$\left(1 - c_1 \left(\frac{r}{l}\right)^d\right)^{n-1} \leq Pr\{\textit{node } i \textit{ is isolated}\} \leq \left(1 - c_2 \left(\frac{r}{l}\right)^d\right)^{n-1}$$

for networks of dimension d, where c_1 and c_2 are constants depending on d.

As a consequence of Theorem 4.2, the asymptotic behavior of "Node i is isolated" is determined by $\lim_{l\to\infty} \left(1 - \left(\frac{r}{l}\right)^d\right)^n$. The values this limit takes are detailed in [67] and they lead to the following:

Corollary 4.1

$$\begin{aligned} r^d n \leq l^d &\Rightarrow \mathcal{G}(r,n) \textit{ is not asymptotically almost surely (a.a.s.) connected} \\ r^d n << l^d &\Rightarrow \mathcal{G}(r,n) \textit{ is a.a.s. disconnected} \end{aligned}$$

A non-strict necessary condition derived from Corollary 4.1 for $\mathcal{G}(r, n)$ to be a.a.s. connected is that $r^d n > l^d$, where l^d is the d-dimensional volume of the region.

4.4.4 Alternatives to employ the critical transmission range

COMPOW: a modular solution

In order to provide a framework for minimum power level adaptability, Narayanaswamy et al. present in [43] "The COMPOW protocol for power control in ad hoc networks: Theory, architecture, algorithm, implementation and experimentation". The authors explain that the proposed architecture:

- Provides modular implementation.
- Guarantees the bidirectionality of links.
- Assures the connectivity of the network.
- Asymptotically maximizes the traffic carrying capacity.
- Provides power aware routes.
- Reduces MAC contention.
- Can be used by any table driven routing protocol.

An important point for them to obtain these features in their architecture is the use of the critical transmission range as the common range of communication. Nevertheless, a reformulation of the problem is needed for their implementation.

The theoretical result shown in equation (4.18) requires, for the practical evaluation of the critical transmission range r_{crit}, knowing both the number of nodes n and the area A of the network. Furthermore, as the range may not be omnidirectional, in practice it is better to deal with the power level P, that, because of hardware construction, can be identical for all the nodes. The problem consists then of finding the power level P_{crit} that guarantees connectivity.

As a set of discrete power levels ($P_{min} = P_0, P_1, \ldots, P_j = P_{max}$) is available at each node (what is a realistic consideration for the current hardware implementations of the wireless cards), the crux of the matter is finding the power level that offers the maximum connectivity (i.e., the maximum reachable set of nodes from one given node in a multihop routing scheme).

The distributed algorithm they use to obtain it is based on parallel modularity in the OSI network layer, as can be seen in figure 4.5. When a table driven routing protocol (see section 5.2) is utilized —a requirement in this architecture— the set of reachable nodes from a distinguished one can be simply examined by looking the number of entries in the routing table.

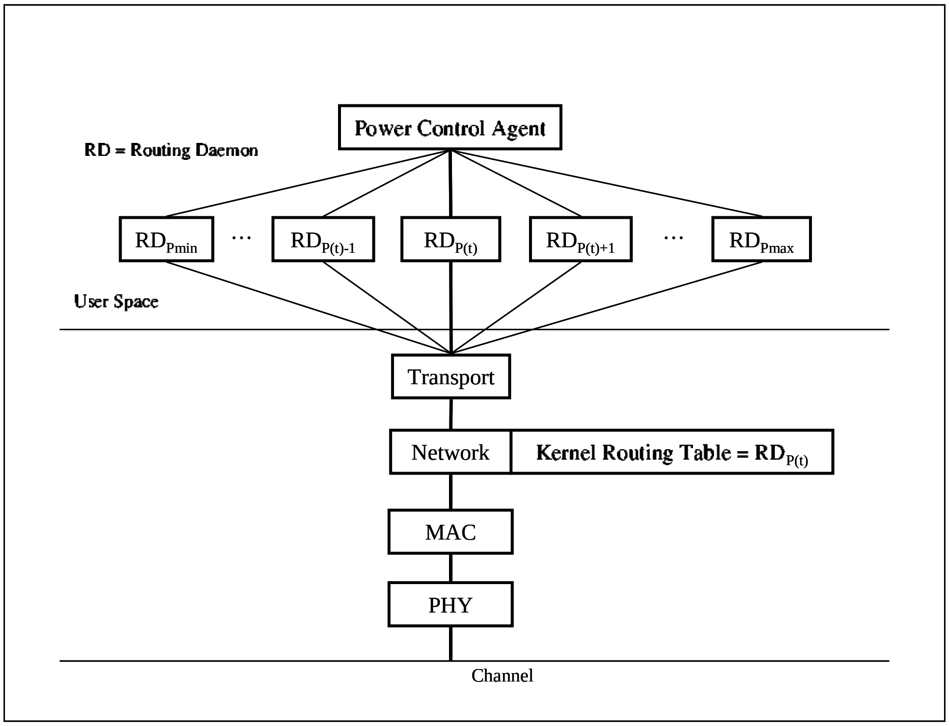

Figure 4.5: Outline of the COMPOW power control scheme

Therefore, COMPOW solution runs (through a Power Control Agent) routing daemons corresponding to the set of power levels to get the different routing tables (only the routing table corresponding to the current power level $P(t)$ is consulted by the DATA packets). In some particular sampling times the Power Control Agent, implemented on top of the Transport Layer —where UDP is used due to is demultiplexing feature— decides the current network power level and the routing table of the corresponding routing daemon is copied to the kernel routing table. Finally, in order to assure that the device drivers establish properly the power level calculated, a power level field is required in the packets.

Another factor to consider is that an excessive number of switchings of the power levels is not desirable, because it brings large latencies. Indeed, working with a suboptimal power level in some occasions becomes more efficient, due to the cost the switchings involve. Also, changes of power level are only allowed when there are no more packets of the current power level waiting for transmission.

A detailed algorithm that manages with graphs

Another solution to make use of the critical transmission range, once the network topology is known, is presented by Sánchez and Manzoni in [69]. Given a network graph $\mathcal{G}$ that represents a set of nodes distributed throughout a certain area, r_{crit} is looked for by finding the *critical link*, since it has the following properties:

- If this link is removed, the network is partitioned.
- The length of this link is the critical transmission range r_{crit}, so the pair of nodes it connects fulfill:

$$(i, j) \quad s.t. \quad |X_i - X_j| = r_{crit}$$

The authors demonstrate the the critical link is a link between two *direct neighbors*. Two nodes X_a and X_b are considered *direct neighbors* if and only if the intersection of the circle centered at X_a with the radius $|X_a - X_b|$ with the circle centered at X_b with the same radius is empty:

$$\neg \exists X_c \quad s.t. \quad |X_a - X_c| < |X_a - X_b| \wedge |X_b - X_c| < |X_a - X_b|,$$

On this basis, the algorithm they propose follows consequently these steps:

1. Mark the nodes that are direct neighbors, obtaining the direct neighbors set.

2. Create a graph by linking the direct neighbors, called by the authors Direct Neighbor Graph (DNG).

3. As the DNG may contain loops and the network connectivity can be maintained without them, some links should be removed, in the order of their length. To do this, another graph is calculated, the Minimum Spanning Tree (MST), that covers all the network, doesn't contain loops and minimizes the sum of the link distances. Two known algorithms can be used for the calculation of MST (Kruskal algorithm [70] and Prim algorithm [71], that are explained in Appendix A).

   In order to reduce the number of calculations, the MST can be directly calculated from the fully connected graph $\mathcal{G}$, so that the DNG doesn't have to be obtained in advance.

4. Select the longest link in the MST, whose length is the critical transmission range r_{crit}.

Chapter 5

Routing in Ad Hoc Networks

In this chapter we introduce the basic concepts concerning routing in ad hoc networks and, after a classification of them, we describe and compare the way of functioning of some of them. Furthermore, we consider the important issue of mobility for the routing protocols, and explain the mobility models that can be taken into account in simulations and finally, describe some protocols designed for high mobility.

5.1 Introduction

The task of routing in AHNs is accomplished in the OSI network layer, although, as we pointed out before, there is a great dependence between the MAC and the network layer in AHNs and, accordingly, issues such as transmission range and power control affect the routing algorithms to a great extent.

The first issue we should consider when analyzing the election of the routing protocol in ad hoc networks is why the conventional routing algorithms are not suitable for them. They are well tested and quite familiar for computer communications community. The problem is that they were designed for static networks; thus, they have problems to converge to a steady state in an AHN with frequently changing topology.

Another feature for conventional protocols like link-state and distance-vector is that they assume bidirectional links (if one node A can hear another node B, then B is also able to hear A), what, unlike in wired networks, is not always true in wireless radio environment.

The count-to-infinity problem of classical Distributed Bellman-Ford algorithm was the first topic to solve and some distributed shortest-path algorithms ([72], [73], [74], [75], [76]) were proposed to eliminate this problem by utilizing information regarding to the length and second-to-last hop (predecessor) of the shortest path to each destination.

5.1.1 Expected properties

Afterwards, a plethora of algorithms have been proposed in order to satisfy some of the properties an ideal algorithm for AHNs [40] should have:

- Have decentralized execution
- Provide loop free routes
- Respond quickly to topology changes
- Adapt to the traffic pattern on a demand basis
- Scale as the network size grows
- Minimize delay
- Present multiple routes to avoid congestion
- Be bandwidth efficient (minimize routing overhead)
- Utilize both unidirectional and bi-directional links
- Act power conservative and allow sleep period operation
- Guarantee security in the network
- Support quality of service and handle message priorities

Besides these characteristics the routing protocols in AHNs should accomplish, there are some relevant features to consider in the design of these algorithms.

For instance, addresses allocation concerns nodes' configuration, but if it is not managed properly —and in AHNs this involves many particularities— the routing protocols cannot function. A comparison of the existing approaches is presented in [77], as well as a recent distributed algorithm to dynamically allocate addresses.

Furthermore, the knowledge of the node locations, through a system like GPS (or, in some years, GALILEO [78], the future European Satellite Navigation System), could help in some occasions to the routing task. Also, it would be interesting find out whether an AHN might support real-time voice and video services, and the requirements this fact would introduce in the system in general and in the routing protocol in particular.

5.1.2 Possible metrics

The RFC 2501 [40] also puts forward some metrics that can be used to assess the performance of the routing protocols:

- End-to-end data throughput: average successful transmission rate.
- End-to-end delay: average time a packet takes from its source to its destination.
- Route acquisition time.
- Percentage out-of-order delivery.
- Efficiency.
- Route diversity: desirable due to both bandwidth and energy constraints.
- Route optimality.

Other possible metrics could be the level of connectivity offered and, of course, the stage of fulfillment of the above cited requirements. In [79], some metrics particulary conceived for power-aware routing (see section 6.3), such as minimizing energy consumed per packet, minimize cost per packet or minimize variance in node power levels, are detailed.

5.1.3 Basic routing protocol schemes

As many of the suggested algorithms in the literature have a traditional routing protocol as underlying algorithm, it is necessary to understand the basic operation for conventional protocols like distance vector, link state and source routing:

- Link state: in link-state routing [15] each router first obtains a view of the complete topology of the network with a cost for each link and then computes the shortest path to every other router by using, for instance, Dijkstra's algorithm.
- Distance vector: in distance vector [15] every node only monitors the cost of its outgoing links and periodically broadcasts an estimation of the shortest distance to every other node in the network. The receiving nodes then use this information to recalculate the routing tables.
- Source routing: in source routing each packet carries the complete path it has to follow around the network, which requires great overhead if the route has many hops. Given that the routing decision is made at the source, it is easy to avoid routing loops.

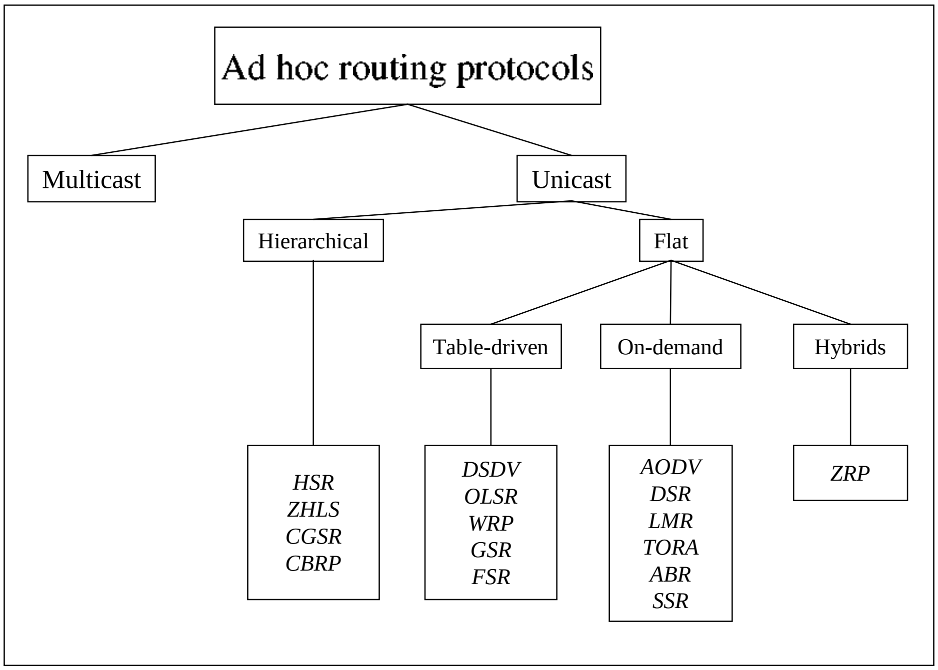

Figure 5.1: Classification of ad hoc routing protocols

5.2 Classification of Routing Algorithms

Depending on the criteria we consider, we can have several classifications of routing algorithms for ad hoc networks.

First, we can clearly distinguish between unicast and multicast routing protocols. Simultaneous communications between one source and a group of destinations in an AHN can be carried into effect by the employment of multicast routing protocols, that have particular design problems (explained in [80]), as well the general problems arisen from the ad hoc context. Some examples of multicast routing protocols for AHNs can be seen in [81].

In [82], three different points of view lead to sort the protocols in the following way:

- A) STRUCTURE:
  - Uniform protocols: all the nodes have the same roles in the routing schemes, so there is a flat routing structure.
  - Non-uniform protocols: to limit routing complexity by decreasing the number of nodes involved in a route computation. They turn out to be an attempt to achieve scalability and reduce overhead.
- B) STATE INFORMATION:

  - Topology-based protocols: the nodes maintain large-scale topology information. Link state based protocols are representative among these routing protocols.
  - Destination-based protocols: the nodes keep some local topology information, like distance vector based protocols do by storing a distance (whatsoever metric employed) and vector (next hop) to the destination.

- C) SCHEDULING:
  - Table driven (proactive) algorithms store the needed information for routing purposes in tables, which are repeatedly updated through control packets that are sent by each node. The updates can also respond to topological changes of the network.
  - On-demand (reactive) protocols, in contrast to table driven routing protocols, compute the route to a specific destination only when needed, so a routing table containing all the nodes as entries does not have to be maintained in each node. When a source wants to send to a destination, it invokes a route discovery mechanism to find the path to the destination. The route remains valid till the destination is reachable or until the route is no longer needed.
  - Hybrids schemes have also been proposed.

Figure 5.1 represents a possible general classification of ad hoc routing protocols.

5.3 Description of some of the most used protocols

5.3.1 DSDV

The Destination-Sequenced Distance-Vector Routing protocol, explained in [83], is grounded on the classical Bellman-Ford routing algorithm, which is used in the Internet under the name of RIP [15] and becomes unacceptable in wireless AHNs due to its long converging time.

DSDV is a distance vector routing protocol where in its routing table, each node maintains, for every available destination, in addition to the number of hops (used as metric) and the next-hop in the route to reach it, a sequence number for each route in order to guarantee loop-freedom (see Table 5.3.1).

The sequence number shows the freshness of a route, since they are increased when the source node detects that a route to some destination

Destination	Next Hop	Number of hops	Sequence number

Table 5.1: An entry of the DSDV routing table

has been broken (in that moment the number of hops for that route is set to infinite). Therefore, higher sequence numbers turn out to be a favorable feature for a route to be chosen. If two routes have the same sequence numbers, then the criteria is the number of hop towards the destination.

The procedure this protocol follows to take care of topology changes is based on two sorts of updating: time-driven updates, which are periodic transmission of nodes routing table; and event-driven updates, which react to failure in the links. When this happens the metric value of the update is one and it is increased while propagated hop-by-hop in the network.

In order to reduce the amount of information in the broadcasting of the routing tables, two different types of update messages are defined: full and incremental dump. A full dump sends the full routing table to the neighbors and can extent many packets whereas in an incremental update only those entries from the routing table that had a metric change since the last update are sent and it must be a single packet. In stable networks, incremental updates are sent to avoid extra traffic and full dumps are quite sporadic, while in fast-changing networks, full dumps become more effective and so, they are more frequent.

The main problem this protocol has arises from the time it needs to converge, because a route cannot be used after some time elapses from the periodic broadcast. This may be unacceptable in a mobile AHN, where the topology is expected to be very dynamic. Moreover, the periodic broadcasts also add a great amount of overhead into the network, a common setback for all the table-driven protocols.

5.3.2 AODV

Ad-hoc On-Demand Distance Vector Routing (AODV) [84] is a one of the most used algorithms for simulations of AHNs and, as considered by IETF a work in progress, is being modified continuously. Each mobile host operates as a specialized router, and routes are obtained as needed (i.e., on demand) with little or no reliance on periodic advertisements. AODV provides loop-free routes even while repairing broken links. Because the protocol does not require global periodic routing advertisements, the demand on the overall bandwidth available to the mobile nodes is substantially less than in those protocols that do necessitate such advertisements. Nevertheless it can still maintain most of the advantages of basic distance-vector routing mechanisms. The algorithm also scales to large populations of mobile nodes wishing to form AHNs.

Nodes that do not lie on active paths neither maintain any routing information nor participate in any periodic routing table exchanges. Furthermore, a node does not have to discover and maintain a route to another node until the two need to communicate, unless the former node is behaving as an intermediate forwarding station to maintain connectivity between two other nodes.

When the local connectivity of the mobile node is of interest, each mobile node can become aware of the other nodes in its neighborhood by the use of several techniques, including local broadcasts known as hello messages. The routing tables of the nodes within the neighborhood are organized to optimize response time to local movements and provide quick response time for requests for establishment of new routes.

The algorithm's primary objectives are:

- To broadcast discovery packets only when necessary.
- To distinguish between local connectivity management (neighborhood detection) and general topology maintenance.
- To disseminate information about changes in local connectivity to those neighboring mobile nodes that are likely to need the information.

AODV uses a broadcast route discovery mechanism, as is also used (with modifications) in the Dynamic Source Routing (DSR) algorithm. Instead of source routing, however, AODV relies on dynamically establishing route table entries at intermediate nodes. This difference pays off in networks with many nodes, where a larger overhead is incurred by carrying source routes in each data packet. To maintain the most recent routing information between nodes, the concept of destination sequence numbers is borrowed from DSDV [83]. Unlike in DSDV, however, each ad hoc node maintains a monotonically increasing sequence number counter, which is used to replace old cached routes. The combination of these techniques yields an algorithm that uses bandwidth efficiently (by minimizing the network load for control and data traffic), is responsive to changes in topology, and ensures loop-free routing.

5.3.3 DSR

Dynamic Source Routing (DSR) [85] is a routing protocol designed specifically for use in mobile AHNs. The protocol allows nodes to dynamically discover a source route across multiple network hops to any destination in the ad hoc network. When using source routing, each packet to be routed carries in its header the complete, ordered list of nodes through which the packet must pass.

A key advantage of source routing is that intermediate hops do not need to maintain routing information in order to route the packets they receive, since the packets themselves already contain all of the necessary routing information. This, coupled with the dynamic, on-demand nature of DSR's Route Discovery, completely eliminates the need for periodic router advertisements and link status packets, significantly reducing the overhead of DSR, especially during periods when the network topology is stable and these packets serve only as keep-alive adverts.

To send a packet to another host, the sender constructs a source route in the packet's header, giving the address of each host in the network through which the packet should be forwarded in order to reach the destination host. The sender then transmits the packet over its wireless network interface to the first hop identified in the source route. When a host receives a packet, if this host is not the final destination of the packet, it simply transmits the packet to the next hop identified in the source route in the packet's header, as well as extracts routes to all downstream nodes.

Once the packet reaches its final destination, the packet is delivered to the network layer software on that host. Each mobile host participating in the ad hoc network maintains a route cache in which it caches source routes that it has learned. When one host sends a packet to another host, the sender first checks its route cache for a source route to the destination. If a route is found, the sender uses this route to transmit the packet. If no route is found, the sender may attempt to discover one using the route discovery protocol.

The route discovery process starts with a route request message broadcast by the source. Each node that receives it adds its unique identification number into the route record till the packet reaches either the destination or an intermediate node that contains its route cache a valid route to the destination. Then, a route reply packet is sent back to the node that initiated the route discovery. To do that, the responding node may use a possible unexpired route it has in its route cache. Otherwise, if bidirectional links are guaranteed, the reverse route can be used back to the source. If symmetric links are not supported, the node may starts its own route discovery process and piggyback the route reply packet on the new route request one.

While waiting for the route discovery to complete, the host may continue normal processing and may send and receive packets with other hosts. The host may buffer the original packet in order to transmit it once the route is learned from route discovery, or it may discard the packet, relying on higher-layer protocol software to retransmit the packet if needed.

Each entry in the route cache has associated with it an expiration period, after which the entry is deleted from the cache. While a host is using any source route, it monitors the continued correct operation of that route. For example, if the sender, the destination, or any of the other hosts named as hops along a route move out of wireless transmission range of the next or previous hop along the route, the route can no longer be used to reach the destination. A route will also no longer work if any of the hosts along the route should fail or be powered off. This monitoring of the correct operation of a route in use can be called route maintenance. When route maintenance detects a problem with a route in use, route discovery may be used again to discover a new, correct route to the destination.

5.3.4 ZRP

The Zone Routing Protocol (ZRP) [86] is a hybrid proactive and reactive protocol for AHNs that tries to combine the advantages of both strategies. It divides the network (as every node sees it) into several routing zones and specifies two separated protocols that operate inside and between the routing zones.

Inside its routing zone, that is defined by the set of nodes around a given node that can be reached with a radius given by a certain number of hops k, each node employs the Intrazone Routing Protocol (IARP). This protocol works in a proactive way, so that routes can be found very fast within the routing zone. However, the particular protocol that has to be used is not specified, so any table-driven protocol is valid (AODV, OSPF, link-state routing ...). This allows different routing zones to use several protocols, what can be a problem in some situations.

When a source doesn't have a IARP route to its destination, it invokes a reactive Interzone Routing Protocol (IERP), that is in charge of finding routes between different routing zones. The protocol employs a service called *bordercasting*, which directs the ROUTE REQUEST from a node to its border nodes (those in the limit of its routing zone, i.e., with a distance of k hops to it) by multicasting. When the border node receives the request, as routing zones of neighboring nodes overlap, it can consult its IARP routing information. If the destination is not in the border node's zone, it adds its identification number to the request and *bordercast* it again. Finally, the ROUTE REQUEST reaches the routing zone that contains the destination, source routing is used, in the sense that each listed node (in the ROUTE REQUEST packet) has a IARP route to the next and previous element in the source route.

Routing zones also help to improve the quality and survivability of discovered routes, by making them more robust to changes in network topology. Once routes have been discovered, routing zones offer enhanced, real-time, route maintenance. Multiple hop paths within the routing zone can bypass link failures. Similarly, suboptimal route segments can be identified and traffic can be re-routed along shorter paths.

The main problem this protocol has is the static nature the choice of the radius of the zones has. The parameter k can be adjusted to the current network operational conditions, but this cannot be done dynamically, what can be a problem with high mobility, as we will discuss later.

5.3.5 CBRP

The basic mechanism CBRP [87] uses is to divide the nodes of an AHN into a number of overlapping or disjoint clusters in a distributed manner. One node is selected as cluster head for each cluster and it maintains the membership information for the cluster. Routes within a cluster are discovered dynamically by using the membership information. By clustering nodes into groups, the protocol efficiently minimizes the flooding traffic during route discovery and speeds up this process as well. Furthermore, the protocol takes into account the existence of unidirectional links and uses theses links for both intra-cluster and inter-cluster routing.

CBRP has the following features:

- Fully distributed operation.
- Less flooding traffic during the dynamic route discovery process.
- Explicit exploitation of unidirectional links that would otherwise be unused.
- Broken routes could be repaired locally without rediscovery.
- Sub-optimal routes could be shortened as they are used.

The operations of CBRP are utterly distributed. The major components of the protocol are: Cluster Formation, Adjacent Cluster Discovery and Routing.

The goal of Cluster Formation is to impose some kind of structure of hierarchy in the otherwise completely disorganized AHN. The algorithm is a variation of the simple lowest ID clustering algorithm in which the node with a lowest ID among its neighbors is elected as the cluster head.

The aim of Adjacent Cluster Discovery is for a cluster to discover all its bi-directionally linked adjacent clusters. For this purpose, each node keeps a Cluster Adjacency Table (CAT) that records information about all its neighboring cluster heads.

Routing in CBRP is based on source routing. It can be viewed as consisting of two phases: route discovery and the actual packets routing. Cluster structure is exploited to minimize the flooding traffic during route discovery phase. Moreover, certain unidirectional links are discovered and used, thus increasing the network connectivity.

5.3.6 Other solutions

As can be seen in Figure 5.1, several protocols have been proposed specifically for AHNs. In addition to the protocols described above, that are some of the most common ones, other solutions appear in the literature.

Thus, within the table-driven flat routing protocols, we also have the Optimized Link State Routing Protocol (OLSR) [88], the Wireless Routing Protocol (WRP) [89], the Global State Routing Protocol (GSR) [90] and the Fisheye State Routing Protocol (FSR) [91]. Also some hierarchical protocols, such as the Hierarchical Routing Protocol (HSR) [92], the Zone-based Hierarchical Link State Routing Protocol (ZHLS) [93] and the Clusterhead Gateway Switch Routing Protocol (CGSR) [94], are table-driven.

On the other hand, the on-demand mechanism for the construction of the routes is employed by Lightweight Mobile Routing (LMR) [95], the Temporary Ordered Routing Algorithm (TORA) [96], the Associativity Based Routing protocol (ABR) [97] and the Signal Stability Routing protocol (SSR) [98], among others.

5.4 Comparison between protocols

In Table 5.2 [99], the five protocols we described above are compared by taking into account some important characteristic of the ad hoc routing protocols, so that it shows how these protocols function, in relation to these features.

As Table 5.2 points out, power conservation or Quality of Service are not supported by these protocols. Anyway, there are other protocols designed specifically to support these features, as we will se in the following chapters. Also, these features will probably be incorporated to some of the solutions that define the work in progress made in the MANET group of IETF [39].

All protocols function in a distributed way, so they can easily adapt to topology changes and singe point failures in the network (as when a link —or a node— comes down) do not become critical.

	DSDV	**AODV**	**DSR**	**ZRP**	**CBRP**
Loop free	Yes	Yes	Yes	Yes	Yes
Multiple routes	No	No	Yes	No	Yes
Distributed	Yes	Yes	Yes	Yes	Yes
Reactive	No	Yes	Yes	Partially	Yes
Unidirectional link support	No	No	Yes	No	Yes
QoS support	No	No	No	No	No
Multicast	No	Yes	No	No	No
Security	No	No	No	No	No
Power conservation	No	No	No	No	No
Periodic broadcasts	Yes	Yes	No	Yes	Yes
Requires reliable or sequenced data	No	No	No	No	No

Table 5.2: Comparison between some significant ad hoc routing protocols

DSDV is the only table-driven protocol in this comparison. It arises from the protocols used in wired networks, but it includes sequence numbers to guarantee loop-free routes. The main problem it has is the convergence time, so that it lasts considerable time to react to topology changes, that may be frequent in a high mobility context. That is why AODV, the reactive version of DSDV was designed taking into account these considerations. Moreover, the authors included multicast capabilities, useful when the communications become point-to-multipoint. Both AODV and DSR employ route discovery mechanisms when a new route has to be found, since they are using a reactive routing solution. The difference is that AODV uses the information in the routing tables of the intermediate nodes to route packets, while DSR is based on source routing, so it will learn more routes than AODV, although it adds important overhead to each and every packet. DSR also has the advantage that it supports unidirectional links.

ZRP and CBRP are very interesting proposals in terms of scalability of the network, since they divide the network into several partitions. Within the zones/clusters they use a proactive scheme and be between the zones/clusters they operate with an on-demand approach, like AODV and DSR. The way the network is partitioned establishes the main difference between both solutions. In ZRP all zones are overlapping and in CBRP clusters can be both overlapping and disjoint.

None of the presented protocols take into account the traffic load when routing, so they do not perform any possible mechanism to adapt themselves to traffic requirements. The criteria employed to define routing metrics consider the shortest number of hops and the quickest responses time to a request.

Beyond this comparative analysis, we can conclude that the continuous improvements in the current ongoing protocols and the new proposals appearing within the AHNs researching community are leading to achieve increasingly sophisticated protocols, as well as the basic routing tasks for AHNs are becoming well understood. However, it is not clear yet whether one protocol can meet the diverse needs of all Ad Hoc Network contexts. As to the direction to take in order to achieve that aim, the solution should probably employ hybrid routing approaches, that seem to be able to adapt better to different situations. Of course, every possibility has to be evaluated in terms of scalability too. Anyway, the development is being done (and it should be done even more) to enhance the performance, the scalability and some particular and important areas such as security, power management and quality of service.

5.5 Mobility

One of the great challenges AHN routing protocols have to overcome is the node mobility, that is one of its natural features and, for some applications —such as disaster recovery or military environment— even a critical parameter. The dynamic change of the network connectivity and, consequently, topology, caused by the movement of the devices within the area covered by the network, has a great impact on the performance of the routing protocols. Indeed, when the node mobility in the network becomes high enough (a large enough percentage of the nodes are moving with high enough speed), the performance of the network is expected to degrade significantly, as can be concluded from the comparison between AHN routing protocols by Broch et al. [100]. Thus, as pointed out in [1], the network resources are consumed by the routing system direct proportionally to the frequency and speed of the nodes' movements.

However, as explained in section 4.3.2, the inherent mobile character of AHN can be employed as a mechanism to increase the capacity of the network.

Furthermore, mobility predictions mechanisms [101] —in addition to reacting to current node movements— can help the routing protocols to minimize disruptions caused by the topology changes and, hence, to consume a minimal amount of network resources while accommodating highly mobile nodes. This can be useful for the performance of critical applications such as voice and video.

5.5.1 Mobility models

As simulations studies are the first step to validate the characteristics and performance evaluation of the routing protocols, a common framework for them is needed. Within it, the mobility is an important issue, so realistic, flexible and easy to implement mobility models have to be established.

In order to do that, some parameters should be considered:

- Environmental conditions: shape and size of the area, presence of obstacles and their locations ...
- Direction of movement
- Speed of users
- Density of users
- Group movement

Another interesting issue that can be introduced in the mobility models is the effects of node semi-presence. This can happen in two situations:

- A node that is moving near the edge of the network.
- A node that works in "sleeping mode" with the aim of saving energy in a network that normally work with power constraints.

It is important to point out that, evidently, the mobility models are application dependent. A battlefield scenario cannot be simulated with the same mobility hypotheses and parameters than a car traffic system. Therefore, first of all, we have to take into account that mobility models designed for cellular network (Random Walk Model [102], Gauss-Markov Model[103]), as they were designed to evaluate specific problem in cellular systems —such as handoff, location management, paging, registration ... —, should not be utilized in the simulation of AHN routing protocols.

Although for the routing task the mobility pattern is a considerable factor, for other purposes this election may not have great relevance. For instance, in the work of Sánchez and Manzoni [69] about the determination of the critical transmission range in AHNs, the authors find that there is not a strong dependence on the mobility models employed in their tests.

Some existing mobility models for AHNs (see [104] and [105]) are:

- Random Walk Model [102]
- Random Waypoint Model [85]

- Random Gauss-Markov Model [103]
- Markovian Model [106]
- Column Model [104]
- Pursue Model [104]
- Reference Point Group Mobility Model [107]

Random Walk Model

The most used mobility models for AHNs are those based on the **Random Walk Mobility Model**. In this model, the speed and direction of the the nodes are memoryless, i.e., those values in a certain instant have no relation to the values in a past time interval, and they are randomly selected from the intervals $[v_{min}, v_{max}]$ and $[0, 2\pi]$.

The characteristic of this model make unrealistic movements such as sudden stops and sharp turns happen. However, it is the base of some variants rather more useful.

For instance, Johnson's **Random Waypoint Model** [85] divides the time in pause and motion periods. When a mobile host starts a motion time interval, it selects a random destination in the simulation space and moves to it at a speed uniformly distributed between the former bounds. Then, it stays there till the following motion period. A problem this model has is that it tends to concentrate the nodes in the center of the simulation area, because, although the final destination of the nodes is random, it is quite probable that, on the way to it, a node has to go through the center.

The **Random Direction Mobility Model** [108] was created to overcome the former situation. In this model, a node travels in a random direction towards the border of the simulation area, then it pauses certain time and choose another direction in the range $[0, \pi]$. Therefore, the nodes always stop in the border (never in the center). The problem this model may have is there may happen an increase in the average number of hops per communication, because the nodes are in the borders with high probability.

Group mobility models

In an AHN there are many situations where considering a certain degree of group mobility is reasonable. For instance, in a situation of disaster recovery or a military deployment, several mobile hosts carried by a group of soldiers, may move to the same physical location, in order to deploy some task, such as destroying land mines, capturing enemy attackers or simply working together in a cooperative way to accomplish a common objective. The model of this kind of situation requires some group mobility model.

Some examples of this kind of models are:

1. **Column model** [104] represents a searching activity. Nodes are distributed initially more or less like a row, and the whole row moves in the same direction, but each node may abandon the perfect line formation and get more or less closer to the others.

2. **Pursue model** [104] can be used for target tracking, which is usually done with some error and randomness.

3. **Exponential Correlated Random (ECR) model** reproduces all possible movements including individual and group by adjusting the parameters –characteristic of each group– of a motion function.

4. **Reference Point Group Mobility Model** [107] allows, as well as ECR, the partition of the network into several groups each with its own behavior. The trajectory of the group is determined by providing a path for the center of the group. Moreover, there is a random motion behavior for each node.

5.5.2 Protocols designed for high mobility

A great quantity and variety of routing protocols have been proposed for AHNs (see section 5.3.6), and, as we have pointed out above, the mobility of the nodes is a factor that affects them a great deal.

The proactive routing protocols (see section 5.2), based on storing of routing information in tables, introduce a great amount of overhead in the system when the mobility rate is high, because the updates of the routing tables may happen very frequently.

On-demand protocols look for the routes only when a packet has to be transmitted, with the subsequent reduce of the overhead. However, the nodes may experience a considerable delay while finding a route to send a packet and, even, there is no guarantee that the route found is valid, since throughout this process, the state of the links may have changed. This problem is intensified when high mobility is present.

ZRP [86] is a hybrid scheme that works locally proactive and globally reactive, and how proactive or reactive it is can be modified by a parameter k (in number of hops) that limits the *intrazone* routing and the *interzone* routing. The election of this parameter depends on both the mobility rate and the frequency of route requests. Nonetheless, this choice is static and ZRP is not able to adapt its behavior dynamically.

That is why the existence of protocols designed to manage and react efficiently to mobility is utterly justified in mobile AHNs. These protocols normally make use of geographic location strategies, such as GPS, to obtain location information of the nodes in the network.

DREAM

The Distance Routing Effect Algorithm for Mobility (DREAM) [109] tries to optimize both the frequency and the quantity of information of the routing updates as a function of the distance between nodes and the node mobility rate. Hence, the greater the distance separating two node is, the slower the nodes appear to move with respect to each other and the less information (in quantity and frequency) is needed to perform the routing task. Equally, the faster a node moves, the more often it must communicate its location. With this idea the location coordinates of a node are transmitted to the rest of the nodes in the network periodically depending on node's velocity and distance to the receivers of this information.

We have to put the accent on the fact that this protocol needs that every node knows its geographic coordinates, so that positioning systems must be used (nowadays GPS). When a node A wants to send a message to node B, it uses the location information of B and send this data packet to its one-hop neighbor in the direction of B. These nodes will forward the packet in the direction of B towards the destination.

LAR

As well as DREAM, the Location-Aided Routing protocol (LAR) [110] requires each node to obtain its geographic location through some external device using GPS. The key of this protocol is the use of this information limit the propagation of the Route Request packets to a smaller *request zone* of the AHN. With that, even the number of these messages is less.

When a node A wants to send a Data packet to a node B, first of all it has to find a route. If, in the current instant t_1, A knows the location L of B in the instant t_0, it can reasonably estimate that in t_1 B will be in a circle of radius $v(t_1 - t_0)$ around L, being v the known average speed D is travelling with. If some more information, like the direction of the movement of B, is available, the *expected zone* can be reduced.

With that estimation of the position of B, A starts a discovery route procedure, but limiting it to a *request zone*, that includes the *expected zone* and other zones. Nodes within the *request zone* forwards the Route Request until the destination is reached.

B-Protocol

Basagni et al. propose in [111] a mobility-adaptive scheme for protocols to provide scalability in AHN. This scheme has to run a *B-Protocol*, that is based on the creation and maintain of a backbone with a small subset of selected nodes that guarantees continually the realization of basic network operation such as forwarding, flooding, etc.

The first step in the way of functioning of this protocols consist on the election of the most suitable nodes (B-node) to take part of the backbone. For this task, every node computes by itself a *weight* —real number ≥ 0— that depends on what is most critical to that node the specific network application (node mobility, its battery life and the number of its neighbors). The highest the weight of a node is, the most suitable it is for being part of the backbone. Once a node is elected as a B-node, its neighbors are neglected to have the same status. The following step is the establishment of the link between B-nodes (B-links), so that the backbone is completely set up.

Chapter 6

Power control in ad hoc networks

This chapter examines the power control mechanisms available in the context of ad hoc networks. It focuses on the protocols and algorithms that appear in the literature in the Medium Access Control and Network layers, as well as the possible power-aware metrics for them.

6.1 Introduction

Managing efficiently power control in wireless ad hoc networks may carry important benefits, mainly because of its impact on battery life of the devices and on the carrying capacity of the network. Thus, the energy savings and the efficient use of the network resources are important reasons that make power management one of the most challenging problems is wireless communications.

As to the first point, the wireless devices' finite power supplies diminish their potential use for "anywhere at anytime" applications, a feature that could be desirable for the ad hoc environment, and, of course, for the WLAN solutions, which are also concerned about power control in the Physical and Medium Access Control Layers, above all. That is why power conservation techniques are common subject in different activities ([112], [113], [114], [115]) and are often employed in the hardware design of wireless systems [116].

Therefore, maximizing the lifetime of mobile wireless terminals by regulating the transmission power level is becoming of great interest, although there are other factors in energy consumption —that will be taken into account in section 6.3—, such as local computations or reception of messages, and others in different network layers ([117]). In fact, the part of a wireless network interface card that consumes most power (almost half of the total and possibly increasing this ratio in the future, by [118]) is the radio-frequency (RF) power amplifier. Hence, there is a considerable energy saving possibility in controlling the RF output power.

The degree of reduction in transmission power (and consequent energy savings) that power control can provide is quantitatively related to the basic path loss model. It determines an attenuation of the signal with the distance d like $1/d^{\alpha}$ [12], where α is the path loss exponent, typically between 2 and 6. Therefore, small difference in transmission ranges will lead to relevant fluctuations in the transmission power needed to keep the same signal quality at the receiver. As a result, either using just enough power to reach the destination or taking advantage of the multihop feature of AHN to route packets are strategies that produce remarkable savings in power consumption of the power amplifier, as shown also experimentally by Monks in [118].

Furthermore, as we analyzed in section 4.4.2, the capacity of the network can be improved by optimizing spatial packing of source-destination pairs (namely, spatial reuse, that consequently implies spectral reuse), so that a greater number of simultaneous transmissions is allowed. First, the optimum —minimum— communication range that guarantees a certain level of connectivity can be selected with this aim in the context of common range transmissions. Moreover, specific Medium Access Control Protocols have been designed to allow for both capacity and energy constraints, by modifying the basic handshaking procedure explained in section 2.3.1, as we will explain in the following section (6.2).

Thus, the power control issue has implications in several network layers, since it influences the range of communication, it determines interferences and it affects routes (because of connectivity). The survey by Jones et al. about energy efficient techniques [117] explains some possible alternatives at all the network layer of the OSI model.

At a first glimpse, the Physical Layer is concerned due to the quality of the receptions, the Network Layer because of the impact of power control on routing and also the Transport Layer is involved because higher power bears on congestion (see [118] and [119]).

Furthermore, the Logical Link Control (LLC) Sublayer, in charge of error control functionality, can be enhanced in terms of energy with several strategies:

- Adaptive Error Control with ARQ (Automatic Repeat Request). In [114], the proposed energy efficient protocol slows down data transmission when the channel conditions become degraded.

- Adaptive Error Control with ARQ/FEC (Automatic Repeat Request / Forward Error Correction). In [115], the authors try to accommodate error control schemes (they combine ARQ and FEC) to traffic requirements and channel conditions in order to achieve more optimal energy saving for each wireless connection.

- Adaptive Power Control and Coding Scheme. A dynamic power control and coding protocol for optimizing throughput, channel quality and battery life is studied in [120] and [121].

At the application layer, besides the wide collection of techniques cited in [117], it is interesting to consider the study by Naik and Wei [122] about software implementation strategies for power-conscious systems, where the authors conclude that by choosing properly the right algorithm for a problem and applying the energy saving techniques, the improvement in energy savings can be higher than 60%.

From a practical point of view, it is also worthy to mention several collateral benefits of power control under certain circumstances. For instance, power control in emergency situations mitigates the multiuser interference and, thus increasing the number of emergency or rescue nodes that may communicate simultaneously. For certain emergency applications of AHNs —for instance, hostage situations or military operations— it is desirable to maintain a low probability of interception and/or a low probability of detection, what can be attained with a low power level.

In the following sections we will focus on the proposed solutions in the MAC and Network layers, taking into account the possibility of considering energy as a metric for AHNs.

6.2 Power controlled MAC Protocols

In this section we describe some MAC protocols designed for power control in wireless multihop AHNs and compare them with the possibilities IEEE 802.11 [5] offers to attain this task.

In general, it has been observed that power control MAC protocols improve performance significantly, but the complexity of their implementation may be very high, involving a lot of signaling overhead or unreasonably high transmit powers at some nodes.

6.2.1 Power control mechanisms in IEEE 802.11

The IEEE 802.11 specification [5], explained in section 2.3.3, allows for a rough implementation of power control although the particular algorithm is left open to the manufacturers. The recommendation of the standard to let a node try to conserve power suggests the mobile devices to switch to sleep mode for some period in that case and inform of this decision to the rest of the nodes.

In the case of AHNs, this procedure is rather complicated, because the synchronization process done through the Time Synchronization Function (TSF) has to be done in a distributed way. In an infrastructure network, an Access Point (AP) is responsible for generating beacons which along with other information contain a valid time stamp that helps the nodes to adjust their local timers. In the absence of a trusted authority, every station is responsible for generation a beacon. After the beacon interval all stations compete for transmission of the "valid" beacon by using the standard backoff algorithm. The station that acquires the floor, keeps on transmitting the beacon and the rest of devices adjust their local timer to the time stamp of the winning beacon.

Therefore, packets that have to be delivered to a station in doze state, have to be buffered by the sender until the end of the beacon interval. They have to be announced using Ad hoc Traffic Indication Maps (ATIMs), which are transmitted in a special interval (the ATIM window) directly after the beacon. ATIMs are unicast frames that must be acknowledged by the receiver. After sending the acknowledgement, the receiver does not keep on in doze state but stays awake and waits for the announced packet (see Figure 6.1).

6.2.2 DBTMA-Enhanced

DBTMA (Dual Busy Tone Multiple Access) [30] is proposed in order to solve the fact that, in RTS-CTS dialogues, when propagation and transmission delays are long, the CTS packets can easily be destroyed.

This protocol splits the channel into two sub-channels at different frequencies: a data channel and a control channel. The control channel is used for the RTS-CTS handshakes and also for the transmission of two narrow-band busy tones, the *transmit busy tone* (BT_t) and the *receive busy tone* (BT_r). A host should not send if it hears any BT_r and should not consent to send if it hears any BT_t.

In [123], DBTMA is enhanced to manage power control to increase the channel reuse, above all in short-range communications. The objective is

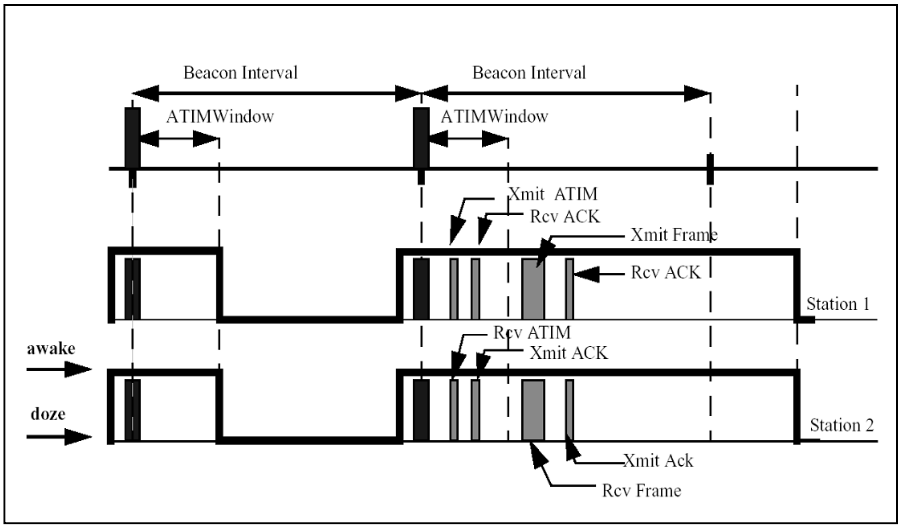

Figure 6.1: Power management for AHN in IEEE 802.11

to enable transmitters in the range of the CTS to transmit at a low enough power that does not interfere the communication of the receiver that sent the CTS. The level of the transmission for the different types of packets are:

- DATA packet and BT_t are transmitted with power control based on the power level of the received CTS.
- CTS and BT_r are transmitted at the normal (largest) power level.
- RTS is transmitted at a power level to be determined based on how strong the BT_r tones are around the requesting host.

Designing those levels in the proper way, and even taking into account the only possibility of having a discrete set of power levels, the channel utilization can be significantly increased, because overlapping between transmissions is reduced.

6.2.3 PCMA

As an attempt to offer a power controlled MAC protocol that follow the same collision avoidance principle IEEE 802.11 uses, Monks et al. propose PCMA (Power Controlled Multiple Access) for wireless packet networks [124]. The modification the authors suggest consists of changing the handshaking procedure so that nodes are allowed to transmit at a low power level.

As the collision avoidance mechanism while acquiring the floor to transmit a packet prevents from hidden and exposed station, its use is a requirement for the efficient operation of wireless medium access. However, it forbids simultaneous transmissions over the neighborhood of the intended receiver and compels to use a common communication range (see section 4.4.1). This implies that the RTS-CTS exchange must acquire the channel over the maximum range over any hidden or exposed station can cause collisions, what is the worst case in terms in power.

The solution PCMA proposes is a generalization of the collision avoidance to power control, so that a more flexible transmission model can be applied. This approach makes use of the signal strength of a received control message as an information tool for the stations to limit the transmission power of the sender of that message. This message is a generalized version of CTS.

In fact, the handshake procedure starts with a REQUEST-POWER-TO-SEND packet (RPTS), that is answered by the receiver with an ACCEPTABLE-POWER-TO-SEND packet (APTS), which is employed to calculate the minimum transmission power that will result in a successful packet reception at the receiver. This exchange precedes the DATA transmissions, that is acknowledged with an ACK packet in order to confirm the successful reception.

The noise tolerance information is used by each active receiver to notice about the maximum additional noise power it can tolerate for itself, given its current received signal and noise power levels. This noise tolerance is advertised periodically through pulses within a busy tone channel, where the signal strength of the pulse indicates the tolerance to additional noise. Therefore, a potential transmitter first observes the medium by listening the busy tones for a while to detect its maximum transmit power for all control (RPTS, APTS, ACK) and DATA packets.

The collision resolution techniques are, as in IEEE 802.11, back-off based, but PCMA employs more sophisticated algorithms.

The performance comparison with IEEE 802.11 done in [118] shows that for dense networks, PCMA exploits significantly better the spatial reuse, so that it provides improvements in terms of throughput while increasing scalability. In terms of delay, PCMA is also considerably better. The average transmission power in PCMA decreases as the load increases, although we may think the opposite, since the background noise level increases due to the higher average number of possible transmitter and a higher transmission power would be necessary. What happens is that some of the longer-range source-destination pairs are blocked by several shorter-range source-destination pairs, what is a consequence of the inherent unfairness of this protocol (and, in general, the power controlled protocols).

6.2.4 PAMAS

PAMAS (Power Aware MultiAccess protocol with Signalling) protocol [125] was designed with the unique objective of being energy conserving. It modifies the MACA protocol described above (section 2.3.1) by providing separate channel for the RTS-CTS control packets and DATA packets. The handshake functions as in MACA and power conservation feature is achieved by forcing devices that are not able to receive and send packets to turn off the wireless interface. The idea is that a data transmission between two nodes need not be overheard by all the neighbors of the transmitter. The separate control channel informs the nodes when and for how long they can power off, that happens only if they are not involved in any current communication.

This approach is suitable when processing a receiving packet is expensive compared to listening to an idle radio medium.

The results from simulation and analysis show that between 10% and 70% power savings can be attained for fully connected topologies.

6.3 Power aware routing

As we have already commented, the routing task in AHNs is performed in a distributed manner, and the nodes cooperate to maintain topology information and use multihop packet routing. The user mobility (section 5.5) allowed for AHNs complicates the routing, since it makes the topology frequently varying. Plenty of protocols (section 5.3) have appeared to partially fulfill the goals AHNs routing design implies (section 5.1.1) and several metrics are suggested in [40] to evaluate these routing protocols (section 5.1.2).

Energy as a metric in AHNs

However, some of those metrics may not lead to power savings, but, on the contrary, to the overuse of energy resources of a small set of mobiles, so that lifetime of the mobiles and network decreases. That is why, Singh and Woo propose in [79] some power-aware metrics that result in energy-efficient routes:

1. *Minimize energy consumed per packet delivery.* Under light loads, this metric will likely result in the shortest-hop path. As network load increases, this is not necessarily certain because the metric will tend to route packets around areas of congestion in the network (and possibly increasing hop-count).

2. *Maximize time to network partition.* This metric is very difficult to

optimize while simultaneously maintaining low delay and acceptable throughput, although it is very important in mission critical applications such as networks within the battlefield. The crux of the matter would be sharing the energy consumption mainly among the nodes whose turning off may derive in network partitioning, in order to maximize the fully connectivity of the network.

3. *Minimize variance in node power levels*. This metric ensures that all nodes remain up and running together for as long as possible.

4. *Minimize cost per packet* can be the employed metric to maximize the life of all mobiles.

5. *Minimize Maximum Node Cost*, where the Node Cost is the cost of routing a packet through that node at a particular instant. The use of this metric delays nodes' failures together with reducing variance in devices' power levels.

The properties of these metrics, the way they can be implemented by routing protocols and their effect on end-to-end delay are also studied in [79]. Results show that the use of power-aware metrics does not introduce extra delays over the traditional shortest-hop metric, mainly because congested paths are often avoided. The improvements in average cost per packet and average maximum mobile cost were significant, specially for large networks and moderately-loaded networks. Compared to shortest-path routing, the cost per packet is reduced by 5-30% (on top of the 40-70% improvement achieved with the use of PAMAS in the MAC level —section 6.2.4—). Therefore, this study utterly justifies the use of more energy efficient routing schemes for wireless ad hoc networks.

A similar study was further done by the same authors for the case of broadcast traffic in [126], where they show the benefits —in terms of power consumption and performance— of using a power-aware broadcast algorithm instead of the flooding algorithm for broadcasting employed in traditional networks. As this traditional flooding algorithm is not suitable for wireless networks, other approaches have been proposed, like in [127]. In [128], the problem of multicasting in AHNs is addressed from the viewpoint of energy efficiency.

In [129], a new metric for energy consumption of ad hoc network protocols is presented. This metric takes also into account the energy spent in local computations and decision-making processes, as well as it considers the available energy of a node and its location to calculate the energy level required for a transmission. It is worthy to notice, as [130] shows, that idle energy dissipation cannot be ignored compared to sending and receiving energy dissipation, so it should be regarded to as a source of energy consumption.

Immediately after, in this chapter, we describe some of the solutions in the field of power-aware routing presented up to date.

6.3.1 GAF, a geographic-informed solution

Geographical Adaptive Fidelity (GAF) [131] algorithm is based on the concept of *routing fidelity*, defined as uninterrupted connectivity between communicating nodes. The aim of GAF is holding the *routing fidelity* approximately constant while adapting node behavior to extend network lifetime.

In GAF, nodes use geographic location information (supplied by some positioning system, such as GPS) to associate themselves into fixed square grids, where all nodes are equivalent with respect to forwarding packets. Nodes in the same grid may switch between sleeping and listening, in a coordinated way that determines who will sleep and for how long. This decision is influenced by application and system information, by mean of some parameters that control, among other things, times of nodes in different states and that may be adapted for high mobility.

In the performance evaluation of the algorithm (that is used under two known routing protocols such as AODV [84] and DSR [85]), it is shown that GAF considerably extends network lifetime —it increases proportionally to node density—, since it is energy conservative, as it performs at least as well as a normal ad hoc routing protocol for packet loss and route latency.

6.3.2 SPAN, a topologically coordinated approach

The design of Span [132], a distributed coordination algorithm for AHNs that reduces energy consumption, is based on the creation and maintenance of a changing backbone of nodes awake. Turning off nodes' radio —when possible— is the main mechanism to save energy in this approach, but this restrains the possibilities of forwarding other nodes' packets around the network, losing certain degree of connectivity. In order to guarantee the maintenance of the necessary cooperation for the delivery of the packets, Span adaptively elects coordinator nodes among the network to form a backbone that rotates its component nodes.

The simultaneous aims of the algorithm, that operates under the routing layer in the OSI model and above the MAC layer, are:

- To allow as many nodes as possible to turn their radio receivers off most of the time.

- To guarantee the forwarding of the packets between any two nodes within the network with minimally more delay than in the case with all nodes awake.
- To provide, by mean of the generated backbone, as much total capacity as the original network, to avoid congestion.

Every node independently decides with a random back-off delay if it is going to become a coordinator node, basing its decision on two factors: the amount of remaining battery energy and the number of pairs of neighbors it can connect together. This assures, with high probability, both a similar energy consumption rate among the nodes and the existence of a capacity-preserving connected backbone.

Span's performance evaluation was done in [132] by using it together with Geographic Forwarding Routing (similar to GLS [57]), and it shows that Span preserves both network connectivity and capacity (better with dense network and higher loads), while it does not notably increase delivery latency and number of hops each packet traverses. Also, it provides significant energy savings, but they increase only slightly as the density increases, because their implementation of Span [132] uses the power saving features of 802.11 (see section 6.2). This makes nodes periodically wake up and listen for traffic advertisements, although it allows non-coordinator nodes to still receive packets when operating in power saving mode.

6.3.3 Other solutions for power-aware routing

From the routing point of view, other power-aware solutions have been proposed in recent years.

PARO (Power-Aware Routing Optimization) [44] tries to take advantage of the fact that, in terms of energy efficiency, it is better to use several short hops than a long one in AHNs routing. This principle contrast to the metrics employed in the classical ad hoc routing protocols (see section 5.3), that normally attempts to minimize the number of hops between source-destination pairs. Thus, PARO reduces the aggregate transmission power consumed by the network by using one or several intermediate nodes —"redirectors"— as packets forwarders.

BECA (Basic Energy-Conserving Algorithm) and AFECA (Adaptive Fidelity Energy-Conserving Algorithm) were presented simultaneously by Xu et al. in [133]. The former establishes the main idea that a node that is not involved in sending, forwarding or receiving data does not need to be listening and consuming power. In AFECA, each node switches between sleeping and listening modes, with randomized sleep times proportional to the number of nearby nodes. As the density increases, more energy can be saved, as generally happens in all the power-aware routing algorithms.

LEAR (Local Energy-Aware Routing) [134] achieves a balanced energy consumption among all participating nodes, by taking into account the remaining battery a node has left to decide whether it participates or not in the selection process of a routing path, when this is needed. It is based on DSR [85] and outperforms it in terms energy usage by 35%, by the simulations done in [134].

Chapter 7

Quality of Service in Ad Hoc Networks

This chapter summarizes the possible strategies for Quality of Service (QoS) support in ad hoc networks.

7.1 Introduction

The concept of QoS delivered in a network refers to "a guarantee by the network to satisfy a set of predetermined service performance constraints for the user" with respect to quantitative parameters such as end-to-end delay, delay variance (jitter), available bandwidth or probability of packet loss. In the networks design, resource management is extremely important to offer to the end users and their applications the adequate conditions for proper transmission of their traffic flows. Evidently, there is a relation between the demand and the resource availability, that has to be handled to provide Quality of Service. The network users would like to experience new applications that require bandwidth and impose constraints in the network, while the network designer is in charge of managing efficiently the resources to make the users' demand feasible.

Since real-time IP applications —such as VoIP (Voice over IP) or video streaming— are becoming popular, many techniques have appeared to support QoS in the context of an IP architecture. Thus, architectures like ATM [136] (Asynchronous Transfer Mode), the IntServ/RSVP [137]/[138] (Integrated Service/Resource reSerVation Protocol), DiffServ [139] (Differentiated Service) or the emerging MPLS [140] (Multi-Protocol Label Switching) incorporate QoS as a central concept. They offer different tradeoffs between complexity and efficiency and, consequently, the have diverse advantages and setbacks that make them suitable in distinct contexts. An overview of these technologies can be found in [141].

The support of quality of service in ad hoc networks is different from the previous cases, although the AHNs can be based on an IP architecture, since the bandwidth restrictions of the wireless medium and the dynamic topology introduce special constraints that make incorporating QoS to the network a challenging task. In fact, all the work done in order to support QoS in the Internet cannot be directly used in AHNs.

Therefore, there is much work to do in this area of research, regardless whether the solution is to adapt the existing mechanisms of QoS to AHNs or to create a new framework of QoS specific to AHNs. A extensive review of the state of the art in 1999 is included in [142].

In any case, a structured vision of how to coordinate the different QoS architecture components is necessary. A interworking approach for heterogeneous network QoS systems is given in [141], where diverse mechanisms are taken into account. Some of them are:

- Traffic engineering
- Signaling
- Admission control
- Packet classification
- Packet scheduling
- Buffer management
- Traffic policing
- Adaptivity

An ad hoc oriented view, given in [143], simplifies the problem and splits the existing research on QoS in AHNs into:

- QoS models
- QoS resource reservation signaling
- QoS routing
- QoS Medium Access Control

The QoS model defines the framework in which certain kind of services could be provided in AHNs. The rest of the components should harmonize their behavior to fulfill the goals intended by the QoS model. The QoS signaling is in charge of the reservation and releasing of the resources. The QoS routing tries to establish paths through the network that meet end-to-end QoS requirements. On the other hand, the QoS support in the MAC layer, in addition to performing the basic wireless MAC functions, such as managing the medium contention and dealing with the hidden/exposed terminal problems, must consider delay characteristics to support real-time traffic. Otherwise, it may happen the MAC protocol is not coordinated in terms of QoS with the upper layer mechanisms, so resources reservation becomes unfeasible.

Other components included in a QoS architecture, such as scheduling and admission control, can be supplied as additional functionalities and, actually, can be utilized for AHNs with little adjustments with respect to their original design for other frameworks, as pointed out in [143].

7.2 QoS models

The general AHNs' characteristics and goals, as well as the particular applications the design of the network is oriented to, should be considered to define a QoS model for them. Thus, dynamic topology and time-varying link capacity condition the model to a great extent. Equally, possible interconnection between an AHN and the Internet would affect the definition of the model.

First of all, we are to comment the limitations of the QoS models utilized for the Internet in the context of ad hoc networking.

The IntServ/RSVP architecture provides quantitative QoS for every individual flow throughout the network. In order to attain that, every router has to maintain a flow-specific information including bandwidth requirement, delay bound, cost, etc. RSVP is the signaling protocol in charge of reserving resources before transmissions. As every node in an AHN acts as a router, the functions of the devices would increase (they should perform admission control, packet classification, packet scheduling ...) and, therefore, the cost in terms of storage and processing that information would be troublesome, above all when the number of flows becomes high, so that scalability problems would appear. Moreover, the RSVP control packets would contend for the bandwidth with the data packet, consuming a considerable percentage of the bandwidth in AHNs. However, this model provides an end-to-end solution for QoS support.

The DiffServ architecture emerged as alternative and complement for the previous IntServ QoS model. It uses the Type Of Service (TOS) bit in the IP header to provide a limited number of aggregated classes. Within this model, QoS can be deployed using implicit admission control with no requirement of end-to-end signaling. As interior routers do not need to keep per-flow state information, it does not burden the routers and, therefore, has better scalability properties. The main problems for the validity of this model in AHNs are two:

1. Defining a node as a interior router or boundary router in a AHN is not straightforward. Source nodes could play the role of the boundary routers, and the rest would be interior routers. However, these roles are continuously changing for every node, which should therefore be able to perform both functionalities, with the corresponding storage cost that it involves.

2. The Service Level Agreement (SLA) necessary for a customer to receive from its Internet Service Provider (ISP) Differentiated Services is not an easy concept to extrapolate to AHNs, because there is no obvious scheme for the mobile nodes to negotiate the traffic rules.

Nevertheless, an architecture and experimental framework for supporting QoS in wireless networks using Differentiated Services has been proposed in [144], where it is shown that there is a need for a signaling protocol which considers low bandwidth and mobility characteristics.

A methodology that identifies the needs of QoS support is shown in [146].

FQMM, a QoS model for AHNs

A first attempt of defining a QoS model for AHNs was done in [145], where IntServ and DiffServ features are also analyzed as to their possible application to AHNs. FQMM (Flexible QoS Model for MANET) is intended for small AHNs (up to 50 nodes) that use a flat topology. It defines three types of nodes, as in DiffServ: ingress nodes (senders), interior nodes (forwarders) and egress nodes (receivers), so that the roles of the nodes within the network can dynamically change.

The determination and allocation of resources at various points in the network employs a hybrid policy in FQMM between per-flow provisioning in IntServ and per-class provisioning in DiffServ. This is done by first dividing traffic into classes. Then, the highest priority class is given per-flow provisioning, while the others are given per-class provisioning.

Moreover, a traffic conditioner at the ingress nodes polices the traffic according to the traffic profile after a valid route is found by the QoS routing protocol. These traffic profiles are differentiated to keep consistency between sessions (that can be per flow or per aggregate of flows) and can be adapted to the dynamics of the network.

The routing protocol utilized within this model should be consistent with the provisioning policy previously defined. Therefore, every route found by the routing protocol used should pass a QoS checking. A more efficient approach for the routing should consider the QoS constraints together with the routing tasks in the protocol design.

However, FQMM does not suggest the ways to implement its QoS model in AHNs, either the QoS routing protocol or the resource management strategies.

7.3 QoS Signaling

The main resources to manage in the context of AHNs are the link bandwidth, the nodal buffers and the battery power. QoS signaling is in charge of reserving and releasing resources, and control the behavior of flows in the network. In order to attain this aim, control signaling information has to be exchanged between the nodes and interpreted properly to trigger the necessary processes.

The interchange of control information between nodes is done by the signaling protocol and it should be carried into effect with the minimum overhead. A first approach consists of piggybacking control information with data packets, what is used in in-band signaling.

This is not always possible due to the complexity of the signaling system. In these cases, when the requirements of the signaling system demand it, explicit control messages have to be sent through the network. This alternative is called out-of-band signaling and, evidently consumes more network bandwidth than in-band signaling. However, it offers more flexibility and allows to perform more sophisticated signaling functions, that are sometimes needed in the QoS support.

Therefore, there exists a tradeoff between supporting complex functionalities and consuming network resources. Evidently, when a complex out-of-band signaling system has to be used, there are many control packets going through the network and competing for the transmission channel with data packets. Consequently, the performance, in terms of throughput, is significantly degraded, what is especially critical in bandwidth constrained networks like AHNs.

7.3.1 dRSVP, a dynamic vision of QoS

Resource reSerVation Protocol (RSVP) [138] is the signaling system used in the Internet. It is an out-of-band signaling system and it is typically related to the IntServ model of QoS.

When a node wants to send information to another node, it sends first a PATH message to the receiver, including the traffic characteristics specifications. The receiver sends back a RESV packet to the source with the resource requirements for the flow. Along the way back to the sender, the RESV packet tries to reserve resources, if possible (otherwise an error message is sent to the receiver), in the routers the packets traverses.

RSVP is not suitable for AHNs mainly because of the heavy overhead that it introduces and its lack of flexibility to adapt to dynamic conditions.

A dynamic conception of the QoS in AHNs is given in [147] by expanding the notion of "reservation". With "Dynamic Qos", a range of values are considered in the resource reservation request, so that the network can provide services within this range. Applications request QoS by specifying the minimum and maximum level of service they can accept and utilize. At the same time, they adapt themselves to the current conditions of QoS support that is offered at some moment by the network. This prospective makes sense when it is assumed that the link layer is able to inform the upper layers of bandwidth changes.

With this adaptive point of view of QoS, the authors of [147] modify RSVP to make it support the resource reservation range, resulting a protocol called "Dynamic RSVP" (dRSVP). When topology changes occur, this signaling protocol does not attempt to guarantee a fixed level of connectivity and, hence, it does not try to make new routes available to support the QoS commitments. Instead, it allows QoS levels to be adjusted accordingly.

As new application flows appear within the network, instead of rejecting them to assure the QoS requirements of other flows, the bandwidth allocation algorithm dRSVP uses tries to adjust the allocation for each flow, so that all flows can be accommodated.

The validity of this view of QoS support is dependent on the adaptability of the users and applications. Applications should be capable of modifying dynamically their behavior to adapt to changing network conditions, so that they allow certain level of degradation in the service.

7.3.2 INSIGNIA, an in-band signaling system

INSIGNIA [148] is an in-band signaling system for QoS support in mobile AHNs. The QoS framework it considers allows packet audio, video and real-time data applications to specify their maximum and minimum bandwidth needs. The architectural components comprised in the INSIGNIA framework can be seen in Figure 7.1.

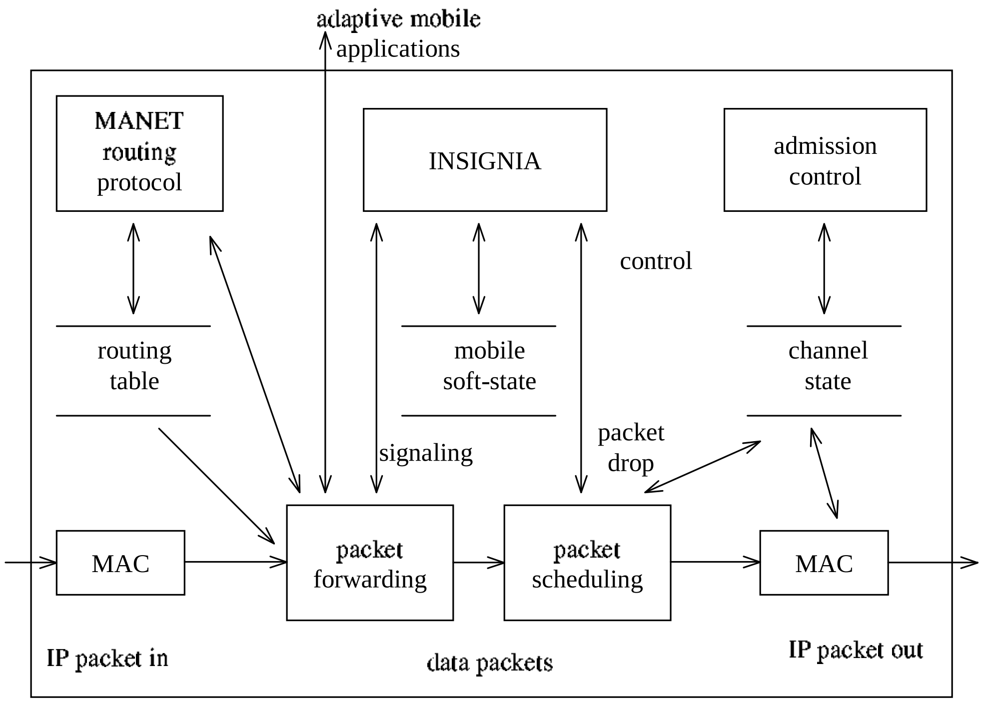

Figure 7.1: Wireless flow management model at a mobile host [148]

The packet forwarding module classifies the incoming packets and redirects them to the corresponding modules (routing, INSIGNIA, local applications and packet scheduling modules). If a received IP packet includes an IP option (called the INSIGNIA option) it is forwarded to the INSIGNIA module, which processes it. Meanwhile, depending on the final destination of the packet, it is sent to a local application —if the mobile host is the destination— or to the packet scheduling module, from where it will be forwarded to the next hop determined by the ad hoc routing protocol. The packet scheduling module is in charge of fairly allocating the resources to different flows.

The INSIGNIA module plays an important role in establishing, adapting, restoring and terminating end-to-end reservations. The signaling system is designed to be lightweight in terms of the amount of bandwidth consumed for network control —it piggybacks QoS signaling messages onto the data packets— and to be capable of reacting to fast network dynamics induced by rapid host mobility or wireless link degradation. It uses QoS periodical reports to advice the source node of the status of real-time flows. Therefore, the source can adapt the flows to the dynamics of the network.

Together with the admission control module, INSIGNIA does the bandwidth allocation for every flow if their QoS resource requirement can be satisfied. Otherwise, the flow is degraded to the best-effort service.

In general, it is an effective signaling protocol for AHNs and can be used to provide QoS for real-time flows. Nevertheless, the scalability problem remains present owing to necessity of keeping the flow state information in the mobile hosts.

7.4 QoS Routing

The QoS routing is in charge of finding routes that satisfy the QoS requirements. By doing it, the efficient use of the network resources can be accomplished and the possible resource reservation made by the QoS signaling component makes sense.

The dynamic feature of the topology in AHNs (because of varying links conditions and mobility of the nodes, mainly) is the principal impediment in the search of a QoS path, due to the difficulty of maintaining the precise link state information. The end-to-end requirements this QoS path has to meet are related to delay, delay jitter bounds, bandwidth demand and, even, multi-metric constraints. It has been proved in [149] that if QoS contains at least two additive metrics (the values can be added along the routing path), then the QoS routing is an NP-complete problem. Suboptimal algorithms can be used, especially for large networks, so that an optimal path based on a single primary metric (for instance, bandwidth) is selected first and, from that one, the next valuable metrics are used. Therefore, an approximate solution is the only possible approach if we want to consider several QoS metrics, as can be the case if we take into account, for instance, link cost and delay.

Thus, QoS turns out to be a challenging task, although there is not yet a convincing solution in the field. Actually, as pointed out in [150], in heavy traffic situations, guaranteeing QoS for lesser priority traffic may be very difficult or even impossible. However, the current ongoing protocols in the MANET working group of the IETF [39] are being improved to support QoS (for instance, AODV or DSR).

Furthermore, other specific approaches have been proposed, as we will describe below.

7.4.1 CEDAR

In [151], a Core-Extraction Distributed Ad hoc Routing (CEDAR) algorithm was proposed for QoS routing in ad hoc network environments. The routes that are searched by the protocol are those that likely satisfy a minimum bandwidth requirement from the source to the destination.

To perform this function, CEDAR has three key elements:

- *Core extraction* selects and maintains a group of nodes that establish virtual links between each other if they are close enough. Each *core* host maintains the local topology of the hosts in its domain, and also performs route computation on behalf of these hosts.
- *Link state propagation* depends in CEDAR on the stability and bandwidth of the links. The bandwidth availability information of the stable links should be made known even to far away located nodes in the network, while information about dynamic or low bandwidth links should remain local. This is implemented by using "increase/decrease" waves (which in turn use a core broadcast mechanism) to react to network changes. Thus, the increase wave is slow-moving (it experiences a delay at each hop), because it notifies stable high-bandwidth links, while the decrease, that indicates unstable or low-bandwidth links, way is fast-moving.
- *Route computation* first establishes a core to the destination. This is done probing on the core, and the resultant core path is cached for future use. The core path provides the directionality of the route from the source to the destination. Using this directional information, CEDAR searches for a stable QoS route. Finally it dynamically recomputes QoS routes upon link failures or topology changes.

With these mechanisms, CEDAR achieves a great level of robustness and a quick adaptability to dynamics in the topology. At the same time, it computes good (but not optimal) routes, and accomplishes bandwidth requirements with high probability as long as the routes it has selected exist.

7.4.2 Ticket-based routing

Chen and Nahrastedt proposed in [152] a ticket-based probing algorithm for QoS routing in AHNs. This algorithms has a multi-path feature, i.e., it searches among multiple paths for a feasible one. However, instead of doing an exhaustive search, the algorithm uses tickets —a ticket is a permission of searching a path— to limit the number of candidate paths, so that some information imprecision in the link state is tolerated.

When a node wants to look for a QoS path to a destination, it issues a probe message with a certain number of tickets, that determines the maximum number of searched paths. Therefore, connections with tighter requirements issue a higher number of tickets. An important issue for the nodes that receive the probe message is where to forward it and how to split the tickets among different paths. In general, links with larger residual bandwidth gets more tickets. Moreover, this algorithms employs techniques of re-routing, path redundancy and path repairing to dynamically maintain the multiple paths.

Another ticket-based routing algorithm for AHNs is proposed in [153], where multi-path routes, defined as those that contain several sub-paths from to the destination, are considered.

7.5 Providing QoS in the MAC layer

The QoS support in AHN has to be carried into effect in different network layers. However, the mechanisms of the upper-layer (QoS signaling and QoS routing) assume the existence of a MAC protocol that, at least, support reliable unicast communication and provide resource reservation for real-time traffic.

Many of the MAC protocols that have been proposed in the AHNs context have been designed mainly to solve the problem of medium contention (see section 2.3), although there are some approaches that are specifically oriented to accomplish certain goals, like energy constrained operation (section 6.2). In that sense, some alternatives have also been considered to provide QoS support in the MAC layer, both for the environment of Wireless LANs ([154] and [155]) and for the AHN context.

Examples of ad hoc oriented QoS MAC protocols are considered in [156], where the Black-Burst (BB) contention mechanism provides QoS guarantees to real-time traffic, and in [157], where MACA/PR is presented. This protocol provides rapid and reliable transmission of non-real-time datagrams as well as guaranteed bandwidth support to real-time traffic.

Chapter 8

Simulations of the connectivity

In this chapter we show the results obtained by simulation of the probability of having a fully connected network as a function of the transmission range.

The goal of this simulation study is to understand what the trend of the network behavior in terms of connectivity is, when we use the transmission range as a parameter.

As we explained in section 4.4.1, a common range of communication is needed if we assume the use of certain handshake procedures in the MAC level —those derived from MACA and IEEE 802.11— and we take into account some probable necessities of routing protocols (acknowledgements). Also, section 4.4.2 analyzed the tradeoffs between connectivity and throughput when the transmission range is a parameter.

In Figure 8.1 we can see some examples of how a random network has different levels of connectivity depending on the transmission range. At some point, there should exist a critical value of this parameter that assures the existence of a route between every two nodes while minimizing the energy consumption (because the lower r is, the less power is needed in every node to use that communication range).

The region where we do the simulations is a square of unit area. Within it, we place randomly a variable number of nodes. Then, we evaluate the connectivity as a function of the transmission range r.

An essential issue in this type of study should be the selection of the metric utilized to define connectivity. In our case we will consider and evaluate only the case of fully connected network, so we are not affected by any criteria in the selection of the metric. Anyhow, in section 4.4.2 some

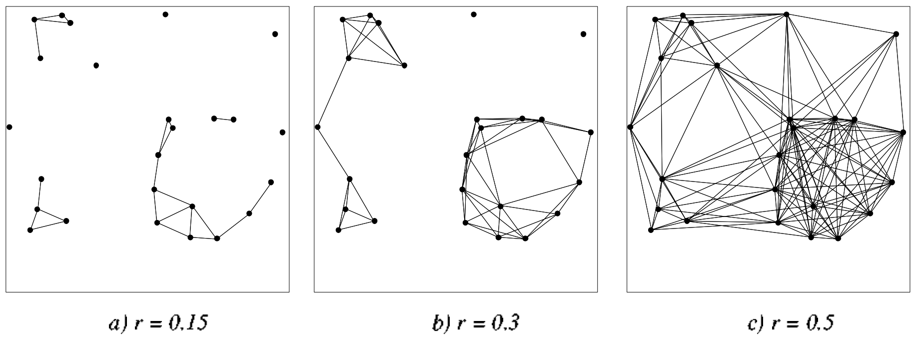

Figure 8.1: The network achieves different level of connectivity depending of the transmission. The graphs correspond to $n = 25$ nodes

possible metrics for connectivity are cited.

For instance, [158] studies the probability of a two-hop connection in a random mobile network when the nodes are randomly distributed, by observing the size of the intersection between the coverage areas of two given nodes and the probability of any other node to be in there.

In our case, the connectivity matrix that determines if there is a route between any two nodes was calculated from the one-hop connectivity matrix by using a Bellman-Ford approach. The algorithm initializes the distance from every node to itself to 0 and to all other nodes to 1 or ∞ depending on whether the nodes are directly connected (one hop) or not. The Bellman-Ford algorithm proceeds by looping through all of the edges in the graph, applying the relaxation operation to each edge. The process of relaxing an edge (X_1, X_2) consists on testing whether we can improve the shortest path estimation to X_2 found so far by going through X_1 and, if so, updating —decreasing, because it starts from a upper bound (∞)— the distance estimation and, if obtaining the route is necessary —that is not our case, because we only want to know the number of hops—, the predecessor field stored in a vector. In the Bellman-Ford algorithm, each link is relaxed several times, in contrast to Dijkstra's algorithm.

There are also other ways to calculate the multihop connectivity, such as the algorithm proposed by Miller in [159], but we have used Bellman-Ford because of simplicity.

The simulations we did with Mathematica 4.1. (we evaluated the connectivity of 5000 random networks for every case of number of nodes) show (see Figure 8.2) the expected behavior and we can see that the greater the number of nodes is, for the same communication range, the higher probability of fully connectivity we achieve with a lower transmission range. This involves, as obvious, that the critical transmission range is lower when there are more nodes in the network.

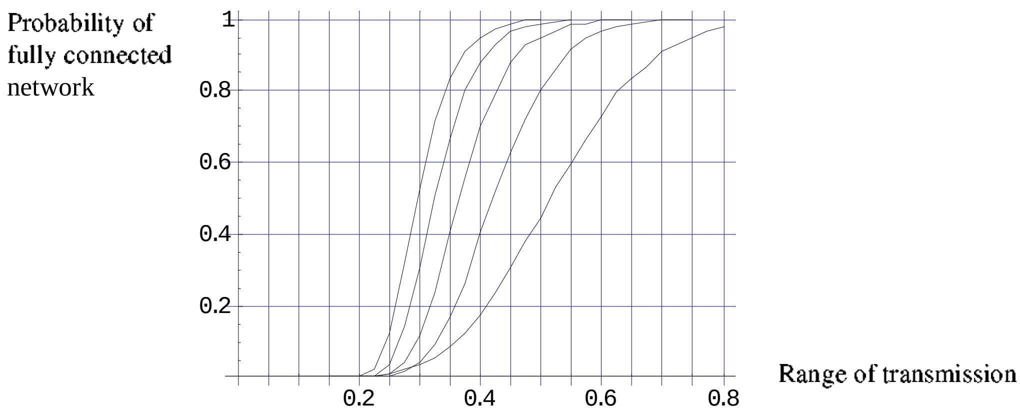

Figure 8.2: Probability of fully connected network as a function of the communication range r. The curves correspond, from right to left, to $n = 5, 10, 15, 20, 25$

It would be interesting to know the asymptotic behavior of the function we represent when the number of nodes is significantly high, in order to deal with scalability issues. So far, we can simulate the network with a maximum number of nodes of 25, due to the quickly increasing simulation times. A higher number of nodes could help us validate the theoretical results of the asymptotic behavior of the critical transmission range, shown in section 4.4.3.

Chapter 9

Conclusions

This work is a literature survey of ad hoc networks, natural inheritors of radio packet networks. Although there are also other technological approaches in the market of wireless communications, ad hoc networks seem to be the most suitable solution for some contexts where their ease of quick deployment and dynamic way of operation become essential. Thus, nowadays, as in the 70's —when interests about packet radio networks started to appear—, military projects play an important role in the research activities in the field, since the applications that can take the best advantage of the features of ad hoc networks are military-oriented, like communications in the battlefield or emergency and rescue operations. Even so, they also appear to be the specific solution for sensor networking, that comprises a set of applications that can respond both to civil and military requirements. Nevertheless, in the future —when these networks are commercially available— the flexibility they offer will make them, useful for other civil applications, like spontaneous meetings or home communications.

A question that is still open in the applicability of ad hoc networks is the sharing of the resources. Sometimes, such a distributed architecture as ad hoc networks have becomes a robust solution, but it also makes the users forming the network employ their own resources to forward other users' packets and to manage the network itself. In an energy and capacity constrained environment, the design should induce or even force the nodes to work most of the time in a collaborative manner. Moreover, the fact that packets are forwarded by nodes which are not the intended receivers makes group membership and, in general, security considerations a great challenge.

The inherent multihop nature of AHNs also has important implications in terms of performance. Indeed, the capacity available per node is conditioned by the fact that every node is not only sending its own data, but also forwarding other nodes' packets throughout the network. It has been shown that the theoretical bounds on capacity per node for a two-dimensional AHN in the best possible conditions —when nodes are optimally placed, range of transmission is optimally chosen and the traffic patterns are known— decreases with the number of nodes n like $1/\sqrt{n}$. This limit is not encouraging in terms of scalability. Actually, it is observed in the performance analysis of ad hoc routing protocols that the capacity is the restraining cause. Consequently, it is worthwhile to examine which factors are affecting it and how to deal with them in order to manage carrying capacity efficiently.

In this work, we have identified three factors that can modify capacity and that lead to several strategies intended to maximize it. Those factors are the traffic pattern, the location and mobility of the nodes and the range of transmission.

Predominantly local traffic patterns constitute the best case in terms of capacity, whereas random traffic pattern is the worst case one may choose for AHNs. Therefore, this conclusion suggests that there are some possibilities to construct large networks with reasonable per node capacities when local traffic patterns are common. Strategies that consider including pure relaying nodes in a static network have been demonstrated not to be worthy, because the necessary number of additional nodes becomes very high to improve significantly the capacity. However, the effects of relaying appear to be interesting in a mobile network, and mechanisms like multiuser diversity, when they properly manage the problem of the high delay, may be, in the future, a source of potential increase of the capacity.

Other techniques, like multipacket reception (MPR), allowed by the continuous advances in signal processing and multiple antenna technologies, have changed some of the underlying assumptions made by conventional MAC protocols, and, indeed, its effects on networking layers other than the physical layer are not still fully understood, as well as their possible impact on the performance and design of MAC protocols. In terms of capacity, it has been shown that MPR does not change the asymptotic capacity law, but it improves its coefficient, especially in dense networks.

The range of transmission is one of the critical parameters in the design of ad hoc networks. It bears on the energy consumption of the mobile devices, as well as on the throughput available per node and the connectivity, so that it regulates the tradeoff between the two latter properties. In this sense, it is possible to estimate, theoretically and practically, the critical transmission range that, while being the least possible to minimize energy consumption and reduce interference (i.e., to increase throughput), guarantees network connectivity. In this work we have considered the two prospectives. For both points of view, an interesting source of future research is to try to face the problem from a probabilistic formulation, that could provide approximate but flexible solutions.

On the other hand, one of the challenges the design of ad hoc networks have is the proper integration of the different network layers and the allocation of responsibilities for each of them. It has been shown that there are several details that obstruct the sheer independent design the OSI model proposes for the network layers. It is still not clear whether some functionalities like power control, security or quality of service should belong completely to one network layer or whether the duty of managing them should be done with a cross-layer prospective. In our opinion, until a framework for independent operation of the network layers appears, it is not possible nowadays to design an ad hoc network without taking into account the cross-layer relationships that some parameters like bandwidth, frequent topology changes and energy constraints induce.

Nevertheless, in the literature several solutions for the different network layers have appeared. As to the Medium Access Control layer, critical in wireless communications, the best available possibility for a real implementation seems to be, up to date, IEEE 802.11, although it was not specifically designed for multihop communications and to support efficently ad hoc networks. That is why many approaches have been presented, and especially interesting are those that pay attention to power control and quality of service, in addition to efficiency and capacity management.

In the routing layer, we can conclude that the continuous improvements in the current ongoing protocols and the new proposals appearing within the AHNs research community are leading to increasingly advanced protocols, as well as the basic routing tasks for AHNs are becoming well understood. However, it is not clear yet whether one protocol can meet the diverse needs of all ad hoc network contexts. As to the direction to take in order to achieve that aim, the solution should probably employ hybrid routing approaches —like ZRP— that seem to be able to adapt better to distinct situations. Of course, scalability should be considered too for evaluating the possibilities. Anyway, the development is being done (and should be done even more) to enhance the performance, the scalability and some particular and important areas such as security, power management and quality of service.

Managing efficiently power control in wireless ad hoc networks may carry important benefits, mainly because of its impact on battery life of the devices and on the carrying capacity of the network. Thus, the energy savings and the efficient use of the network resources are important reasons that make power management one of the most challenging problems is wireless communications. Several mechanisms for power control have been proposed in different network layers, since it influences the range of communication, determines interferences and affects routes (because of the connectivity). Our attention was mainly focused on the MAC and routing protocols that were designed specifically to take into account the energy constraints AHNs have. The complexity of their implementation may be high and even they may involve additional overhead, but it has been observed that these mechanisms improve the performance, above all in terms of some power-aware metrics that can be defined.

Equally, mechanisms to support QoS in AHNs become necessary when the applications that run within the networks need certain delay or bandwidth requirements. The dynamic behavior of AHNs, in terms of varying link state and topology changes, complicates to a great extent the QoS support and makes it a great challenge that implies the coordination of the routing and MAC network layers with a signaling system in charge of the resources reservation.

All these issues and features of ad hoc networks, most of the times related to each other in the design of the network, involve a great incentive for the research community to provide of a flexible, efficient and stable architecture for ad hoc networks, which will certainly revolutionize the wireless communications in the near future.

References

[1] S. Ramanathan and M. Steenstrup. *A Survey of Routing Techniques for Mobile Communications Networks*. Mobile Networks and Applications, 6(1): 89-104, 1996.

[2] IEEE 802.15 Working Group for WPANs, http://grouper.ieee.org/groups/802/15

[3] http://www.bluetooth.com

[4] The Wireless LAN Association (WLANA), http://www.wlana.com

[5] IEEE Computer Society LAN MAN Standards Committee, *Wireless LAN medium access control (MAC) and physical (PHY) layer specifications*. IEEE Standard 802.11-1997. The Institute of Electrical and Electronics Engineers, New York, NY, 1997.

[6] http://www.etsi.org/bran

[7] http://www.hiperlan2.com

[8] Per H. Lehne and Magne Pettersen. *An Overview of Smart Antenna Technology for Mobile Communications Systems*. IEEE Communication Surveys, Vol. 2, No. 4, 1999.

[9] RadioLAN, http://www.radiolan.com

[10] R.C. Dixon. *Spread Spectrum System with Commercial Applications*, 3rd Edition. John Wiley & Sons, INC. USA, 1994.

[11] J.J. Garcia-Luna-Aceves and Jyoti Raju. *Distributed Assignment of Codes for Multihop Packet-Radio Networks*. Proceedings of IEEE MILCOM'97, Monterey, California. November 2-5, 1997.

[12] Jerry D. Gibson. *The Mobile Communications Handbook*. CRC Press, INC. USA, 1995.

[13] Chen, K.-C. *Medium access control of wireless LANs for mobile computer*. IEEE Network, Vol.8, Issue 5, pg. 50-63.

[14] Ajay Chandra, V. Gummalla, and John O. Limb. *Wireless Medium Access Protocols*. IEEE Communications Surveys and Tutorials. Second Quarter, 2000 Vol. 3, No 2.

[15] Tanenbaum, Andrew S., *Computer Networks*, Prentice Hall Inc, USA, 1996

[16] B.M. Leiner, D.L. Nielson, and Tobagi F.A. *Issues in packet radio network design*. Proceedings of the IEEE, vol. 75, No 1. January 1987.

[17] P. Karn. *MACA a new channel access method for packet radio*. Proceedings of the 9th ARRL/CRRL Amateur Radio Computer Networking Conference, September 1992.

[18] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. *MACAW: A media access protocol for wireless LAN's*. In Proceedings of the SIGCOMM '94, pages 212–225, August 1994.

[19] M. Gerla, R. Kapoor, M. Kazantzidis, P. Johansson. *Ad hoc Networking with Bluetooth*. Proceedings of WMI at Mobicom, Rome, Italy, July 2001.

[20] C. Fullmer and J. Garcia-Luna-Aceves. *Floor acquisition multiple access (FAMA) for packet-radio networks*. Proceedings of ACM SIGCOMM '95, 1995.

[21] J.J. Garcia-Luna-Aceves, Chane L. Fullmer. *Performance of Floor Acquisition Multiple Access in Ad-Hoc Networks*. Proc. Third IEEE Symposium on Computers and Communications (ISCC'98), Athens, Greece, June 30-July 2, 1998.

[22] F. Talucci, M. Gerla and L. Fratta. *MACABI (MACA By Invitation): A Receiver Oriented Access Protocol for Wireless Multiple Networks*. In PIMRC '97, Helsinki, Finland, September 1-4, 1997.

[23] Z. Tang and J. J. Garcia-Luna-Aceves. *Hop Reservation multiple access (HRMA) for multichannel packet radio networks*. Proceedings of the IEEE IC3N'98, Seventh International Conference on Computer Communications and Networks, October 1998.

[24] A. Tzamaloukas and J. J. Garcia-Luna-Aceves. *Channel-Hopping Multiple-Access*. In Proceedings IEEE International Communications Conference (ICC '00), New Orleans, Louisiana, June 2000.

[25] M. Joa-Ng and I. Lu. Spread Spectrum Medium Access Protocol with Collision Avoidance in Mobile Ad-Hoc Wireless networks. In Proc. IEEE INFOCOM 99, April 1999.

[26] I. Chlamtac et al. *ADAPT: A Dynamically Self-Adjusting Media Access Control Protocol for Ad Hoc Networks*. Proc. IEEE GLOBECOM'99, pp.11-5, Rio De Janeiro, Brazil, Dec. 1999.

[27] I. Chlamtac et al. *An Adaptive Medium Access Control (MAC) Protocol for Reliable Broadcast in Wireless Networks*. Proc. IEEE ICC'00, vol. 3, pp. 1692-6, New Orleans, LA, June 2000.

[28] Z. Tang and J.J. Garcia-Luna-Aceves. *A Protocol for Topology-Dependent Transmission Scheduling*. Proc. IEEE Wireless Communications and Networking Conference 1999 (WCNC 99), New Orleans, Louisiana, September 21-24 1999.

[29] C. Zhu and M. S. Corson. *A Five Phase Reservation Protocol (FPRP) for Mobile Ad Hoc Networks*. Proceedings IEEE INFOCOM, October 1998.

[30] Deng and Z. Haas. *Dual busy tone multiple access (DBTMA): A new medium access control for packet radio networks*. IEEE ICUPC, pp. 973–7, Oct. 1998.

[31] R. Rozovsky and P. R. Kumar. *SEEDEX: A MAC protocol for ad hoc networks*. Proceedings of The ACM Symposium on Mobile Ad Hoc Networking and Computing. MobiHoc 2001, Long Beach, Oct. 4-5, 2001.

[32] Asimakis Tzamaloukas. *Sender- and Receiver-Initiated MAC Protocols for Ad-Hoc Networks*. PhD Thesis, Computer Engineering, University of California, Santa Cruz, CA 95064, December 2000.

[33] Asis Nasipuri, Shengchun Ye, and Robert E. Hiromoto. *A MAC Protocol for Mobile Ad Hoc Networks Using Directional Antenna*. Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC 2000), Chicago, September 2000.

[34] Asis Nasipuri, Jothsna Mandava, Hanumantha Manchala, and Robert E. Hiromoto. *On-Demand Routing Using Directional Antennas in Mobile Ad Hoc Networks*. Proceedings of the IEEE International Conference on Computer Communication and Networks (ICCCN2000), October 2000, Las Vegas.

[35] Xu, S.; Saadawi, T. *Does the IEEE 802.11 MAC protocol work well in multihop wireless ad hoc networks?* IEEE Communications Magazine, Vol. 39 Issue 6, June 2001 pages 130 -137.

[36] http://www.uta.fi/EGEDL/outline/packetradiotechnology.html

[37] Robert E. Kahn, Steven A. Gronemeyer, Jerry Burchfiel, and Ronald C. Kunzelman. *Advances in Packet Radio Technology*. Proceedings of the IEEE, vol 66, no 11, November 1978.

[38]] John Jubin and Janet D. Tornow. *The DARPA Packet Radio Network Protocols*. Proceedings of the IEEE, vol 75, no 1, January 1987.

[39] http://www.ietf.org/html.charters/manet-charter.html

[40] M.S. Corson and J. Macker. *Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations*. Request For Comments 2501, Internet Engineering Task Force, January 1999.

[41] M. Navarro. *Optimal routing of multicast traffic in ad hoc networks*. Master's Thesis of Networking Laboratory. Helsinki University of University, 2001.

[42] The Mobile Ad Hoc Networks Study Group (GEDOC), http://www.siam.dcc.ufmg.br/gedoc

[43] Swetha Narayanaswamy, Vikas Kawadia, R. S. Sreenivas and P. R. Kumar. *The COMPOW protocol for power control in ad hoc networks: Theory, architecture, algorithm, implementation, and experimentation*. May 16, 2001.

[44] Javier Gmez, Andrew T. Campbell, Mahmoud Naghshineh and Chatschik Bisdikian. *PARO: Conserving Transmission Power in Wireless ad hoc Networks*. IEEE 9th International Conference on Network Protocols (ICNP'01), Riverside, California. November 2001.

[45] C.K. Toh. *Maximum Battery Life Routing to Support Ubiquitous Mobile Computing in Wireless Ad Hoc Networks*. IEEE Communications, June 2001.

[46] J. P. Monks, J.-P. Ebert, A. Wolisz, and W. W. Hwu. *A Study of the Energy Saving and Capacity Improvement Potential of Power Control in Multi-hop Wireless Networks*. Proc. of Workshop on Wireless Local Networks, Tampa, Florida, USA, November 2001.

[47] J. Lundberg. *Routing Security in Ad Hoc Networks*. Helsinki University of Technology, 2000.

[48] A.Vanhala. *Security in Ad-hoc Networks*. Helsinki University of Technology, 2000.

[49] K. Weniger und M. Zitterbart. *IPv6 Autoconfiguration in Large Scale Mobile Ad-Hoc Networks*. Proceedings of European Wireless 2002, Florence, Italy, Feb. 2002

[50] L. Joe and Philip M. Feldman. *Fundamental Research Policy for the Digital Battlefield*. RAND (http://www.rand.org), 1998.

[51] Humbad, Shailesh Niranjan. *Suitability of Distributed Mobile Wireless Networking for Urban Traffic Congestion Mitigation*. Master in City Planning, Department of Urban Studies and Planning, MIT, June 2001.

[52] Jinyang Li, Charles Blake, Douglas S. J. De Couto, Hu Imm Lee and Robert Morris. *Capacity of Ad Hoc Wireless Networks*. Proceedings of the 7th (ACM) International Conference on Mobile Computing and Networking, Rome, Italy. August 2001.

[53] P. Gupta and P.R. Kumar. *The capacity of wireless networks*. IEEE Transactions on Information Theory, 46 (2). March 2000.

[54] S. Toumpis and A. Goldsmith. *Ad Hoc Networks Capacity*. Signals, Systems and Computers, 2000. Conference Record of the 34th Asilomar Conference on, Vol. 2, pg. 1265 - 1269.

[55] P. Gupta, R. Gray, and P. R. Kumar. *An experimental scaling law for ad hoc networks*. Univ. of Illinois at Urbana-Champaign, May 2001.

[56] L. E. Miller. *Distribution of Link Distances in a Mobile Network*. Published in the proceedings of Journal of Research of the National Institute of Standards and Technology, March-April, 2001.

[57] Jinyang Li, John Jannotti, Douglas S. J. De Couto, David R. Karger, and Robert Morris. *A Scalable Location Service for Geographic Ad Hoc Routing*. Proc. ACM/IEEE MobiCom, pages 120-130, August 2000.

[58] Matthias Grossglauser and David Tse. *Mobility Increases the Capacity of Ad-hoc Wireless Networks*. In Proc. IEEE Infocom, April 2001.

[59] L. Tong, Q. Zhao and G. Mergen *Multipacket Reception in Random Access Wireless Networks: From Signal Processing to Optimal Medium Access Control*, to appear in IEEE Communications Magazine.

[60] G. Mergen and L. Tong. *Capacity of Regular Ad Hoc Netowrks with Multipacket Reception*. Allerton Conference Communications, Control, and Computing, 2001.

[61] G. Mergen and L. Tong. *Receiver Controlled Medium Access in Multihop Ad Hoc Networks with Multipacket Reception*. Proceedings of 2001 MILCOM, Washington, DC, Oct., 2001.

[62] Q. Zhao and L. Tong. *A Multi-Queue Service Room MAC Protocol for Wireless Networks with Multipacket Reception*. Submitted to IEEE/ACM Transactions on Networking, Feb., 2001.

[63] R. Knopp and P. A. Humblet. *Information Capacity and power Control in Single-Cell Multiuser Communications*. In Proc. Int. Conf. on Communications, Seattle, June 1995.

[64] Richard Frenkiel, B.R. Badrinath, Joan Borras and Roy Yates. *The Infostations challenge: Balancing cost and ubiquity in delivering wireless data*. Preprint, August 1999.

[65] M. Adler, and C. Scheideler.*Efficient communication strategies for ad-hoc wireless networks*. In Proc., 10th ACM Symp. Parallel Algorithms and Architectures, Puerto Vallerta, Mexico, 1998, pp. 259–268.

[66] P. Gupta and P. R. Kumar.*Critical Power for Asymptotic Connectivity in Wireless Networks*.In Stochastic Analysis, Control, Optimization and Applications, Eds. W.M.McEneany et al., Birkhauser, Boston, p. 547-566, 1998.

[67] P.Santi, D.M.Blough, F.Vainstein. *A Probabilistic Analysis for the Range Assignment Problem in Ad Hoc Networks*. In Proc. ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc), Long Beach, October 2001, pp. 212-220.

[68] R. Ramanathan and R. Rosales-Hain. *Topology control of multihop radio networks using transmit power adjustment*. In Proceedings of IEEE Conference on Computer Communications (INFOCOM), Mar. 2000, vol. 2, pp. 404–413.

[69] Miguel Sánchez ,Pietro Manzoni and Zygmunt J. Haas. *Determination of Critical transmission Range in Ad-Hoc Networks*. In Proccedings of Multiaccess Mobility and Teletraffic for Wireless Communications 1999 Workshop (MMT'99), October 6th-8th, Venice, Italy.

[70] J.B. Kruskal. *On the shortest spanning tree of a graph and the traveling salesman problem*. Proceedings of the American Mathematical Society, 7:48-50, 1956.

[71] R.C. Prim. *Shortest connection networks and some generalizations*. Bell Systems Technology Journal, 36:1389-1401, 1957.

[72] C. C. Cheng, et. al. *A Loop-Free Extended Bellman-Ford Routing Protocol Without Bouncing Effect*. ACM Sigcomm 1989, pp. 224-236.

[73] J.J. Garcia-Luna-Aceves. *A Fail-Safe Routing Algorithm for Multihop Packet-Radio Networks*. In INFOCOM. IEEE, Miami, Florida, April 1986.

[74] J. Hagouel. *Issues in Routing for Large and Dynamic Networks*. PhD thesis, Columbia University, May 1983.

[75] P. Humblet, *Another Adaptive Shortest-Path Algorithm*. IEEE Trans. Commun. (June 1991). Helsinki University of Technology. Networking Laboratory.

[76] B. Rajagopalan and M. Faiman. *A Responsive Distributed Shortest-Path Routing Algorithm with Autonomous Systems*. Internetworking: Research and Experience, 2:51–69, March 1991.

[77] S. Nesargi and R. Prakash. MANETconf: Configuration of Hosts in a Mobile Ad Hoc Network. In Proceedings of INFOCOM 2002.

[78] http://www.europa.eu.int/comm/energy__transport/en/gal__en.html

[79] Singh, S., Woo, M., and Raghavendra, C.S. *Power aware routing in mobile ad hoc networks*. In Proc. of MOBICOM, 1998, pp. 181–190.

[80] K. Obraczka and G. Tsudik. *Multicast routing issues in ad hoc networks*. In Proc. of IEEE ICUPC, Oct 1998.

[81] Sung-Ju Lee. *Routing and Multicasting Strategies in Wireless Mobile Ad hoc Networks*. University of California, Los Angeles, Computer Science Department, September 2000.

[82] Feeney L.M. *A Taxonomy for Routing Protocols in Mobile Ad Hoc Networks*. SICS Technical Report T99:07. Swedish Institute of Computer Science, Kista, Sweden, October 1999.

[83] C.E. Perkins and P. Bhagwat. *Highly Dynamic Destination-Sequenced Distance Vector Routing (DSDV) for Mobile Computers*. Proceedings of ACM SIGCOMM'94, London, UK, Sep. 1994, pp. 234-244.

[84] Charles E. Perkins, Elizabeth M. Royer, and Samir Das. *Ad Hoc On Demand Distance Vector (AODV) Routing*. IETF Internet draft, draft-ietf-manet-aodv-09.txt, November 2001 (work in progress).

[85] Johnson, D.B., and Maltz, D.A. *Dynamic Source Routing in Ad-Hoc Wireless Networks*. Mobile Computing, edited by T. Imielinski and H. Korth, chapter 5, pp. 153-181, Kluwer, 1996.

[86] Haas, Z.J. *A Routing Protocol for the Reconfigurable Wireless Networks*. IEEE ICUPC'97, San Diego, CA, October 12-16, 1997.

[87] M. Jiang, J. Li, and Y.C. Tay. *Cluster-Based Routing Protocol (CBRP)*. draft-ietf-manet-cbrpspec-01.txt, Internet Draft, IETF, Aug. 1999.

[88] P. Jacquet and T. Clausen. *Optimized Link State Routing Protocol*. IETF Internet draft, draft-ietf-manet-olsr-06.txt, September 2001 (work in progress)

[89] S. Murthy and J.J Garcia-Luna-Aceves. *An Efficient Routing Protocol for Wireless Networks*. ACM Mobile Networks and Applications Journal, Oct. 1996.

[90] T.-W.Chen and M.Gerla. *Global State Routing: A New Routing Scheme for Ad-hoc Wireless Networks*. In Proc. of IEEE ICC'98, Atlanta, GA, pages 171-175, Jun. 1998.

[91] G. Pei, M. Gerla and T.-W. Chen. *Fisheye State Routing in Mobile Ad Hoc Networks*. In Proc. of the 2000 ICDCS Workshops, Taipei, Taiwan, Apr. 2000, pp. D71-D78.

[92] B. A. Iwata, C.-C. Chiang, G. Pei, M. Gerla, and T.-W. Chen. *Scalable Routing Strategies for Ad Hoc Wireless Networks*. IEEE Journal on Selected Areas in Communications (JSAC). Issue on Wireless Ad Hoc Networks, pages 1369 -1379, Vol. 17, No. 8, August 1999.

[93] M. Joa-Ng and I.-T. Lu, A Peer-to-Peer zone-based two-level link state routing for mobile Ad Hoc Networks, IEEE Journal on Selected Areas in Communications, Vol. 17(8), Aug. 1999, pp. 1415-1425.

[94] Ching-Chuan Chiang, Hsiao-Kuang Wu, Winston Liu, and M. Gerla. *Routing in Clustered Multihop Mobile Wireless Networks with Fading Channel*. In Proc. of IEEE Singapore International Conference on Networks (SICON'97).

[95] M.S. Corson and A. Ephremides. *A Distributed Routing Algorithm for Mobile Wireless Networks*. ACM/Baltzer Wireless Networks, vol.1, no.1, pp.61-81, February 1995.

[96] V. Park and S. Corson. *Temporally-Ordered Routing Algorithms (TORA) Version 1 Functional Specification*. IETF Internet Draft, draft-ietf-manet-tora-spec-01.txt, Aug. 1998 (work in progress).

[97] C.K. Toh. *Associativity-Based Routing for Ad-Hoc Mobile Networks*. Wireless Personal Communications, Vol. 4, No. 2, pp. 1-36, Mar. 1997.

[98] R. Dube, C.D. Rais, K.-Y. Wang, S.K. Tripathi. *Signal stability-based adaptive routing (SSA) for ad hoc mobile networks*. IEEE Personal Communications, vol.4, no.1, pp.36-45, February 1997.

[99] T. Larsson and N. Hedman. *Routing Protocols in Wireless Ad-Hoc Networks – A Simulation Study*. Unpublished master's thesis. Technical report, Lulea University, 1998.

[100] J. Broch, D.A. Maltz, D.B. Johnson, Y.C. Hu, J. Jetcheva .*A performance comparison of multi-hop wireless ad hoc network routing protocols*. In Proc. MOBICOM, 1998, 85-97.

[101] W. Su, S-J. Lee and M. Gerla. *Mobility Prediction and Routing in Ad Hoc Wireless Networks*. International Journal of Network Management, Wiley and Sons, 2000.

[102] M. M. Zonoozi and P. Dassanayake. *User mobility modeling and characterization of mobility patterns*. IEEE J. Select. Areas Commun., vol. 15, pp. 1239-1252, Sept. 1997.

[103] B. Liang and Z. Haas. *Predictive distance-based mobility management for PCS networks*. In Proc. of IEEE INFOCOM'99 (1999).

[104] http://www.disca.upv.es/misan/mobmodel.htm

[105] T. Camp, J. Boleng, and V. Davies. *Mobility Models for Ad Hoc Network Simulations*. Submitted 2001.

[106] C.-C. Chiang. *Wireless Network Multicasting*. PhD dissertation, University of California, Los Angeles, Department of Computer Science, May 1998.

[107] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang. *A Group Mobility Model for Ad Hoc Wireless Networks*. In ACM/IEEE MSWiM, pages 53–60, 1999.

[108] Elizabeth M. Royer, P. Michael Melliar-Smith, and Louise E. Moser. *An Analysis of the Optimum Node Density for Ad hoc Mobile Networks*. In Proceedings of the IEEE International Conference on Communications, Helsinki, Finland, June 2001.

[109] S. Basagni, I. Chlamtac, V.R. Syrotiuk, and B.A. Woodward. *A distance routing effect algorithm for mobility (DREAM)* In Proc. MobiCom '98, Dallas, TX, pp. 76-84, October 25-30, 1998.

[110] Young-Bae Ko and Nitin H. Vaidya. *Location-Aided Routing (LAR) in Mobile Ad Hoc Networks*. In ACM/Baltzer Wireless Networks (WINET) journal, Vol.6-4, 2000 - Extended version of the Mobicom'98 paper.

[111] S. Basagni, D. Turgut and S.K. Das. *Mobility-Adaptive Protocols for Managing Large Ad hoc Networks*. In Proceedings of IEEE International Conference on Communications (ICC), Helsinki, Finland, June 11-14, 2001, pp.1539-1543.

[112] T. Imielinski, S. Viswanathan, and B. R. Badrinath. *Energy Efficient Indexing on Air*. In Proceedings of the ACM Conference on Management of Data (SIGMOD) , pages 25–36, 1994.

[113] I. Chlamtac, C. Petrioli, J. Redi. *Analysis of Energy-conserving Access Protocols for Wireless Identification Networks*. In Proceedings of the 5th International Conference on Telecommunication Systems (ITC-97), March 20 - 23, 1997.

[114] M. Zorzi, R.R. Rao. *Error Control and Energy Consumption in Wireless Communications Channels*. In Proc. 34th Annual Allerton Conference,Oct. 1996. Also submitted to IEEE Trans. on Computers.

[115] C. Fragouli P. Lettieri and M. Srivastava. *Low power error control for wireless links*. In Proc. Third ACM/IEEE MobiCom Conference, pages 139–150, Budapest,Hungary, Sept. 1997.

[116] A. P. Chandrakasan, S. Sheng and R. W. Brodersen. *Low-Power CMOS Digital Design*. IEEE Journal of Solid-State Circuits, Vol.27, No.4, pp. 473, 1992. And Norwell, MA: Kluwer Academic, 1995.

[117] C. E. Jones, K. M. Sivalingam, P. Agrawal, and J. C. Chen. *A survey of energy efficient network protocols for wireless networks*. Wireless Networks, 7(4):343–358, July 2001.

[118] Jeffrey P. Monks. *Transmission Power Control for Enhancing The Performance of Wireless Packet Data Networks*. PhD Thesis, Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL, March, 2001.

[119] P. Sinha. *Routing and Transport Layer Protocols for Wireless Networks*. Ph.D. Thesis, Computer Science, University of Illinois. May 2001.

[120] Agrawal P., Chen J-C, Kishore S., Ramanathan P., Sivalingam K. *Battery power sensitive video processing in wireless networks*. In Proceedings IEEE PIMRC'98, Boston, September 1998.

[121] B. Narendran, J. Sienicki, S. Yajnik, and P. Agrawal. *Evaluation of an adaptive power and error control algorithm for wireless systems*. IEEE ICC, vol. 1, pp. 349–355, June 1997.

[122] K. Naik, D. S. L. Wei. *Software implementation strategies for power-conscious systems*. ACM Journal on Mobile Networks and Applications 6 (2001), 291-305.

[123] S.-L. Wu, Y.-C. Tseng, and J.-P. Sheu. *Intelligent Medium Access for Mobile Ad Hoc Networks with Busy Tones and Power Control*. IEEE Journal on Selected Areas in Communications, 18(9):1647-1657, Sep. 2000.

[124] Jeffery P. Monks, Vaduvur Bhargavan, and WenMei W.Hwu. *A power controlled multiple access protocol for wireless packet networks*. In Proceedings of INFOCOM, 2001, pp. 219–228.

[125] Suresh Singh and C.S. Raghavendra. *Power efficient MAC protocol for multihop radio networks*. In The Ninth IEEE International Symposium

on Personal, Indoor and Mobile Radio Communications, 1998, pp. 153–157.

[126] S. Singh, C. S. Raghavendra, and J. Stepanek. *Power-Aware Broadcasting in Mobile Ad Hoc Networks.* In Proceedings of IEEE PIMRC'99, Osaka, Japan, Sep. 1999.

[127] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides. *On the Construction of energy-Efficient Broadcast and Multicast Trees in Wireless Networks.* IEEE Infocom'2000, Tel-Aviv, Israel.

[128] J. E. Wieselthier, G. D. Nguyen, A. Ephremides. *Algorithms for energy-efficient multicasting in static ad hoc wireless networks.* ACM Journal on Mobile Networks and Applications 6 (2001), 251-263.

[129] Rajesh Bhairampally. *Energy Complexity: A Metric for Energy Consumption of Ad Hoc Network Protocols.* Submitted for Mobihoc 2002.

[130] Mark Stemm and Randy H Katz. *Measuring and reducing energy consumption of network interfaces in hand-held devices.* IEICE Transactions on Communications, vol. E80-B, no. 8, pp. 1125–1131, Aug. 1997.

[131] Ya Xu, John Heidemann, and Deborah Estrin. *Geography-informed Energy Conservation for Ad Hoc Routing.* In Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking, pp. 70-84. Rome, Italy, USC/Information Sciences Institute, ACM. July, 2001.

[132] B. Chen, K. Jamieson, H. Balakrishnan, R. Morris. *Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks.* In Proc. 6th Annual ACM/IEEE MobiCom Conference, 2001.

[133] Ya Xu, John Heidemann, and Deborah Estrin. *Adaptive Energy-Conserving Routing for Multihop Ad Hoc Networks.* Research Report527, USC/Information Sciences Institute, October, 2000.

[134] K. Woo, C. Yu, H. Y. Youn, and B. Lee. *Non-Blocking, Localized Routing Algorithm for Balanced Energy Consumption in Ad Hoc Networks of Mobile Hosts.* International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication System (MASCOTS), 2001, pp. 117-124.

[135] Q. Li and J. Aslam and D. Rus. *Online power-aware routing in wireless ad-hoc networks.* In Proc. 6th Annual ACM/IEEE MobiCom Conference, 2001.

[136] http://www.atmforum.org

[137] R. Braden, D. Clark, and S. Shenker. *Integrated services in the internet architecture: an overview*. Request For Comments 1633, Internet Engineering Task Force, June 1994.

[138] Braden B., Zhang L., Berson S., Herzog S., and Jamin S. *Resource Reservation Protocol (RSVP) - Version 1 Functional Specification*. Request For Comments 2205, Internet Engineering Task Force, September 1997.

[139] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. *An Architecture for Differentiated Services*. Request For Comments RFC 2475, DARPA, December 1998.

[140] http://www.mplsforum.org

[141] Jens Burkhard Schmitt. *Heterogeneous Network Quality of Service Systems*. Kluwer Academic Publishers. June 2001

[142] Chen, S. *Routing Support for Providing Guaranteed End-To-End Quality of Service*. PhD Thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 1999.

[143] K. Wu and J. Harms. *QoS in Mobile Ad Hoc Networks*. Crossing Boundaries– the GSA Journal of University of Alberta, Vol. 1, No. 1, Nov. 2001, pp.92- 106.

[144] Indu Mahadevan and K. M. Sivalingam. *Architecture and Experimental Framework for Supporting QoS in Wireless Networks Using Differentiated Services*. In ACM/Baltzer Mobile Networks and Applications Journal, vol. 6, No. 4, pp. 385-395, 2001.

[145] H. Xiao , W.K.G. Seah, A. Lo, and K. C. Chua. *A Flexible Quality of Service Model for Mobile Ad-Hoc Networks*. In the proceedings of IEEE Vehicular Technology Conference, 15-18 May 2000, Tokyo, Japan, pp445-449.

[146] Cansever, D., A. Michelson, and A. Levesque. *Quality of Service Support in Mobile ad-hoc IP Networks*. Proceedings of MILCOM 1999, October 1999.

[147] Mirhakkak, M., Schult, N., Thomson, D. *Dynamic quality-of-service for mobile ad hoc networks*. Mobile and Ad Hoc Networking and Computing, 2000. MobiHOC. 2000 First Annual Workshop on , 2000 Page(s): 137 -138

[148] S. Lee and A. Campbell. *INSIGNIA: In-band Signaling Support for QOS in Mobile Ad Hoc Networks*. Proc of 5th International Workshop on Mobile Multimedia Communications (MoMuC,98), Berlin, Germany, October 1998.

[149] Z. Wang and J. Crowcroft. *QoS routing for supporting resource reservation.* IEEE J. on Select. Areas in Comm., Vol.14(7), pp.1228-1234, September 1996.

[150] Chakrabarti, S. and Mishra. A. *QoS issues in ad hoc wireless networks.* IEEE Communications Magazine , Volume: 39 Issue: 2 , February 2001 Page(s): 142 -148

[151] P. Sinha, R. Sivakumar, and V. Bharghavan. *Cedar: Core extraction distributed ad hoc routing.* In Proc. of IEEE INFOCOM, 1999.

[152] Shigang Chen, Klara Nahrstedt. *Distributed Quality-of-Service Routing in Ad-Hoc Networks.* IEEE Journal on Special Areas in Communications, Vol. 17, No. 8, August 1999, pp. 1-18.

[153] W.-H. Liao, Y.-C. Tseng, S.-L. Wang, and J.-P. Sheu. *A Multi-Path QoS Routing Protocol in a Wireless Mobile Ad Hoc Network.* Telecommunication Systems (special issue of best papers from ICN2001).

[154] Giammarco, C.; Malick, K.; Morreale, P. *Wireless quality of service assurance for network survivability.* In Military Communications Conference Proceedings, 1999.

[155] A. Muir and J.J. Garcia-Luna-Aceves. *An Efficient Packet Sensing MAC Protocol for Wireless Networks.* To appear in The Journal of Mobile Networks and Applications

[156] J. Sobrinho, A.S. Krishnakumar. *Quality of service in ad hoc carrier sense multiple access wireless networks.* IEEE Journal on Selected Areas in Communications, vol. 17, no. 6, pp. 13531368, Aug. 1999.

[157] C. R. Lin and M. Gerla. *MACA/PR: An Asynchronous Multimedia Multihop Wireless Network.* Proceedings of IEEE INFOCOM '97, March 1997.

[158] L. E. Miller. *Probability of a Two-Hop Connection in a Random Mobile Network.* 35th Conf. on Info. Sci. and Syst. (CISS '01), Baltimore, 21-23 March 2001.

[159] L. E. Miller. *Multihop Connectivity of Arbitrary Networks.* March 2001.

Appendix A

Algorithms

A.1 Algorithms to compute the MST

A.1.1 Kruskal

Kruskal algorithm starts with a list of links among the nodes of the graph, sorted by cost. A link is appended to the tree under construction (Minimum Spanning Tree or MST) —the links are taken in ascendent cost order— only if:

- It adds a new node to the current tree.
- It does not create a loop in the tree under construction.

The algorithm ends when all the nodes are included in the MST.

input : the set E of the sides of tree of n nodes and the cost $c[e_{ij}]$ of the links, with $i, j \in \{1, \ldots, n\}$

output : the set $E(1)$ of sides of the MST

$E(1) \leftarrow 0$
$E(2) \leftarrow E$ $\{E(2)$ is the set of the remaining sides$\}$
while *($E(1)$ contains less then $n-1$ sides) AND ($E(2) \neq \emptyset$)* **do**
  Choose the link e_{ij} with minimum cost within $E(2)$
  $E(2) = E(2) - e_{ij}$
  if X_i, X_j *does not belong to the same tree* **then**
    Unite the trees X_i and X_j belongs to one tree $E(1)$
  endif
endw

Figure A.1: Kruskal algorithm for the calculation of the MST

A.1.2 Prim

The construction of the MST tree by using the Prim algorithm starts by choosing an random node (for simplicity the first one), from which the tree will be built. To this initial node new links are added, once at a time. The order followed to include new links to the tree has two criteria:

1. The link selected connects some of the nodes of the tree under construction with a node that it is not in this tree.

2. Among those links that fulfill the first criteria, the one with minimum cost is selected.

The algorithm ends when all the nodes are included in the MST.

input : n nodes and the cost $c[e_{ij}]$ of the links, with $i, j \in \{1, \ldots, n\}$

output : p_j, with $j \in \{2, \ldots, n\}$, pointer of node X_j to its father in the MST

$O \leftarrow \{1\}$ $\{X_1$ root of the T tree$\}$
$P \leftarrow \{2, \ldots, n\}$
for $j \in P$ **do**
  $e_j \leftarrow c[e_{j1}]$
  $p_j \leftarrow 1$
endfor

$e_j = c[e_{j1}]$ $\{e_j$ represents the minimum cost of a link from X_j to a node in $O\}$
$p_j = 1$ {all peaks initially connected to the root}

while $P \neq \emptyset$ **do**
  Choose node $X_k | e_k \leq e_j \forall j \in P$. If tight choose the minimum k
  $O \leftarrow O \cup \{k\}$
  $P \leftarrow P - \{k\}$
  for $j \in P$ **do**
    if $e_j > c[e_{kj}]$ **then**
      $e_j \leftarrow c[e_{kj}]$
      $p_j \leftarrow k$
    endif
  endfor
endw

Figure A.2: Prim algorithm for the calculation of the MST