Università degli Studi di Roma "La Sapienza"

Dottorato di Ricerca in Ingegneria dell'Informazione e della Comunicazione

XVII Ciclo – 2004

Location-aware, power-efficient MAC and routing strategies for UWB wireless communications

Luca De Nardis

Università degli Studi di Roma "La Sapienza"

Dottorato di Ricerca in Ingegneria dell'Informazione e della Comunicazione

XVII Ciclo - 2004

Luca De Nardis

Location-aware, power-efficient MAC and routing strategies for UWB wireless communications

Advisor

Professor Maria-Gabriella Di Benedetto

Author's address:
Luca De Nardis
Dipartimento di Scienza e Tecnica dell'Informazione e della Comunicazione
Università degli Studi di Roma "La Sapienza"
Via Eudossiana, 18 I-00184 Roma, Italy
e-mail: lucadn@newyork.ing.uniroma1.it

To Anna

Acknowledgment

I wish to thank my advisor, professor Maria-Gabriella Di Benedetto, for her fundamental support and guidance during this three-years-long journey.

Contents

Chapter 1

Introduction

In the last few years the concept of ad-hoc networking gathered an increasing attention, thanks to its promise of flexible, self-organizing networks, allowing for the definition of new network scenarios and applications precluded to traditional wireless networks, which rely on a fixed infrastructure in order to operate [1].
Effective deployment of ad hoc networks requires however to address a whole new set of design issues, involving almost all aspects of a communication system, ranging from hardware issues up to application design. Among the new requirements posed by ad hoc networking an extremely challenging one is the introduction of energy-aware solutions, capable of extending the life of a network under conditions of a limited power supply available in each terminal.
Energy-awareness design assumes a overwhelming relevance for a peculiar class of ad-hoc networks, in which terminals are able to detect specific external events in the area of deployment of the network, typically by monitoring the variations in the value of one or more physical parameters (e.g. humidity or temperature). Monitoring of such parameters requires from network terminals the capability of sensing the environment: for this reason, such specific ad-hoc networks are referred to as sensor networks.
Typical application scenarios of sensor networks make an efficient management of available energy a key requirement, given the typical size of such networks (hundreds of terminals), that makes battery replacement in all network terminals almost impossible.

In this framework, Ultra Wide Band (UWB) radio emerged as an appealing solution for the deployment of ad-hoc and sensor networks [2].
Several features of UWB make this technology particularly attractive for indoor and outdoor wireless networks:

- the extremely high bit rates that UWB is expected to provide, in the

order of hundreds of Mb/s or even Gb/s.

- the high temporal resolution inherent to UWB, that provides high robustness in presence of multipath, thus allowing communications even in presence of several obstacles and in conditions of Non-Line-Of-Sight (NLOS) propagation.

- the accurate ranging capability, also provided by the high temporal resolution of UWB signals, offering distance information that can be used to derive information on physical position of terminals in the network.

The above features derive from the key characteristic of UWB signals, i.e. the use of a bandwidth that spans over several GHz in the range of frequencies going from 0 to 10 GHz. As a consequence of such large bandwidth occupation, UWB emissions cover a large portion of the frequency spectrum, and must in principle coexist with other Hertzian waveforms propagating over the air interface. The principle of coexistence imposes thus upper bounds on UWB power emission so as to limit interference on existing narrowband services.
In April 2002 by releasing UWB radio emission masks the Federal Communications Commission (FCC) in the USA [3] opened the way for the concept of coexistence with traditional and protected radio services. The emission masks issued by the FCC regarding indoor UWB systems strongly limit however operation to a bandwidth lying between 3.1 and 10.6 GHz, and set very stringent limits on out of band emission. Such masks are presented in fig. 1.1 and 1.2 for indoor and outdoor UWB devices respectively. The FCC emission masks serve at present as a reference for UWB system design within and outside the USA. As far as Europe is concerned, for example, the European Radio Organization (ERO) issued in July 2003 a tentative definition of UWB emission masks which very closely followed the FCC settings [4] [1].

The severe power emission constraints set by FCC lead to the conclusion that, although UWB has the potential of allowing simultaneous communication of a large number of users at high bit rates [2, 6, 7], the available power levels will allow for application of UWB to two specific cases, corresponding to transmissions at either high bit rates over short ranges or low bit rates over medium-to-long ranges.
The high bit rate/short range case includes Wireless Personal Area Networks (WPANs) for multimedia traffic, cable replacement applications (such as wireless USB), and wearable devices (e.g. wireless Hi-Fi headphones). In particular, UWB was selected as the enabling technology in the IEEE 802.15.3a

[1] In November 2004 the Electronic Communications Committee (ECC) of the European Conference of Postal and Telecommunications Administrations (CEPT) released a new draft of UWB emission masks ([5]), with emission levels significantly lower than FCC levels.

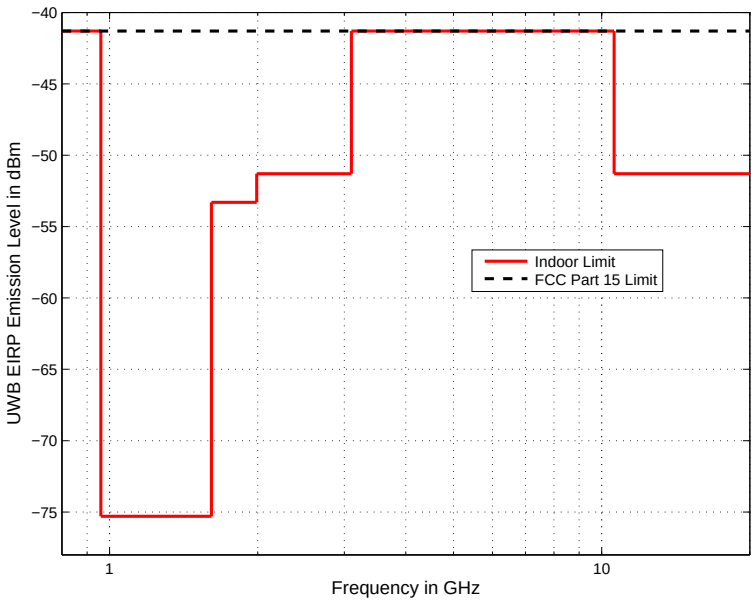

Figure 1.1: FCC emission masks for UWB indoor devices.

WPAN standard, aiming at providing very high bit rates (above 500 Mb/s) on a 10 meters range. The selected UWB implementation, based on Direct Sequence Impulse Radio, can in fact provide a bit rate of up to 640 Mb/s [8]. It should be noted however that, although UWB is adopted as the physical layer of the IEEE 802.15.3a standard, the specific features of UWB are not exploited at the MAC layer, since the MAC adopted in 802.15.3a was inherited without modifications from the parent IEEE 802.15.3 standard, originally defined for a narrow-band physical layer [9], and based on a centralized network organization inspired to the solution adopted in Bluetooth [10].
An enhanced MAC protocol, based on the original 802.15.3 MAC, was developed within the framework of the U.C.A.N. project [11]. The U.C.A.N. MAC introduces primitives for retrieving distance information between devices in a piconet and allows for relaying within a piconet, the choice of the relaying device being based on the distance between devices, with the final goal of extending coverage within the piconet while minimizing the emitted power [12].

The low bit rate/medium-to-long range case applies to long-range sensor networks (e.g. indoor/outdoor distributed surveillance systems), non-real-time

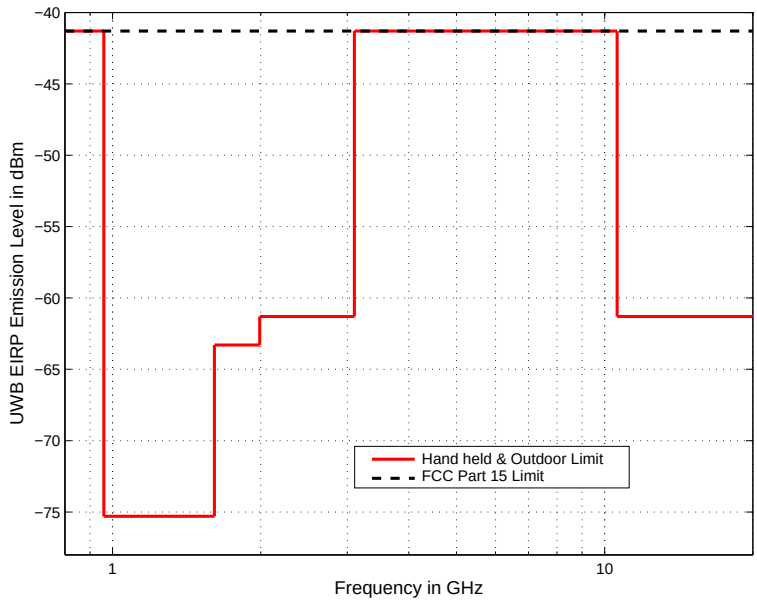

Figure 1.2: FCC emission masks for UWB outdoor and hand held devices.

data applications (e.g. e-mail and instant messaging), and in general all data transfers compatible with a transmission rate in the order of 1 Mb/s over several tens of meters.
In particular, UWB is considered as a candidate solution in the deployment of sensor networks in challenging environments, such as mixed indoor/outdoor scenarios and high multipath scenarios. The interest toward UWB in these scenarios is motivated by both the high robustness to multipath and the accurate ranging capability.
Although robustness to multipath is a valuable feature, the key aspect that makes UWB an interesting candidate for sensor networks deployment is its ranging capability. Accurate ranging is in fact the first step for retrieving position information within a sensor network. Such information is regarded as fundamental in several application scenarios for sensor networks, in particular all those scenarios where the usefulness of the information provided by a sensor is tightly related to the spatial position where such information is gathered. This is the case for example of fire monitoring in forests, or information on the content of packages in a large storehouse.
Based on the above considerations, the IEEE created a new Working Group

aiming at the definition of a new standard, named 802.15.4a, dealing with indoor/outdoor sensor networks with accurate positioning capability. Although the choice on the physical layer for this new standard is still open, UWB is a strong candidate for the final definition of the 802.15.4a standard. It should be noted that, as much as in 802.15.3a, the new standard will be derived from the already existing IEEE 802.15.4 standard, that forms the basis for the industrial standard ZigBee [13]. The new standard will initially adopt the MAC defined for 802.15.4, but it is expected that a new definition of the MAC will be eventually required, in order to efficiently exploit the ranging information provided by the new 802.15.4a physical layer.
Research activity in the field of low data rate UWB networks is also carried out in Europe, in particular within the Integrated Project P.U.L.S.E.R.S. [14], dealing with the design of advanced MAC and networking strategies for both low data rate and high data rate UWB networks.

Within the above framework, the present work focuses on the design of a new MAC protocol specific for low data rate UWB systems, with the goal of taking into account UWB specific features in MAC design and enabling new strategies for resource management based on the ranging information provided by UWB. To this aim, Chapter 2 introduces the key concepts on UWB, and analyzes specific aspects of this transmission technique, such as Multi User Interference (MUI) and pulse shaping. Chapter 3 presents the key issues in MAC design for UWB systems, and proposes a MAC protocol specifically tailored for low data rate UWB systems, capable of exploiting UWB multiple access and ranging capabilities. The possibility of using such ranging information for obtaining position information in the network is analyzed in Chapter 4, where the problem of positioning in wireless networks is addressed. Chapter 5 investigates how the distance and position information provided by the proposed MAC protocol can be exploited for designing innovative routing strategies for multi-hop UWB networks in order to address the energy-awareness issue posed by low data rate and sensor networking. The performance of the proposed MAC and routing strategies are finally evaluated in specific test cases in Chapter 6, while Chapter 7 draws conclusions.

Chapter 2

Ultra Wide Band Radio

2.1 Ultra Wide Band radio definition and characteristics

The most common adoption of the term "Ultra Wide Band" (UWB) comes from the UWB radar world and refers to waveforms that are characterized by an instantaneous fractional energy bandwidth greater than about 0.20-0.25. The energy bandwidth of a waveform can be defined as follows [2].
Let E be the instantaneous energy of the waveform; the energy bandwidth is then identified by the frequencies f_L and f_H, which delimit the interval where most of E (for example over 90%) falls. We call the width of the interval [f_L, f_H], i.e. $f_H - f_L$ the energy bandwidth.
Given the lower limit f_L and the higher limit f_H of the Energy Spectral Density (ESD), the center frequency of the spectrum is given by $(f_H + f_L)/2$. The fractional bandwidth is defined as the ratio between the energy bandwidth and the center frequency and is expressed as:

$$\text{fractionalbandwidth} = \frac{(f_H - f_L)}{\left(\frac{f_H + f_L}{2}\right)} \tag{2.1}$$

If the fractional bandwidth is greater than 0.20-0.25, we define the signal as UWB. A signal with an energy bandwidth of 2 MHz, for example, can be considered UWB only if the center frequency of its spectrum is lower than 10 MHz. Note that the definition of an UWB signal can be given only taking into account the center frequency of the signal itself.
In the UWB emission masks released by FCC in U.S.A. in 2002 ([3]), f_L and f_H are defined in less general terms as the lower and upper frequencies of the -10 dB emission points. Always in [3], it is stated that a signal can be assumed to be UWB if its bandwidth at -10 dB emission points exceeds 500 MHz, independently of the fractional bandwidth value.

The 500 MHz minimum bandwidth is used by the FCC in order to define UWB signals in combination with a threshold at 2.5 GHz. Signals below the threshold are in fact considered UWB if their fractional bandwidth exceeds 0.20, while for signals above the threshold the condition for being considered as UWB is to have a bandwidth exceeding 500 MHz.

In the radar field, UWB emission is obtained by radiating waveforms formed by sequences of very short pulses, considering thus a duration of the pulse that is typically about a few hundred picoseconds. In these systems, the information to be transferred is represented in digital form by a binary sequence, and the information of each bit (0 or 1) is transferred using one or more pulses by means of repetition codes, where the use of several pulses for a single bit aims at increasing robustness in the transmission of each bit. The pulsed transmission principle can be applied in a straightforward manner to communications, and forms the basis for the most common UWB transmitter scheme in communication systems.

The most common and traditional way of emitting an UWB signal for communications is in fact by radiating pulses that are very short in time. This transmission technique goes under the name of Impulse Radio (IR). The information data symbols can modulate the pulses in different ways; Pulse Position Modulation (PPM) and Pulse Amplitude Modulation (PAM) have been proposed as modulation schemes ([15], [16]). In addition to modulation and in order to shape the spectrum of the generated signal, pseudorandom or pseudonoise (PN) codes are used to encode the data symbols. In the most common approach, encoded data symbols introduce a time dither on generated pulses leading to the so-called Time-Hopping UWB (TH-UWB). Another scheme often proposed is Direct-Sequence Spread Spectrum (DS-SS) where encoded data symbols modulate the amplitude of basic pulses; this technique is referred to as Direct-Sequence UWB (DS-UWB) in the case of IR UWB signals ([17], [18], [19], [8]).

The UWB definition released by the FCC [3] does not limit, however, the generation of UWB signals to IR and opens the way, at least in the United States, for alternative (nonimpulsive) schemes. An ultra wide bandwidth, say 500 MHz, might be produced by a very high data rate, independently of the characteristics of the pulses. The pulses might, for example, satisfy the Nyquist criterion at an operating pulse rate 1/T, which would require a minimum bandwidth of $B = 1/(2T)$ and thus be limited in frequency, but unlimited in time having the classical raised-cosine infinity-bouncing shape with nulls at multiples of 1/T. Systems with an ultra wide bandwidth of emission due to high-speed data rate rather than pulse width, provided that the fractional bandwidth or minimum bandwidth requirements are verified at all times of the transmission, are not precluded. Methods such as Orthogonal Frequency Division Multiplexing (OFDM) and Multi-Carrier Code Division Multiple Ac-

cess (MC-CDMA) are capable of generating UWB signals at appropriate data rates. Recent proposals in the United States, and in particular in the IEEE 802.15.TG3a Working Group, refer to a multi-band (MB) alternative to DS-UWB [8] in which the overall available bandwidth is divided into sub-bands of at least 500 MHz ([20], [21], [22]).
Since the focus in this work is not on UWB signals generation, we will consider in the following the most common definition of the UWB signal, i.e. the Impulse Radio UWB signal combined with Time Hopping and PPM. The transmitted signal s(t) can be written as follows ([2]):

$$s(t) = \sum_{j=-\infty}^{+\infty} p\left(t - jT_s - c_jT_c - a_j\varepsilon\right) \tag{2.2}$$

The signal presented in eq. (2.2) can be generated starting from the binary sequence to be transferred by using a repetition coder (i.e. a special case of channel coder) that associates to each binary digit N_s bits with the same value, generated at a rate of $1/T_s$ bits/s. We will indicate the j-th coded bit as a_j. The bit interval, or bit duration, T_b is then given by the time required to transmit the N_s coded bits: $T_b = N_sT_s$.
Next, an integer-valued code $\mathbf{c} = (\ldots, c_0, c_1, \ldots, c_j, c_{j+1}, \ldots)$ is applied to the coded bits. Usually, $\mathbf{c}$ is assumed to be a pseudorandom code, and each element c_j of the code satisfies the condition $0 \leq c_j \leq N_h - 1$. The generic element of $\mathbf{c}$ is used to introduce a random delay in the transmission of each pulse, and for this reason $\mathbf{c}$ is referred to as Time Hopping (TH) code. The delay introduced by the j-th value of the code is given by c_jT_c where T_c is the so-called chip time.
The sequence enters a PPM modulator, which generates a sequence of Dirac pulses $\delta(t)$ at a rate of $1/T_s$ pulses/s. These pulses are shifted in time from nominal positions jT_s by $c_jT_c + a_j\epsilon$ where ϵ verifies the condition $c_jT_c + \epsilon < T_s$ for all c_j. The shift introduced by the PPM modulator, $a_j\epsilon$, depends on the value of the coded bit a_j.
Finally, the sequence of delayed pulses enters in a pulse shaper filter with impulse response p(t) such that the signal at the output of the pulse shaper filter is a sequence of strictly non-overlapping pulses. Note that the choice of the pulse is crucial in an IR system, since the Energy Spectral Density (ESD) of the pulse determines the envelope of the Power Spectral Density (PSD) of the corresponding UWB signal; the problem of selecting good pulse shapes will be analyzed in section 2.2.

When N_u transmitters are active at the same time, the received signal at the reference receiver will be given by the combination of the corresponding

transmitted signals:

$$s_{rec}(t) = \sum_{k=1}^{N_u} \sum_{j=-\infty}^{\infty} p^{(k)} \left(t - jT_s - c_j^{(k)} T_c - a_j^{(k)} \varepsilon - \delta^{(k)} \right) + n(t) \tag{2.3}$$

In eq. (2.3) above, the terms $p^{(k)}$ and $\delta^{(k)}$ take into account the effect of the propagation channel on each signal and the asynchronism between the different users. The problem of modeling the effect of multiple interferring users on the reference receiver will be addressed in the next subsection.

2.1.1 Multi User Interference Modeling

In a TH-UWB system, the TH principle allows simultaneous access to the network by different users. Pulses belonging to different transmissions will eventually collide. The problem of modeling Multi User Interference (MUI) in UWB systems is still open, since the traditional approach, based on the assumption that the interference noise can be modeled as Gaussian is far from realistic in the case of UWB systems, in particular when low bit rates are considered [23]. In the following, we will adopt the Pulse Collision interference model introduced in [24], that will be used in testing the MAC protocol proposed in Chapter 3. This model proved in fact to be more close to simulation results for low rate scenarios, as shown in [24] and [2].
The time occupied by one single UWB pulse indicated by T_M is defined here as the time interval typically centered on the main lobe in which most of the energy of the pulse at the receiver is concentrated. Typical values for T_M lie between 70 psecs and 20 nsecs depending upon transmitted pulse shape and channel behavior. Under the simplifying hypothesis for asynchronous networks that the pulse inter-arrival process follows a Poisson distribution, the probability that one or more pulses collide with the useful pulse when P_U packets are transmitted over the air interface by N_U active users can be expressed as:

$$Prob_{PulseCollision} = 1 - e^{\left(-2 \cdot (P_U - 1) \cdot \frac{T_M}{T_F}\right)} \tag{2.4}$$

P_U depends upon N_U, packet length L, data rate R, and packet generation rate G. Its average instantaneous value is:

$$P_U = \frac{N_U \cdot L \cdot G}{R} \tag{2.5}$$

Since a detailed evaluation of the effect of a collision would require a detailed model of the receiver that will not be introduced here, we assume that a pulse collision, independently from the percentage of overlapping between

colliding pulses, causes a random decision at the receiver, so that the pulse error probability can be expressed as:

$$Prob_{PulseError} = 0.5 \cdot Prob_{PulseCollision} \tag{2.6}$$

Considering that each bit is encoded into N_S pulses we assume an error on the bit when more than $N_S/2$ pulse errors occur. This corresponds to assuming a hard receiver detection. Bit error probability is thus expressed by:

$$Prob_{BitError} = \sum_{i=\left\lceil \frac{N_S}{2} \right\rceil}^{N_S} \left(\begin{array}{c} N_S \\ i \end{array} \right) \cdot Prob_{PulseError}^{i} \cdot (1 - Prob_{PulseError})^{N_S - i} \tag{2.7}$$

As already mentioned above, this MUI model will be used in the analysis of MAC performance in Chapter 3.

2.2 Pulse Shaping

As mentioned in the previous section, the choice of the impulse response of the pulse shaper filter is crucial since it affects the PSD of the transmitted signal [6]. In this section we will propose an approach for the selection of pulse shape aiming at the maximization of the transmitted power, given the emission limits set by the FCC.
The pulse used in UWB communications is often referred to with the term monocycle instead of pulse, because the pulse-modulated sinewave used in conventional radars is formed by one cycle of a sinewave, or several cycles, in which case it is called a polycycle.
Although often called a monocycle, the adopted pulse in UWB communication systems is rarely a cycle of a sine wave. It is in fact less difficult and less expensive to generate non-sinusoidal pulses than pulse-modulated sinewaves. Generating pulses of duration on the order of a nanosecond with an inexpensive technology (CMOS chips) has become possible after the introduction of UWB Large Current Radiator (LCR) antennas by Harmuth [25]. LCR antennas are driven by a current, and radiate a power that is proportional to the square of the derivative of current. If a step function current, for example, is applied to the antenna, a pulse is generated: the steeper the step function current, the narrower the generated pulse [26].
The pulse shape that can be generated in the easiest way by a pulse generator actually has a bell shape such as a Gaussian. A Gaussian pulse p(t) can be described by the following expression:

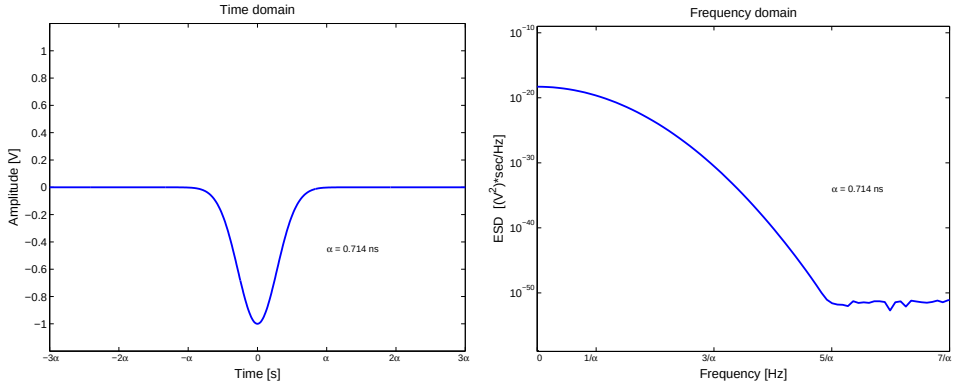

Figure 2.1: Example of a Gaussian pulse: (a) waveform, and (b) corresponding Energy Spectral Density ESD.

$$p(t) = \pm \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\left(\frac{t^2}{2\sigma^2}\right)} = \pm \frac{\sqrt{2}}{\alpha} e^{-\frac{2\pi t^2}{\alpha^2}} \tag{2.8}$$

where $\alpha^2 = 4\pi\sigma^2$ is the shape factor and σ^2 is the variance. The waveform and corresponding Energy Spectral Density (ESD) of the pulse of eq. (2.8) with a minus sign are shown in fig. 2.1.

To be radiated in an efficient way, however, a key feature of the pulse is to have a zero dc offset. Several pulse waveforms might be considered, provided that this condition is verified. This is the case for Gaussian derivatives, that are suitable waveforms for UWB pulse shaping. Actually, the most currently adopted pulse shape is modeled as the second derivative of a Gaussian function [7], described by:

$$p''(t) = \left(1 - 4 \cdot \pi \cdot \frac{t^2}{\alpha^2}\right) \cdot e^{-\frac{2\pi t^2}{\alpha^2}} \tag{2.9}$$

Note that this pulse has energy $E_{p''} = 3\alpha/8\ V^2 s$. The second derivative Gaussian pulse of eq. (2.9) is usually referred to as the pulse at the receiver, that is, after passing through the transmitter and receiver antennas. Ideally, a second derivative Gaussian pulse can be obtained at the output of the transmitting antenna if the antenna is fed with a current pulse shaped as the first derivative of a Gaussian waveform (and thus zero dc current), the radiating pulse being proportional to the derivative of the drive current in an ideal antenna [27].

Other pulse shapes have also been proposed such as the Laplacian [28], compositions of Gaussian pulses having same length and reversed amplitudes with a fixed time gap between the pulses [29], and Hermite pulses [30].

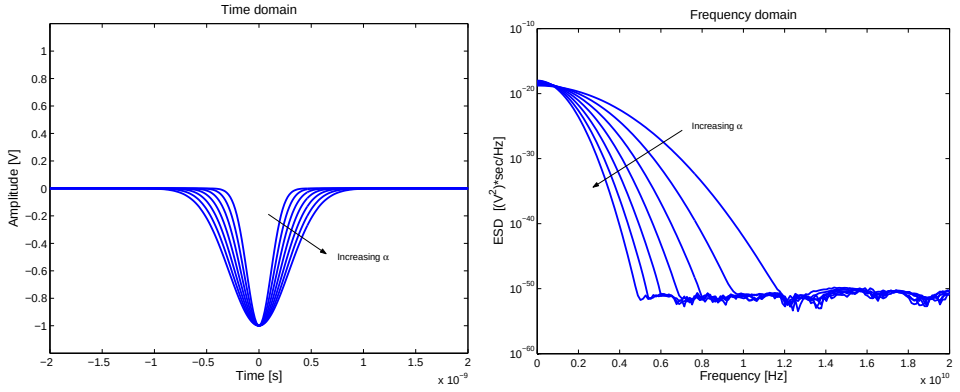

Figure 2.2: The Gaussian pulse case: effect of α on (a) pulse duration and (b) ESD.

Shaping the spectrum by changing the pulse waveform is an interesting feature of IR. Basically, the spectrum may be shaped in three different ways: pulse width variation, pulse differentiation, and a combination of base functions.
In the following, we will consider the Gaussian pulse as a case study. The Gaussian pulse is well suited for our analysis, since its shape can be modified in a straightforward way by modifying the shape factor α, and infinite waveforms can be obtained by differentiating the original pulse.
In the two next subsections, we will first analyze the effect of pulse width variation and differentiation on pulse shape and corresponding ESD characteristics and we will then analyze the possibility of combining different Gaussian waveforms, such as different derivatives and different pulse widths, for the purpose of approximating the FCC emission masks introduced in Chapter 1.

2.2.1 Pulse width variation and pulse differentiation

Pulse width is tightly related to the shape factor α. Reducing the value of α shortens the pulse, and thus enlarges the bandwidth of the transmitted signal. As a consequence, the same waveform can be used to occupy different bandwidths by adjusting the value of the shape factor. As an example, fig. 2.2 shows the effect of alpha on both time and frequency occupation of the Gaussian pulse for α varying in the range 0.414 - 1.014 ns.

Note that the Gaussian pulse has infinite duration leading to unavoidable overlap between pulses and ISI. As already done in section 2.1.1, however, it is reasonable to consider for the Gaussian pulse a limited duration T_M by limiting the cutout energy below a given threshold. Under this assumption, an upper limit for α is given by pulse duration T_M, which cannot exceed chip duration T_c, while a lower limit is given by technological limitations in gener-

ating extremely short pulses.
Differentiation of the Gaussian pulse influences the ESD as well; both peak frequency and bandwidth of the pulse vary with increasing differentiation order. In particular, it is possible to find a general relationship between the peak frequency f_{peak}, the order of differentiation k, and the shape factor α by observing that the Fourier transform of the k-th derivative has the property:

$$X'_k(f) \propto f^k \cdot e^{-\frac{\pi \cdot f^2 \cdot \alpha^2}{2}} \tag{2.10}$$

which leads to a peak frequency $f_{peak,k}$ for the k-th derivative:

$$f_{peak,k} = \sqrt{k} \cdot \frac{1}{\alpha\sqrt{\pi}} \tag{2.11}$$

Equation (2.11) shows that Gaussian derivatives of higher order are characterized by higher peak frequencies. Differentiation is thus a way to move energy to higher frequency bands.
Let us consider the first 15 derivatives of the Gaussian pulse, shown in fig. 2.3. Figure 2.4 shows the effect of differentiation on the the value of peak frequencies for the 15 derivatives as a function of the shape factor α based on computation of eq. (2.11).

The overall effect of differentiation of the Gaussian pulse is a shift of the ESD to upper frequencies, as shown in fig. 2.5, with a slight increase of the -10 dB bandwidth of the ESD itself (fig. 2.6).

2.2.2 Meeting the emission masks

It was shown in the previous section that both differentiation and pulse width variation affect the ESD of the Gaussian pulse and can be used to shape the transmitted waveform. In most cases, however, the flexibility in shaping the spectrum guaranteed by a single waveform may not be sufficient to fulfill specific requirements. Meeting emission masks set by regulation authorities is a typical task demanded of the pulse shaper.
In particular, the release of the FCC emission masks for UWB devices triggered research in finding pulse shaping techniques capable of approximating the masks with high efficiency, that is, with the maximum possible transmit power given the FCC limits. In this section, we will analyze the possibility of tuning the ESD of a generated pulse by combining a few single reference pulse waveforms as for example the Gaussian pulse and its derivatives to adjust the ESD to the limits imposed by the masks. Note that by doing so, we implicitly assume that the generated waveform can be assimilated to one pulse instead of several. It can be shown, however, that the assumption is always satisfied, in the sense that, despite modulation and coding, the emission masks are always

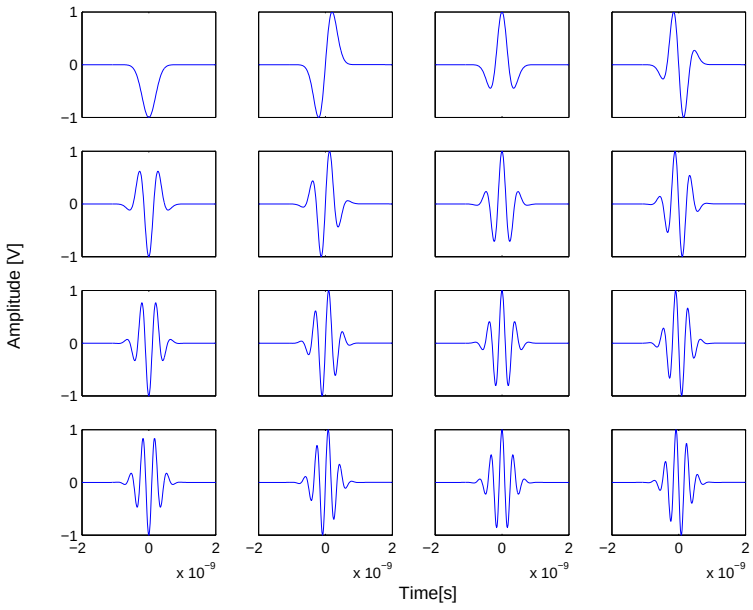

Figure 2.3: Waveforms of the Gaussian pulse and its first 15 derivatives.

met provided that this holds true for the single pulse.
The topic of meeting the emission masks by a properly selecting the pulse shape has been addressed in previous work. In [31], an algorithm is proposed for generating a family of pulse waveforms capable of meeting the FCC masks and suitable for multi-user scenarios, thanks to the short time duration and orthogonality properties between different pulses in the family. The performance of PPM-TH systems adopting this pulse family was analyzed in [32].
In a recent solution based on Gaussian waveforms ([33]), an algorithm to select the best pulse differentiation order and shape factor value for fitting the mask in the bandwidth 3.1-10.6 GHz was proposed. A limitation of this approach is in the difficulty of meeting the mask outside the above bandwidth by using a single derivative.
In this section, we will therefore investigate the possibility of obtaining the optimal waveform as a combination of different derivative functions of the Gaussian pulse, to approximate the mask at all frequencies, including the 0-0.96 GHz band. A possible approach is to use linear combinations of N Gaussian derivatives, each being characterized by a given α value (different derivatives may have different α values), which can be thought as independent

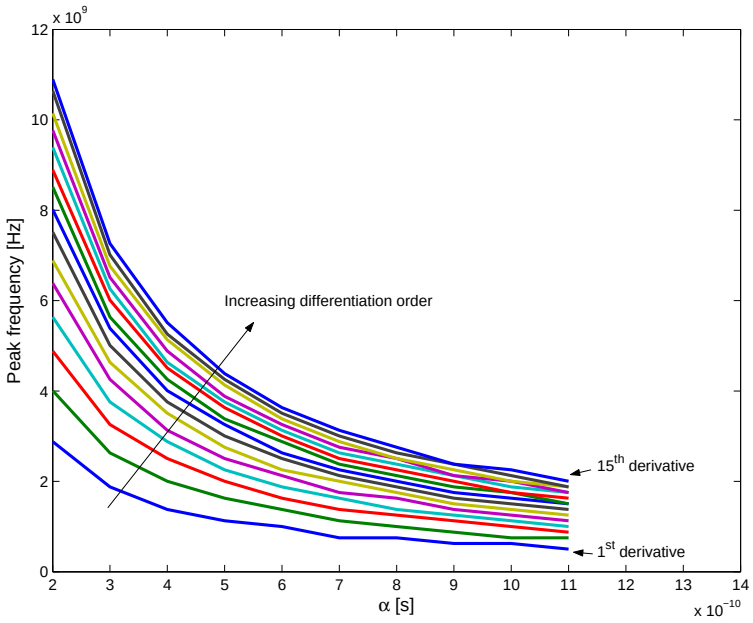

Figure 2.4: Variation of peak frequency with α for the first 15 derivatives of the Gaussian pulse of eq. (2.8), according to eq. (2.11).

Base Functions (BFs) in a space of N dimensions. The choice of the coefficients of the linear combination must be made depending on a design objective such as meeting a given PSD.

A procedure for selecting the coefficients can be described as follows:

1. Choose a set of BFs.
2. Generate in a random way a set of coefficients, named S.
3. Check if the PSD of the linear combination of the functions obtained with coefficients S satisfies the emission limits.
4. If the emission limits in step 3 are met and this is the first set S verifying the limits, then initialize the procedure by setting $S_B = S$. If the emission limits in step 3 are met and the procedure was already initialized, then compare S with S_B; if S leads to a better waveform than S_B according to well-defined distance metrics, set $S_B = S$.

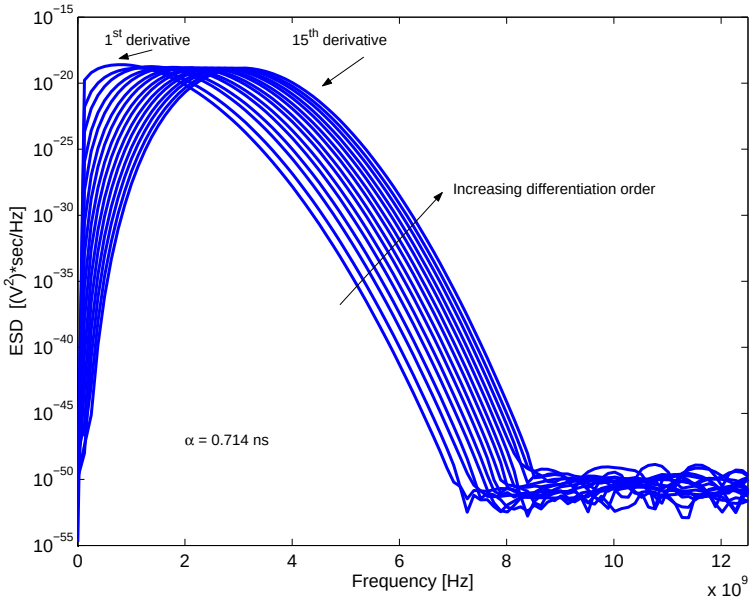

Figure 2.5: ESD of the first 15 derivatives of the Gaussian pulse.

5. Repeat steps 1-3 either until the distance between the mask and PSD of the generated waveform falls below a fixed threshold, or for a predetermined number of iterations.

Note that the combination of N derivatives and the possibility of choosing different α values for different derivatives provides a high degree of flexibility in the generation of pulse waveforms. The algorithm may, however, require a high number of iterations before converging to a solution that is capable of guaranteeing the requested distance between the synthetic and reference target.

For example, let us consider the first 15 derivatives of the Gaussian pulse. Figure 2.7 shows the PSD of a waveform obtained by linear combination of the above BFs plotted against the FCC emission mask, according to the algorithm proposed in this section, for the case of α equal to 0.714 ns for all derivatives, $T_s = 10^{-7}$ and 100 iterations. Figure 2.7 shows that the combination of several BFs leads to a good approximation of the emission mask with the adopted α, in particular in the band 0.96-3.6 GHz. Outside this band, power is less

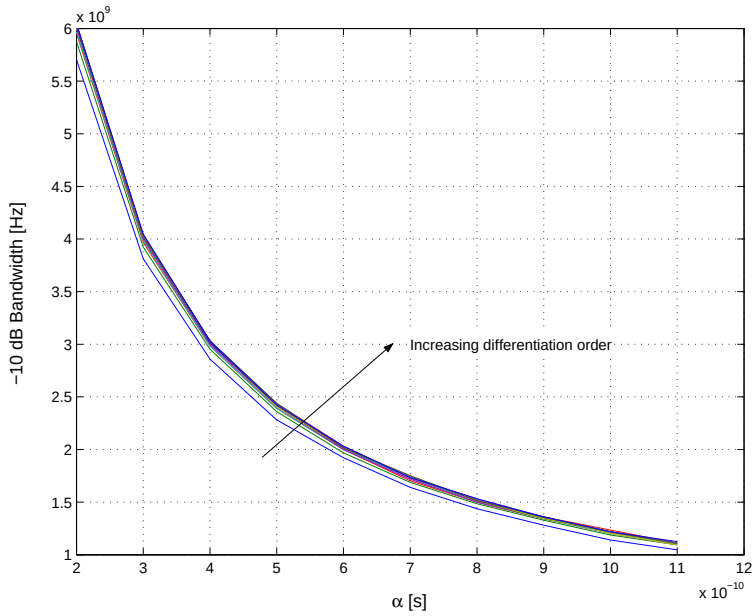

Figure 2.6: Variation of -10 dB bandwidth with α for the first 15 derivatives of the Gaussian pulse.

efficiently used. Based on the results presented in section 2.2, the adoption of different α values can increase performance. For example, one can choose a high value of α (1.5 ns) for the first derivative and smaller values (0.314 ns) for the other derivatives. This alternative choice of α values leads to the PSD shown in fig. 2.8. Note the relatively narrow shape of the first derivative on fig. 2.8 (curve with the first maximum at the very left of the plot) due to the high α value.

This alternative selection of α values leads to a better approximation of the mask than the first one, thanks to the larger bandwidth of the higher derivatives. Note that the selection of a relatively high α for the first derivative improves efficiency of power utilization in the low frequency band. An upper bound for α is, however, given by waveform duration, and determined by the chip time T_c.

Random selection is only one of the possible strategies for the set of coefficients in the linear combination. A more systematic way of selecting such coefficients is to apply standard procedures for error minimization such as

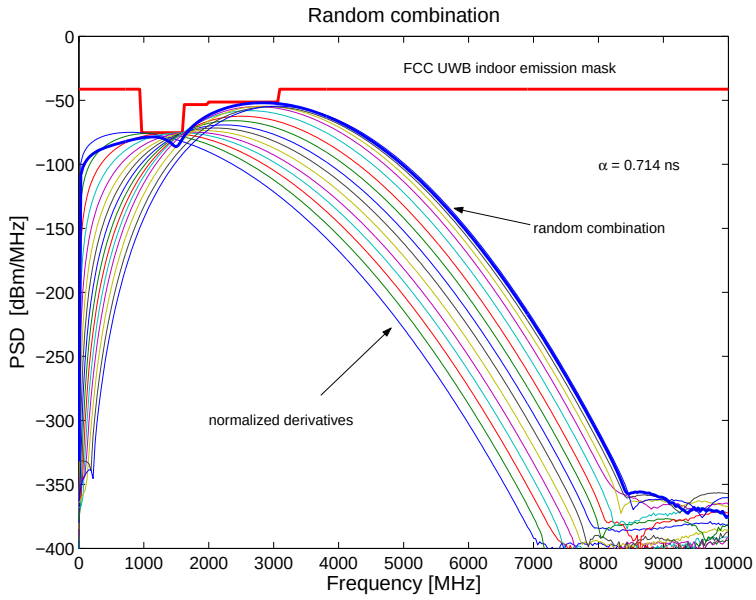

Figure 2.7: PSD of the base functions (thin plots) and of the combined waveform (thick plot) with $\alpha = 0.714$ ns for all derivatives, $T_s = 10^{-7}$ and 100 iterations.

the Least Square Error (LSE), in which the following error function must be minimized:

$$e_s(t) = \int_{-\infty}^{+\infty} |e(t)|^2 \, dt = \int_{-\infty}^{+\infty} \left| f(t) - \sum_{k=1}^{N} a_k \cdot f_k(t) \right|^2 dt \tag{2.12}$$

In eq. (2.12) $f(t)$ is the target function. Note that since requirements are specified in terms of meeting a PSD, the error of eq. (2.12) rewrites as follows:

$$E = \int_{-\infty}^{+\infty} |P_M(f) - F(f)|^2 df \tag{2.13}$$

where $P_M(f)$ represents the emission mask and $F(f)$ the PSD of the linear combination. Equivalently, one can consider the corresponding autocorrelation functions $R_M(t)$ and $R_F(t)$ and minimize an error defined as follows:

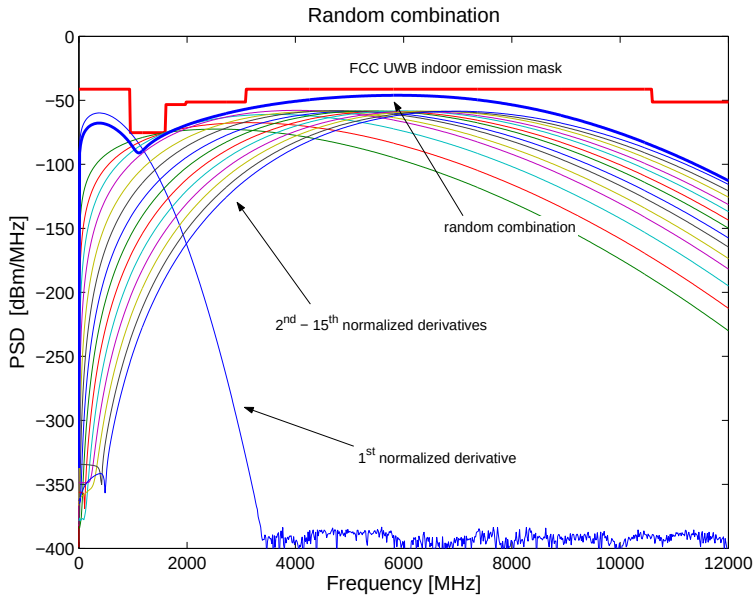

Figure 2.8: PSD of the base functions (thin plots) and of the combined waveform (thick plot) with $\alpha = 1.5$ ns for the first derivative and $\alpha = 0.314$ ns for higher derivatives, for all derivatives, $T_s = 10^{-7}$ and 100 iterations.

$$E = \int_{-\infty}^{+\infty} |R_M(t) - R_F(t)|^2 dt = \int_{-\infty}^{+\infty} \left| R_M(t) - \left[\sum_{k=1}^{N} a_k^2 \int_{-\infty}^{+\infty} f_k(\xi) f_k^*(\xi + t)\, d\xi \right] \right|^2 dt \tag{2.14}$$

Note that both eq. (2.13) and eq. (2.14) approximate the PSD of the transmitted signal by considering the envelope of the PSD. This approach is justified by the fact that the shape of the PSD, and in particular its maximum value is mainly determined by the Fourier transform of the impulse response of the pulse shaper $P(f)$. Using $P(f)$ in place of the exact PSD simplifies the optimization problem if one computes the voltage m(t), which represents the mask bound in the time domain.
The target emission voltage mask can be obtained by dividing the power emission mask normalized by $1/T_s$ by the freespace impedance, taking the square root, and applying the Fourier anti-transform. In this case, the error is then defined as follows:

$$E = \int_{-\infty}^{+\infty} \left| m(t) - \sum_{k=1}^{N} a_k f_k(t) \right|^2 dt \tag{2.15}$$

As a general remark, note that the LSE method cannot guarantee by itself compliance with the emission mask. The optimization procedure is based on an average quadratic distance and does not impose bounds on a frequency-by-frequency basis. This is shown in fig. 2.9.

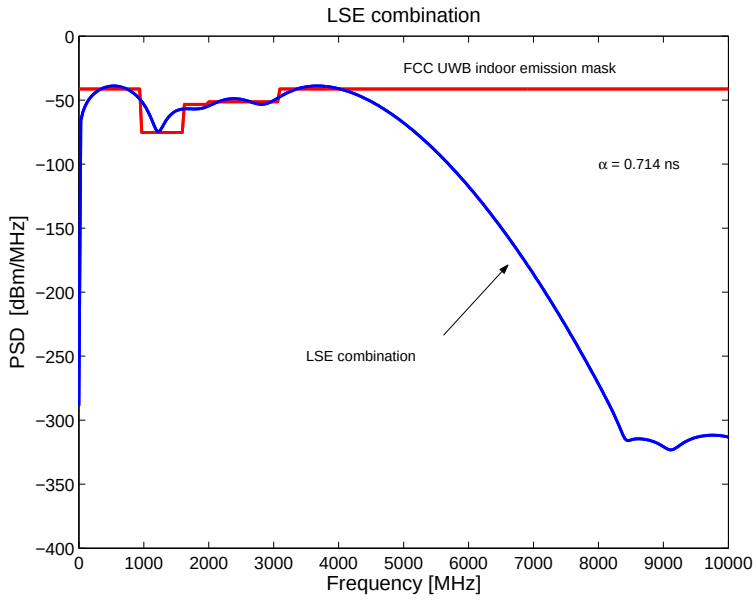

Figure 2.9: Envelope of the PSD of the linear combination of Gaussian waveforms vs FCC indoor emission mask, for $\alpha = 0.714$ ns for all derivatives.

Note that a flexible pulse shaping function would also allow to improve MUI robustness in overlapping UWB networks, by adopting different pulse shapes in the two networks. In UWB networks based on TH-CDMA MUI is the key factor in determining the maximum achievable bit rate, in particular for medium and high rate transmissions. Pulse shaping can be a powerful tool in reducing the negative effect of MUI. A MAC protocol capable of selecting different pulse shapes, for example, can optimize network organization by assigning different waveforms to different groups of terminals, thus reducing the MUI noise suffered by each terminal and increasing network performance.

This solution was tested in a scenario characterized by two disjoint UWB networks in the same physical area. Each of the two networks, referred to as N1 and N2, was composed by 24 transmitting devices at a rate $R_b = 50$ Mb/s. The effect of MUI noise generated by interfering devices in both N1 and N2 on a useful link in N1 was analyzed by measuring the Bit Error Rate (BER) of the link as a function of the E_b/N_0 ratio. The effect of pulse shaping was analyzed as follows. In all simulations the waveform considered at the output of the transmitting antennas in N1 was the second derivative of the Gaussian pulse, while the waveform adopted in N2 varied for each run of simulations.

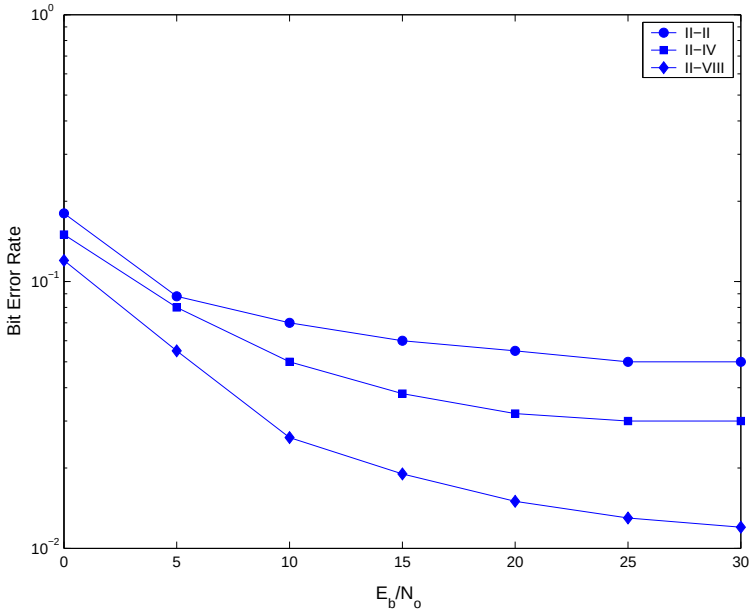

Figure 2.10: Bit Error Rate as a function of Eb/N0 for different waveform selection in two overlapping UWB networks N1 and N2. Circles: both N1 and N2 use the II derivative of the Gaussian pulse; Squares: N1 and N2 use the II and IV derivative of the Gaussian pulse, respectively; Diamonds: N1 and N2 use the II and VIII derivative of the Gaussian pulse, respectively.

Results are presented in fig. 2.10 for the cases of second, forth and eighth derivative of the Gaussian pulse adopted as pulse waveform in N2. Results show that the adoption of different waveforms in the two networks reduces the BER, thanks to the different bandwidths occupied by the corresponding

PSDs. This is also confirmed by the fact that higher derivatives adopted in N2 lead to lower BER, due to the effect of differentiation on bandwidth and peak frequency, shown in section 2.2.1. Note that the strategy of assigning different waveforms to different networks can be applied to linear combinations of base functions as well, guaranteeing at the same time low MUI interference and good approximation of the emission masks.

Chapter 3

Medium Access Control design in UWB systems

3.1 Introduction

The design of a Communication System traditionally proceeds along the principle of a layered architecture. The aim is to design each layer of the system independently of the internal structure of the lower layers, which are considered as black boxes offering a service to upper layers [34]. Rules of interaction between different layers are defined by interfaces which determine the requested inputs and corresponding outputs. Within this framework, the MAC is generally considered as the bottom part of the Data Link Control (DLC) layer. The service offered by the MAC to the upper DLC is to provide a bit pipe, preventing or resolving contentions in the access to the medium. Following the layered approach, the functions executed in the MAC should be defined without taking into account the underlying physical layer. The design of an efficient MAC often requires however an accurate knowledge of the physical layer, and in most existing systems specific properties of the transmission technique are exploited in order to reduce the effect of multiple access interference. In the case of UWB systems, this is a crucial issue where UWB potentials (for example precise ranging capability) may enable the definition of novel MAC functions, as well as lead to a drastically different implementation of more traditional MAC functions.

As a general principle, the role of the MAC layer is to allow multiple users to share a common resource. The definition of resource, and of the procedures by which access to the medium is granted, depends on the adopted transmission and multiple access techniques. As discussed above, the evaluation metrics for a MAC protocol should be however defined independently of the physical layer.

A tentative set of key parameters can be given as follows [35]:

- *Throughput*: percentage of channel capacity in use during data transmission
- *Delay*: average time spent by a packet in the MAC queue
- *Fairness*: equal opportunity for terminals to access the medium

Other parameters which are not listed above are usually related to a specific MAC protocol or a specific scenario. In a centralized architecture, for example, an evaluation criterion may consist in the degree of asymmetric flexibility in the bandwidth allocation to the downlink and uplink streams for a given user [36].

Accordingly, key MAC design objectives should be: i) to maximize throughput, ii) to guarantee an acceptable delay, and iii) to grant fair access. The above goals should be fulfilled in a dynamic environment, i.e. under variable channel conditions, traffic characteristics, and local network topologies. Flexibility is thus an additional feature which an advanced MAC should incorporate.

In this chapter, we will analyze how the above objectives can be pursued in the case of UWB wireless networks. It should be noted that several issues being raised in the design of a MAC for UWB networks are not novel, and that similar problems have been solved in the past for conceiving the MAC of wireless networks. The aim is thus twofold. First, we will identify the key areas to be investigated in MAC design, and second, we will try to understand which areas can benefit from existing solutions and which oppositely require innovation in order to take into account UWB specificity. In order to reach these goals, we will review, for each key area, several examples of the solutions adopted in existing wireless networks. Such review, although not exhaustive, will provide the background required to analyze UWB specific topics in MAC design. It will be then shown that different areas are affected in different degrees by the adoption of UWB, ranging from cases in which existing solutions can be adopted without modifications to cases in which UWB is the key enabling technology for new MAC capabilities and innovative design, typically based on ranging and positioning.

3.2 MAC Design Guidelines

In this section, we will identify the key MAC design areas, corresponding to a set of MAC functions to be implemented in order to reach the key objectives discussed in section 3.1.

In the following, the general case of a distributed network architecture will be considered. The focus being on MAC organization, however, both centralized and distributed MAC solutions will be considered. Each of the MAC functions described below, in fact, can be implemented either in a centralized or in a distributed fashion independently on the overlaying network architecture.

MAC functions which in general need to be implemented in most systems are as follows:

a) Medium sharing - This function determines how terminals access the medium in order to transmit packets.

b) MAC Organization - This function deals with the organization of the network at the MAC level, i.e. how terminals coordinate themselves in resource sharing.

c) Admission control - In Quality of Service (QoS) aware networks, this function is used to regulate the access of traffic sources to the network, avoiding congestion.

d) Packet scheduling - When multiple traffic flows are present in the same terminal, packet scheduling is used to select the next packet to be transmitted.

e) Power control - Power control aims at optimizing power utilization in the network.

Overlaid over the above functions is QoS management. QoS involves most of the functions defined above and can be seen as a horizontal function. For this reason, we will first address the general topic of QoS in distributed wireless networks. Then, a description of the basic MAC functions listed above will be given, and several examples of their implementation in existing wireless networks will be presented. Improvements and difficulties in the realization of such functions due to specific UWB characteristics will be highlighted.

3.2.1 QoS management at the MAC Layer

Technological progress and social modifications make data networks more and more appealing as a universal way to transfer all kinds of information. Thus a modern data network must be capable to deliver at the same time data, voice, multi-medial (e.g. streaming video), and real-time-critical traffic by adapting its behavior to different user requirements and traffic characteristics.

Voice and multi-media traffic, in particular, are characterized by requirements which are not present in non-real-time data traffic, e.g. the necessity of

transferring bit streams at a minimum bit rate (determined by the application generating the traffic) with an upper bound on the end-to-end delay. The fulfillment of the above requirements guarantees that the end user perceives the offered service with the requested quality: QoS defines thus the performance which must be guaranteed by the network in order to meet user expectations.

As a consequence, the adoption of strategies that modify network behavior depending upon traffic characteristics and QoS constraints is commonly considered as a natural evolution of data networks.

The first step in the design of such strategies is the definition of a set of parameters defining QoS. Note that although each different service is characterized by its own application-level QoS parameters (e.g. resolution, frame rate for video services, sample rate and sample size for audio services), these are mapped onto a unique set of network-level QoS parameters, which can be listed as follows: bandwidth, end-to-end delay, jitter, bit error rate and packet loss.

Typical values of these parameters depend upon the corresponding service. A few examples are reported in Table 3.2.1 [37].

Service	QoS Parameter	Range
Audio (Telephone Speech)	Bandwidth	16 kbps
	End-to-end delay	400 ms
	Packet loss	10^{-2}
Video (HDTV, lossy compression)	Bandwidth	20 Mbps
	End-to-end delay	250 ms
	Bit error rate	10^{-6}
	Packet loss	10^{-11}
Data	Bandwidth	0.2 - 10 Mbps
	End-to-end delay	1 s
	Packet loss	10^{-11}

Table 3.1: Mapping of Services on Qos Parameters.

Different solutions have been proposed to introduce QoS in data networks; all of these solutions rely however on the definition of a set of service classes, identifying the different levels of QoS which can be guaranteed by the network [38]. Such solutions for QoS support are defined as a component of the Network layer. Nevertheless, the effective deployment of QoS is heavily affected by the underlying Data Link and Physical layers. In the case of wireless networks, the radio transmission medium has a heavy impact on the QoS offered by the network. Network layer QoS mechanisms are in fact based on the reliability of the physical medium, something which cannot be easily guaranteed in the case of mobile terminals. Thus, the QoS concept needs to be

adapted to this hostile environment.

As a matter of fact, it is impossible to guarantee at any time the fulfillment of QoS requirements at the physical layer. Any transmission medium is in fact characterized by an outage probability, defined as the out-of-service probability of the physical medium, which is different from zero. The key difference between wired and wireless networks is in the value assumed by such outage probability; in wired networks the probability of a link down is low enough to allow the upper layers to simply overlook this event. In the case of radio networks, instead, link failures are frequent enough to impact the link and network layers, and in particular the MAC sublayer, leading to the necessity of introducing mechanisms to rapidly recover errors on the link, e.g. Forward Error Correcting (FEC) codes or Automatic Repeat on reQuest (ARQ) protocols. In spite of these mechanisms, however, the fulfillment of QoS requirements cannot be guaranteed deterministically, but only with a given probability. In this perspective, the outage probability of the physical medium translates at the MAC layer in an out-of-service probability. This probability of failure, i.e. missing the QoS requirements, can be reduced by correctly designing and tuning the FEC and ARQ mechanisms cited above. As the channel behavior worsens, however, the fulfillment of the requirements becomes less and less realistic. The MAC may simply have not enough resource to compensate the channel outage probability.

In such a condition of scarce resource, priorities are required to obtain a fair resource sharing, at the MAC layer. Priorities can be defined at two levels:

1. priority between different users/terminals
2. priority between different traffic types (real-time/voice traffic, data traffic)

As a consequence, the introduction of QoS management involves several MAC functions, from admission control and packet scheduling, to power control and MAC organization. Examples of the impact of QoS management on the functions defined at the beginning of section 3.2 will be given throughout the following subsections.

3.2.2 Medium Sharing

Most of the existing MAC protocols for distributed networks are based on the key hypothesis that users share a single channel. From a resource sharing point of view, this implies that the resource to be shared is the radio access itself.

Two possible choices are available for resource management: terminals may contend in order to gain channel control (random access), or channel control

may be granted by a control unit based on a specific resource assignment protocol (scheduled access). While the random access approach is appropriate for bursty traffic, scheduling allows a more efficient utilization of the channel when continuous streams of data packets must be transferred. Even in the case of a scheduled approach, however, a random access phase is requested, since the scheduling sequence is typically unavailable at network startup. In this section we will thus focus on available solutions for random access, while examples of scheduled access will be moved to section 3.2.3.

Random access typical solutions for wireless networks are Aloha, Carrier Sensing Multiple Access (CSMA), and Out-of-Band signaling [35].

Aloha main advantage is simplicity. The Aloha protocol only foresees in fact a CRC field to be added to data packets before transmission. If a collision occurs, a backoff procedure is activated in order to schedule retransmission of the corrupted packet. Aloha has been proven to well behave when low traffic load is offered to the network, while performance decreases abruptly as traffic load increases and packet length grows [34]. Slotted Aloha, in which a slotted time axis is adopted, and terminals start transmission attempts only at the beginning of a time slot, improves performance without really solving this issue. For this reason, Aloha was proposed for the specific case of short, rare packet transmission (e.g. control packets), i.e. when the transmission time is low enough to mitigate the effect of collisions. Under the condition of a high traffic load, a higher throughput can be obtained by means of CSMA which is based on a channel sensing period performed by each terminal before starting transmission. The performance obtained by CSMA is however heavily affected by two phenomena, the well known "hidden terminal" and "exposed terminal" problems. In order to solve the hidden and exposed terminal problems, alternative solutions to CSMA have been proposed. The Multiple Access with Collision Avoidance (MACA) protocol [39] for example replaces the carrier sensing procedure with a three-way hand-shake between transmitter and receiver. Following this approach, further modifications of the MACA protocol have been developed, such as MACAW [40] and MACA-By Invitation (MACA-BI) [41].

Practical implementations of MAC protocols combine hand-shake and carrier sensing, as proposed in the Floor Acquisition Multiple Access (FAMA) protocol [42]. These protocols are commonly referred to as CSMA with Collision Avoidance (CSMA-CA). An example of CSMA-CA is the Distributed Foundation Wireless MAC (DFWMAC) which has been adopted for the MAC layer of the 802.11 IEEE standard [43]. 802.11 adopts a Clear Channel Assessment (CCA) function which performs channel sensing in two different ways, either by measuring the received power and comparing it with a threshold, or by performing a true carrier sensing by detecting another 802.11 signal on the same channel.

An alternative solution to CSMA-CA is offered by the Out-of-Band signaling protocol [44]. This solution splits the bandwidth available for communication into two channels: a data channel used for data packet exchange, and a narrowband signaling channel on which sinusoidal signals (referred to as busy tones) are asserted by terminals which are transmitting and/or receiving in order to avoid interference produced by hidden terminals. In a distributed network, this would require each terminal which detects a transmission to transmit a busy tone to block all nodes in an area of radius 2*R around the transmitting node, R being the radio range, with the consequence of amplifying the exposed terminal problem (fig. 3.1) [35], [44].

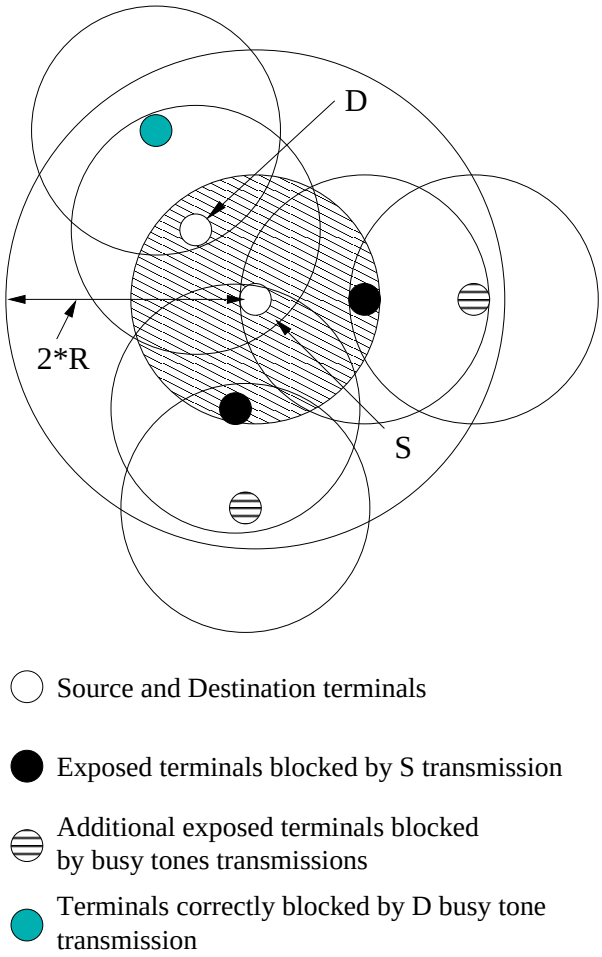

Figure 3.1: Amplification of the exposed terminal problem with Out-of-Band signaling.

In order to reduce the number of exposed terminals, the use of two different Busy Tones for transmitting and receiving terminals was proposed [45].

3.2.3 MAC Organization

In a distributed scenario, independent terminals cooperate in order to build the network. Two main approaches are possible for network self-organization at the MAC layer: Domain-dependent (clustered) and Domain-independent (flat).

Most of the MAC protocols proposed in the literature and adopted in WLAN standards rely on the explicit definition of a MAC Domain leading to a clustered network architecture, where each cluster corresponds to a MAC domain. A clustered architecture simplifies resource management within each cluster by allowing a centralized approach. Two examples of Domain-dependent MAC protocols are Bluetooth and IEEE 802.15.3.

In Bluetooth [10] a Frequency Hopping - Code Division Multiple Access (FH-CDMA) scheme is adopted, and each MAC Domain is associated with a FH sequence. Terminals in a given area self-organize into MAC Domains (piconets) composed by up to eight terminals, and a centralized resource management is performed within each piconet. The set of independent piconets in the area is called scatternet.

The MAC Domain setup is started by a terminal following a neighbor discovery phase, achieved by means of a dedicated scan procedure. The terminal assumes the role of piconet Master; It determines the FH sequence, and pages one of the discovered neighbors (Slave) to start the piconet. Other devices can be included in the piconet, with individual paging by the Master. Alternatively, a device can join an active piconet by paging the Master. Since in Bluetooth the paging device always assumes the Master role, if the new paging device wants to assume a Slave role in the piconet, a Master-Slave switch procedure is required. The Master-Slave switch procedure is executed on request. No explicit criterion about when the switch should be performed is defined.

The IEEE 802.15.3 standard [9] is a second example of Domain-dependent MAC. The standard was originally developed for traditional, narrowband physical layers in the ISM band, but as it was briefly explained in Chapter 1, a modified version of this standard, adopting an UWB physical layer, is currently in the standardization process under the name IEEE 802.15.3a. In 802.15.3, as much as in Bluetooth, the medium access is controlled in a centralized fashion within each MAC piconet. A piconet is controlled by a PicoNet Controller (PNC) which emits a periodic beacon. The channel associated with the piconet is selected based on a scanning procedure, which determines the channel subject to lower interference. No specific device is targeted in the piconet setup. It is up to the neighboring devices to join the new

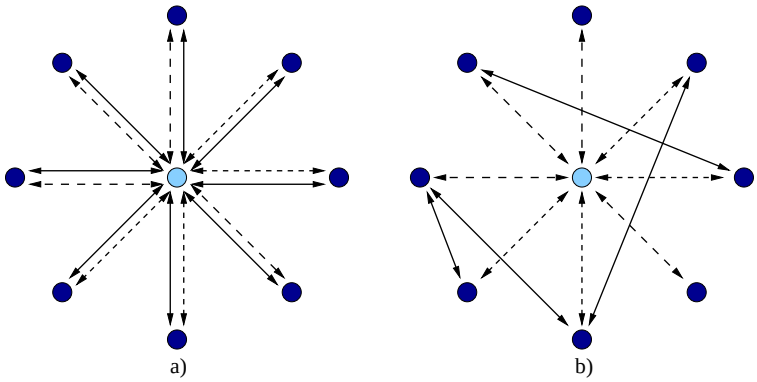

Figure 3.2: Logical piconet topologies: control traffic (dashed arrows) and data traffic (filled arrows).

piconet by synchronizing to the beacon and sending an association request to the PNC by means of random packets either in a CSMA or slotted Aloha fashion. Associated devices ask for Channel Time Allocation (CTA), i.e. time slots, also by means of a CSMA protocol.

Differently from the management of the Master role in Bluetooth, the PNC role in 802.15.3 is always assigned to the device with highest PNC capabilities. Each device is requested to provide its capabilities to play the PNC role during the association procedure. If a newly associated device is better suited than the current PNC, a PNC handover is performed.

The PNC and Master carry out similar tasks in the two standards as regards piconet management, including traffic scheduling and piconet synchronization. A key difference between the two systems however is the way the data traffic flows through the piconet. In Bluetooth, direct communication between two different devices is not allowed, and the Master is in charge of relaying all traffic through the network. The piconet topology is in this case a typical star topology, as depicted in fig. 3.2 a). In 802.15.3, on the contrary, the PNC only schedules CTAs to the devices, without being involved in the data packets exchange. Thus, while the piconet management is fully centralized, data transfer is performed on a peer-to-peer basis, i.e. in a pure ad-hoc manner (fig. 3.2 b)).

The two systems, however, share the same problem of how to allow interpiconet communication. This becomes a major issue in scalability, especially regarding traffic scheduling and routing in networks composed of a large number of piconets.

The adoption of a Domain-independent architecture avoids the problem of interpiconet communications. In a single channel scenario, a Domain-

independent organization can be achieved by adopting any of the random access MAC protocols presented in section 3.2.2. UWB radio can however provide multiple communication channels by assigning different TH codes to different users. As a consequence, protocols designed for multiple channel networks, which have been shown to achieve better throughput in comparison with single channel solutions based on Aloha or CSMA [46], are a viable solution for medium access in UWB systems.
Multiple channel solutions have been typically developed for DS-CDMA. As we will see in the following sections, however, key concepts are portable to UWB case.
As a general comment, the definition of a multiple channel MAC protocol is strictly related to the hardware capability of the terminals: if a receiver is complex enough to be capable of listening to several codes simultaneously, the complexity in designing a multiple channel MAC is significantly reduced.

The definition of MAC organization is crucial in the design of a specific MAC for UWB networks. The adoption of a Domain-based structure is a potential solution, since it addresses management of multiple Time Hopping codes. On the other hand, a multiple channel MAC could significantly increase network throughput by exploiting the inherent multiple channel UWB capability. This aspect is discussed further in section 3.3.

3.2.4 Admission control

Admission control is required when congestion must be avoided in order to meet network performance requirements. Admission control is mandatory in QoS-aware networks in which unregulated access might easily provoke violation of performance guarantees. Best effort networks do not require admission control, but can benefit from its introduction. Admission control is typically implemented with centralized schemes, as in cellular networks [47] and centralized wireless networks [48]. A few proposals for distributed schemes are however also available in the literature. These schemes rely on the cooperation of terminals in evaluating the impact of additional traffic flows on network throughput, and rejecting requests that would cause unacceptable performance degradation, typically due to MUI generated by such potential new entries [49], [50], [51].

This approach is suitable for UWB networks, which rely on TH-CDMA for multiple access. A distributed admission control scheme for an UWB network is proposed in [52], based on Bambos [49].

3.2.5 Packet Scheduling

The packet scheduling algorithm determines the order in which buffered packets are selected for transmission. In wired networks, this function has two main objectives: 1) to guarantee a fair access to all flows to the available capacity, and 2) to support QoS if different traffic classes are present. The simplest solution is the First Come First Serve (FCFS) algorithm in which packets are sent in the same order in which they are buffered. This solution, however, provides no protection against ill-behaving sources, which can capture any percentage of the available bandwidth by increasing their packet emission rate. In order to increase fairness, a Round Robin scheme adopted to serve each traffic flow is proposed in [53]. Fair access however is not guaranteed since packets of different lengths can be present in each queue. The Weighted Fair Queueing algorithm [54] addresses this issue by assigning a weight to each queue with the aim of emulating a bit-per-bit Round Robin between different flows. In this case the introduction of QoS in the scheduling strategy is straightforward since the weights can be easily adjusted in order to take into account the QoS classes.

Efficient packet scheduling in wireless networks cannot ignore the status of the wireless channel. Several wireless scheduling algorithms, which are sensitive to channel status, have been proposed. These are based on either a simple on-off Markov channel model [55] or on more sophisticated channel models leading to accurate evaluation of the Signal-to-Noise-Ratio (SNR) [56] and external interference [57].

To this respect UWB does not present any relevant difference from other radio transmission techniques and the above protocols are directly applicable to the specific UWB case.

3.2.6 Power Control

Due to the broadcast nature of the wireless medium, the achievable performance in wireless networks strictly depends on the capability of minimizing the undesidered effects of each radio transmission on neighboring receivers. Power control leads thus to optimization of emitted power levels and achieves three desirable effects [49]: 1) Minimization of power consumption, leading to longer autonomy, 2) Reduction of interference, and 3) Adaptation of emitted power to link variations due to channel modifications and mobility.

Power control received significant attention in the last few years, in conjunction with the introduction of 3G cellular networks based on CDMA, since it mitigates the near-far phenomenon, in which a transmitter close to the receiver shadows the signal of a further transmitter. The centralized structure of cellular networks, however, simplifies the solution to this problem since the

presence of a base-station significantly helps in the development of efficient power control algorithms. The issue is far more complicated in a distributed network architecture in which several independent links may be set up at the same time without any central controller. Nevertheless, power control should be a key property of distributed MAC protocols since it allows a significant increase in network capacity [58]. A distributed power control protocol for CDMA ad-hoc networks jointly with a power-related admission control function is proposed in [49].

Power control is important in the case of UWB networks as well, at least for two reasons: 1) UWB networks are affected by the near-far effect, although it can be expected that the high processing gain provided by the TH-IR can partially mitigates this phenomenon, and 2) the low power levels allowed for UWB communication networks impose efficiency in the use of power.

3.3 The UWB Case

The key functions required in the MAC design of a distributed network were described in section 3.2. For a few of these, the use of UWB radio was investigated, either by identifying a solution proposed for the UWB case (admission control) or by explaining how UWB can benefit from existing solutions (packet scheduling). A more detailed analysis is required for other functions, which can be heavily affected by the adoption of UWB, such as Medium Sharing and MAC Organization. Such analysis is the main topic of section 3.3.1. Section 3.3.2 deals with new potential MAC functions enabled by the UWB technique.

3.3.1 Medium Sharing and MAC Organization in UWB networks

The review of section 3.2 shows that the selection of channel access protocols and definition of MAC Organization are tightly related. The first step in analyzing potential MAC organization for UWB networks is thus to evaluate the applicability of existing channel access protocols to this specific case.

Several UWB definitions have been proposed in the last few years, ranging from Time Hopping Impulse Radio (TH-IR) to Direct Sequence UWB (DS-UWB), and to Multi-band UWB. In the following, we will consider TH-IR, which is the most common definition of UWB radio. In this case, the most intuitive solution is to identify each channel by a different Time Hopping code.

Independently of a selected MAC organization, that is clustered or flat, the following questions must be addressed:

- How is the channel defined for a TH-IR system?

- How is the channel accessed by terminals?
- How do terminals manage multiple channels?

It is worth noting, first, that if a Domain-based organization is adopted, the definition of intra-Domain procedures is almost independent on the channel definition. If for example a TDMA scheduled access scheme is considered, as in IEEE 802.15.3, procedures for CTA request and allocation are the same independently on the definition of the channel. Oppositely, procedures for inter-Domain communications, or channel selection, are highly dependent on channel definition.

The 802.15.3 MAC standard defines a procedure for channel selection aiming at minimizing inter-piconet interference since many interfering devices are expected to be present in the 2.4 GHz ISM band. The adoption of a UWB physical layer should amplify this issue since the UWB signal spreads over a much larger bandwidth than ISM and partially overlaps with a large number of narrowband systems. As a consequence, an accurate channel monitoring will be required in order to meet the severe coexistence issues imposed to UWB and to allow such a system to reach the requested performance. As an example of how the adoption of TH-IR would impact such procedure, if the channel is defined by means of a TH-code, coexistence with a narrowband system could be achieved by choosing a code which introduces a notch in the UWB signal Power Spectral Density in the band occupied by the narrowband system.

The channel, as defined above, must be accessed by terminals in order to exchange data and control information. The selection of the protocol to access the channel should then consider the TH-IR characteristics. The Aloha protocol requires no specific actions to be performed by the transmitter before emitting a packet. Its application to UWB is thus straightforward. As explained in section 3.2.2 the main concern about this protocol is its poor performance in heavy traffic load conditions. It should be noted, however, that the evaluation of such performance is carried out under the hypothesis of destructive collisions, which is quite realistic in the case of narrowband signals, characterized by a high duty cycle. TH-IR signals, on the contrary, can achieve low duty cycles, and could thus offer a higher protection in case of packet collisions. Further research is necessary in order to better characterize UWB interference and correctly evaluate the effect of packet collisions.

The CSMA/CA protocol requires the capability of sensing the channel in order to understand if a transmission can be started. As a general remark, CSMA protocol is only suited for spread spectrum signals with low processing gain. In fact, spread spectrum systems with high processing gain do not experience significant performance increase by switching from Aloha to CSMA based on power measurement because of the lack of correlation between the

interfering power measured at the transmitter and the interfering power suffered at the receiver [59]. Furthermore, true carrier sensing (i.e. the identification of another transmission) is complicated by the spreading itself, which makes it difficult to detect a spread signal, if the synchronization preamble is missed [60]. The extremely high processing gain guaranteed by TH-IR is expected to amplify these drawbacks, leading to the conclusion that CSMA is most likely not suitable for UWB systems. Let us consider for example a scenario in which two devices, A and B, need to transmit data to the same receiver C. B is already transmitting, while A is performing a carrier sensing procedure. Due to the spatial positions of A and B, A does not receive any of the pulses transmitted by B, and considers the channel as clear (fig. 3.3).

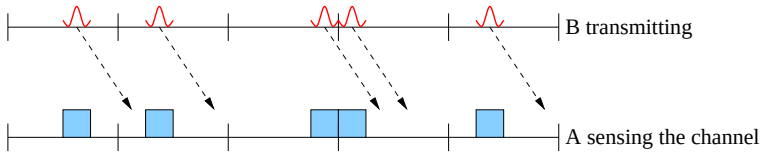

Figure 3.3: Example of error in Carrier Sensing procedure in a TH-IR system.

In the worst case, A will start transmitting with the same phase as B (i.e. the same code value) and with a delay equal to the difference between the propagation delays from A and B respectively to C: this will lead to systematic collisions at C when A starts transmitting (fig. 3.4).

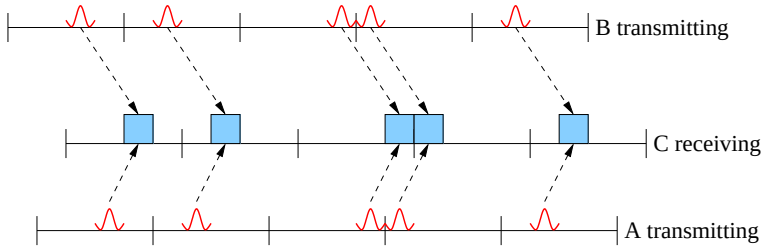

Figure 3.4: Collision at the receiver due to error in Carrier Sensing procedure in a TH-IR system.

In general, the number of collisions will depend on both the relative delays between the two transmissions and the autocorrelation properties of the Time Hopping code. Simulations and measurements are required to evaluate if a simpler protocol (e.g. Aloha) can guarantee the requested performance by relying on the temporal diversity properties of TH-IR UWB and thus skipping the carrier sense procedure.

The third random access protocol analyzed in section 3.2.2 is the Out-of-Band signaling. This solution has however several drawbacks which make its adoption in UWB networks unlikely. First, it requires additional hardware to generate one or more narrowband tones, leading thus to higher cost. Second, and most important, the propagation characteristics for narrowband tones and UWB emissions are likely to be completely different, leading to unpredictable protocol behavior due to different ranges for control and data information.

The above considerations indicate Aloha to be the best solution to allow channel access in low data rate UWB networks. This technique could be adopted in a Domain-based architecture, in which random access is only used to transfer control information and scheduled access is adopted for data transmissions. It is worth noting however that existing standards, such as 802.15.3 and Bluetooth, do not define procedures for the interconnection of independent MAC Domains. As a consequence, the maximum size of a piconet is a first bound for the maximum network size and network scalability. The same issue of course raises in the case of UWB networks, if a MAC Domain is defined. TH-IR, however, provides a built-in multiple access scheme based on TH-codes, so that the adoption of a MAC Domain is not mandatory in the design of the MAC for an UWB network. In fact, a completely distributed MAC organization can be foreseen, in which each link is activated on a different TH-code. Protocols described in section 3.2.2 for code assignment and multiple channel MAC management in CDMA networks can be adapted to the UWB case, with a few additional considerations.

First, UWB systems differ from DS-CDMA, since spectrum spreading is achieved thanks to the impulsive nature of the UWB signal, without requiring a high chip rate for obtaining a high processing gain. This means that even if the same TH code is selected for two simultaneous transmissions, most probably, this will not lead to excessive interference because temporal separation will avoid systematic collisions. As a consequence, in network scenarios characterized by low or medium terminal densities and low bit rates, a solution based on a single TH code may lead to good throughput. In this case, the overhead due to a code assignment protocol may be avoided.

Second, TH-IR systems may require higher time lags for synchronization than existing DS-CDMA systems. Even if it will not hinder point-to-point communications, it should be taken into account when defining MAC and higher layer protocols. Most protocols rely in fact on the availability of a broadcast channel heard by all terminals; Each terminal must be able to synchronize to this channel within an acceptable time. This condition may not be met in TH-IR systems. In this case, alternative solutions should be adopted, either based on the absence of a broadcast channel, or based on a network reference clock maintained by all terminals in order to reduce the synchronization time to the common channel.

3.3.2 UWB novel functions

In the above sections, we analyzed how UWB may influence the implementation of traditional MAC functions. UWB unique characteristics should however enable the definition of new functions as well that are specifically designed to exploit this technology.

The main innovation offered by UWB is the capability of achieving high accuracy ranging. It should be noted that this characteristic is typical of spread spectrum signals in general. Time of Arrival (TOA) estimations for example can be obtained in DS-CDMA systems by evaluating time shifts between the spreading code in the receiver and the same code in the received signal. The ranging precision thus depends upon the capability of determining this time shift, and is directly related to the adopted chip rate, i.e. the spread signal bandwidth. GPS system, for example, relies on this technique, and guarantees an accuracy in TOA estimation of 100 ns, corresponding to an accuracy in the order of meters in distance estimation [61].

The key advantage offered by UWB is thus the ranging accuracy. In fact, errors in the order of centimeters can be guaranteed, much better than the precision achievable by DS-CDMA systems, thanks to a time accuracy of less than 100 picoseconds. This precision is useful in the short range scenarios (tens of meters) expected for UWB networks where positioning is effective only if high accuracy can be achieved.

Ranging information can be exploited in several ways in resource management. Examples are: a) Definition of distance-related metrics for both MAC and higher layers, enabling the development of power-aware protocols, e.g. [62]; b) Evaluation of initial transmission power levels, required in distributed power control protocols [63]; c) Introduction of distributed positioning protocols in order to build a relative network map starting from ranging measurements. This map can enable location-based enhancements in several MAC and network functions, such as position-based routing, and position-aware distributed code assignment protocols in multiple channel MAC, in order to minimize MUI.

3.4 $(UWB)^2$: Uncoordinated, Wireless, Baseborn medium access for Ultra Wide Band communication networks

The analysis carried out in the previous sections focused on the definition of MAC design issues in UWB wireless networks.
First, available solutions were analyzed with respect of requirements of UWB

networks. The areas in which design can benefit from existing solutions and those which, oppositely, require dedicated solutions for UWB were identified. In particular, it was shown that issues related to admission control, packet scheduling and power control can be addressed by adopting similar approaches to those proposed for existing wireless networks. On the other hand, medium sharing and MAC organization require specific design in order to take into account the peculiar characteristics of UWB, such as high processing gain and high synchronization latencies.
Second, the accurate ranging capability was identified as the key feature of Ultra Wide Band that can enable novel MAC functions.

The results of such analysis form the basis for the definition of a MAC protocol suitable for UWB systems that is specifically designed for the special case of low data rate UWB networks. The characteristics of such protocol will be presented in this section.

$(UWB)^2$ takes advantage for data transmission of the multiple access capabilities warranted by the TH Codes, and relies for the access to the common channel on the high MUI robustness provided by the processing gain of UWB. The proposed protocol also takes into account synchronization requirements. $(UWB)^2$ is a multi-channel MAC protocol. Multi-channel access protocols have been widely investigated in the past, since the adoption of multiple channels may significantly increase the achievable throughput [17]. In multi-channel protocols the overall available resource is partitioned into a finite number of elements. Each element of the resource partition corresponds to a channel. According to the definition of resource, a channel can therefore correspond to:

1. A time slot, as in Time Division Multiple Access (TDMA);
2. A frequency band, as in Frequency Division Multiple Access (FDMA);
3. A code, as in Code Division Multiple Access (CDMA).

The design of an UWB MAC may adopt any of the above solutions. As described in section 3.2.3, the IEEE 802.15.3a standard for example proposes a TDMA MAC for UWB [9]. TH-IR UWB, however, provides a straightforward partition of the resource in channels, each channel being associated to a TH code. The design of a multi-channel CDMA MAC protocol forms therefore the natural basis for the design of a MAC in TH-IR UWB. As already mentioned in section 3.2.3, Multi-channel CDMA MAC algorithms, commonly referred to as multi-code, have been intensively investigated for Direct Sequence (DS) CDMA networks. Among all we cite random CDMA access [64], and, more recently, Multi-Code Spread Slotted Aloha [65]. Note, however, that although in the last years most of the research efforts were focused on DS CDMA,

Frequency Hopping (FH) CDMA and TH CDMA also provide viable solutions. Performance of multi-code MAC protocols are limited by two factors:

1. MUI, caused by the contemporary transmission of different packets from different users on different codes;

2. Collisions on the code, caused by the selection of the same code by two different transmitters within radio coverage.

Robustness of the system to MUI is determined by the cross correlation properties of the codes; The lower the cross correlation between different codes, the higher the number of possible simultaneous transmissions. The effect of code collisions can be mitigated by adopting appropriate code selection protocols. The task of assigning codes to different transmitters in the same coverage area is a challenging issue in the design of distributed networks. Within this framework, Sousa and Silvester ([66]) provided a thorough overview of possible code assignment solutions:

1. Common code: all terminals share the same code, relying on phase shifts between different links for avoiding code collisions.

2. Receiver code: each terminal has a unique code for receiving, and the transmitter tunes on the code of the intended receiver for transmitting a packet.

3. Transmitter code: each terminal has an unique code for transmitting, and the receiver tunes on the code of the transmitter for receiving a packet.

4. Hybrid: a combination of the above schemes.

The Common code scheme is sort of a limit case for a multi-code protocol, since no real multi-code capability is exploited. If phase shifts are too small, this solution collapses into the single Aloha channel. Note however that, in the case of very low data rate UWB networks, even the Common code can be an appealing solution, since the processing gain guaranteed by the low duty cycle of UWB can provide by itself enough protection from MUI to avoid the additional complexity of multi-code management.

The Receiver code scheme has the main advantage of reducing receiver complexity, since a terminal must only listen to its receiving code. On the other hand, multiple transmissions involving the same receiver may likely result in collisions, since the same code is adopted by all transmitters.

Oppositely, the Transmitter code scheme avoids collisions at the receiver, since each transmitter uses its own code and thus two transmissions directed

to the same receiver use different codes. On the other hand, the adoption of a Transmitter solution requires in principle a receiver to be capable of listening to all possible codes in the network.

Hybrid schemes allow a trade-off between the above conditions. A Hybrid scheme may foresee the use for signaling of either the Receiver or Common code schemes over which the receiver can read the information about the code which will be used for data. A Transmitter code scheme may then be used for data. When the set of codes is limited, however, the Transmitter code scheme may be subject to collisions due to reassignment of the same code. In this case, a code assignment protocol is required for optimizing the use of the limited set of available codes. An example of such a protocol is presented in [67]. The solution proposed in [67] is a distributed assignment protocol for CDMA multihop networks: it guarantees that, if code C is used by terminal T, code C is never selected within a two-hops range from T, avoiding thus the occurrence of collisions.

The $(UWB)^2$ protocol applies the multi-code concept to the specific case of a TH-IR UWB system. $(UWB)^2$ adopts a Hybrid scheme based on the combination of a common control channel, provided by a Common TH code, with dedicated data channels associated to Transmitter TH codes. The adoption of a Hybrid scheme can be motivated as follows:

1. It simplifies the receiver structure, since data transmissions (and corresponding TH codes) are first communicated on the control channel;

2. It provides a common channel for broadcasting. This is a key property for the operation of higher layers protocols. Broadcast messages are for example required for routing and distributed positioning protocols.

Note that the use of a Common code at the beginning of each transmission also allows an easy transition to the adoption of a Common code solution, whenever the bit rate and the offered traffic are low enough to allow the generated MUI noise to be managed in each receiver with the UWB processing gain alone. As regards code assignment, a unique association between MAC ID and Transmitter Code can be obtained by adopting, for example, the algorithm described in [68] which avoids implementing a distributed code assignment protocol.

$(UWB)^2$ is specifically designed for low data rate networks; as a consequence, it does not assume that synchronization between transmitter and receiver is available at the beginning of packet transmission, because clock drifts in each terminal may lead to complete loss of alignment between two devices in the average time between two DATA packets. As a consequence, a

synchronization trailer long enough to guarantee the requested synchronization probability is added to the packet. The length of the trailer depends on current network conditions, and is provided to the MAC by the synchronization logic.
$(UWB)^2$ also exploits the ranging capability offered by UWB. Distance information between transmitter and receiver is in fact collected during control packets exchange. Such information can enable optimizations of several MAC features, and allow the introduction of new functions, such as distributed positioning. Procedures adopted in $(UWB)^2$ for transmitting and receiving packets are described below. The procedures have two main objectives:

1. To exchange information such as the adopted synchronization trailer, i.e. hopping sequence and length;

2. To perform ranging. Since no common time reference is available, a two way handshake is required to collect distance information by estimating the round-trip-time of signals in the air.

In the following it is assumed that, at each terminal T, MAC Protocol Data Units (MACPDUs) resulting from the segmentation/concatenation of MAC Service Data Units (MACSDUs) are stored in a transmit queue. The segmentation/concatenation block is also in charge of determining the amount of error protection to be added to each PDU by means of a PDU trailer.
It is also assumed that T is able to determine how many MACPDUs in the queue are directed to a given receiver R.

3.4.1 Transmission procedure

Figure 3.5 contains the flow chart of the transmission procedure.

Terminal T periodically checks the status of the transmit queue. Detection of one or more MACPDUs triggers the transmission procedure which can be described as follows:

1. The ID of the intended receiver R is extracted from the first PDU in the queue;

2. T determines the number N_{PDU} of MACPDUs in the queue directed to R;

3. T checks if other MACPDUs were sent to R in the last T_{ACTIVE} seconds. If this is the case, T considers R as an Active receiver, and moves to step 5 of the procedure;

4. If R is not an Active receiver, T generates a Link Establish (LE) PDU. The LE PDU, shown in fig. 3.6, is composed by the following fields:

- SyncTrailer - Used for synchronization purposes
- TxNodeID - The MAC ID of transmitter T
- RxNodeID - The MAC ID of receiver R
- TH_{Flag} - This flag is set to true if the standard TH code associated to TxNodeID will be adopted for transmission of DATA PDUs. The flag is set to false if a different TH code is going to be adopted.
- TH Code (optional) - If the TH_{Flag} is set to false, the information on the TH-code to be adopted is provided in this field.
- FEC/CRC - Bits for error correction/revelation

5. Terminal T sends the LE PDU and waits for a Link Confirm (LC) response PDU from R.

6. If the LC PDU is not received within a time T_{LC}, the LE PDU is retransmitted for a maximum of N_{LC} times, before the transmission of the MACPDU is assumed to be failed.

7. After receiving the LC PDU, T switches to the TH code declared in the LE PDU and transmits the DATA PDU. The DATA PDU, shown in fig. 3.7, is composed of the following fields:

   - SyncTrailer - Used for synchronization purposes
   - Header, including the fields TxNodeID, RxNodeID, PDU_{Number} and N_{PDU}
   - ACK-flag - Used to inform the receiver R if an ACK PDU should be sent in order to inform the transmitter T on the result of the transmission
   - Payload - Containing data information
   - FEC/CRC - Bits for error correction/revelation

8. Once the transmission is completed, T checks again the status of the data queue, and repeats the procedure until all MACPDUs in the transmit queue are served.

When the ACK-flag field is set to 1 in the DATA PDU, the transmitter expects an ACK PDU to be sent by the receiver, in order to schedule a retransmission of a packet if its reception was corrupted by noise or interference, following a predefined backoff scheme. The effect of the selected backoff scheme on performance will be analyzed in section 3.5, where an evaluation of $(UWB)^2$ performance will be presented. As regards the transmission of the ACK PDU, two solutions are possible: either the receiver R transmits such

PDU on the Common TH-Code, or it transmits the ACK PDU on a Receiver specific TH-Code, at the price of an additional overhead required for communicating such code to the transmitter T in the case such code cannot be derived from the MAC ID of R.

Note furthermore that when the MACSDU is constituted by a broadcast packet (e.g. a routing control packet), the MAC will adopt a simplified transmission procedure, where the DATA PDU that encapsulates the MACSDU is directly transmitted on the Common TH-Code, without performing the LE/LC exchange. The broadcast nature of such PDU would in fact make impossible the reception of a LC PDU by all interested terminals. Furthermore, for this kind of PDUs the ACK-flag will be automatically set to 0 in order to avoid the transmission of several ACK PDUs by each neighbor of T receiving the broadcast PDU. A broadcast ID known to all terminals is set as Receiver ID in these PDUs in order to inform neighbors of the broadcast nature of the transmission.
Such simplified procedure guarantees of course a lower protection of broadcast PDUs from interference; on the other hand, it makes possible for the upper layers to have a straightforward mean to communicate broadcast information. Furthermore, the potential loss of a control broadcast packet is usually much less critical than the loss of a DATA packet since updated control information is usually retransmitted either on a periodic basis or within a short time.

3.4.2 Reception procedure

Figure 3.8 contains the flow chart of the reception procedure.

A terminal R in Idle state listens to the Common TH code, indicated as TH-0. When a SyncTrailer is detected, R performs the following procedure.

1. R checks the RxNodeID field. If the value in the field is neither the MAC ID of R nor the broadcast ID, the reception is aborted and the reception procedure ends;

2. Since in the following we are not considering broadcast packets, let us assume that the RxNodeID contains the MAC ID of R. In this case, since R is assumed in Idle state, MACPDUs directed to this terminal will necessarily be LE PDUs;

3. Following the reception of a LE packet, R creates a LC PDU, shown in fig. 3.9. The LC PDU is structured as follows:

   - SyncTrailer - Used for synchronization purposes
   - TxNodeID - The MAC ID of T
   - RxNodeID - The MAC ID of R

- FEC/CRC - Bits for error correction/revelation

4. R sends the LC PDU and moves in the Active state, listening on the TH code indicated in the LE PDU. If no DATA PDU is received within a time T_{DATA} the receiver falls back to Idle state and the procedure ends.

5. When a DATA PDU is received, R processes the payload, and extracts N_{PDU} from the header. If the ACK-flag is set to 1, R generates and sends an ACK PDU with the structure presented in fig. 3.10 reporting the status of the transmission. Next, if $N_{PDU} > 0$, R remains in Active state, since at least N_{PDU} more DATA PDUs are expected to be received from T. If $N_{PDU} = 0$, R goes back to the Idle state.

It should be noted that the above procedures are related to the setup of a single link. During the reception procedure for example R also keeps on listening to the common code. It is assumed in fact that a terminal can act as a receiver on one or more links while acting as a transmitter on another link.

Finally, note that the exchange of LE/LC PDUs can also be triggered on a periodic basis for the purpose of updating distance information. This is likely to be the case for example if a distributed positioning protocol is adopted which relies on up-to-date distance estimations to build a network map.

3.5 Simulation results

The performance of the $(UWB)^2$ was analyzed by means of simulations in order to evaluate its behavior in terms of throughput and delay.
The simulation scenario consisted in N terminals, randomly located in an area of 80x80 m^2 size. Each terminal was characterized by a radio transmission range of 120 meters in order to guarantee almost full connectivity between terminals. Each terminal generated MACPDUs to other terminals in the network following a Poisson process characterized by an average inter-arrival time T_{PDU}. The size of each MACPDU, with the format reported in fig. 3.7, was set to L = 2000 bits. As regards UWB physical layer parameters, the pulse rate was set to $1/T_s = 10^6$ pulses/sec, $N_s = 1$, and $T_M = 1$ ns. In the simulations we assumed all terminals to adopt the same synchronization sequence of length $L_{sync} = 100$ pulses. Performance of the $(UWB)^2$ protocol was evaluated for a number of terminals N varying between 25 and 50, and for T_{PDU} values in the interval [1.25, 0.039063] secs, corresponding to data rates between 1600 bits/s and 51200 bits/s, respectively.
Following the approach adopted in section 2.5, no correction capability was considered, and we assumed that all bits in a packet to be correct, for a packet to be correct. During simulations, a real-time evaluation of P_U rather than the average value P_U of eq. (2.5) was adopted for computing the probability of

pulse collision. As a consequence, the Packet Error Probability was evaluated as follows:

$$PEP = 1 - \prod_{i=0}^{L-1} \left(1 - Prob_{BitError}\left(i\right)\right) \tag{3.1}$$

where $Prob_{BitError}\left(i\right)$ is the error probability for the i-th bit in the packet, as defined in eq. (2.7).

As already stated, two performance indicators were considered. Throughput, defined as the ratio between received MACPDUs and transmitted MACPDUs, and delay, both evaluated in presence of retransmissions, i.e. with the ACK-flag set to 1 in all DATA PDUs.
Note that all results presented in the following take into account the control traffic consisting in the LE/LC PDUs exchanged to setup DATA PDU transfers and perform ranging.
As anticipated in section 3.4.1, retransmissions are scheduled by a transmitter following a backoff algorithm. In evaluating the performance of the $(UWB)^2$ protocol, two different backoff algorithms were considered:

Immediate retransmit - In this algorithm retransmissions are performed as soon as the information of the transmission error is sent back by means of the ACK packet.

Binary Exponential Backoff (BEB) - In this algorithm retransmissions are performed after a random delay. The average delay before attempting a retransmission for the r-th time is equal to N_r times the transmission time of a DATA PDU; the value of N_r is randomly extracted in the interval $[2^0, 2^{min(r,r_{max})}]$. In our simulations, we chose $r_{max} = 10$.

The introduction of a random element in the retransmission policy avoids the problem of systematic collisions that would occur when two devices collide and keep on re-scheduling the transmission of colliding packets at the same time. It should be noted however that in the traffic scenarios considered for Low Data Rate UWB networks, the event of collision is expected to be quite rare; furthermore, a collision between two PDUs P1 and P2 will be destructive, i.e. it will lead for example to corruption of PDU P1, only when the power of the colliding PDU P2 is sufficiently high to overcome the MUI resiliance at the intended receiver of P1 guaranteed by the high processing gain of the UWB signal. As a consequence, in most cases PDU P2 will be received correctly, since it is characterized by a higher power level, and will not hinder the correct reception of P1 retransmission, even if it is retransmitted immediately after

the reception of a negative ACK.
This motivated the idea of comparing the standard BEB algorithm with the immediate retransmission of corrupted PDUs.
The measured values for throughput and delay are presented in figs. 3.11 and 3.12, respectively.

Figure 3.11 shows that measured throughput was higher than 0.985 in all simulation cases. Furthermore, the two backoff schemes considered led to comparable values in all simulations, highlighting the fact that most of PDUs collisions are not destructive thanks to the MUI resiliance guaranteed by UWB.
This conclusion is confirmed by fig. 3.12, showing that the average delay is only slightly increased as the number of offered packets increase, and is in all cases close to the minimum value given by the transmission time of a MACPDU at the bit rate of 1 Mb/s. Furthermore, the adoption of the Binary Exponential Backoff scheme led to higher delays, since in the rare cases where a destructive collision occur, transmitters are forced to wait in average a longer time before attempting a retransmission.
Noticeably, the simulation results are in good agreement with theoretical results obtained in [64] and [69] for Spread Spectrum Aloha networks based on Direct Sequence. Although in fact a direct comparison of theoretical results and data obtained by simulations is not possible due to the hypothesis of a slotted time axis assumed in [69], the values of throughput and delay predicted by the theory for a processing gain of 30 dB are close to the results obtained in our simulation where the duty cycle of the signal was set to $T_M/T_s = 10^{-3}$, corresponding to approximately 30 dB of processing gain.

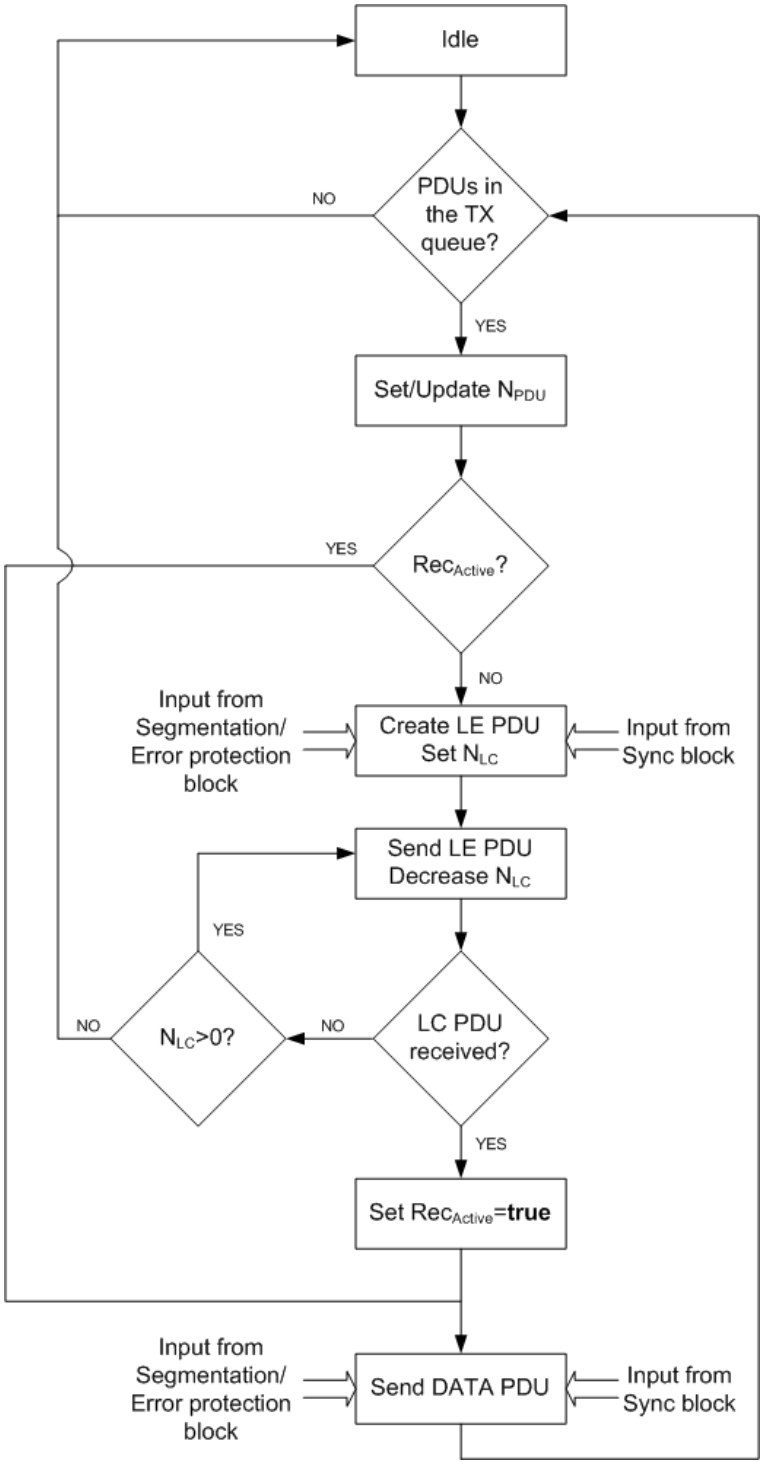

Figure 3.5: Transmission procedure in $(UWB)^2$.

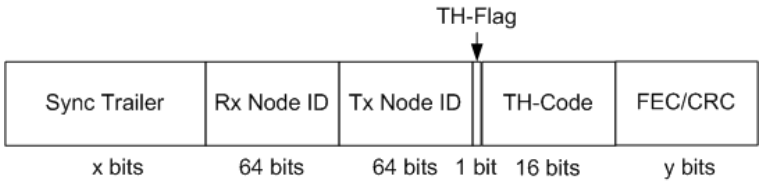

Figure 3.6: Structure of the Link Establish PDU in $(UWB)^2$.

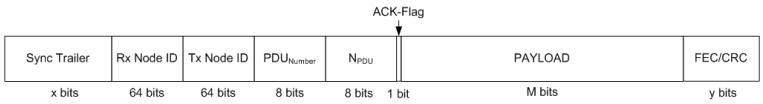

Figure 3.7: Structure of the DATA PDU in $(UWB)^2$.

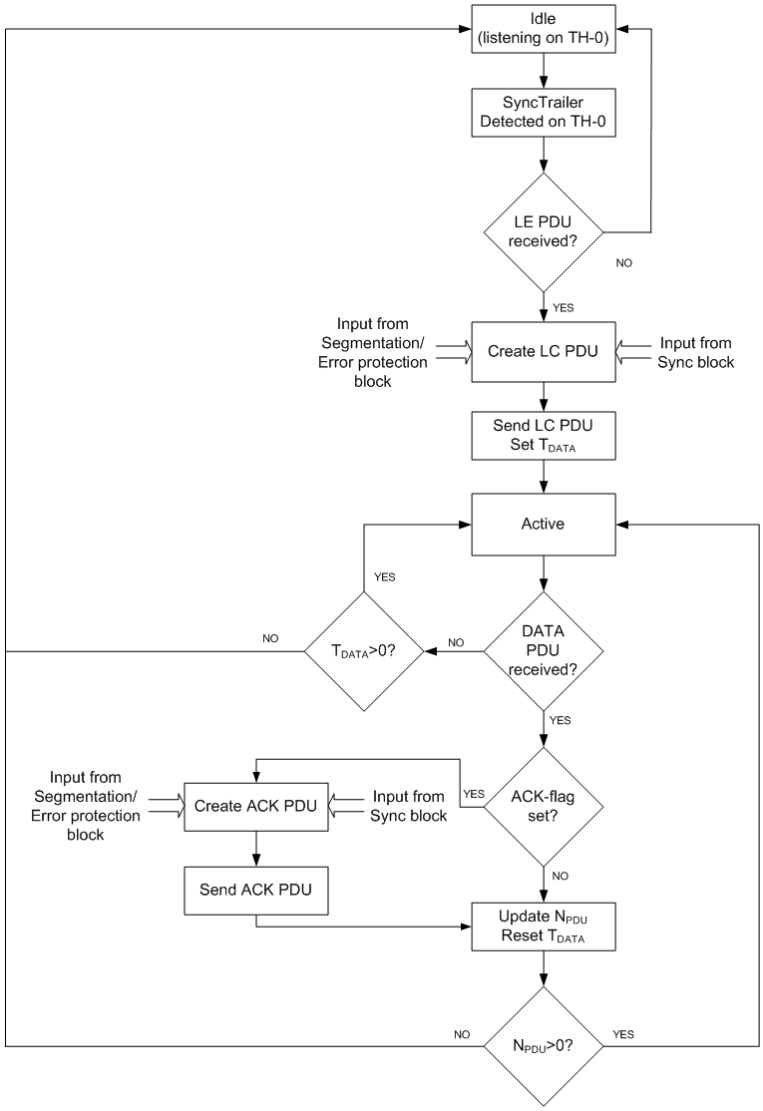

Figure 3.8: Reception procedure in $(UWB)^2$.

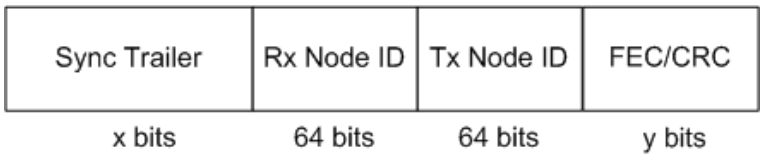

Figure 3.9: Structure of the Link Confirm PDU in $(UWB)^2$.

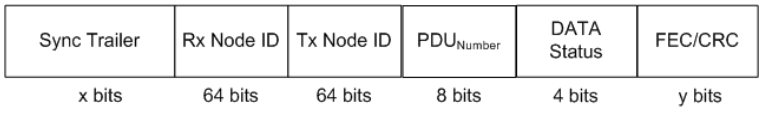

Figure 3.10: Structure of the ACK PDU in $(UWB)^2$.

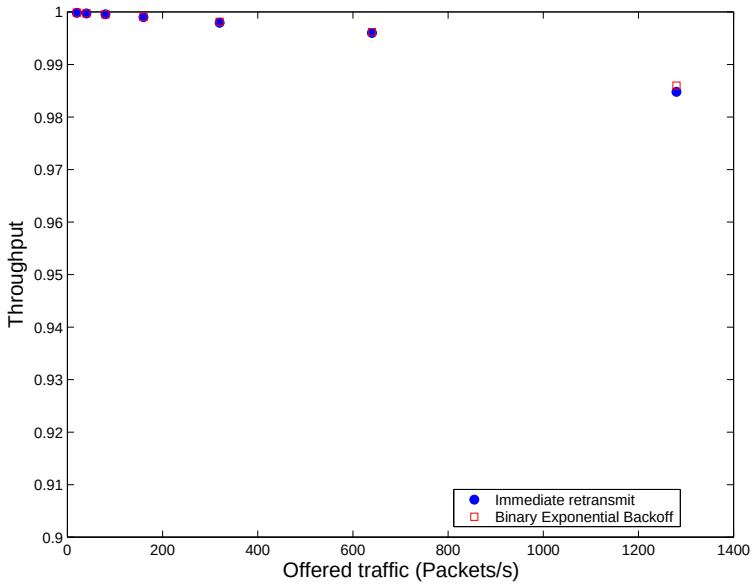

Figure 3.11: Throughput as a function of the offered traffic expressed in Packets/s (Empty squares: Binary Exponential Backoff scheme, Filled circles: Immediate retransmit scheme).

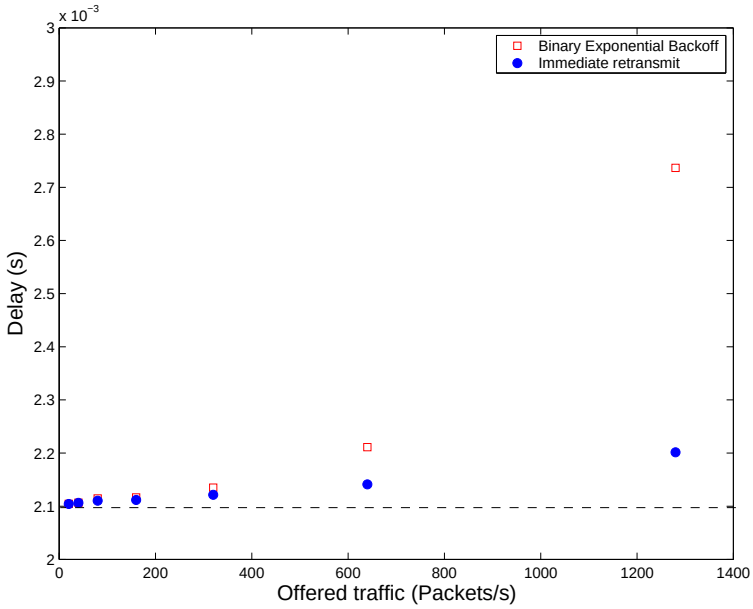

Figure 3.12: Delay as a function of the offered traffic expressed in Packets/s - The dashed line shows the delay due to transmission time of a MACPDU at the bit rate of 1 Mb/s (Empty squares: Binary Exponential Backoff scheme, Filled circles: Immediate retransmit scheme).

Chapter 4

Positioning in ad-hoc wireless networks

The terms ranging and positioning, as well as localization, are used in current literature in a very flexible manner, i.e. there is no common understanding or convention on the exact meaning of each of these terms. It is important thus to establish a set of definitions which will be used from this moment on regarding, in particular, ranging, node-centered positioning, relative positioning, and absolute or geographical positioning.

Ranging is defined as the action of computing distance of a target node from a reference node. A reference node wanting to obtain ranging information regarding a target node in the network can acquire this information by establishing a peer-to-peer communication link with the target node. This communication link is used for evaluating parameter values which are used at the reference node for estimating its distance from the target node. The parameters are typically based either on the evaluation of channel attenuation or on delay of propagation, as will be further described in section 4.1. Figure 4.1 shows an example of ranging, and introduces the following notation: assuming for example $N2$ as the reference node, we indicate the ranging information for target nodes $N1$ and $N3$ as $RANG_{N2}(N1)$, $RANG_{N2}(N3)$.

Node-centered positioning is defined as the action of computing the positions of a set of target nodes with respect to a reference node. Node-centered positioning can also be obtained based on peer-to-peer connections, provided that both distance and angle information about each target node is obtainable at the reference node. Note that any node can play the role of reference node and compute the position of other nodes in its own reference system, but that each node associates a different set

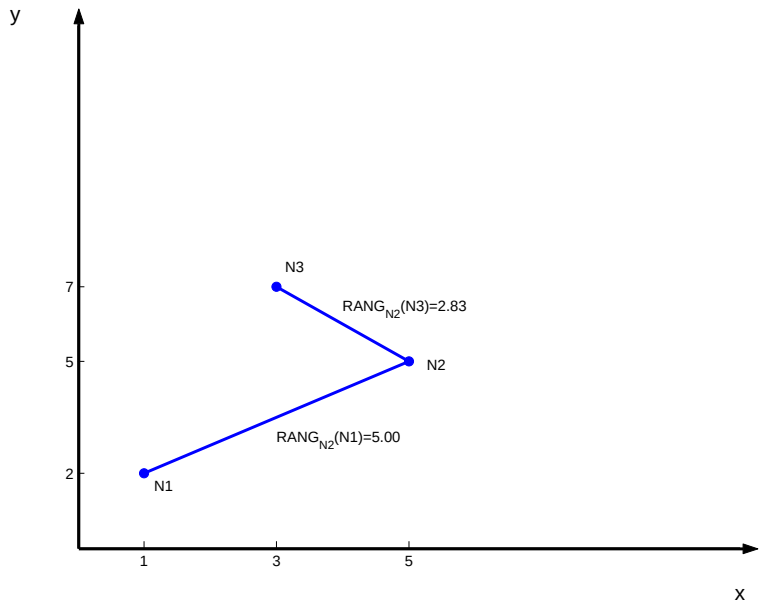

Figure 4.1: Example of ranging with node $N2$ as reference node.

of coordinates to the same target node. Figure 4.2 shows an example of node-centered positioning, with two different reference nodes.

The following notation is used in fig. 4.2: assuming for example $N2$ as reference node, we indicate the node-centered positioning information related to target nodes $N1$ and $N3$ as $POS_{N2}(N1)$, $POS_{N2}(N3)$.

Relative positioning indicates the action of computing the position of a set of nodes with respect to a common system of coordinates. The key difference with node-centered positioning is that all nodes share the same reference system and thus each node is unequivocally associated with a unique set of coordinates. The adoption of the same coordinate system in all nodes requires to organize the set of nodes into a network allowing information exchange among more than two nodes, following the rules set by dedicated algorithms for the selection of the common coordinate system, and the translation of the coordinates of each node, as it will be further explained in sections 4.3.1 and 4.3.2. Figure 4.3 gives an example of relative positioning.

The origin of the reference coordinate system can be chosen arbitrarily, and may not coincide with the position of a node in the network. In most cases, however, the reference coordinate system is obtained starting from

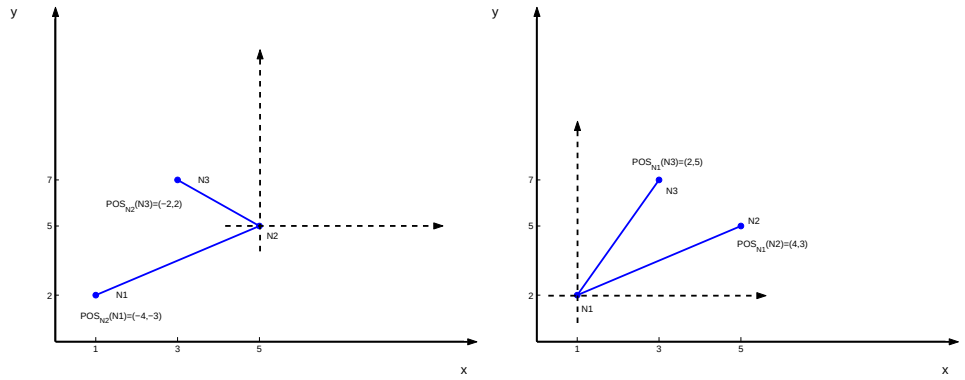

Figure 4.2: Example of node-centered positioning, with $N1$ and $N2$ as reference node on the left and right plot, respectively.

the node-centered coordinate system. As an example, in fig. 4.3 the reference coordinate system coincides with the node-centered coordinate system of node N1, which assumes coordinates $POS(N1) = (0, 0)$. A special case of relative positioning corresponds to the adoption of a reference coordinate system coincident with the global coordinate system, with coordinates given in terms of latitude and longitude; This specific solution is referred to as **absolute (or geographical) positioning**, since the coordinates associated to each node are unique worldwide.

Both node-centered and relative positioning require a prior ranging procedure for retrieving distances. The degree of accuracy in distance estimation has thus an impact on the accuracy with which positioning can be achieved. The distance estimation technique must be then selected according to requirements imposed by the application layer.

4.1 Ranging

Ranging consists in estimating the distance D between a transmitter and a receiver. Given a transmitted signal $s(t)$, the corresponding received signal writes:

$$r(t) = h(t) * s(t) + n\,(t) \tag{4.1}$$

where $h(t)$ is the channel impulse response and $n(t)$ is thermal noise. Let us assume for now that the signal propagates over a perfect channel, characterized by an impulse response given by:

$$h\,(t) = A\,(D)\,\delta\,(t - \tau\,(D)) \tag{4.2}$$

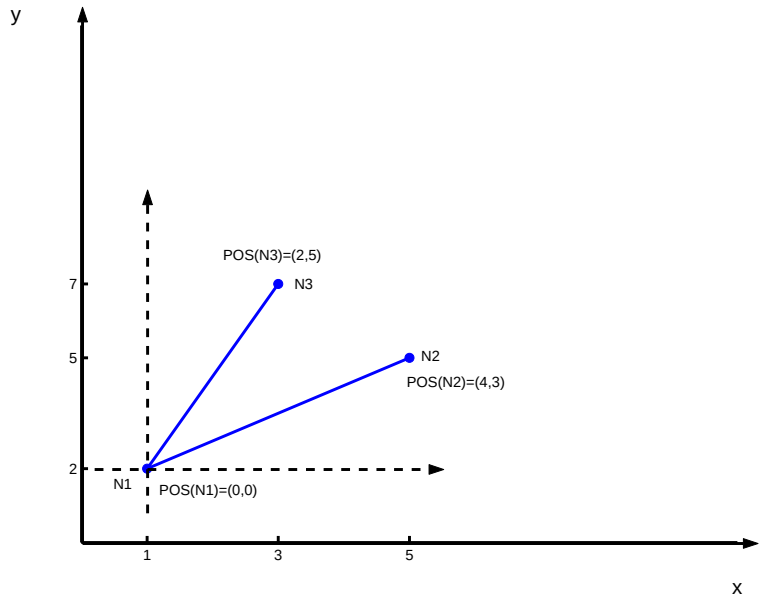

Figure 4.3: Example of relative positioning with a reference coordinate system centered in $N1$.

The received signal can be written as follows:

$$r(t) = A(D)\, s(t - \tau(D)) + n(t) \tag{4.3}$$

Equation 4.3 indicates that distance D can be estimated from either attenuation $A(D)$ or delay $\tau(D)$. The use of $A(D)$ vs. $\tau(D)$ determines the method for estimating distance: Received Signal Strength Indicator (RSSI) vs. Time Of Arrival (TOA).

RSSI is based on the emission at the transmitter side of a signal using a fixed reference power which is known at the receiver. The receiver measures the power of the received signal, and derives distance from measured attenuation. Since the relation between distance and attenuation depends on channel behaviour, an accurate propagation model is required in order to reliably estimate distance. Terminal mobility and unpredictable variations in channel behaviour can be a real problem [70]. An additional source of error with RSSI is introduced by the limitations of the hardware. The transmitter, in particular, must be finely tuned on signal emissions at the reference power; A discrepancy between transmitted and reference power reflects in a systematic bias in distance estimation

[71]. In conclusion, RSSI is not a very accurate method, and its adoption is confined to applications requiring coarse ranging.

TOA computes distance based on the estimation of the propagation delay between transmitter and receiver. TOA is the most commonly used distance estimation method in the radar field, and for this reason the terms TOA and ranging are often interchanged. Delay estimation is a key topic in wireless communications, since it is required for achieving symbol synchronization between transmitter and receiver, and several solutions to this problem are available in the literature. Most of these derive from the Maximum Likelihood (ML) estimator, which is defined as follows.
Under the hypothesis of Additive White Gaussian Noise (AWGN) n(t), the ML estimate of delay τ, called $\hat{\tau}_{ML}$, corresponds to the τ value which minimizes the ML function. The $\hat{\tau}_{ML}$ can be expressed as follows [72]:

$$\hat{\tau}_{ML}(r) = \underset{\tau \in \Re}{argmin} \left(e^{-\frac{1}{N_0} \int_{T_{obs}} (r(t) - s(t-\tau))^2 dt} \right) \tag{4.4}$$

where N_0 is the bilateral Power Spectral Density (PSD) of the noise and T_{obs} is the observation interval over which the estimation is performed.

A third way of obtaining distance information is the *Angle of Arrival (AOA)*: in this case the distances between terminals are reconstructed from the angle between them; this technique requires the adoption of arrays of directional antennas, and the results are highly sensitive to multi-path, NLOS effect and array precision.

4.2 Positioning

The ranging procedure provides an estimation of distances between pairs of nodes of a given network. Let us assume that each node Ni knows its distance from all the other nodes. Among these, Ni can choose k reference nodes ($N1$, ..., Nk) to form a reference system, in which estimating its position. How to estimate position from a set of relative distances is the object of this section. In the ideal case of free-of-error distance estimations, the *Spherical positioning* technique, also known as *TOA positioning*, is a viable solution. Here, based on the observation that in a tridimensional space (x, y, z) each distance $RANG_{Nj}(Ni)$ between Ni and the reference Nj determines a sphere of radius $D_{ji} = RANG_{Nj}(Ni)$ centered in Nj, position (X_i, Y_i, Z_i) = $POS(Ni)$ of Ni is determined by the intersection of the k spheres of radii (D_{1i}, ..., D_{ki}) centered in the reference nodes ($N1$, ..., Nk). Since the intersection of

four spheres is required for determining a single point in the tridimensional space, at least four reference nodes are required in tridimensional positioning. Note that the introduction of additional reference nodes is not necessary for position computation under the hypothesis of perfect distance estimation, but proves to improve performance in the non-perfect case, since redundancy may reduce the impact of faulty ranging measurements. The intersection between the k spheres can be computed by solving the following system of equations:

$$\left\{ \begin{array}{c} \sqrt{(X_1 - X_i)^2 + (Y_1 - Y_i)^2 + (Z_1 - Z_i)^2} \\ \sqrt{(X_2 - X_i)^2 + (Y_2 - Y_i)^2 + (Z_2 - Z_i)^2} \\ \cdots \\ \sqrt{(X_k - X_i)^2 + (Y_k - Y_i)^2 + (Z_k - Z_i)^2} \end{array} \right\} = \left\{ \begin{array}{c} D_{1i} \\ D_{2i} \\ \cdots \\ D_{ki} \end{array} \right\} \tag{4.5}$$

with $k \geq 4$. The same approach can be applied in a bidimensional space (x, y) i.e. a plane. In this case the position of node Ni (X_i, Y_i) = $POS(Ni)$ is determined by the intersection of three circles and can be computed by solving the following system of equations:

$$\left\{ \begin{array}{c} \sqrt{(X_1 - X_i)^2 + (Y_1 - Y_i)^2} \\ \sqrt{(X_2 - X_i)^2 + (Y_2 - Y_i)^2} \\ \cdots \\ \sqrt{(X_k - X_i)^2 + (Y_k - Y_i)^2} \end{array} \right\} = \left\{ \begin{array}{c} D_{1i} \\ D_{2i} \\ \cdots \\ D_{ki} \end{array} \right\} \tag{4.6}$$

with $k \geq 4$. Figure 4.4 shows an example of spherical positioning in a bidimensional space.

Spherical positioning can be used only when a common time reference is available to Ni and all reference nodes, that is perfect distance estimation can be obtained. This is unfortunately not the case in many practical situations, where misalignments and clock drifts introduce random delays between clocks. Computing exact positioning is still possible provided that a common time reference is available among at least the k reference nodes. The *Hyperbolic positioning* technique, also known as *Time Difference Of Arrival (TDOA)*, determines the position of Ni based on the difference between times of arrival from the k reference nodes and Ni. Let us assume that the k reference points share a common time reference, and that the clock at Ni is delayed by an unknown time δ with respect to the common time reference. The key point here is that subtraction between times of arrival from different reference nodes removes the delay δ. This can be shown by observing that for a pair (Nn, $N(n-1)$) of reference terminals one has:

$$D_{ni} - D_{(n-1)i} = c\,(\tau_{ni} + \delta) - c\left(\tau_{(n-1)i} + \delta\right) = c\left(\tau_{ni} - \tau_{(n-1)i}\right) \tag{4.7}$$

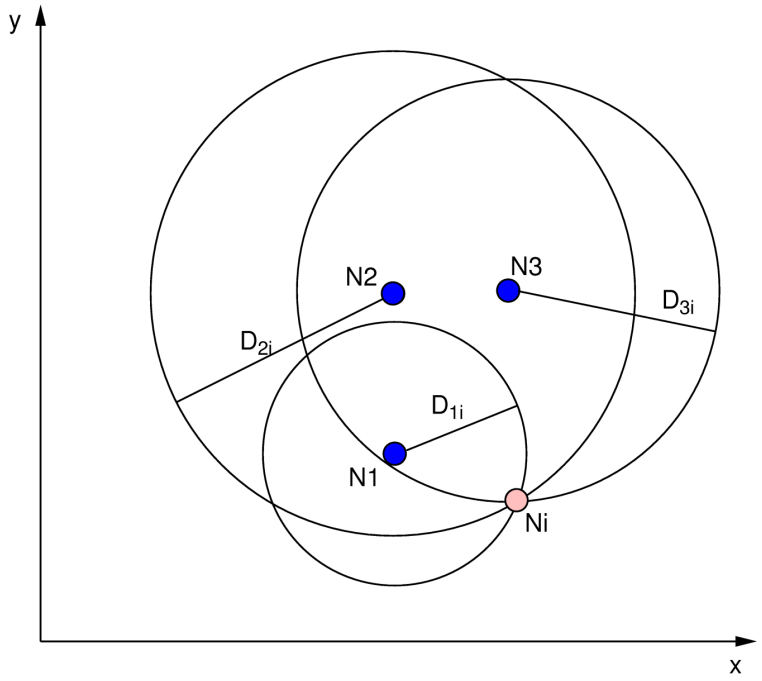

Figure 4.4: Example of spherical positioning of Ni in a bidimensional space with reference nodes $N1$, $N2$, and $N3$.

The position of Ni in a tridimensional space is then determined as the intersection of hyperboloids in space, as described by the following equations:

$$\left\{ \begin{array}{c} \sqrt{(X_2-X_i)^2+(Y_2-Y_i)^2+(Z_2-Z_i)^2}-\sqrt{(X_1-X_i)^2+(Y_1-Y_i)^2+(Z_1-Z_i)^2} \\ \sqrt{(X_3-X_i)^2+(Y_3-Y_i)^2+(Z_3-Z_i)^2}-\sqrt{(X_2-X_i)^2+(Y_2-Y_i)^2+(Z_2-Z_i)^2} \\ \cdots \\ \sqrt{(X_k-X_i)^2+(Y_k-Y_i)^2+(Z_k-Z_i)^2}-\sqrt{(X_{k-1}-X_i)^2+(Y_{k-1}-Y_i)^2+(Z_{k-1}-Z_i)^2} \end{array} \right\} = \left\{ \begin{array}{c} D_{2i}-D_{1i} \\ D_{3i}-D_{2i} \\ \cdots \\ D_{ki}-D_{(k-1)i} \end{array} \right\} \tag{4.8}$$

with $k \geq 4$.

In a bidimensional space, one has:

$$\left\{ \begin{array}{c} \sqrt{(X_2-X_i)^2+(Y_2-Y_i)^2}-\sqrt{(X_1-X_i)^2+(Y_1-Y_i)^2} \\ \sqrt{(X_3-X_i)^2+(Y_3-Y_i)^2}-\sqrt{(X_2-X_i)^2+(Y_2-Y_i)^2} \\ \cdots \\ \sqrt{(X_k-X_i)^2+(Y_k-Y_i)^2}-\sqrt{(X_{k-1}-X_i)^2+(Y_{k-1}-Y_i)^2} \end{array} \right\} = \left\{ \begin{array}{c} D_{2i}-D_{1i} \\ D_{3i}-D_{2i} \\ \cdots \\ D_{ki}-D_{(k-1)i} \end{array} \right\} \tag{4.9}$$

with $k \geq 3$. Figure 4.5 shows an example of hyperbolic positioning in a bidimensional space.

Hyperbolic positioning requires an accurate common time reference between the reference nodes, but does not rely on precise synchronization between reference nodes and target node. It is particularly indicated for networks based on an infrastructure, in which coordination between infrastructure nodes is relatively easy to achieve and maintain. For this reason, this technique was proposed for positioning purposes in cellular networks, where base-stations can play the role of reference nodes and allow a mobile node to derive its own position [73].

The spherical and hyperbolic positioning techniques presented above require an error-free ranging information in order to provide a solution. Thermal noise can however introduce errors in ranging, leading to imperfect estimations of distance between nodes. In this case the analytical solution of eqs. (4.5) and (4.8) may not exist, as shown in fig. 4.6 for spherical positioning. The effect of errors in ranging estimations on the accuracy of positioning can be reduced

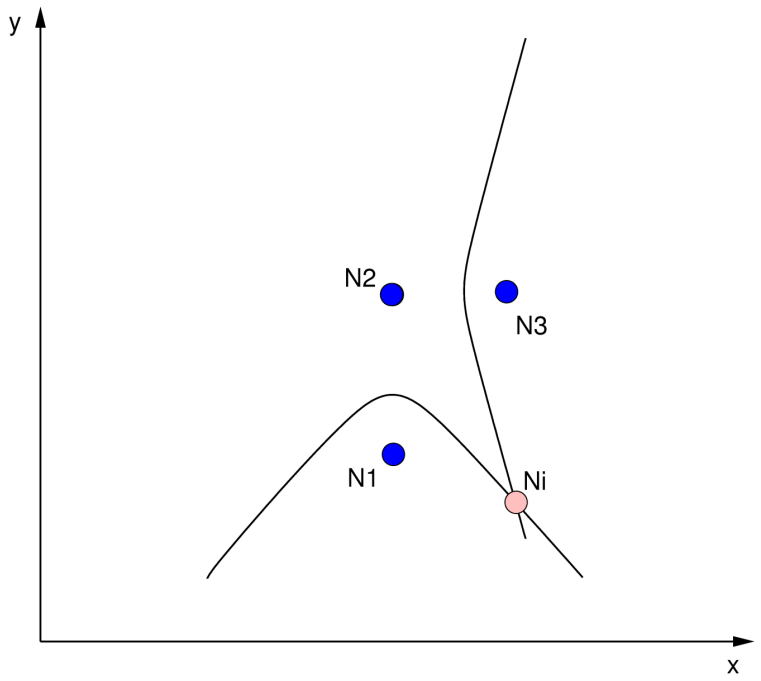

Figure 4.5: Example of hyperbolic positioning of node Ni in a bidimensional space with reference nodes $N1$, $N2$, and $N3$.

by adopting minimization procedures such as the Least Square Error (LSE). For this purpose, following the approach presented in [74] it is convenient to rewrite eq. (4.5) into a set of equations which are linear in $POS(Ni)$, i.e.:

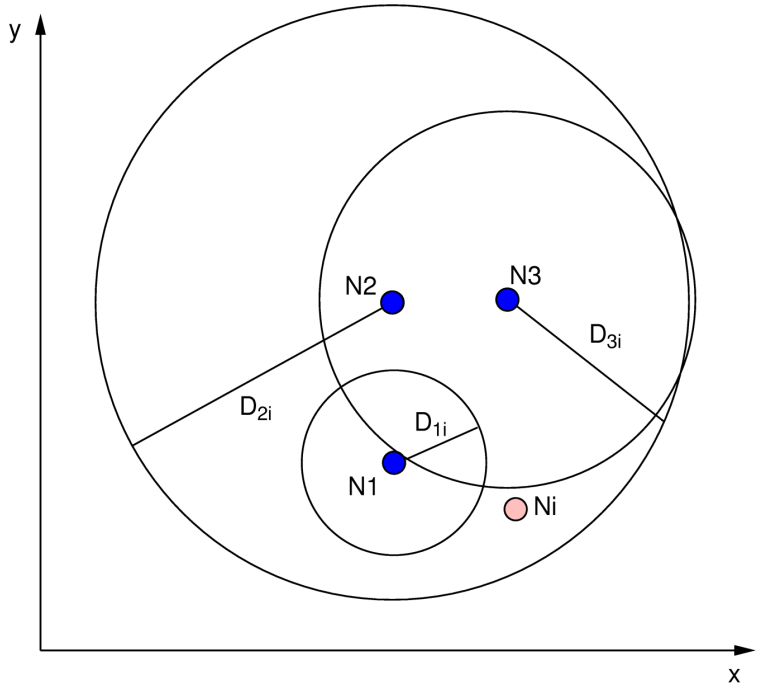

Figure 4.6: Effect of ranging error on spherical positioning of node Ni in a bidimensional space with reference nodes $N1$, $N2$, and $N3$.

$$\bar{A}\bar{I} = \bar{b} \tag{4.10}$$

with:

$$\bar{A} = -2 \begin{bmatrix} (X_1 - X_k) & (Y_1 - Y_k) & (Z_1 - Z_k) \\ (X_2 - X_k) & (Y_2 - Y_k) & (Z_2 - Z_k) \\ \cdots & \cdots & \cdots \\ (X_{k-1} - X_k) & (Y_{k-1} - Y_k) & (Z_{k-1} - Z_k) \end{bmatrix} \tag{4.11}$$

$$\bar{I} = \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} \tag{4.12}$$

and

$$\bar{b} = \begin{bmatrix} D_{1i}^2 - D_{ki}^2 - X_1^2 + X_k^2 - Y_1^2 + Y_k^2 - Z_1^2 + Z_k^2 \\ D_{2i}^2 - D_{ki}^2 - X_2^2 + X_k^2 - Y_2^2 + Y_k^2 - Z_2^2 + Z_k^2 \\ \cdots \\ D_{(k-1)i}^2 - D_{ki}^2 - X_{(k-1)}^2 + X_k^2 - Y_{(k-1)}^2 + Y_k^2 - Z_{(k-1)}^2 + Z_k^2 \end{bmatrix} \quad (4.13)$$

with $k \geq 4$.

The system defined by eq. (4.10) can be then solved in the sense of LSE minimization.

As already anticipated, in the case of a positioning with errors in estimation of ranges, the adoption of a redundant set of range measurements is helpful in reducing the variance of the positioning error. As an example, let us consider the problem of determining the position of a node Nx in a network composed of 10 nodes.

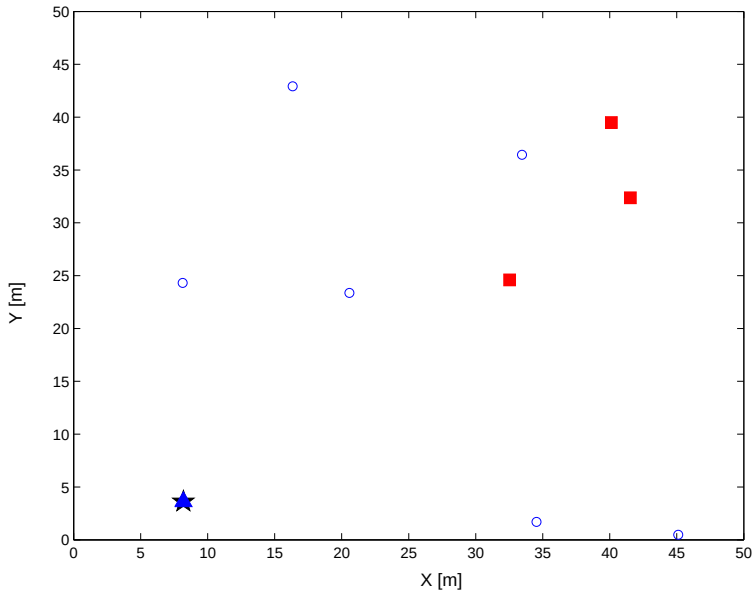

Figure 4.7: Position estimation by means of LSE without ranging errors. Squares: reference nodes; Circles: other nodes; Star: estimated and effective position of target node.

Let us assume that each ranging measurement is affected by an error modeled as a Gaussian process with 0 mean and variance σ^2. If we consider the ideal case in which $\sigma^2 = 0$, k = 3 reference nodes are enough in order to determine the position of Nx without error, as shown in fig. 4.7.

Let us analyze the effect of errors in the range measurements.

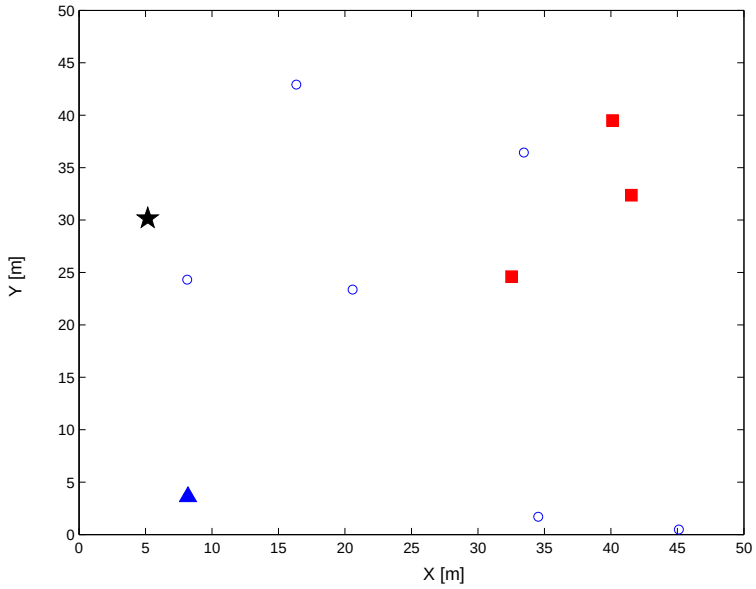

Figure 4.8: Position estimation by means of LSE with ranging errors, $\sigma^2 = 5$. Squares: reference nodes; Circles: other nodes; Star: estimated position of target node; Triangle: effective position of target node.

Figure 4.8 shows the results obtained with the same set of nodes when an error with variance $\sigma^2 = 5$ is introduced. The effective position of target node is represented by a triangle.

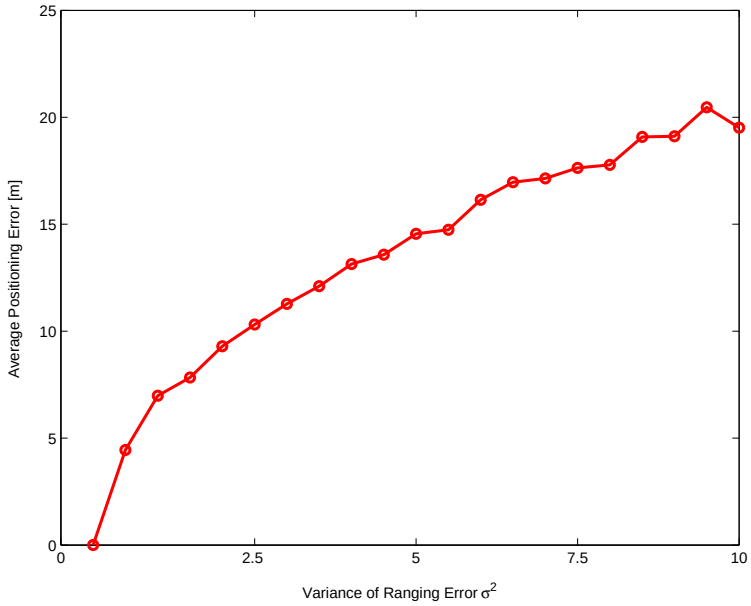

Figure 4.9: Average positioning error as a function of the variance of ranging errors.

Figure 4.8 shows that the introduction of a ranging error caused a positioning error of about 25 m. It can be expected that, as the ranging error increases, the positioning error increases too: fig. 4.9, obtained for k = 3 and σ^2 varying between 0 and 10 with steps of 0.5, for 10000 averages for each σ^2 value, shows indeed that the higher the ranging error (expressed by its variance), the higher the positioning error (expressed as the distance between estimated position and real position).

As already anticipated, however, the error in position estimation can be reduced by introducing redundancy in the number of estimated distances. The redundancy can be introduced by increasing the number of reference nodes k, leading thus to an over-determined minimization problem. The effect of an increased k on the positioning error is shown in fig. 4.10 for $\sigma^2 = 5$.

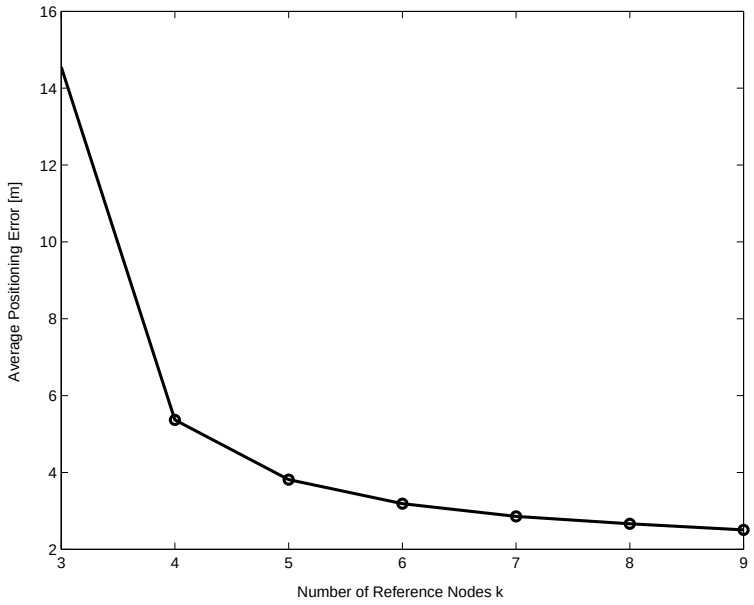

Figure 4.10: Average positioning error as a function of the number of reference nodes k, for variance of ranging errors $\sigma^2 = 5$.

4.3 Positioning in wireless mobile networks

Several solutions are available in literature to the problem of positioning in wireless networks formed by mobile terminals. Positioning protocols typically use one of the approaches described in the previous sections for ranging and positioning, and include procedures for spreading distance and position information throughout the network. Positioning protocols can be subdivided in two categories:

1. GPS-based positioning protocols;
2. GPS-free positioning protocols.

The GPS-based approach offers several advantages:

- Absolute positioning;
- Reduced signalling traffic, as only positions of the terminals have to be exchanged in the network, and no other information is needed to build a precise network topology;
- Synchronized clock in all terminals.

On the other hand, a GPS-free approach can solve some of the issues posed by GPS-enabled terminals:

- Capability to operate in indoor environments, in which GPS hardware is useless;
- Better miniaturization, which is not addressable for GPS devices due to the antenna size.

This approach, however, requires the exchange of more information between terminals (which are not aware of their own spatial position and only have measurements of distances from their neighbors) and a dedicated algorithm for building a reliable network topology based on the distance information. Such a topology will be obviously formed by relative positions of the nodes in the network; this knowledge is however sufficient for applying location-based routing and resource management optimizations. A key issue in the adoption of this class of protocols is the availability of as many accurate measurements of distance between terminals as possible; these measurements can be obtained by the terminals in different ways, as it was explained in section 4.1.
Once that the range measurements between n reference terminals and each terminal i with unknown position are available, the topology is obtained by solving the trilateration problem, described by eqs. (4.5) and (4.6) for tridimensional and bidimensional positioning, respectively.

In the following sections we will analyze the possible solutions to the positioning problem in an ad-hoc scenario: as a consequence, we will focus on distributed algorithms. In scenarios where ranging data can be concentrated in a single terminal, however, centralized algorithms are a valid alternative: as an example, in [75] a GPS-based centralized solution is proposed, which formulates a linear problem basing on a few nodes with known position and the constraints (physical connections available) between all nodes in the network, and solves this problem in a centralized terminal which collects all information from other nodes.

4.3.1 GPS-based positioning protocols

The PicoRadio Approach

In the framework of the PicoRadio project a protocol for positioning information has been proposed ([76], [74]) which relies on GPS only in a low percentage of network terminals (as an extreme case only four, in order to get a tridimensional topology), which act as anchor nodes and send through the network their absolute position.
At local level (i.e. assuming full connectivity in the network) the solution of the trilateration problem through a LSE approach is proposed, using redundant

range measurements in order to mitigate the effect of the error introduced by the RSSI measurement technique adopted in the protocol. Simulations show that the adoption of many (up to 35) independent range measurements with mean error of 50% of the transmission range of each terminal leads to a position estimation with low (variance ≤ 0.02) error.

At the same time, a simple Topology Discovery procedure is proposed, which is called Assumption Based Coordinates (ABC) that iteratively builds the network topology and works as follows. Consider four terminals T_0, T_1, T_2 and T_3 and suppose known the relative distances $d_{T_1T_0}$, $d_{T_2T_0}$, $d_{T_3T_0}$, $d_{T_3T_1}$, and $d_{T_3T_2}$. Assuming T_0 as a reference (for example because it is an anchor node) we set $X_{T_0} = 0$, $Y_{T_0} = 0$ and $Z_{T_0} = 0$; next, we choose the line passing through T_1 and T_0 as x axis by setting $X_{T_1} = d_{T_1T_0}$, $Y_{T_1} = 0$ and $Z_{T_1} = 0$. If we set $Z_{T_2} = 0$, thus choosing the plane including T_0, T_1 and T_2 as the $x - y$ plane, the coordinates of terminal T_2 are then given by:

$$\begin{aligned} X_{T_2} &= \frac{d^2_{T_0T_1} + d^2_{T_0T_2} - d^2_{T_1T_2}}{2 \cdot d_{T_0T_1}} \\ Y_{T_2} &= \sqrt{d^2_{T_0T_2} - x_2^2} \end{aligned} \tag{4.14}$$

with the arbitrary choice $Y_{T_2} > 0$ (that imposes a direction to the y axis). In a similar way we obtain for T_3:

$$\begin{aligned} X_{T_3} &= \frac{d^2_{T_0T_1} + d^2_{T_0T_3} - d^2_{T_1T_3}}{2 \cdot d_{T_0T_1}} \\ Y_{T_3} &= \frac{d^2_{T_0T_3} - d^2_{T_2T_3} + X^2_{T_2} + Y^2_{T_2} - 2 \cdot X_{T_2} \cdot X_{T_3}}{2 \cdot Y_{T_2}} \\ Z_{T_3} &= \sqrt{d^2_{T_0T_3} - X^2_{T_3} - Y^2_{T_3}} \end{aligned} \tag{4.15}$$

assuming $Z_{T_3} > 0$ (i.e. imposing a direction to the z axis and supposing T_3 not belonging to the $x - y$ plane).
Note that if T_0 is an anchor node, this algorithm leads to a correct reconstruction of relative network topology but with random orientation with respect to the global system of coordinates.
When no global connectivity in the network is available, not all nodes are able to get a direct range measurement from four anchor nodes, and a Cooperative ranging protocol is proposed to solve this problem. In the protocol the global positioning problem is subdivided in many local positioning procedures in which every node plays the same role, executing the following cycle of functions:

1. Receive ranging and localization information from neighbors
2. Solve the local positioning problem
3. Send the results to all its neighbors

The Cooperative ranging is divided in two phases: start-up and update. Two alternative algorithms are proposed for the start-up phase: Global Topology Discovery and TERRAIN.

Global Topology Discovery - This solution is based on the ABC algorithm: each terminal assumes to be the center of the coordinates, and starts a local topology discovery; the anchor nodes do the same, but at the same time they force all neighboring nodes to linearly transform (transposing and rotating it) their system of coordinates in order to map it on the system started by the anchor. This information is propagated through the network, each anchor node removing a degree of freedom, until the entire network converges to a unique global system of coordinates. The authors report that this algorithm was tested by simulation, but it showed a low robustness to range measurements errors, leading to large positional errors [76].

TERRAIN - The TERRAIN (Triangulation via Extended Range and Redundant Association of Intermediate Nodes) algorithm also relies on the ABC algorithm. In TERRAIN, however, ABC is only started by the anchor nodes: a terminal that is not aware of its position does not start any local topology discovery but, oppositely, it waits for the information from four different anchor nodes to reach it and then performs a simple trilateration. Note that in this way the system of coordinates adopted by each node is the same (because it is the global one offered by anchor nodes), and so no linear transformation is requested.

Once that an initial topology is obtained through either the Global Topology Discovery or the TERRAIN algorithm, it is updated and refined by iteratively repeating local trilaterations in which each node periodically recalculates its own coordinates using the most recent information from its neighbors.

The protocol was designed for sensor networks, characterized by limited computational resources, limited power and slow/no mobility: for this reason it is structured in such a way to require really simple computations in each node, but at the same time it is based on an iterative refinement of position estimates which is only compatible with a really low mobility scenario. It is important to note that this protocol does not guarantee the convergence of the solution in any case: depending on the initial position estimates and on the range measurement errors the solution for one or more nodes may diverge

leading to large positioning errors. In particular, simulations performed by the authors showed that the problems of convergence of the algorithm are amplified by the choice of RSSI as the technique for ranging.

Ad Hoc Positioning System (APS)

The APS protocol ([77]), like the PicoRadio approach described in the previous subsection, relies on a set of GPS-enabled terminals (called *landmarks*) for achieving a global positioning system. The APS protocol addresses the general case of no full connectivity, and proposes three methods to propagate the information given by the landmarks through the network, in order to allow also nodes which are not directly connected to any GPS terminal to infer their distances from three or more landmarks and, as a consequence, to evaluate their position. The approach proposed is similar to the one used in distance vector routing protocols: each node communicates with its one-hop neighbors, exchanging distance estimates to landmarks and measuring one-hop distances by means of RSSI; the information on global position starts from the landmarks and propagates from them with a controlled flooding. The three proposed propagation methods are the following.

DV-Hop - This is a pure distance vector approach: each node maintains a table containing the coordinates (Xi, Yi) and the distance in hops h_i for each landmark i. The tables are exchanged hop-by-hop, until a landmark j receives a table containing a distance in hops by another distance landmark i: at this point j computes the mean distance covered by each hop, and propagates this information as a correction factor for non-GPS-enabled nodes. The protocol tries to propagate the corrections only in the direction from which the table arrived, in order to take into account also network anisotropies which may lead to different distances for each hop in different directions. Note that no measurements are used in this solution, and so the DV-Hop scheme is immune from range errors; on the other hand its results may be not reliable in the case of anisotropic networks, in which the node disposition is not homogeneous.

DV-distance - This method is similar to the DV-Hop described above, with the main difference that each hop is measured with RSSI, and each node transmits to its neighbors the distance to all known landmarks, obtained as the sum of the distance covered in each hop. This strategy is obviously more accurate, since it does not assume all hops to have the same length, but it is subject to errors in RSSI measurements.

Euclidean - In this scheme the true Euclidean distance between landmarks is propagated, potentially leading to a more precise distance evaluation

compared to DV-Hop and DV-Distance strategies; the evaluation of the distance is performed as shown in fig. 4.11.
Supposing that A needs to evaluate the distance to L, the distances AB and AC are directly evaluated by A, while distances BL, CL and BC are transmitted to A from B and C; A can evaluate the distance AL because all other distances in the polygon ABCL are known. Note that also the position A' satisfies the conditions, but in most of cases A will be able to find its true position thanks to conditions given by other neighbors. The result is subject to measurement errors, but the landmarks can mitigate the effect of errors by relaying the true, GPS-obtained, distance between them instead of the distance obtained by the other nodes with RSSI.

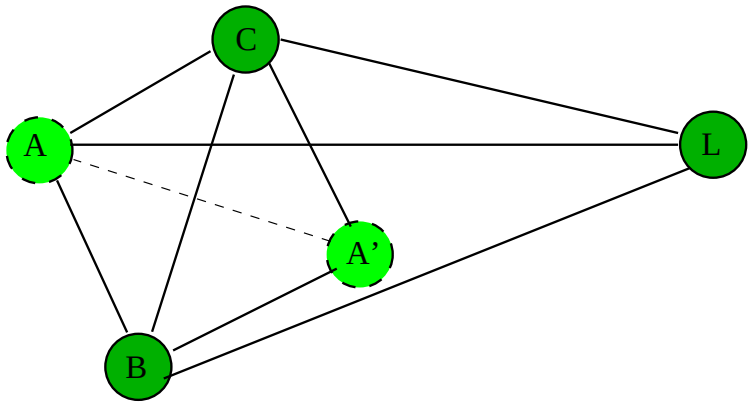

Figure 4.11: Euclidean distance evaluation Node A is able to determine its own position based on distances AB, AC, BL, CL and BC; note that the position represented as A' would satisfy the conditions too, so that additional information is required to solve the ambiguity.

The different schemes were compared by the authors in terms of location error and coverage (i.e. the percentage of nodes which receives information from three different landmarks and is able to perform a position estimation). Simulation results show that the Euclidean scheme shows a higher accuracy in position estimation and higher robustness to range measurement errors, at the cost of lower coverage (especially when few landmarks are available) and higher signalling overhead.

4.3.2 GPS-free positioning protocols

The GPS-based algorithms can provide a solution to the positioning problem when three or more terminals are able to obtain their own position informa-

tion (either via GPS or with position data pre-loaded in the terminal) and can thus act as anchors for all terminals in the network.
On the other hand, in scenarios in which none of the terminals is initially informed of its own position, an alternative approach is required. This is the case for example of a network of sensors deployed in random position. Sensors are typically low-cost devices not equipped with GPS, and their position is not known a-priori, so that pre-loading position information is not an option. The additional constraints posed on the positioning algorithm by such network scenario leads to the following solution to the positioning problem: a distributed protocol in which every terminal starts the positioning procedure by assuming itself as the origin of its own coordinate system. A similar protocol has been proposed in the framework of Terminodes project [78].
The protocol, based on range measurements obtained by means of a Time Of Arrival (TOA), is described in the two following subsections.

The Self Positioning Algorithm

The Self-Positioning Algorithm (SPA) [79] leads to a relative network topology, and uses a TOA technique in order to obtain range measurements. The algorithm builds a node-centered coordinate system (called Local Coordinate system) for each node in the network, and forces all nodes to converge to a unique relative coordinate system (called Network Coordinate system).

In order to build its own Local Coordinate system, each node i performs the following actions:

1. Detect its set of one-hop neighbors K_i; This phase is accomplished by using beacons, in order to maintain an up-to-date map of one-hop neighbors

2. Evaluate the set of distances from its neighbors K_i; The distance measurement from each one-hop neighbor is obtained by means of TOA estimation

3. Send D_i and K_i to its one-hop neighbors.

In this way each node i will know directly its distances from all its one-hop neighbors, the IDs of its two-hop neighbors, and a subset of the distances from its one-hop neighbors to its two-hop neighbors, as shown in fig. 4.12.

The determination of the local coordinate system in a 2D scenario requires the selection of two additional terminals p, q in the K_i set. p and q must not lie on the same line with i. Furthermore, since all reciprocal distances between i, p and q must be known, an additional constraint in the choice of p and q is set: $p \in K_q$ (or, equivalently, $q \in K_p$). The local coordinate system is defined

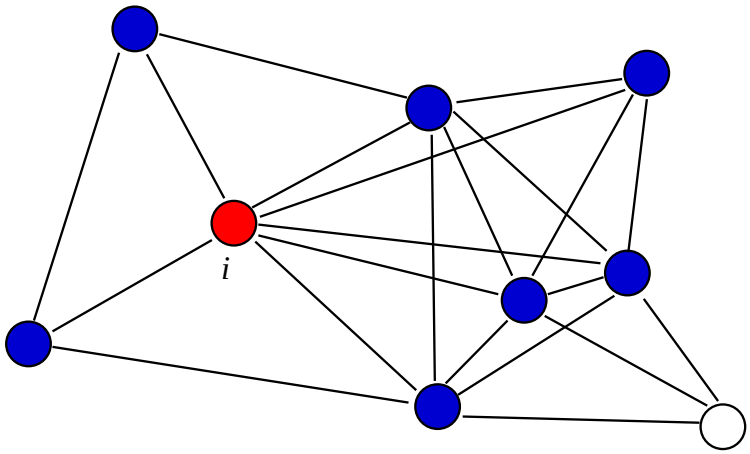

Figure 4.12: Example of network knowledge available at terminal i - red circle represents terminal i, blue circles represent one-hop neighbors of i, white circle represent a two-hop neighbor of i. Lines between nodes represent known distances.

as follows: p lies on the positive x axis, while q has a positive y component. As a consequence, coordinates of i, p and q are:

$$\begin{cases} X_i = 0,\ Y_i = 0 \\ X_p = d_{ip},\ Y_p = 0 \\ X_q = d_{iq} \cdot \cos\gamma,\ Y_q = d_{iq} \cdot \sin\gamma \end{cases} \tag{4.16}$$

where γ is the angle $(p,\widehat{i,q})$, defined as:

$$\gamma = \arccos \frac{d_{iq}^2 + d_{ip}^2 - d_{pq}^2}{2 \cdot d_{iq} \cdot d_{ip}} \tag{4.17}$$

Given the terminals i, p and q the position of a fourth terminal j can be evaluated by trilateration if the distances d_{ij}, d_{pj} and d_{qj} are known. In fact, we obtain:

$$\begin{aligned} j_x &= d_{ij} \cos\alpha_j \\ j_y &= \begin{cases} d_{ij} \sin\alpha_j & if\ \beta_j = |\alpha_j - \gamma| \\ -d_{ij} \sin\alpha_j & if\ \beta_j = \alpha_j + \gamma \end{cases} \end{aligned} \tag{4.18}$$

where γ is the angle $(p,\widehat{i,j})$ and β_j is the angle $(j,\widehat{i,q})$, given by:

$$\begin{aligned} \alpha_j &= \arccos \tfrac{d_{ij}^2 + d_{ip}^2 - d_{pj}^2}{2 \cdot d_{ij} \cdot d_{ip}} \\ \beta_j &= \arccos \tfrac{d_{iq}^2 + d_{ij}^2 - d_{qj}^2}{2 \cdot d_{ij} \cdot d_{iq}} \end{aligned} \tag{4.19}$$

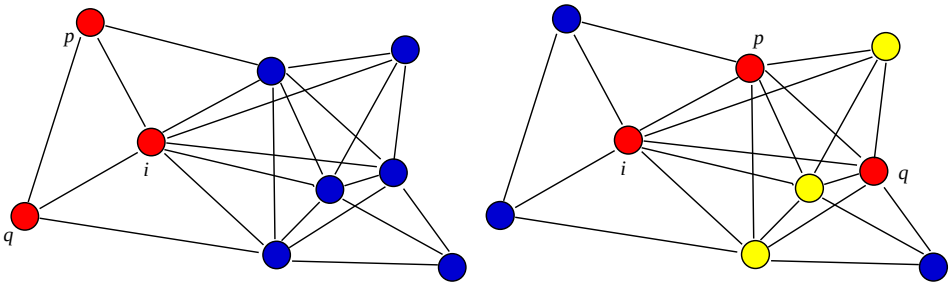

Figure 4.13: Effect of local coordinate system selection on Local View Set - red circles represents terminals i, p and q, yellow circles represent terminals positioned in the local coordinate system and blue circles represent terminals in unknown position.

Depending on the selected couple p, q, i will be able of evaluating the position of only a subset of its one-hop neighbors. This subset is called Local View Set LVS_i, and it is a subset of K_i: the couple p, q should be chosen in order to maximize the size of LVS_i. Note that a terminal can be included in the LVS_i either based on its distances from i, p and q, or based on its distances from i and two other terminals already included in the LVS_i. If we define $C_i = \{(p,q) \in K_i \ \ such \ \ that \ \ p \in K_q\}$ as the set of all possible couples ps and qs, the best couple will be:

$$(p,q) = \arg \max_{(p_k,q_k)\in C_i} |LVS_i(p_k,q_k)| \tag{4.20}$$

Figure 4.13 shows the comparison of two different choices of p and q, and the corresponding size of LVS_i.
Figure 4.13 clearly shows that a smart choice of terminals p, q will lead to a much higher percentage of positioned terminals.

Following the above strategy, each node will define a coordinates system in which it occupies the position $(0,0)$: in order to define a global network topology, all node-centered systems of coordinates must be linearly transformed in order to have a unique orientation (i.e. the same direction for x and y axes of all nodes) and thus obtain the Network Coordinate system. If we consider two terminals i and k the linear transformation can be performed if the two following conditions are met:

- $i \in LVS_k$ and $k \in LVS_i$
- $\exists j \neq i,k$ such that: $j \in LVS_i$ and $j \in LVS_k$;

Note that if there is no a common one-hop neighbor j, the linear transformation cannot be performed. Figure 4.14 presents an example of terminal j

which satisfies the condition.

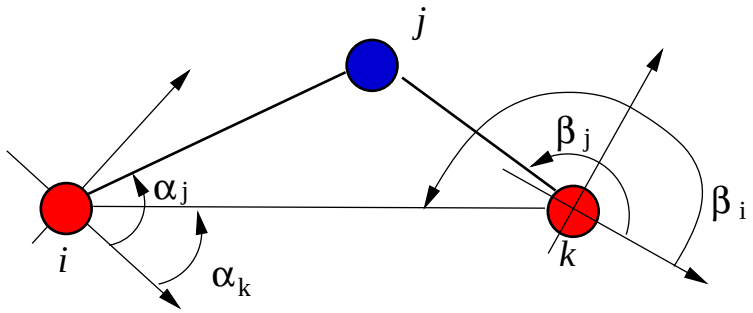

Figure 4.14: Example of common neighbor j satisfying the condition for linear transformation of reference systems of nodes i and k.

Assuming that k must modify its system to get the same orientation of system centered on i, two possible scenarios can be foreseen, depending on local coordinate systems of both k and i:

1. k must rotate its local system of a quantity called correction angle;
2. k must rotate its local system of a quantity called correction angle and must mirror the system over x or y axis.

The node j is used to understand if the system has simply to be rotated or it has also to be mirrored, by evaluating the following conditions:

1.
$$\begin{array}{l} if\ \{[(\alpha_j - \alpha_k) < \pi\ and\ (\beta_j - \beta_i) < \pi]\} \\ or\ \{[(\alpha_j - \alpha_k) > \pi\ and\ (\beta_j - \beta_i) > \pi]\} \end{array} \tag{4.21}$$

2.
$$\begin{array}{l} if\ \{[(\alpha_j - \alpha_k) > \pi\ and\ (\beta_j - \beta_i) < \pi]\} \\ or\ \{[(\alpha_j - \alpha_k) < \pi\ and\ (\beta_j - \beta_i) > \pi]\} \end{array} \tag{4.22}$$

When condition (4.21) is met, mirroring is required, and the correction angle is $\beta_i + \alpha_k$. Oppositely, when condition (4.22) is met, no mirroring is required, and the correction angle is $\beta_i - \alpha_k + \pi$.

An example of the two scenarios is represented in fig. 4.15.

Once that k obtains the same orientation as i, all nodes in LVS_k can do the same, because they know their position in the system centered on k and also the position of node k in the system of node i. Furthermore, the two systems have the same orientation, and this means that for each node $l \in LVS_k$ the position in the new system is simply given by the sum of two vectors.

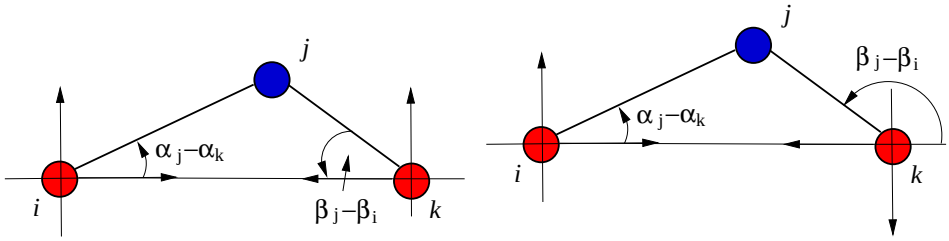

Figure 4.15: Example of coordinate system conversion: left - Mirroring is requested, right - Mirroring is not requested.

Terminals belonging to LVS_l will use the same approach, as the coordinate system adopted by l will now have the same orientation of the coordinate system adopted by k, and the network coordinate system will propagate throughout the network. Terminals that were not able to build their own local coordinate system can still obtain their position if they are able to communicate with three terminals which already received the network coordinate system.

Location Reference Group

The computation of positions in a network coordinate system relies on the selection of a local coordinate system of a terminal i as the reference coordinate system. Nevertheless, if terminal i moves, all terminals must re-compute their position in the new coordinate system, leading to a significant waste of computing power and energy. In order to reduce the effect of mobility on stability of network coordinate system, in [79] the concept of Location Reference Group (LRG) is introduced. The Location Reference Group consists in a set of terminals, which should include the most stable terminals in the network (fixed or slowly mobile, with high autonomy). Furthermore, since broadcasting is used to maintain LRG, the set should include only terminals close enough to avoid packets going through the whole network at every update. The center of the LRG is computed based on the coordinates of its terminals (through a broadcast approach) and it is assumed as network center. The network system orientation is given by the mean value of the orientations of the terminals belonging to the LRG. Such a center is recomputed every time a terminal included in the LRG moves. The average moving speed of the LRG center is expected to be lower than the moving speed of a single terminal, thus assuring a higher stability to network coordinate system in case of high mobility. Intuitively, the larger the LRG, the more stable will be the network system: on the other hand a larger LRG leads to higher overheads due to the broadcast procedure involving a higher number of terminals.

4.4 Positioning in UWB systems

The performance of a distributed positioning algorithm, such as those presented in the previous section, can be defined as the effectiveness of the positioning algorithm in building a global coordinate system and provide each terminal with its own position. Performance is thus determined by two main factors:

1. the network connectivity degree, that determines the average number of neighbors that each terminal can physically reach

2. the ranging accuracy

The network connectivity degree is mainly determined by the considered network scenario, that defines the density of the network, i.e. the number of terminals per unit area, and the average transmission power determining the transmission range of each terminal.
The ranging accuracy, on the other hand, is strictly related on the adopted transmission technique, and we will show in this section that UWB radio can provide a ranging performance far higher than any conventional narrowband or wideband transmission technique, such as those available in existing wireless LANs and cellular networks.
UWB radio, thanks to its GHz-wide bandwidth, is particularly suited for being adopted as a basis for TOA based ranging, as proposed for example in [80]. The accuracy of the TOA estimation expressed by the variance of the TOA estimation error $\sigma_{\hat{\tau}}^2$ is in fact related to the bandwidth of the signal and SNR at the receiver. According to the general theory of ML estimators, in fact, the lower limit for $\sigma_{\hat{\tau}}^2$ in presence of Additive White Gaussian Noise is given by the Cramer-Rao lower bound [81]:

$$\sigma_{\hat{\tau}}^2 = \frac{N_0}{2\int_{-\infty}^{+\infty} (2\pi f)^2 |G(f)|^2 \, df} \tag{4.23}$$

The maximum theoretical ranging performance made available by UWB signals can be thus evaluated by calculating such lower bound for an UWB pulse g(t) that fully exploits the energy made available by the FCC emission masks presented in Chapter 1. Let us suppose that the pulse g(t) has a constant bilateral Power Spectral Density $|G(f)|^2$, i.e.:

$$|G(f)|^2 = \begin{cases} G_0 & for \quad f \in [f_L, f_H] \cup [-f_H, -f_L] \\ 0 & outside \end{cases} \tag{4.24}$$

Eq. (4.23) can be thus written as follows:

$$\sigma_{\hat{\tau}}^2 = \frac{N_0}{8\pi^2 \int_{-\infty}^{+\infty} f^2 |G(f)|^2 df} = \frac{N_0}{8\pi^2 2 \int_{f_L}^{f_H} f^2 G_0 df} = \\ = \frac{N_0}{8\pi^2 2 G_0 \left[\frac{f^3}{3}\right]_{f_L}^{f_H}} = \frac{N_0}{\frac{8}{3}\pi^2 2 G_0 \left(f_H^3 - f_L^3\right)} \tag{4.25}$$

And one can write:

$$\sigma_{\hat{\tau}}^2 = \frac{N_0}{\frac{8}{3}\pi^2 2 G_0 (f_H - f_L)\left(f_H^2 + f_H f_L + f_L^2\right)} = \\ = \frac{N_0}{\frac{8}{3}\pi^2 2 G_0 B \left(f_H^2 + f_H f_L + f_L^2\right)} \tag{4.26}$$

Eq. (4.26) shows that the variance in delay estimation is inversely proportional to the signal monolateral bandwidth occupation B and to a term which depends on the lower and upper frequencies f_H and f_L. It can be easily shown that, for a fixed bandwidth B, this term increases as f_H increases.
Let us consider an IR-UWB signal which fully exploits the frequency band [3.1 - 10.6] GHz with the maximum PSD allowed by the FCC. We obtain the following values: $B = 7.5$ GHz, $f_H = 10.6$ GHz, $f_L = 3.1$ GHz, $2G_0 = 7.413 \cdot 10^{-14}$ W/Hz and $N_0 \cong 4 \cdot 10^{-21}$ W/Hz for a noise temperature $T_s = FT_0$ with F = 7 dB. The limit given by eq. 4.26 then writes:

$$\sigma_{\hat{\tau}}^2 = 8.82 \cdot 10^{-39} \tag{4.27}$$

This corresponds to a lower bound on average distance estimation error equal to $c\sigma_{\hat{\tau}} = 2.82 \cdot 10^{-11}$ m. The above result provides only a theoretical bound for delay estimation error. Receiver hardware limitations, reduced efficiency in the generation of the transmitted signal, and the presence of multipath and MUI lead to a far lower accuracy in delay estimation and thus in ranging procedure [82]. Nevertheless, UWB can be expected to provide accuracies in the order of tens of centimeters, that will thus allow for a positioning accuracy well below 1 meter in outdoor and indoor environments.
The feasibility of accurate positioning based on UWB was recently demonstrated as described in [83], presenting a TDOA-based UWB positioning system. The system is able to recover position of UWB tags with an accuracy in the order of 50 centimeters in indoor environments.

As an example of the effect of both network connectivity degree and ranging accuracy on positioning accuracy, we analyzed by means of simulations the SPA protocol presented in [79].
We considered 20 mobile terminals distributed in an area of $80x80m^2$. In order to show the effect of network connectivity, simulations were performed for varying transmission ranges, starting from 20 meters (low network connectivity) up to 60 meters (high network connectivity), and assuming perfect

ranging, i.e. no error in distance estimation.
Results of this simulations are presented in fig. 4.16, showing the percentage of terminals that were able to evaluate their own position in the global coordinate system.

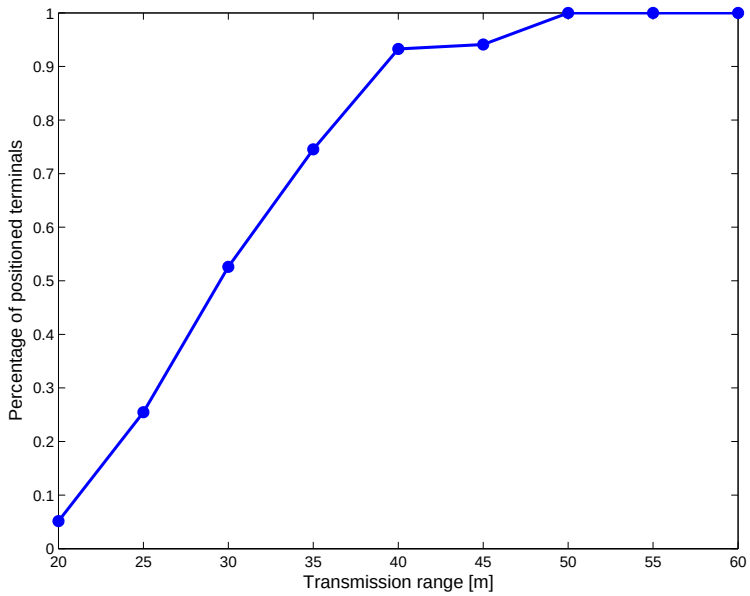

Figure 4.16: Percentage of terminals sharing the same coordinate system as a function of the transmission range.

Interestingly, having a connected network, as presented in fig. 4.17, does not guarantee that each terminal in the network is able to obtain the global coordinate system. Higher connectivity is required for a terminal in order to join the network coordinate system, since each terminal needs either to build a Local View Set in order to get a local coordinate system and then rotate it to meet the global system, or to obtain its position in the global system by trilateration from three other terminals.

On the other hand, fig. 4.16 shows that in conditions of medium to high network connectivity (i.e. with range $\geq$ 40 meters) almost all terminals are able to get their own position in a global coordinate system, and can thus exploit such information for resource allocation and routing.
The results presented above highlight that network density is fundamental in determining the positioning protocol performance and, as a consequence, different network densities lead to dramatically different percentages of positioned terminals. It should be noted however that in most cases terminal

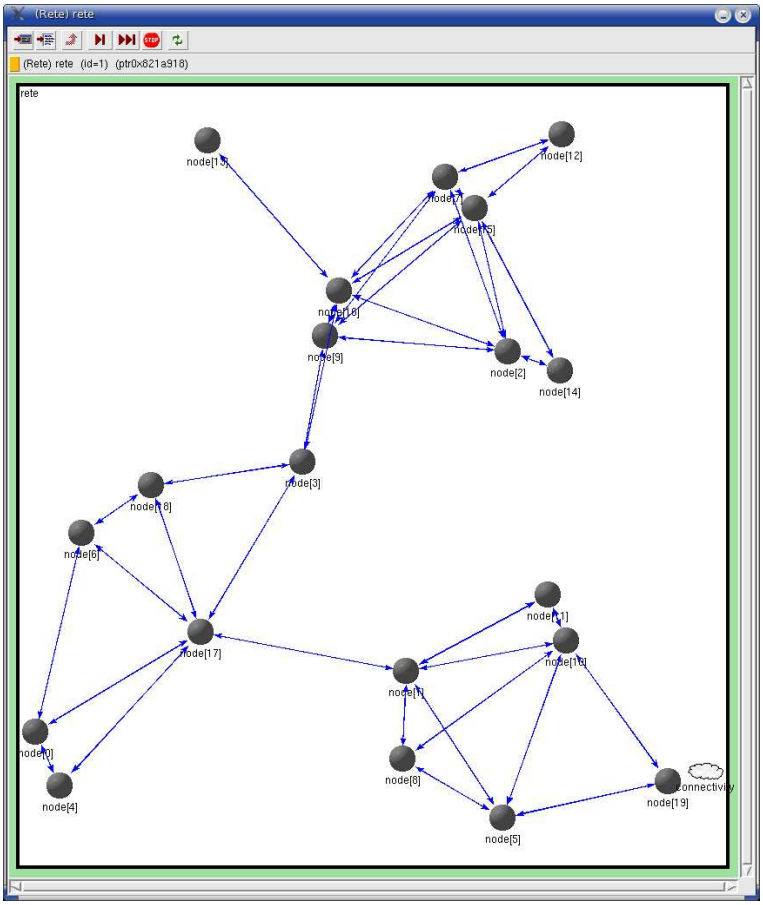

Figure 4.17: Example of connected network with low connectivity (transmission range = 25 m).

density will vary during network lifetime as a consequence of mobility and variations of number of terminals. This is particularly true for networks of terminals characterized by limited energy: in this case in fact terminals will eventually run out of energy, affecting thus the performance of positioning protocol. As a consequence, it is interesting to investigate how many terminals are required, for given transmission ranges and area size, in order to achieve a given percentage p of positioned terminals. Figure 4.18 presents the percentage of positioned terminals as a function of the number of terminals for three different values of the transmission range. If we assume that the cause of the variation in number of terminals in a network is battery exhaustion, we can estimate what will be the evolution of positioning protocol performance during network lifetime by reading fig. 4.18 from right to left.

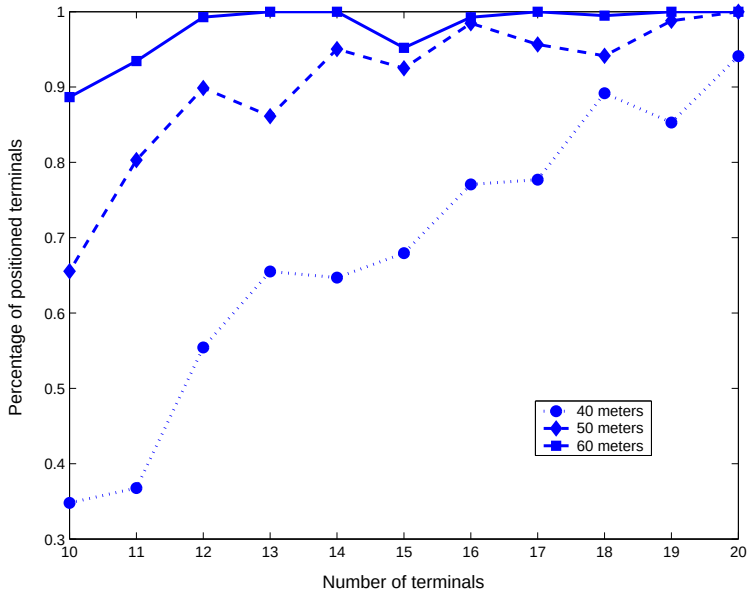

Figure 4.18: Percentage terminals sharing the same coordinate system as a function of the number of terminals for three values of transmission range (circles: range = 40 m, diamonds: range = 50 m, squares: range = 60 m).

As already said, however, distance estimates provided to the positioning protocol will be in most cases subject to errors, the amount of the error depending on the adopted transmission technique.
As a consequence, we analyzed the effect of ranging error on positioning accuracy in conditions of high connectivity (transmission range = 60 m). We considered three different cases, characterized by a ranging error with uniform distribution in the intervals [-0.1m, 0.1m], [-1m, 1m], [-10m, 10m] respectively.

Results of simulations are presented in fig. 4.19.

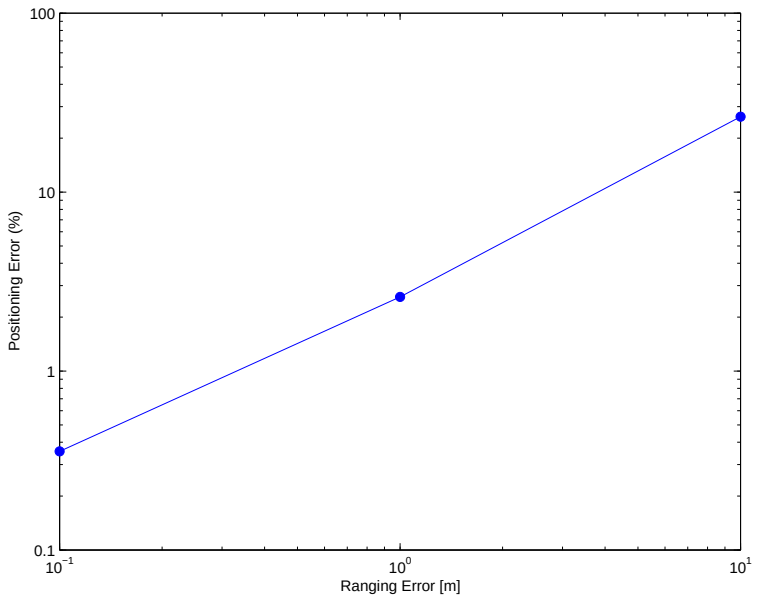

Figure 4.19: Percentage positioning error as a function of maximum ranging error.

Simulation results highlight that large ranging errors heavily impact positioning accuracy in terms of relative error in estimated positions: fig. 4.19 shows in fact that ranging errors above 1 meter lead to positioning errors that are not compatible with applications requiring accurate position for working properly. The ranging accuracy provided by UWB is thus a key element for guaranteeing the positioning accuracy required by such applications.
An example of the effect of position information on higher layers will be discussed in the next chapter, dealing with a power-aware, location-based MAC and routing strategy.

Chapter 5

UWB-based location-aided, power-aware routing

It was observed in Chapter 1 that the requirements posed by actual UWB regulation on UWB power efficiency coincide with the increasing interest in applications based on the deployment of ad hoc and sensor networks, that triggered significant research efforts regarding the introduction of the energy-awareness concept in the design of solutions for such networks.
In this chapter we will show how the distance information made available by the UWB technology, and the position information derived from distances, can be exploited to achieve low power levels and increased network lifetime on the long term, while providing an adequate network performance (in terms of data throughput) on the short term.

5.1 Power-efficient MAC strategies

As far as power-aware MAC design is concerned, most of the research activity took the moves from existing solutions for medium access in wireless networks, and in particular CSMA and Collision Avoidance.
CSMA-CA is by far the most common approach adopted in narrowband wireless LANs. The CSMA-CA family of protocols suffers, however, of several drawbacks in terms of power consumption: packet collisions, which cause a waste of power due to re-transmission, terminals blocked in idle state by active transmissions, and overhearing, i.e. power consumption in the reception of packets by non-intended destinations. Several protocols aiming at the solution of such drawbacks were proposed. Among all it is worth mentioning the Power-Aware Multiple Access protocol with Signalling for ad-hoc networks (PAMAS) [84]. This protocol combines the CSMA approach with the out-of-band principle, in order to minimize the time a terminal spends in the idle

state without neither transmitting nor receiving. The protocol foresees in fact a control channel on which the handshake between transmitter and receiver takes place, and a data channel on which data packets are exchanged. The control channel allows terminals to determine when they may safely switch to sleep mode, saving thus power, without affecting data throughput or end-to-end delay. The main idea behind PAMAS is that if a node detects the channel as busy, it goes in sleep mode rather than waste power in idle mode without being able of exchanging data packets. The protocol defines dedicated handshakes, which allow terminals to determine for how long they can keep the radio interface switched off. Simulations show that this approach leads to a power save of up to 70% compared to the standard MACA protocol. It should, however, be noted that part of the dramatic advantage shown by PAMAS over the MACA protocol is intrinsically due to the assumption of adopting a CSMA-CA approach at the MAC level. PAMAS exploits, in fact, the presence of "pause" periods in the terminal lifetime, due to busy channel or lost contention, to save energy without affecting delay and throughput. If a different solution, for example Aloha without neither Carrier Sensing nor Collision Avoidance, is adopted, the PAMAS-like approach is no longer a suitable option.

5.2 Power-efficient routing

Power-efficient routing is another topic which is of great interest in relation to distributed ad hoc and sensor networks. Several solutions have been proposed, either pursuing power efficiency on top of existing routing protocols ([85], [86]) or proposing completely new routing schemes [87]. An approach of particular interest in this field is the definition of power-efficient routing metrics, taking into account power in the evaluation of the cost of a route between source and destination, final aim being the selection of source-destination paths that minimize power consumption. To this regard, [88] provides a thorough analysis of the effect of power-aware metrics on network lifetime and fairness in energy consumption between different terminals in the network. The metrics taken into account in the analysis include transmission power, cumulative transmitted and received power, residual power in each node, and their combinations. Interestingly, simulation results presented in [88] show that a straightforward minimization of transmitted power does not necessarily lead to a longer lifetime for each terminal in the network. The adoption of a power-aware metric may in fact give rise to paths composed by a high number of hops, involving thus a higher number of terminals in each communication. This strategy provides however fair power consumption between different terminals, increasing thus the network operation time before the first terminal runs out of power

with the consequence of potentially causing a network partition.

For the specific case of UWB, a method for setting up connections by optimizing a power-dependent cost function has been introduced in [89], and further refined in [90],[62]. In [89], a communication cost is attached to each path, and the cost of a path is the sum of the costs related to the links it comprises. The cost of a link is expressed as the sum of two components as follows:

$$C(x, y) = \delta \cdot C_0 \cdot d^{\alpha} + C_1 \cdot R \cdot d^{\alpha} \tag{5.1}$$

The first component takes into account the synchronization cost for setting-up a new link. If two nodes already share an active link, $\delta = 0$ and there is no synchronization cost. If two nodes do not share an active link, $\delta = 1$ and a synchronization cost is added. The second component takes into account the cost for transmitting data, and depends upon the requested data rate R. Both terms are related to power consumption, and therefore depend upon the distance d between two nodes. Note that the evaluation of such a distance relies on the precise ranging capabilities offered by the UWB technique. The parameter α is related to channel propagation characteristics and has commonly a value between 2 and 4. Constants C_0 and C_1 are used to weight the synchronization and transmission components.

In [90] the proposed strategy is compared against traditional routing in a scenario characterized by fixed terminals and full network connectivity. Results show that the power-saving strategy, as expected, leads to multi-hop communication paths between terminals within reach of each other (physical visibility) and by this way increases network performance by reducing average emitted power and thus interference levels [90]. In [62] an improved version of the cost function was proposed, in order to introduce additional parameters in the route selection metrics. The general form of this cost function is given by:

$$C(x, y) = C(power) + C(sync) + C(interference) + C(quality) + C(delay) \tag{5.2}$$

where the first two terms, related to power and synchronization respectively, resemble the two terms defined in eq. (5.1).

Taking the moves from this work, an advanced UWB routing metric was defined. Such metric is described in the next section.

5.3 UWB cost function

As regards the routing criteria, the key factors which should be taken into account in the selection of a multi-hop route in an UWB network can be listed

as follows:

- Synchronization: as already anticipated in Chapter 3, the assumption of an uncoordinated MAC protocol leads to a significant synchronization overhead. In particular routing control packets present the highest overhead, since synchronization must be acquired between terminals which, in the worst case, are not aware one of each other. On the other hand, transmission of DATA packets over active connections will in general require a lower overhead, since transmitter and receiver will preserve at least a coarse synchronization information between two consecutive DATA packets, thus allowing a shorter synchronization header; as already discussed in section 3.4, however, in the case of low data rate networks a full synchronization header will be also required in each DATA packet, due to the large average time between two consecutive DATA packets.

- Power: UWB networks will require a smart management of available power in order to optimize network performance while meeting the emission limits for UWB devices. As a consequence, power considerations should have a predominant role in the selection of a route, in order to efficiently exploit the available power. As already mentioned in the previous section, the concept of power-aware routing for ad-hoc networks has been widely analyzed in the literature, and several power-efficient solutions were defined. The most straightforward approach is to minimize the emitted power by selecting minimum power routes. The effect of such a choice is exemplified in figs. 5.1 and 5.2, showing the effect on emitted power levels of hops minimization (fig. 5.1) vs power minimization (fig. 5.2) in a network of 10 devices in fixed positions, where the emitted power is measured in each point of the network over the entire simulation time. In this example power minimization was obtained by using the cost function defined in eq. (5.1) with $C_0 = 0$ and $C_1 = 1$.

- Multi-User Interference: The selection of power optimized routes, by itself, is not enough to guarantee an efficient use of available power at network level. MUI has a key role in determining the overall performance: the selection of a route in a high density region, in fact, causes an increase of the power required to achieve an acceptable Packet Error Rate (PER) on all active links in such region due to the higher interference, and thus leads to an inefficient use of available power. As a consequence, MUI must be taken into account in route selection; This can be achieved by considering the local network topology in route selection, as shown in figs. 5.3 and 5.4. Figure 5.3 shows the minimum energy route between source and destination, which is likely to cause

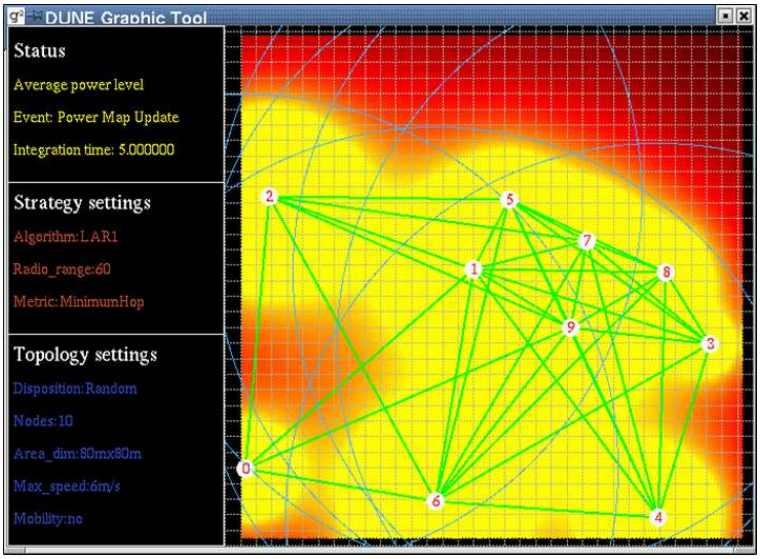

Figure 5.1: Effect of route selection based on minimum number of hops on average emitted power levels. In the plot, dark colors represent low power levels (normalized to the maximum transmit power) while light colors represent high power levels.

high interference due to the high network density. Oppositely, fig. 5.4 shows an alternative route which takes into account network topology and thus avoids the high density region in which many nodes could be affected by the interference caused by the new links.

- Link reliability: node mobility and variable network conditions (due to link set-up and releases, nodes switching on and off) may cause high instability in selected routes, leading to frequent Route Reconstruction procedures, and thus significantly reducing routing protocol efficiency. Furthermore, a low link reliability can easily lead to an insufficient QoS level for traffic classes characterized by hard limits on PER.
  In order to reduce such instability, link reliability should be evaluated in the route selection phase. Since in most cases a node has no information on the mobility degree of its neighbors, the reliability of its links can be only evaluated on the basis of the number of correctly received packets with each of its neighbors in a given time interval. A couple of neighboring nodes with low or no mobility, in fact, will be able to exchange a higher number of packets, being only influenced by thermal noise, channel and MUI introduced by other neighboring terminals. Note that in this case also network topology should be taken into account, since

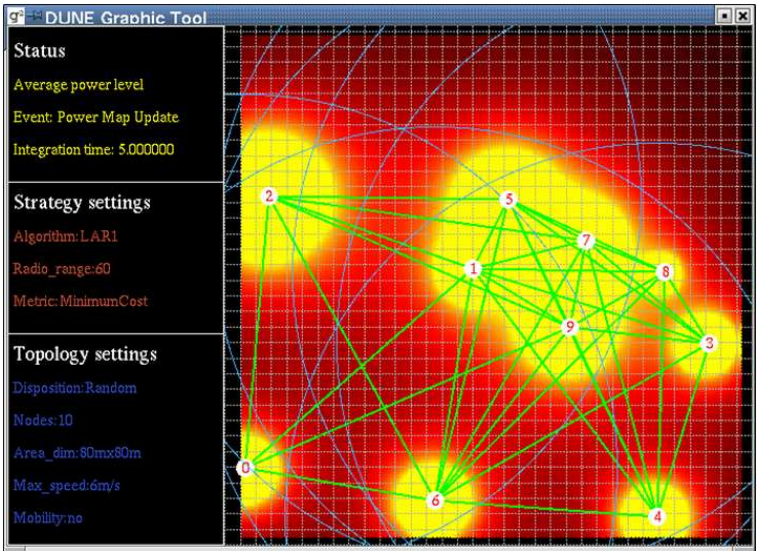

Figure 5.2: Effect of route selection based on minimum power on average emitted power levels. In the plot, dark colors represent low power levels (normalized to the maximum transmit power) while light colors represent high power levels.

the presence of many neighbors at close distance will likely cause high interference levels, thus reducing link reliability.

- Traffic load: Each of the criteria defined above for route selection can potentially lead to an unbalanced distribution of routes, causing a terminal which particularly fits one or more criteria to be selected more frequently than others. As an example, a fixed terminal can guarantee a higher reliability, and will thus be selected in most cases if reliability is the main criterion in route selection. Such terminal will then experience a higher traffic, with the consequence of a reduced autonomy. This negative effect can be avoided if the traffic load of each terminal is taken into account in route selection: a terminal will be less and less appealing for new routes as its traffic increases, thus leading to a fair distribution of routes all over the network.

- End-to-end delay: It was observed above that link reliability is fundamental in order to meet QoS requirements for traffic classes in which data integrity is fundamental, such as ftp or http transfers. On the other hand, in many situations guaranteeing a low end-to-end delay is far more important than delivering all packets: this is the case of voice

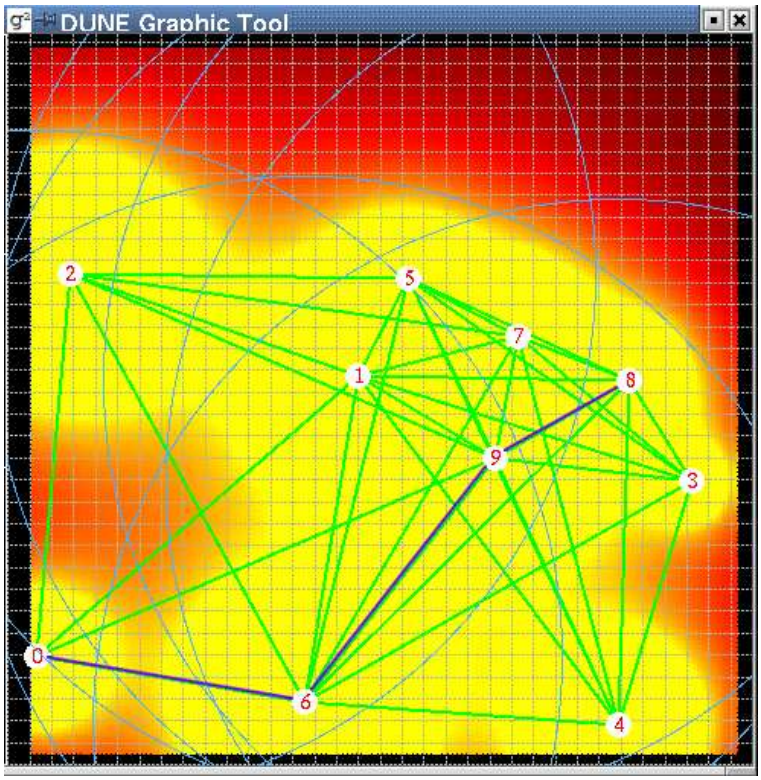

Figure 5.3: Example of minimum energy route between nodes 0 and 6, subject to high interference in 9, due to nodes 1, 7 and 5. In the plot, dark colors represent low power levels (normalized to the maximum transmit power) while light colors represent high power levels. Selected route is formed by links represented in dark color.

and multimedia traffic. As a consequence, delay must also be taken into account in route selection, in order to assure acceptable delays for time-sensitive traffic classes.

The list of criteria defined above shows that route selection is in most cases the result of a trade-off between opposing requirements: as an example, power minimization will lead to routes composed by many small distance hops, which guarantee the lowest emitted power levels. On the other hand, the delay experienced over such a route could be unacceptable for time-sensitive traffic, thus requiring a compromise between power efficiency and end-to-end delay.

Based on the above analysis, an advanced UWB cost function was defined. The UWB cost function is defined as an additive function. Each term in the sum takes into account one of the routing criteria defined above and is weighted

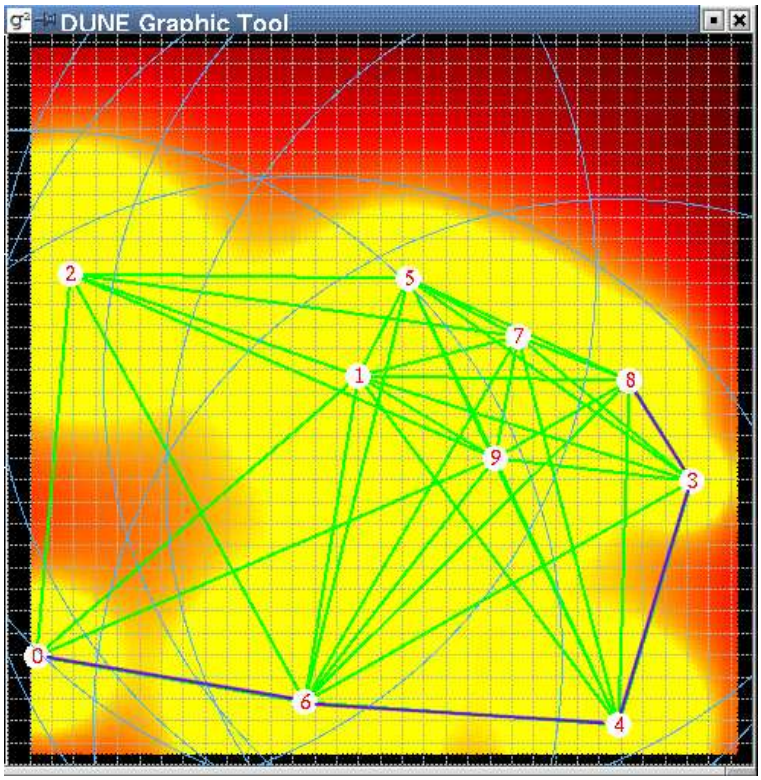

Figure 5.4: Example of interference-aware route between nodes 0 and 6, avoiding the high-density region between source and destination. In the plot, dark colors represent low power levels (normalized to the maximum transmit power) while light colors represent high power levels. Selected route is formed by links represented in dark color.

by a coefficient, so that the cost function for a link between two generic nodes x and y takes the general form:

$$\begin{aligned} UWB_{Cost}(x,y) = {} & c_{Sync} \cdot Sync(x,y) + c_{Power} \cdot Power(x,y) + \\ & + c_{MUI} \cdot MUI(x,y) + c_{Reliability} \cdot Reliability(x,y) + \\ & + c_{Traffic} \cdot Traffic(y) + c_{Delay} \cdot Delay(x,y) \end{aligned} \tag{5.3}$$

In eq. (5.3), all terms are related to the specific interaction between nodes x and y, with the only exception of the Traffic term, which only depends on the receiver; The traffic load already carried out by the receiver is in fact independent on which transmitter is considered. In the remaining part of this

section, each term will be described in detail, by giving its analytical formulation and by highlighting the information required to compute its value. The specification of each term carried out in the following is based on the hypothesis that the MAC protocol is capable of providing distance information, and in particular assuming the $(UWB)^2$ MAC proposed in Chapter 3. It should be noted however that the cost function can be applied with minor changes to different scenarios, e.g. characterized by a centralized MAC. The last part of this section will analyze the degree of flexibility of the UWB cost function guaranteed by a fine tuning of the coefficients presented in eq. (5.3).

5.3.1 Synchronization term

The synchronization term is defined as follows:

$$Sync\,(x,y) = \delta\,(x,y) \tag{5.4}$$

Where the $\delta(x,y)$ is 0 if nodes x and y already share an active connection, and 1 otherwise.

As a consequence of the hypothesis of having a fully distributed MAC, the synchronization between transmitter and receiver must be acquired from scratch for all packets. This synchronization cost is thus common to all paths during the path search procedure, and is not considered in the cost function. Additionally, a synchronization maintenance overhead is also present during the transmission of DATA packets over an active connection. This cost is taken into account in the Synchronization term, by assuming that if x and y already share active connections, the synchronization maintenance overhead for the new connection will be negligible; Oppositely, if the new connection is going to be the first one between x and y, the synchronization maintenance overhead will be included in the evaluation of the overall cost of the route. As already stated, this assumption is reasonable for high data rate networks, but it may not hold for low data rate networks, where synchronization is lost between two subsequent DATA packets. In this case this term can be neglected, since all paths will have the same synchronization cost.

5.3.2 Power term

The power term writes:

$$Power\,(x,y) = \left(\frac{d\,(x,y)}{d_{\max}}\right)^{\alpha} \tag{5.5}$$

where $d(x,y)$ is the distance between x and y, d_{max} is the maximum transmission distance as estimated by x, and α is the exponent of power attenuation

with distance. This term takes into account the power required to transmit over the link between x and y with a given SNR, normalized by the maximum transmit power. In fact we can write the power over the link (x, y) as follows:

$$SNR = \frac{P_T/A(d)}{P_N} = \frac{P_T/\left(A_0 \cdot d^{\alpha}\right)}{P_N} \Rightarrow P_T\left(x, y\right) = SNR \cdot P_N \cdot A_0 \cdot d\left(x, y\right)^{\alpha} \tag{5.6}$$

In the same conditions (i.e. same SNR and same bit rate), the maximum power can be then expressed in terms of the distance d_{max} with the same formula:

$$SNR = \frac{P_T/A(d)}{P_N} = \frac{P_T/\left(A_0 \cdot d^{\alpha}\right)}{P_N} \Rightarrow P_{\max} = SNR \cdot P_N \cdot A_0 \cdot d_{\max}^{\alpha} \tag{5.7}$$

And finally we have:

$$\frac{P_T\left(x, y\right)}{P_{\max}} = \frac{SNR \cdot P_N \cdot A_0 \cdot d\left(x, y\right)^{\alpha}}{SNR \cdot P_N \cdot A_0 \cdot d_{\max}^{\alpha}} = \left(\frac{d\left(x, y\right)}{d_{\max}}\right)^{\alpha} \tag{5.8}$$

The computation of the power term of the cost function requires the receiver node y to have an estimation of the distance $d(x, y)$, which is expected to be provided by the UWB ranging capability. Furthermore, an estimation of the maximum power P_{max} of node x is also required; in the following we assume that all terminals share the same maximum power, so that an explicit communication of such value is not necessary.

5.3.3 MUI term

The analytical formulation for the MUI term is:

$$MUI\left(x, y\right) = \frac{1}{N_{neighbors,y}} \cdot \sum_{n=0}^{N_{neighbors,y}-1} \left(1 - \frac{d_{\min}}{d\left(x, n\right)}\right)^2 \tag{5.9}$$

Where:

- $N_{neighbors,y}$ is the number of neighbors known to x, excluding y;
- n is the generic neighbor, excluding y;
- d_{min} is the distance between x and its closest neighbor, excluding y.

This term takes into account the effect of a new transmission of node x on its neighbors; Since the closest neighbor will suffer the highest impact from the emission of terminal x, a small d_{min} will increase the value of this term, thus

leading to a higher impact of MUI term on the overall cost. The evaluation of the MUI term requires x to build and maintain a distance estimation vector, containing an estimation of distances from x to all its neighbors; Furthermore, since the MUI term value calculated in x must be available in y to calculate the cost of the (x, y) link, the value for each of the $N_{neighbors,y+1}$ neighbors must be included by x in its outgoing packets. As an alternative, x could include its distance estimation vector in the packets, allowing each receiver to evaluate the MUI cost term. Such a solution should be preferred, since in this case the information to be sent would be exactly the same required by the distributed positioning protocol, as explained in Chapter 4.

5.3.4 Reliability term

The reliability term has the following expression:

$$\frac{1}{2} \cdot \left[\frac{1}{N_{packets,x}} + \frac{1}{N_{neighbors,x}} \cdot \sum_{n=0}^{N_{neighbors,x}-1} \left(1 - \frac{d_{\min}}{d\,(y,n)} \right)^2 \right] \tag{5.10}$$

where:

- $N_{packets,x}$ is the number of packets y received from x in the last observation interval; $N_{neighbors,x}$ is the number of neighbors known to y, excluding x;
- n is the generic neighbor, excluding x;
- d_{min} is the distance between y and its closest neighbor, excluding x.

This term takes into account two factors which influence the reliability of link (x, y):

1. The stability of the link, expressed by the number of packets that y received from x in a given time; This implicitly takes into account nodes mobility, since a highly mobile node will likely experience a low link stability with its neighbors.
2. The expected MUI which will affect the receiver y. The evaluation of this factor is the same as defined in subsection 5.3.3 for the MUI term, but referred here to the potential receiver y.
   The evaluation of this term requires the same information as the MUI term i.e. the maintenance of the distance estimation vector defined in the previous subsection.

5.3.5 Traffic term

The analytical expression for the Traffic term writes:

$$Traffic(y) = \frac{1}{B_{\max,y}} \cdot \sum_{i=0}^{N_{active,y}-1} B_i \tag{5.11}$$

where:

- $B_{max,y}$ is the maximum overall rate which can be guaranteed by node y;
- B_i is the rate of the i-th active connection in y;
- $N_{active,y}$ is the total number of the active connections in y.

As anticipated in section 5.3, this term is introduced to the aim of avoiding an unfair selection of routes in the network, by increasing the cost of routes including nodes already involved in many active connections.

5.3.6 Delay term

The value of the delay term is given by:

$$Delay(x, y) = 1 \tag{5.12}$$

We assume in first approximation that end-to-end delay can be considered to be proportional to the number of hops in the path; as a consequence, this term is modeled with a constant value, in order to balance the effect of the Power term defined in subsection 5.3.2 and avoid routes with excessive number of hops.

5.3.7 Tuning of cost function coefficients

All UWB cost function terms defined in the previous subsections are designed in order to take values in the interval [0,1]. This allows a straightforward tuning of the function by means of the set of coefficients $\{C_{Sync}, C_{Power}, C_{MUI}, C_{Reliability}, C_{Traffic}, C_{Delay}\}$. Each coefficient also takes values in the interval [0,1], and different coefficient sets can be adopted to support at the same time traffic with different characteristics. As an example, non-interactive data traffic, such as an ftp transfer, requires a high degree of data integrity but can tolerate high end-to-end delays. The cost function could be customized for this traffic class by increasing the relative weight of the Reliability cost term, while reducing the weight of the Delay term. Oppositely, voice-like traffic can tolerate a relatively high PER, but poses strong constraints on the end-to-end delay. In this case the role of the Reliability and Delay terms is inverted, with

	A	B	C
C_{Sync}	0	0	0
C_{Power}	0	1	0.1
C_{MUI}	0	0	1
$C_{Reliability}$	0	0	0
$C_{Traffic}$	0	0	0
C_{Delay}	1	0	0

Table 5.1: Definition of UWB Cost function coefficient sets.

the latter term characterized by a much higher relative weight than the former one.

As an example of how coefficient selection can influence the performance of the network, let us consider a comparison between three sets of coefficients, shown in Table 5.1. We considered a network of 20 fixed terminals, generating random connection requests; route selection is performed with a standard source routing protocol ([1]) combined with the UWB Cost function.
Coefficient set A only considers the delay term, and thus leads to a straightforward hops minimization; set B, oppositely, only considers the Power term, and leads thus to the lowest emitted power; set C, finally, takes in to account the Power term, but gives a higher weight to the MUI term. Figure 5.5 shows that coefficient set C has an intermediate behavior in terms of hops between set A (minimum number of hops) and set B (minimum emitted power); nevertheless, the performance of set C is the highest in terms of throughput, as shown in fig. 5.6, thanks to the better management of MUI.

5.4 Location-aware routing protocols

The UWB ranging capability is the basis for the definition of the power-aware cost function defined in the previous section; as already stated, however, ranging information can be even more efficiently exploited to build a network map through a distributed positioning protocol, providing thus the input for a location-based routing algorithm. Distributed positioning protocols have been addressed in Chapter 4, while in this section, we will analyze the concept of location-based routing. First, we will review location-aware routing protocols available in literature, with focus on power-efficient protocols. Next, we will address the topic of how location information can be gathered and exchanged in a UWB network, and we will analyze the effect of UWB-based positioning on the location-aware routing protocols originally designed for GPS-enabled terminals.

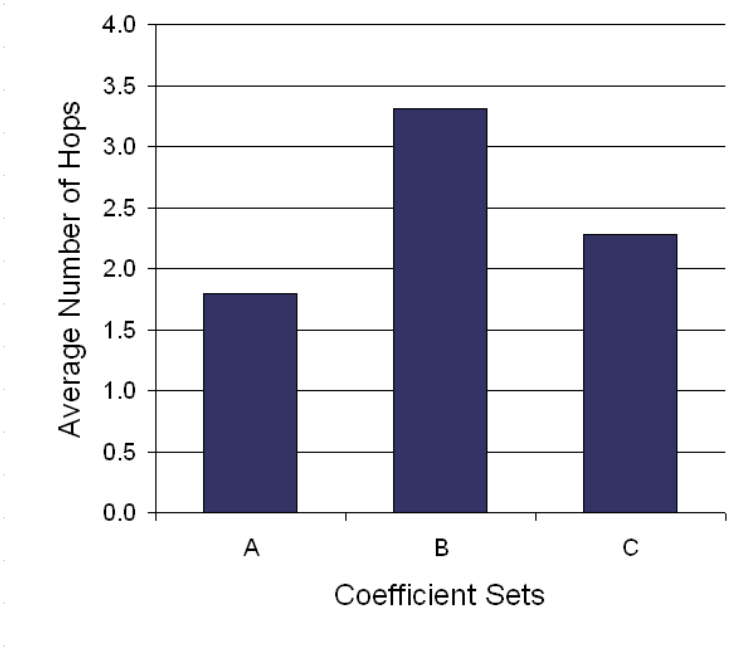

Figure 5.5: Average number of Hops for the coefficient sets defined in Table 5.1.

As a general concept, all location-aware routing protocols pursue an improvement of network performance by taking into account location information in the route selection process. Depending on how this information is used, different aspects of the routing performance can be optimized. Subsection 5.4.2, for example, shows how location information can be used for increasing route stability. The main focus being here on power efficiency, however, the remaining part of this Section addresses routing protocols that exploit location information for reducing routing overhead, and thus power consumption.

5.4.1 Greedy Perimeter Stateless Routing

The Greedy Perimeter Stateless Routing (GPSR) [91] protocol uses location information to reduce protocol overhead and obtain a good scalability when both terminal mobility and network size increase. GPSR adopts positional information as the key metric in packet forwarding, and uses the following simple "greedy" forwarding strategy:

1. Each packet is marked by the source terminal with the latest information about the location of the destination

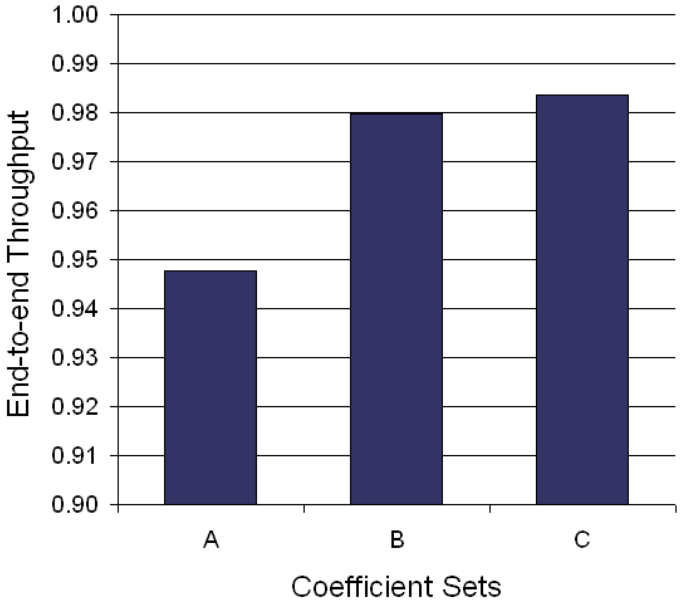

Figure 5.6: Average end-to-end throughput for the coefficient sets defined in Table 5.1.

2. Each intermediate node forwards the packet to the neighboring visible node that is closest to the location of the destination stored in the packet itself

Note that greedy forwarding does not guarantee that a path between source and destination is always detected, even if it exists. An example of greedy forwarding failure is presented in fig. 5.7. Node x, after receiving a packet from source S intended for destination D, cannot find a neighbor which satisfies the greedy forwarding rule, since both neighbors y and z are further than x from the destination D.

When greedy forwarding fails, the GPSR protocol switches from greedy forwarding to perimeter forwarding, in which a terminal is allowed to forward the packet to a neighbor which is further than itself from the destination, in order to solve the stall caused by greedy forwarding. The perimeter forwarding strategy uses the so-called right-hand rule to forward packets around the area in which the greedy approach fails, and can be described as follows (see fig. 5.8).

Suppose that x is the first terminal in which greedy forwarding fails.

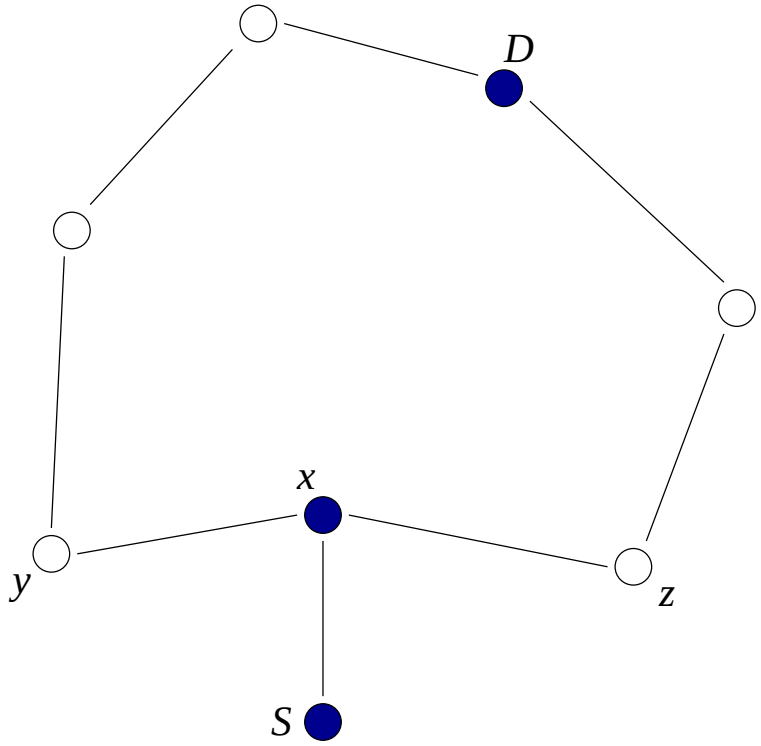

Figure 5.7: Example of greedy forwarding failure. Arcs in the figure represent physical visibility between nodes. The greedy forwarding rule fails at node x, since both neighbors y and z are further than x from the destination D.

1. Terminal x records its position in the apposite L_f field in the packet;
2. Terminal x uses the information on its location and the location of destination D recorded in the packet to determine the $\overline{xD}$ line;
3. Terminal x, based on the position of its neighbors, determines which is the first edge in counterclockwise direction from the $\overline{xD}$ line (edge $\overline{xy}$ in fig. 5.8).
4. Terminal x forwards the packet to terminal y;
5. Terminal y checks if it is closer to destination D than the position recorded in the L_f field. If this is the case, y switches back to greedy forwarding, otherwise repeats the procedure from step 2.

In the example in fig. 5.8 x enters the perimeter forwarding, and w switches back to the greedy forwarding. The grey area defines the perimeter over which

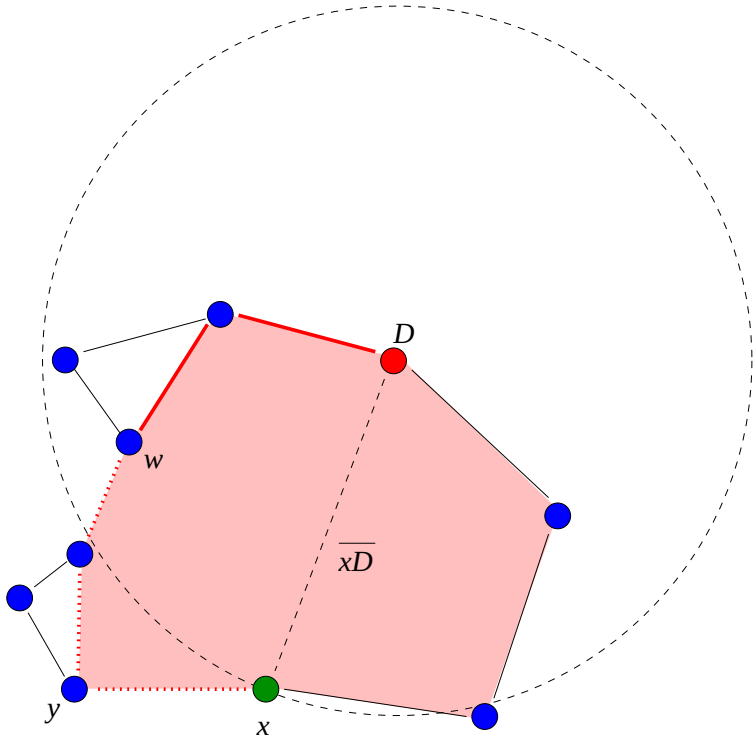

Figure 5.8: Combination of perimeter forwarding and greedy forwarding. Arcs in the figure represent physical visibility between nodes. The selected path is represented by a thick line (dotted links: perimeter forwarding, filled links: greedy forwarding).

the algorithm switches from greedy forwarding to perimeter forwarding.

5.4.2 Location-aware Long-lived Route Selection in Wireless Ad Hoc Network

The Location-aware Long-lived Routing (LLR) protocol [92] belongs to the family of source-initiated, on-demand routing protocols. The protocol exploits location information aiming at the minimization of route failures and, as a consequence, of route reconstruction procedures. As in traditional source routing protocols, a source terminal S needing for a route to a destination broadcasts Route Discovery packets to its neighbors, which will forward packets until the destination is reached.

In LLR, the source S includes two additional information fields in each

generated packet: its own position (X_S, Y_S) and radio transmission range R_s. Each terminal A in physical connectivity with S that receives the packet uses such information and its own position (X_A, Y_A) and transmission range R_A to evaluate two figures:

1. Forward Movement Limit (FML), that measures the maximum relative movement between A and S which can be tolerated before the distance between A and S is larger than the transmission range of A, defined as:

$$FML = R_A - distance(A, S) = R_A - \sqrt{(X_S - X_A)^2 + (Y_S - Y_A)^2} \tag{5.13}$$

2. Backward Movement Limit (BML), that measures the maximum relative movement between A and S which can be tolerated before the distance between A and S is larger than the transmission range of S, defined as:

$$BML = R_S - distance(A, S) = R_S - \sqrt{(X_S - X_A)^2 + (Y_S - Y_A)^2} \tag{5.14}$$

FML and BML are then used to evaluate the Normalized Movement Limit (NML), which provides a single figure averaging FML and BML (note that $FML \neq BML$ when $R_S \neq R_A$):

$$NML = \frac{FML \cdot BML}{FML + BML} \tag{5.15}$$

Terminal A updates the packet by including the NML and substituting (X_A, Y_A) and R_A to (X_S, Y_S) and R_S, respectively, and forwards it to its neighbors.
The NML is an indicator of link stability: the higher the NML value, the lower the probability of link failure due to terminal mobility. When the destination D receives several Route Discovery packets originated from the source S, it selects the route characterized by the highest NML, in order to minimize the probability of a route failure. In [92] two alternative ways are foreseen to evaluate the end-to-end reliability of each route, depending on the way a Route Discovery packet is updated:

Additive NML : each terminal adds the NML value relative to the last hop to the value carried by the packet, and the best path is characterized by the maximum overall NML

Min_Max ***NML*** : each terminal compares the NML just evaluated with the value recorded in the packet: if the current value is lower than the old one, the packet is updated by overwriting the NML field, otherwise the field is left unchanged. In this case the best route is determined by the worst link in each route (which determines the NML value in the Route Discovery packet)

5.4.3 Distance Routing Effect Algorithm for Mobility

The Distance Routing Effect Algorithm for Mobility (DREAM) algorithm [93] proposes the idea of using positional information to reduce the amount of routing overhead. The protocol combines both proactive and reactive approaches, by relying on both periodic updates by each terminal for the dissemination of location information and a flooding-like procedure for sending a packet to the targeted destination.
When a source terminal S starts the procedure at $t = t_1$ it is supposed to have the following information:

- Its own position (X_S, Y_S)
- The positions of its one-hop neighbors
- The position of the destination D, (X_D, Y_D), at a given time $t_0 < t_1$
- The maximum speed v of the destination, or at least a probability density function of the speed, $p(v)$

The source, based on this information, defines a geographical region in which the routing packets should be forwarded, and determines the subset of neighbors that are positioned inside this region, as shown in fig. 5.9. The forwarding region is determined by the angle α.

Depending on the information on destination speed, two possibilities exist:

1. v is known: in this case the distance x traveled by the destination in the period $t_1 - t_0$ is given by $x = v * (t_1 - t_0)$, and, given a distance r between S and D, one has:

$$\alpha = \arcsin \frac{x}{r} \tag{5.16}$$

2. $p(v)$ is known: in this case in order to determine a forwarding region including D with probability p, the corresponding value of α can be derived from ([93]):

$$\int_{\frac{r \cdot \sin \alpha}{t_1 - t_0}}^{+\infty} p\,(v) dv \geq p \tag{5.17}$$

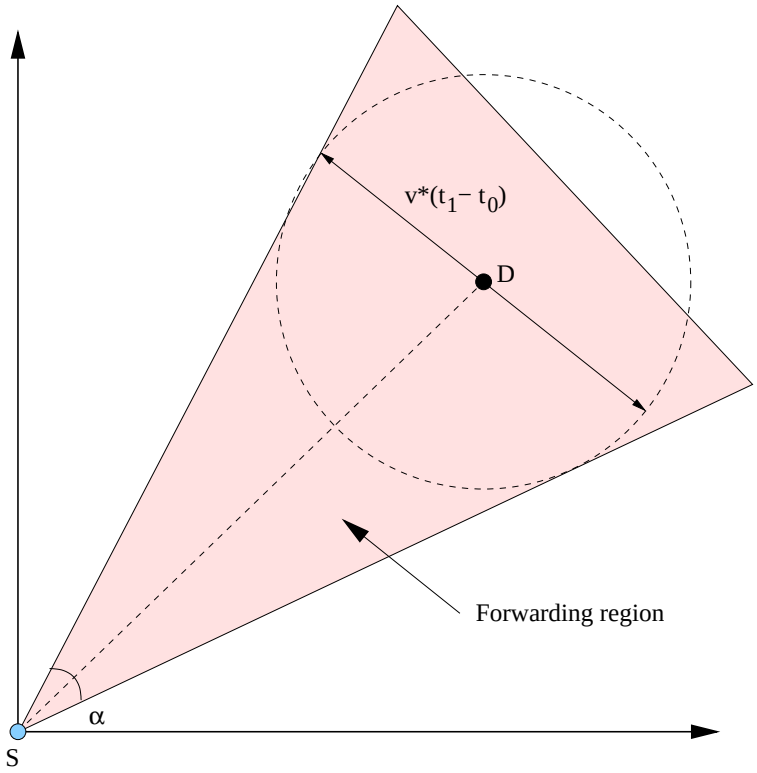

Figure 5.9: Definition of the forwarding region in the DREAM protocol.

Once α has been determined, the terminal sends the packet indicating the neighbors lying within the forwarding region. Each intermediate terminal, when receiving a packet, evaluates if the packet should be forwarded by defining its own forwarding region, and by checking if any neighbor lies within it. If no neighbor lies within the forwarding region, packet is not forwarded, reducing thus the overhead by avoiding packet retransmissions in wrong directions. As stated before, the procedure requires positional information regarding not only the source, but also its neighbors and destination. This information is exchanged through a distributed dissemination algorithm, which constitutes the proactive part of the protocol.
Each terminal A periodically broadcasts update packets containing its own position; two types of update packets are defined in DREAM:

- *short lived packets*, for updating the location tables in A's neighbors
- *long lived packets*, for updating the location tables in terminals that are not in direct connectivity with A

Note that the lifetime of each packet is defined in terms of physical distance reached from the terminal originating the packet: each terminal B receiving an update packet generated by terminal A checks both A's position (recorded in the packet) and its own position: if the distance $A - B$ is higher than the packet lifetime the packet is discarded, otherwise it is forwarded. Short lived and long lived packets are emitted by each terminal with different frequencies, taking into account the so-called distance effect, which is represented in fig. 5.10 and can be described as follows.
Given a terminal A moving at speed v, and two terminals B and C that at $t = t_0$ are positioned at distances D_{AB} and $D_{AC} \gg D_{AB}$ from A respectively, terminal B will experience a faster variation of the forwarding region to destination A than terminal C, due to A's movement in the time interval $t_1 - t_0$ (i.e. the angle α_B will increase faster than the angle α_C to keep A in the forwarding region.) As a consequence, B will need much more frequent updates than C regarding A's position: such updates will be provided by means of short lived packets that will reach B without increasing the routing overhead all over the network. C will receive updates at lower frequency by means of long lived packets.

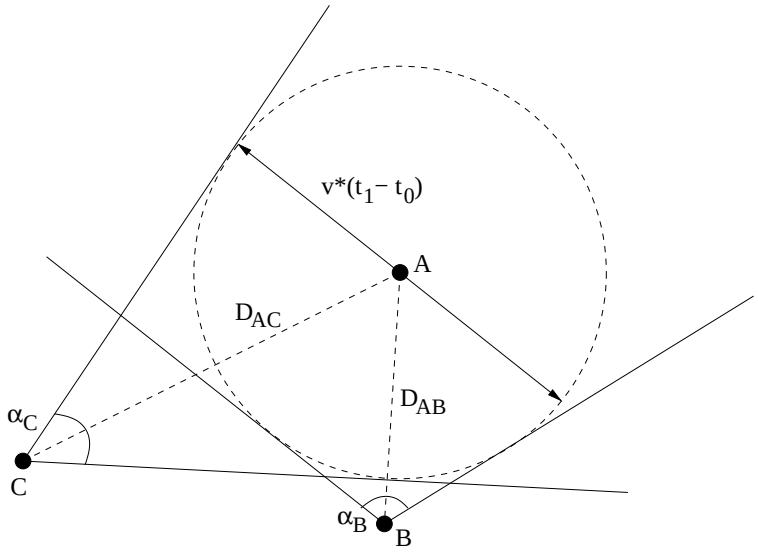

Figure 5.10: Example of distance effect. Given two nodes B and C at distances D_{AB} and D_{AC}, with $D_{AC} > D_{AB}$, the angles α_B and α_C as derived either from eq. (5.16) or eq. (5.17) are such that $\alpha_B > \alpha_C$, and B needs thus more frequent updates than C on the position of A, in order to keep A within the forwarding region.

Each terminal is thus expected to adapt the emission frequencies of both long lived and short lived packets to its own speed: fixed terminals will use the lowest frequencies, while highly mobile terminals will send frequent updates.

5.4.4 Location Aided Routing

The Location Aided Routing (LAR) protocol [94, 95] is a typical on-demand routing protocol. In order to find a route between source and destination terminal, it relies on a flooding-based Route Discovery procedure originated by the source by means of a broadcast RouteReQuest (RRQ) packet. The packet is forwarded by other nodes all over the network until either it reaches the destination (in which case the connection enters in the *Found* status) or a timeout expires leading to a Route Discovery failure. When the destination receives a RRQ packet from the source, it replies with a RouteRePly (RRP) packet which proceeds backward on the selected path, until it reaches the source, that can then start the transmission of DATA packets. A node which detects a connection failure while sending or receiving RRP or DATA packets starts an alarm procedure based on the transmission of broadcast Route ReConstruction (RRC) packets. When an RRC packet reaches the source of the connection, the transmission of DATA packets is stopped, and the source decides either to drop the connection or to start searching for a new path to destination.
The major drawback of a flooding-based on-demand protocol is the huge amount of routing overhead generated during path search procedures. The Location-Aided Routing exploits location information in order to reduce the amount of routing overhead, although in a different way from the DREAM algorithm described in subsection 5.4.3. The LAR protocol uses in fact the location information during connection setup in order to reduce the number of control packets, while no position information is used during DATA packets transmission, since they are sent along the path found during the setup phase. The basic location information required by LAR consists in:

- Source position
- Destination position
- Maximum terminal speed

Such information is exploited during the Route Discovery procedure as follows. Suppose that a terminal S starts a Route Discovery procedure to destination D at time $t = t_1$, and that the last information update regarding D's location was received by S at $t = t_0$. Based on the estimation of the maximum speed v of terminal D, S can evaluate the maximum distance traveled by D since the last location update. Such a distance is given by $d_{MAX} = v \cdot (t_1 - t_0)$. As

a consequence, the current position occupied by D lies in a circular region of radius d_{MAX} centered on $(x_D(t_0), y_D(t_0))$, referred to as the *Expected zone*, represented in fig. 5.11.

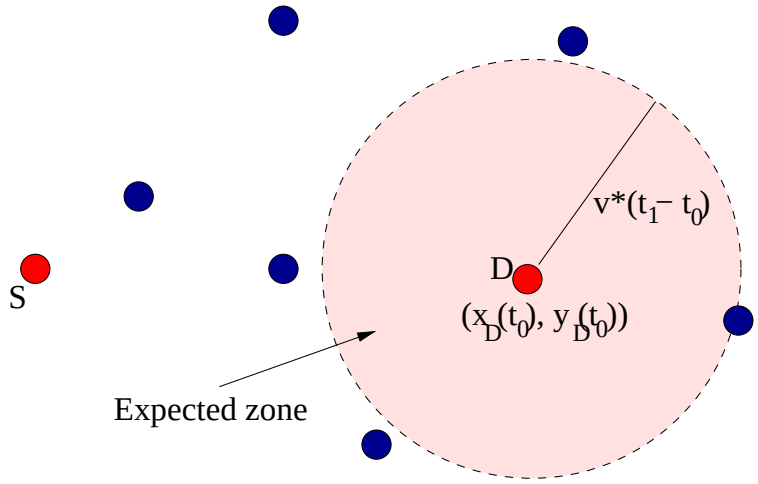

Figure 5.11: Definition of *Expected zone* in the LAR protocol. If source S knows the position of destination D in $t = t_0$, $(x_D(t_0), y_D(t_0))$, and the maximum speed v of D, it can define the region in which D can lie within in $t = t_1$ (*Expected zone*) as a circle of radius $d_{MAX} = v \cdot (t_1 - t_0)$ centered in the position of D in $t = t_0$.

The Expected zone indicates which zone of the network should be reached by RRQ packets. The key idea in LAR is to exploit this information to reduce the amount of RRQ packets flooding through the network, by allowing forwarding of packets generated by the source only in the direction of the Expected zone containing the destination. The region of the network in which forwarding is allowed is referred to as *Request zone*. An intermediate terminal is allowed to forward an RRQ packet only if it lies within the Request zone defined by the source of the connection request.
Note that the Request zone may be defined in several ways, the only constraint being that such zone must include both the position of the source S and the Expected zone. A smaller Request zone leads to a stronger reduction of the routing overhead, and can thus achieve a higher power efficiency; on the other hand, it leads to a lower number of neighbors involved in the Route Discovery procedure, and may cause a lower percentage of successful connections. Figures 5.12 and 5.13 show two different definitions of the Request zone, and the effect on the number of neighbors involved in the Route Discovery procedure. In fig. 5.12 a conic Request zone is considered, similar to the forwarding region

defined in the DREAM protocol, while a spherical Request zone is presented in fig. 5.13.

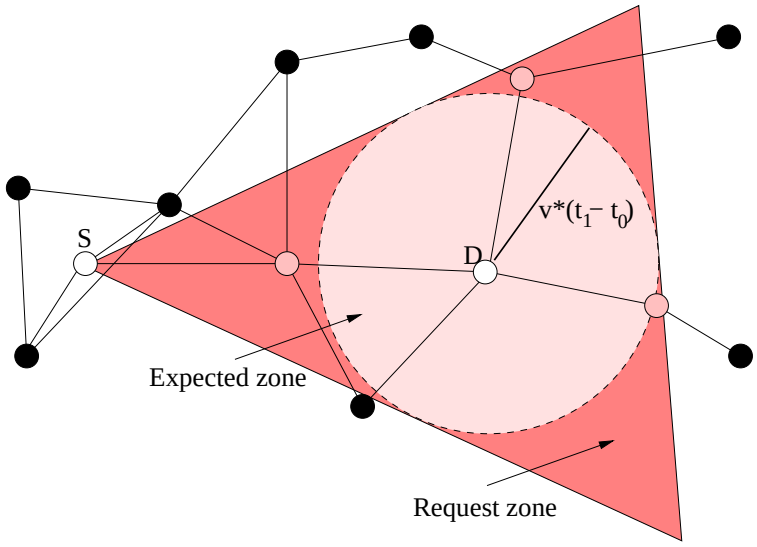

Figure 5.12: Definition of conic Request zone in the LAR protocol. The nodes that fall within the Request Zone defined by source S and destination D are represented in grey, while nodes that are not involved are represented in black.

In both figures nodes that are involved in the Route Discovery procedure are represented in grey, while nodes which are not involved are black. It is evident that the conic Request zone reduces the overhead much more effectively than the spherical one, since a lower number of nodes forwards the RRQ packets. On the other hand, for particular topologies the conic Request zone may lead to connection set-up failures which could be avoided with a larger Request zone. An example of Route Discovery failure is presented in fig. 5.14, showing a situation in which no intermediate nodes in physical connectivity with the source fall within the Request zone.

The choice of size and shape of the Request zone is thus the result of a trade-off between effectiveness in overhead reduction and probability of successful connection set-up.
A more advanced trade-off can be obtained by allowing the source node to perform several attempts of Route Discovery to a given destination and adopting in each attempt an increasingly large Request zone: the result would be a higher percentage of successful connection set-ups, at the price of a higher overhead, obtained as the sum of the routing packets transmitted in each attempt.

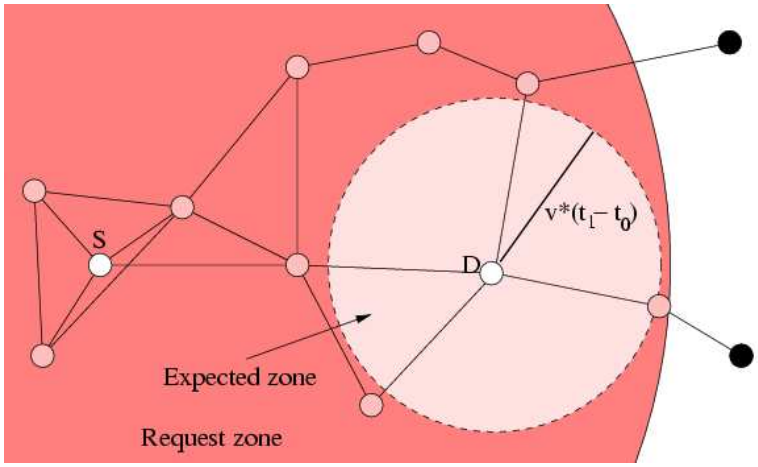

Figure 5.13: Definition of spherical Request zone in the LAR protocol. The nodes that fall within the Request Zone defined by source S and destination D are represented in grey, while nodes that are not involved are represented in black.

An example of the effect of such trade-off is presented in figs. 5.15 and 5.16. Figure 5.15 presents a failure of the LAR protocol adopting a conic Request zone. In this figure, yellow links represent physical connectivity between nodes, the purple area represents the Request zone originating in the source node, and the red node is the intended destination of the connection request. Similarly to the situation presented in fig. 5.14, no neighbor of the source terminal falls within the Request zone, and the connection request fails.
If a second attempt is now performed with an enlarged Request zone, the connection request is successful, at the price of a much higher control overhead; this case is presented in fig. 5.16, highlighting that the price for reaching the destination is a much higher number of nodes involved in the path search procedure and thus an overhead similar to the one obtained if a standard flooding was adopted.
Furthermore, this approach neglects the effect on the latency in connection set-up: since each attempt would require a significant amount of time, the interval between connection request and set-up may turn out to be unacceptable in case of delay-sensitive applications.

As in the case of the DREAM protocol, information about the position of the destination is required at the source in order to reduce routing overhead, and thus location information must be disseminated through the network.

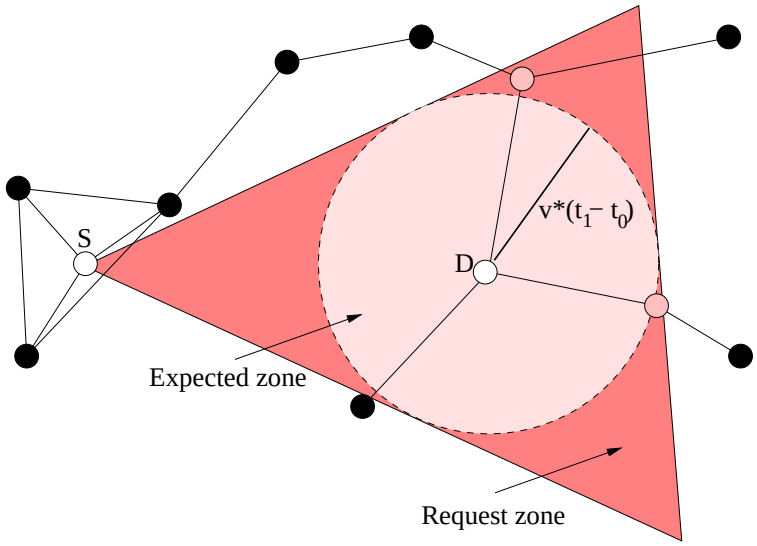

Figure 5.14: Route Discovery failure due to a conic Request zone in the LAR protocol. The nodes that fall within the Request Zone defined by source S and destination D are represented in grey, while nodes that are not involved are represented in black. Note that a Route Discovery procedure generated in S and directed to destination D fails, since none of the grey nodes is in physical connectivity with the source S. A larger Request Zone, involving black nodes, would instead lead to a success.

Oppositely to DREAM, however, in LAR such dissemination is performed by piggybacking location information in all routing packets, without any additional control packet. At the beginning of network operations, terminals will be thus forced to find routes in the absence of location information. In such a situation basic flooding is adopted, following an approach similar to Dynamic Source Routing [1].

5.5 From GPS to UWB positioning

The protocols described in subsection 5.4 share the common assumption that each terminal retrieves its own location information from GPS. As a consequence, protocols focus on how to disseminate location information, assuming the problem of retrieving it as solved.
The adoption of the UWB technology in place of GPS as the basis for retrieving position information opens new doors to location-based applications (such as indoor deployment), but also poses new challenges.

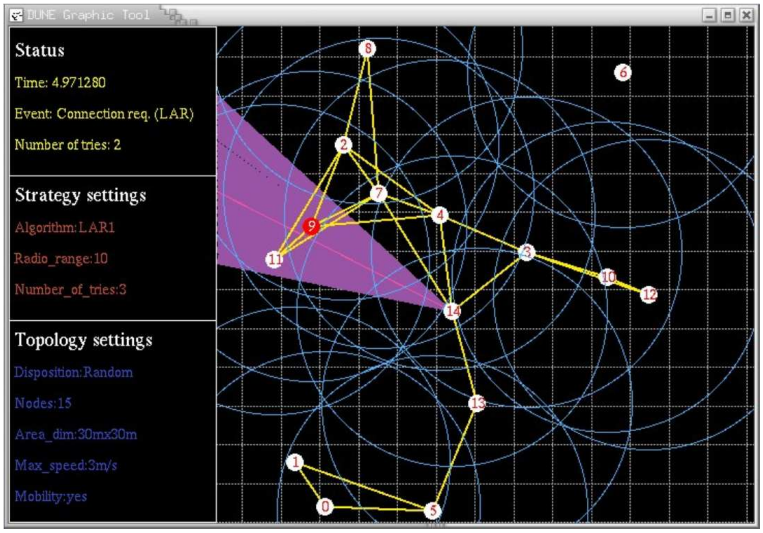

Figure 5.15: Route Discovery failure due to a conic Request zone in the LAR protocol. Yellow links represent physical connectivity between nodes, the purple area represents the Request zone originating in the source node, and the red node is the intended destination of the connection request.

Since UWB can only provide ranging information, in fact, distributed processing of ranging measurements is required in order to build a network map. As analyzed in the previous Chapter, the problem of how to build such a map is not trivial, in particular in the case of a pure ad-hoc network, where no fixed reference points are available. UWB-based positioning poses thus additional requirements, with respect to GPS, to location aided routing protocols:

- Determination of a network-wide coordinate system may require a long time. Furthermore, occasional large errors in position information due to lack of connectivity or unfavorable topology may occur. The routing algorithm must be then capable to find routes and establish connection even in conditions of positioning information incomplete or absent.

- Estimation of absolute terminal speeds is in general not available.

The GPSR protocol requires a location service in order to provide the position of the destination to the source, in order to mark the packet with such information. When this location service is not available from start, as in UWB-based positioning, each data packet must be forwarded by means of flooding, with a high routing overhead.

LLR can be adapted to a UWB-based positioning system with minor changes, but the protocol is inherently tailored for scenarios characterized

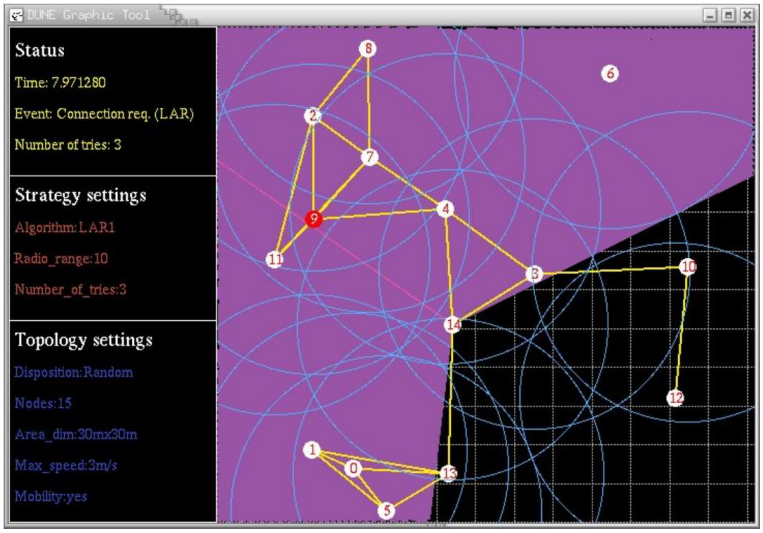

Figure 5.16: Route Discovery success thanks to an enlarged Request zone. Yellow links represent physical connectivity between nodes, the purple area represents the Request zone originating in the source node, and the red node is the intended destination of the connection request.

by terminal mobility (in which route stability is the main concern), and does not offer major advantages in terms of power efficiency in networks composed of still or slowly mobile terminals.

In the DREAM protocol, both selection of routes and dissemination of information rely on positioning to work properly: in particular, the proactive dissemination algorithm exploits the capability of a terminal to determine its physical distance from terminals which are not reachable, to evaluate the validity of an update packet. If location information is not available or subject to errors, both the proactive and reactive parts of the protocol are affected. Furthermore, lack of speed information poses an additional issue to the correct behavior of the protocol.

Oppositely, LAR protocol inherently offers a backup solution when no positioning information is available. In fact, in the case of absence of information regarding the position of the destination, the protocol switches to a basic flooding scheme, resembling a Dynamic Source Routing [1] protocol. In this case the lack of positioning information results in a reduced efficiency of the protocol.

The above considerations led to the choice of adopting LAR as the routing protocol for the integrated power-efficient and location-aware solution presented in the next Chapter.

Chapter 6

Performance analysis

In this chapter we will first summarize the proposed solution for UWB low data rate networks, and then evaluate the performance of this solution in three different network scenarios.

6.1 Summary of the proposed solution

The proposed solution for the deployment of a power-efficient, location aided MAC and routing strategy for UWB low data rate networks is composed of four elements:

The $(UWB)^2$ MAC protocol - introduced in Chapter 3, the MAC is designed for retrieving bidirectional distance information between a couple of terminals, and provide this information to the positioning algorithm.

The Self Positioning Algorithm - the SPA algorithm was selected since it guarantees the capability of obtaining position information in all network scenarios, including those characterized by the absence of anchor nodes. In order to implement this algorithm and combine it with the selected MAC and routing protocols, the first step was the identification of the information units (i.e. the packets) that must be exchanged between the terminals adopting the algorithm. It can be derived from the description of the algorithm provided in section 4.3.2 that a node should exchange three distinct kinds of packets:

- ranging packets, used for estimating the distance between neighbors
- distance update packets, used by a node for sending to the neighbors its up-to-date distance estimation and thus allow the update of the Local Coordinate systems

- reference system update packets, used by a node for sending to the neighbors its coordinate system and allow the creation of a Network Coordinate system.

The exchange of such packets introduces an overhead that, if not properly managed, can reduce or even eliminate at all the advantage of exploiting the position information at the routing layer. In order to minimize such overhead, the original positioning algorithm was modified in order to obtain a strong integration with the underlying $(UWB)^2$ MAC. Such integration is obtained in two ways:

1. the ranging packets, as already explained in Chapter 3, are directly exchanged at the MAC layer, since ranging is performed every time the LE/LC PDU exchange is performed in order to setup the transfer of DATA PDUs;
2. the distance update and the reference system update information are never sent in stand-alone packets; oppositely, when the positioning protocol has information to be transferred to neighboring nodes, it passes this information to the MAC, that buffers this information until a DATA PDU is scheduled for transmission. The information related to the positioning protocol is then piggybacked in the DATA PDU, significantly reducing the overhead caused by the protocol.

Note that the modifications introduced above have the only effect of delaying the transmission of the information generated by the positioning algorithm; such delays, however, do not have a significant impact on the performance of the positioning algorithm itself, since such delays are in most cases negligible when compared to the time intervals between two consecutive distance updates or reference system updates. The comparison performed by means of simulations between the original, stand-alone positioning algorithm and the modified version, integrated with MAC and routing protocols, confirmed that the performance of the positioning algorithm is not affected by the delays introduced to reduce the control overhead.

The LAR routing algorithm - as already stated at the end of the previous Chapter, the LAR protocol was selected because it is capable of operate as well also when no position information is available, falling back to source routing. In order to adopt this protocol in an UWB network, however, some modifications to the original LAR protocol were required. First, the management of position information in the protocol was modified in order to take into account the presence of an UWB-based dis-

tributed positioning algorithm in place of the GPS-based positioning originally foreseen in the protocol. The main difference to be taken into account is that GPS-equipped nodes always share a reference coordinate system, provided by the GPS hardware. Oppositely, in a distributed coordinate system, two terminals may have different coordinate reference systems, and thus may attribute different coordinates to the same node. As a consequence, when a terminal piggybacks its information on the positions of other terminals, it must include in such information the reference system in which the coordinate sets are valid. The knowledge of the reference system will allow terminals that receive the position information to understand if they can use this information to update their own position database (i.e. if they share the same reference system with the terminal sending the information) or if the information must be ignored.
Second, in order to allow a higher reduction of control overhead, the conic Request zone presented in fig. 5.12 was adopted; in order to avoid any additional control overhead, only one connection attempt was allowed, thus not adopting the policy of allowing a new connection request with a larger Request zone after an unsuccessful connection attempt.
Third, in the original LAR protocol a terminal only forwards the first RRQ packet received for a given connection request, in order to avoid loops and further reduce the control overhead; this behavior, however, would not allow the adoption of a metric different from the number of hops, and in particular of the metric described in section 5.3. For this reason, a different approach was adopted during a connection request: an intermediate terminal I, after receiving the first RRQ packet for a connection request, sets a timeout T. Another RRQ packet for the same connection request is only forwarded if the two following conditions are met:

1. The timeout is not expired;
2. The cost attached to the RRQ packet is lower than the cost of the last RRQ packet forwarded for the same connection request.

If an RRQ packet meets the conditions, its attached cost is saved to be compared with the cost of subsequent RRQ packets for the same connection request, and the cost attached to the RRQ packet is updated before it is forwarded, by adding the cost of the last hop, evaluated by I based on the adopted cost function.
Similarly, when the destination terminal D receives the first RRQ packet, it waits for a time T before sending back an RRP packet, and selects the path stored in the RRQ at lowest cost between all RRQ packets received

before the timeout expires.
Note that the behavior of the protocol when the metric is the number of hops is unchanged in the new approach. In fact, in a low data rate scenario, in which traffic congestion is not likely to happen, the first RRQ packet to reach a terminal will be the one that traveled the lowest number of hops from the source, and will be thus the only one to be forwarded.

The UWB cost function - the cost function defined in section 5.3 can be adapted to any traffic scenario by selecting the proper set of coefficients. As a consequence, in the selection of the coefficient set it was taken into account that the target was constituted by low data rate networks, and that the main goal was to minimize power consumption in order to extend network lifetime. This lead to the selection of a set of coefficients that gives the highest relevance to the minimization of emitted power, corresponding to the set B defined in Table 5.1.

In the next sections we will analyze the performance of the above solution. The analysis will focus on two main aspects:

- Effectiveness of the proposed solution in terms of network lifetime and throughput
- Effect of mobility on system performance

To this aim, we first address the problem of mobility modeling, by introducing the mobility models used for simulations and briefly describing their effect on network topology. Next, we define three test cases, each of them with different node mobility characteristics. The behavior of the proposed MAC and routing strategy within each test case is finally studied.

6.2 Mobility models

Introducing mobility within the simulated scenarios allows to test the proposed protocols under more realistic conditions. This is particularly true for routing protocols, where performance greatly varies with the topology of the network. Since different mobility models lead to different topologies, however, it is important to select the mobility model that best fits the reference scenario, that is the typical mobile user that each node in the network represents.

Different mobility models are presented in literature. An interesting subclass is composed by *group* mobility models. This class of models was introduced to emulate particular node behaviors, such as crowds moving towards a common destination or rescue (and similar) squads, where each node is bound

to show some degree of mobility similarity with other nodes belonging to the same group.

6.2.1 Mobility models in literature

Several mobility models for ad-hoc network are proposed in the literature. Among them, we mention:

- the *random mobility* or *Brownian* model [96], with no relation between speed and direction of the node in two subsequent timeslots;
- the *random direction* model [97], with nodes keeping the same velocity during the whole simulation;
- the Ko's mobility model used for evaluating the LAR protocol [94], according to which the path walked by the node is formed by sections with exponentially distributed length and random direction;
- the *Markovian* model [98], which has many moving states and a transit matrix used to determine whether the node should keep or change the current motion direction;
- the *random waypoints* [1], whose movement paths are composed by segments with random speed, direction, and duration, separated by standstill periods, simulating pausing intervals;
- the *Inertia* [99], in which the node moves along a random direction, with random speed, for a random time, and then decides if keeping the same movement characteristic for the next movement segment (the node's inertia), or selecting a new direction, speed, and duration.

All the previous models are, clearly, inadequate to simulate a group movement, since each node has its own movement pattern, which has no relationship with those exhibited by other nodes; in fact, these models where mainly developed with the aim of testing traffic loads offered to the system, rather than reproducing realistic behaviors. In [100] simulation results are reported for the two following non-group mobility models: the *Random waypoints* ([1]), and the *Inertia* ([99]) models.

In the *Random waypoints* model, each node selects a random direction, speed, and movement duration; at the end of the movement, the node stands still for a random time; after the pause time, the node selects a new random direction, speed, and movement duration, and starts walking along the new path. The *Random waypoints* model shares a particular mobility pattern with other models adopting the random direction approach: the "density wave" effect, with a greater concentration of the nodes in the central area.

According to the *Inertia* model, after the end of its movement segment, a node must choose the direction for the next segment. The two alternative choices for the next segment are: (1) the node keeps the current direction, with probability ρ, or (2) the node selects randomly a new direction, with probability $1 - \rho$; the weight of the two alternatives can be different, with an higher probability of keeping the current direction ($\rho > 1/2$). The name *Inertia* refers indeed to the node property of seldomly changing its direction.

The need for group mobility models, showing correlated movements among the nodes and offering flexibility for the implementation of particular behaviors, has lead to the design of other models; among these, the Exponential Correlated Random (ECR) model [101]. Other approaches are proposed in literature, in order to bypass the limitation of ECR; Those reviewed here are the *Reference Point Group Mobility* model ([102]), and the *Reference Velocity Group Mobility* model ([103]), and *Kerberos* ([100]).

- The *Reference Point Group Mobility* (RPGM) model, presented by Hong *et al.* in [102]. To represent the group mobility behavior of the mobile nodes, the model defines, for each group, a logical *reference point*, whose movement is followed by all nodes in the group. The path the reference point moves along defines the entire group mobility behavior, including location, speed, direction, acceleration, etc. Thus the group trajectory is determined by providing a path for the reference point, with any node in the group randomly placed in the neighborhood of its reference point.

  Once next position (x, y) of the reference point is determined (the path of the reference point can be selected with any non group-oriented mobility model), each other node position is calculated adding a random motion vector $\overrightarrow{RM}$ to (x, y).

  The main RPGM characteristic is that the groups are characterized by *physical proximity*, that is all the nodes are close to the reference point.

- The *Reference Velocity Group Mobility* (RVGM) model, presented in [103] as an evolution of RPGM. The fundamental concept proposed is that movement similarity is a more fundamental characteristic for group mobility than physical proximity; thus nodes being part of the same group should show close velocities and directions, rather than physical proximity.

  Each group, therefore, is characterized by a *group velocity*. The nodes member of the group have velocities close to the characteristic group velocity, with slight deviations. Hence, the characteristic group velocity is also the mean group velocity. The distributions of both the group velocity and the velocity deviation can be of any type, in order to model the various possible mobility patterns.

- *Kerberos*, presented in [100]. The basic concept in *Kerberos* mobility model is that each node in a group is allowed to move around freely, provided it keeps contact with the other nodes in its group.

  This basic concept is translated into the adoption of a condition of physical proximity between the given node and the other nodes of its group: when this condition is met, the node moves according to a non-group mobility model (e.g., *Inertia*), while, when the node is far from its fellows, it is obliged to move closer to them.

  Kerberos requires that, after each mobility update, the node must check how many nodes of its own group it is connected to, either directly or through other nodes in the group: this set of nodes forms the *fellowship* of the node. If the number of nodes in the fellowship is greater than half the number of nodes forming the group, the node is allowed to move freely; otherwise, the node is separated by the main chunk of group (if any exists), and is compelled to move towards the closest group node not belonging to its fellowship.

The above models are compared in [100] in terms of the following metrics, to be applied to the evaluation of mobility models impact on routing performance:

- **Number of link changes**: is the number of transitions from the state "connected" to the state "disconnected" and vice versa, for the link between any pair of nodes in the network. Its average over the number of links in the network is the **average Number of Link Changes**, and is calculated separately for links between nodes of the same group and for links between nodes of different groups;

- **Link Duration**: is the average duration of the link between two nodes, that is the average time that two nodes remain continuously within range. Its average over the number of links in the network is the **average Link Duration**, and is calculated separately for links between nodes of the same group and for links between nodes of different groups;

- **Path Availability**: is the fraction of time during which a path of links within transmission range is present between two nodes in the network. Its average over the number of links in the network is the **average Path Availability**, and is calculated separately for paths between nodes of the same group and for paths between nodes of different groups.

- **Link Availability**: is the percentage of time during which link between two given nodes is active, i.e. the two links are within transmission range. Its average over the number of links in the network is the **average Link**

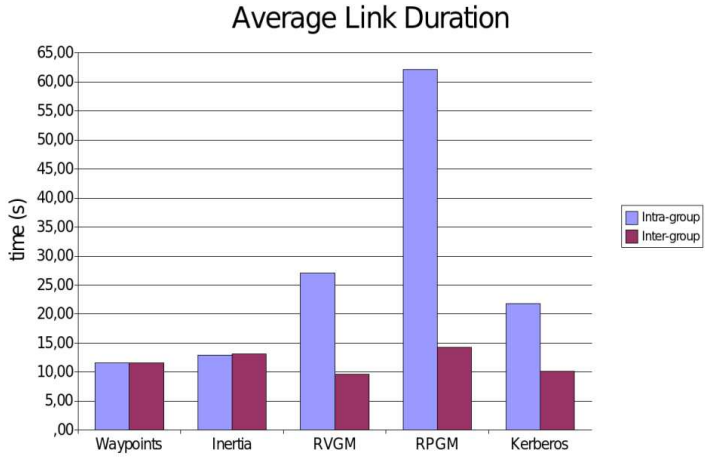

Figure 6.1: Average link duration, in seconds, for different mobility models.

Availability, and is calculated separately for links between nodes of the same group and for links between nodes of different groups.

In the following, we will report results regarding a subset of the above metrics that will be useful in explaining the behavior of the MAC and routing strategies presented in the next section.

Link Duration

The *average Link Duration* of each mobility model is shown in fig. 6.1, for Intra- and Inter-group links; the values are expressed in seconds.

This metric tells apart the group by the non-group mobility models: the formers show a high difference in average Link Duration between intra- and inter-group links. In the case of *Kerberos* and, above all, RPGM, the *physical proximity* property is what make the link more stable: the nodes forming the intra-group link are very likely to remain close in the next future. On the other hand, *Kerberos* shows the lowest *average Link Duration* among group mobility models. Between the non-group mobility models, Inertia has a slightly higher *average Link Duration* than Random Waypoints. This is due to the fact that the two nodes composing the link will be more likely to keep on moving in the same direction if they are modeled with Inertia, rather than with Random Waypoints.

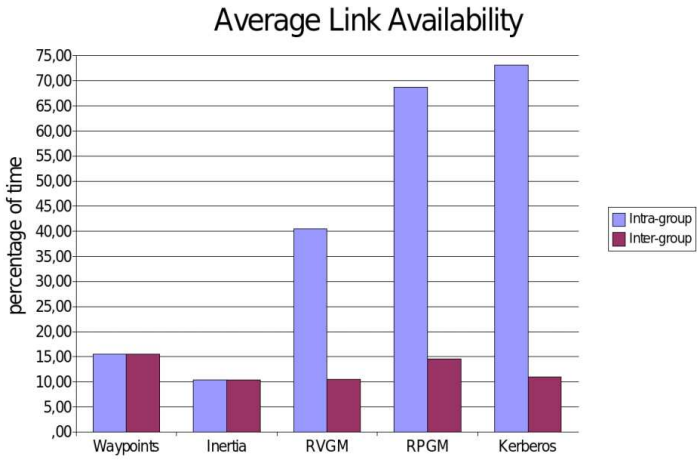

Figure 6.2: Average link availability, in percentage of time, for different mobility models.

Link Availability

Figure 6.2 shows the results for the *average Link Availability* metric, expressed in percentage of time in which a given link is established.

Results in fig. 6.2 show that group mobility models, in general, guarantee a higher probability of having a link between nodes in the same group, while no clear difference can be observed between group and non-group mobility models when Inter-group links are considered.

Path Availability

Simulation results for the *average Path Availability* are shown in fig. 6.3. The Intra-group paths are those in which the starting and the ending point of the path belongs to the same group.

The difference between Random Waypoints and Inertia is given by the fact that Random Waypoints scatters the nodes all over the simulation area: the probability of network partition (which implies a dramatic decrease in *average Path Availability* values) is therefore high, but is is lower than in Inertia, where some nodes tend to be separated by the remaining network for long times. As regards the group mobility models:

- *Kerberos* shows the highest *average Path Availability* for Intra-group paths. With the adoption of this mobility model, the network shows highly connected groups that span over wide areas. Intra-group links are less frequent, but they allow a large number of path to exist;

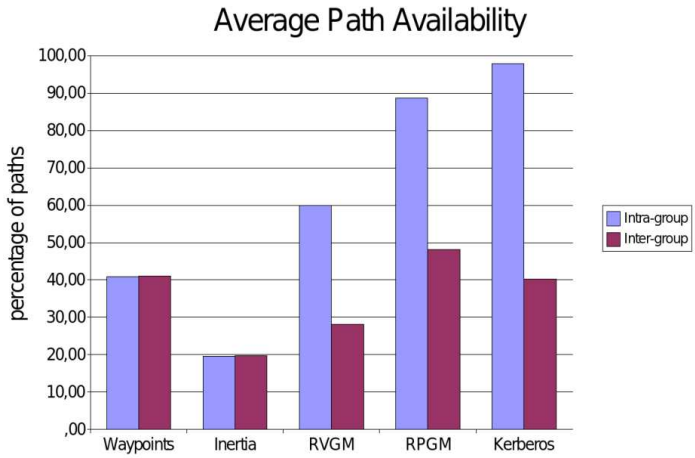

Figure 6.3: Average path availability, in percentage of paths, for different mobility models.

- RPGM shows a lower *average Path Availability* for Intra-group paths, but higher than *Kerberos* for Inter-group ones. Physical proximity requirement allows for a few group partitions, but the groups cover wider areas than *Kerberos*, thus increasing the probability of an Inter-group link, which sustains a number of Inter-group paths;
- RVGM results are due to the aforementioned property of similarity in movements, rather than in physical proximity, for the nodes of a group. The resulting values are similar to an average between the non-group and the remaining group mobility models.

Simulation results show that different mobility models have quite different properties, as regards the *Path Availability*. Non-group mobility models show no difference between Intra- and Inter-group paths, since all the nodes are moved independently from their group fellows, while the opposite is true for group mobility models: comparing, for example, the Inertia and the kerberos, the latter shows a higher availability for Intra-group path than the former, while the opposite holds for Inter-group paths.
The results above show the effect of different mobility models on topology-related parameters. The actual effect of mobility on network performance depends however also on the traffic pattern in the network, as it will shown in the next section.

6.3 Test cases

We will investigate three different test cases, which are characterized by the following common scenario: a network of N terminals deployed in initial random positions. Each terminal periodically generates a connection request to a random destination, following a Poisson distribution with average time λ between two subsequent requests. Each connection request is characterized by a constant bit rate R_c and a total number of bits to be transferred which is randomly selected in the interval [1, $MaxDimBit$]. Each terminal is furthermore characterized by a total amount of energy E_{TOT}, which is reduced after each packet transmission or reception based on the following energy model [104]:

$$E_{TX} = E_{start} + L \cdot (E_{Tx\_bit\_rate}(R_b) + E_{Tx\_bit\_prop}(R_b) \cdot d^{\alpha}) \tag{6.1}$$

$$E_{RX} = E_{start} + L \cdot (E_{Rx\_bit\_fixed} + E_{Rx\_bit\_rate}(R_b)) \tag{6.2}$$

where L is the length of the packet.
Note that, although this model is not specific for UWB, it addresses a class of devices which is close, in terms of achieved bit rate and complexity, to the UWB devices foreseen for low bit rate networks.

All test cases share furthermore the settings adopted for positioning and routing protocols, in particular data and control packet sizes. Finally, the initial amount of energy E_{TOT} for each node is the same in all the three test cases, although its actual value depends on the transmission range.

The three test cases differ however in the mobility pattern of the terminals:

Test case 1 Network of still terminals in random positions

Test case 2 Network of mobile terminals in initial random positions, following the *Inertia* mobility model

Test case 3 Network of mobile terminals in initial random positions, following the *Kerberos* mobility model

The initial selection of terminal positions can heavily influence the network performance, in particular in the scenario considered in Test case 1: as a consequence, in order to determine the average performance of the strategies, N_{runs} simulation runs were executed, with terminals deployed in random positions in an area of $A_{side}xA_{side}$ m^2. Given a set of terminals in random coordinates, however, network topology is also determined by the transmission range, which influences the network connectivity. For this reason, tests were executed for three different transmission ranges, R_{TX_low}, R_{TX_med} and R_{TX_high}, leading to low, medium and high network connectivity, respectively.

Parameter	Value
R_b	1 Mb/s
R_c	50 Kb/s
λ	50 s
$MaxDimBit$	1 Mbit
N	20
α	4
E_{start}	$2.76 \cdot 10^{-5} J$
$E_{Tx\_bit\_rate}$	$3.25 \cdot 10^{-7} J$
$E_{Tx\_bit\_prop}$	$1.25 \cdot 10^{-11} J$
$E_{Rx\_bit\_fixed}$	$1.13 \cdot 10^{-7} J$
$E_{Rx\_bit\_rate}$	$2.79 \cdot 10^{-7} J$
A_{dim}	80 m
N_{runs}	10
R_{TX_low}	20 m
R_{TX_med}	40 m
R_{TX_high}	60 m
RRQ size	760 bit
RRP size	760 bit
$DATA$ size	5000 bit
$ACKDATA$ size	250 bit
RRC size	250 bit

Table 6.1: Simulation settings common to all test cases.

The common settings to all tests cases are presented in Table 6.1, while Table 6.2 presents the initial energy per node for each transmission range.

In each test case four different strategies were compared, which combine in different ways the two key components of the proposed solution: location aided routing algorithm and power-aware routing metric. The distinction between the routing protocol and the routing metric allows in fact to evaluate the effect of each of the two components on system performance.
The four strategies are as follows:

- Dynamic Source Routing with hop minimization (DSR+Hop)
- Dynamic Source Routing with cost function minimization (DSR+Cost)
- Location Aided Routing with hop minimization (LAR+Hop)
- Location Aided Routing with cost minimization (LAR+Cost)

Range	**Energy** E_{TOT}
R_{TX_low}	25 J
R_{TX_med}	250 J
R_{TX_high}	800 J

Table 6.2: Initial energy amount per node for each of the three transmission ranges considered in simulations.

As previously stated, the performance of each strategy was analyzed under conditions of limited available energy taking into account two aspects: long term system performance, i.e. how long the network is alive and thus capable of transferring data, and short term system performance, i.e. how network behaves during its life. The following indicators were selected to measure long term and short term performance:

- Number of Found Connections
- Number of DATA packets delivered to destination
- Percentage of Found Connections
- End-to-end throughput for DATA packets

A Found Connection is defined as a connection in which the destination terminal, following the reception of a RouteReQuest (RRQ) packet, is able to send back to the source terminal a RouteRePly (RRP) packet. The number of total number of Found Connections is a good indicator of the long term behavior of the network since, given the constant rate λ of connection requests generation, a longer network lifetime will lead to a higher number of Requested Connections and, eventually, of Found Connections. Since each connection, however, is set up in order to transfer DATA packets, we also consider the total number of DATA packets transferred during network lifetime, in order to determine which is the impact of each strategy at packet level.
The percentage of Found Connections, oppositely, allows to measure the network behavior on the short term, by measuring how good each connection is served by the network. The end-to-end throughput provides the same kind of short term information at the packet level.
Note that the evaluation of the short term performance behavior is fundamental in the comparison of the different strategies since, from a theoretical viewpoint, a strategy could lead to an overall high number of Found Connections, thanks to a long lifetime which allows for a high number of Requested Connections, but at the price of a low percentage of Found Connections.

6.3.1 Test case 1

This test case analyzed a scenario with terminals in fixed locations, which are unknown at beginning of operations.
Since terminals could not move, network topology only changed when a terminal ran out of energy: as a consequence, in this scenario, performance was significantly affected by the transmission range. Furthermore, as we will see, different strategies were affected in a different way by variations in network connectivity.
Figure 6.4 presents the number of Found connections under conditions of low network connectivity. Note that in fig. 6.4 only results related to DSR-based strategies are presented: the reason for this choice is the fact that, as shown in section 4.4, in conditions of low network connectivity most terminals will not be able to achieve position information, and the LAR will thus fall back to the standard DSR approach. For the same reason, also in Test Cases 2 and 3 only DSR strategies were considered for the case of low connectivity.
The performance of the two strategies is quite similar, since the main factor in determining network behavior is the degree of network connectivity. This conclusion is also supported by the other indicators, i.e. the number of DATA packets, the percentage of Found connections, and the DATA throughput, presented in figs. 6.5, 6.6 and 6.7.

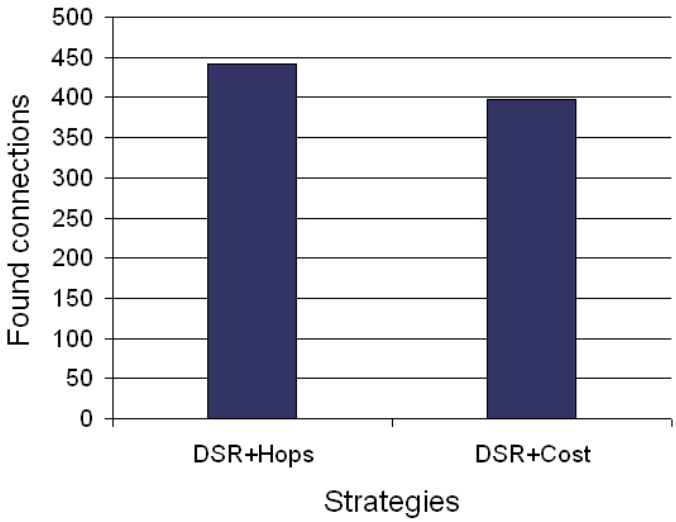

Figure 6.4: Found connections for Test Case 1 with R_{TX_low}.

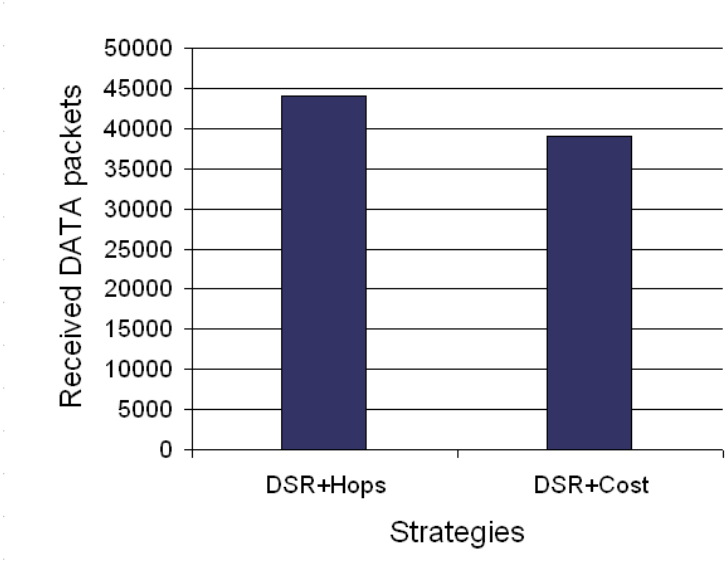

Figure 6.5: Received DATA packets for Test Case 1 with R_{TX_low}.

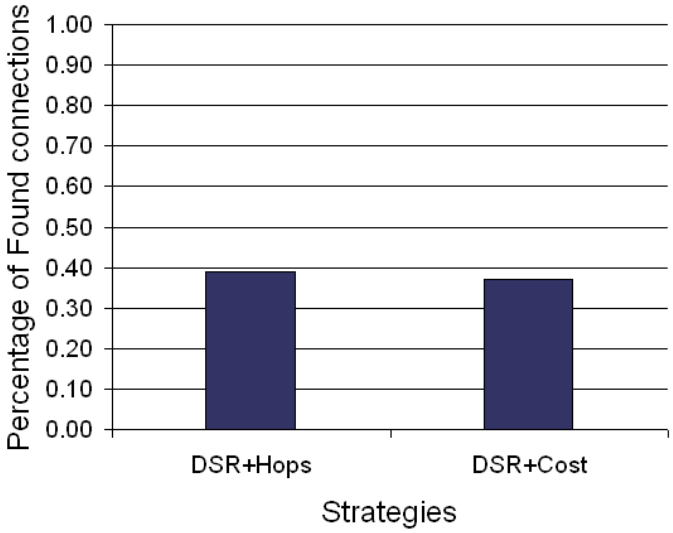

Figure 6.6: Percentage of Found connections for Test Case 1 with R_{TX_low}.

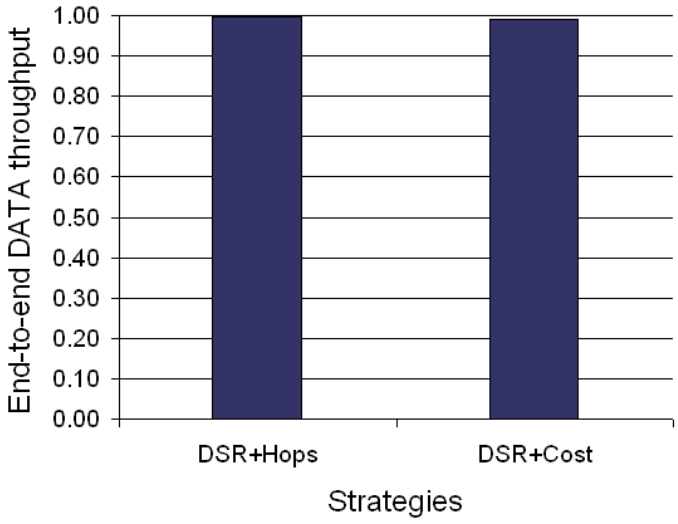

Figure 6.7: End-to-end DATA throughput for Test Case 1 with R_{TX_low}.

Interestingly, the adoption of the power-aware cost function in place of number of hops leads to slightly worse performance in terms of number of Found Connections and of delivered DATA packets, indicating a shorter network lifetime. This can be explained by observing that when the transmission range is low, the energy costs in transmission and reception which are not related to propagation are more relevant in the overall energy consumption, thus eliminating the advantage of increasing the number of hops per connection. As an example, let us consider the case of the transmission of a DATA packet over a distance $d = 10$ m. The overall energy cost of the transmission (E_{TX}+E_{RX}), based on the model in eqs. 6.1 and 6.2 is given by:

$$\begin{aligned} E &= 2 \cdot E_{start} + L \cdot (E_{Rx\_bit\_fixed} + E_{Rx\_bit\_rate}(R_b) + E_{Tx\_bit\_rate}(R_b)) + \\ &\quad + L \cdot (E_{Tx\_bit\_prop}(R_b) \cdot d^{\alpha}) \\ &= 5.52 \cdot 10^{-5} + 5000 \cdot (7.17 \cdot 10^{-7}) + 5000 \cdot (1.25 \cdot 10^{-7}) \\ &= 4.27 \cdot 10^{-3} J \end{aligned} \tag{6.3}$$

If two hops over $d = 5$ m are used in place of the single hop, we get the overall cost:

$$
\begin{aligned}
E &= 4 \cdot E_{start} + 2 \cdot L \cdot (E_{Rx\_bit\_fixed} + E_{Rx\_bit\_rate}(R_b) + E_{Tx\_bit\_rate}(R_b)) + \\
&\quad + 2 \cdot L \cdot (E_{Tx\_bit\_prop}(R_b) \cdot d^{\alpha}) \\
&= 1.104 \cdot 10^{-4} + 2 \cdot 5000 \cdot (7.17 \cdot 10^{-7}) + 2 \cdot 5000 \cdot (7.8125 \cdot 10^{-9}) \\
&= 7.35 \cdot 10^{-3} J
\end{aligned} \tag{6.4}
$$

which is higher than the cost of the single hop at longer distance.

As already said, the adoption of LAR, based on UWB-positioning, would not lead in these conditions to any advantage due to the failure of the positioning protocol. It should be noted however that even in the case of GPS-based positioning (i.e. assuming that each terminal knows its own position) the performance increase guaranteed by the adoption of LAR would be far from dramatic, as confirmed by simulations performed under this hypothesis. This can be explained by observing that the main advantage of LAR is the reduction of overhead by avoiding that broadcast packets are forwarded in wrong directions. Under condition of low connectivity, the number of broadcast packets is inherently limited by the low number of links available in the network, thus achieving the same effect pursued by LAR, at the price of a low percentage of Found connections.
If we consider a higher transmission range, on the other hand, the above considerations are no longer valid, and we would thus expect an advantage in terms of network lifetime by adopting the power-aware cost function and the LAR protocol.

The results for transmission range set to R_{TX_med} are presented in figs. 6.8 and 6.9 for Found Connections and delivered DATA packets, respectively. Results show that in both cases the LAR+Cost strategy leads to the best performance, increasing by a factor of 2 both the number of Found connections and the number of DATA packets. Furthermore, the adoption of the power-aware cost function leads to better performance independently of the selected routing algorithm.

Noticeably, also the short term performance indicators, presented in figs. 6.10 and 6.11, are improved by the adoption of LAR routing and power-aware cost function.

The advantage of the LAR+Cost solution is confirmed when the R_{TX_high} transmission range is considered: fig. 6.12 shows that the number of Found connections is increased by a factor of 2 by adopting the LAR+Cost strategy. The increase in Found connections obtained by adopting the Cost function

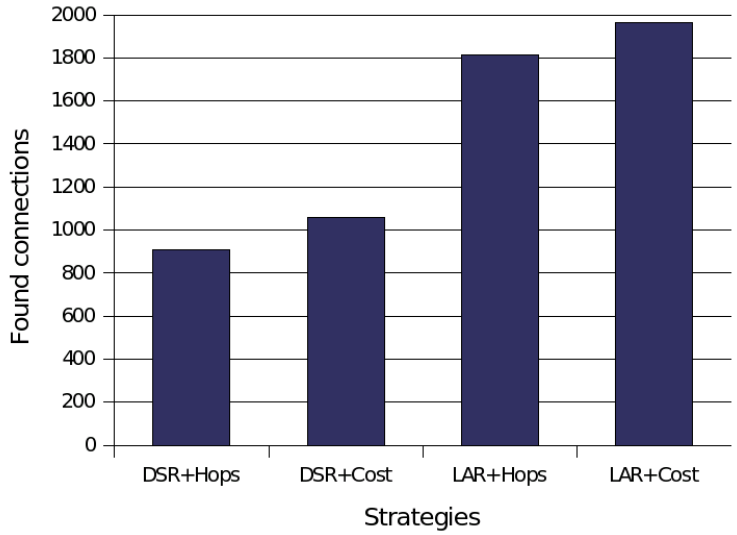

Figure 6.8: Found connections for Test Case 1 with R_{TX_med}.

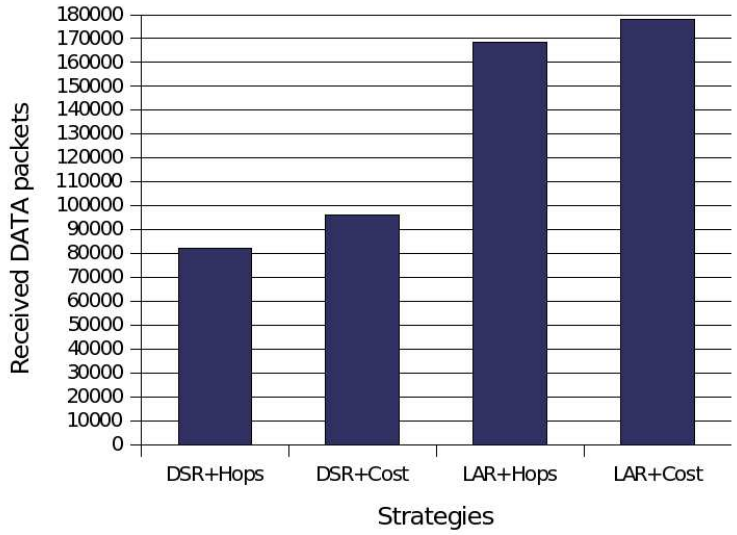

Figure 6.9: Received DATA packets for Test Case 1 with R_{TX_med}.

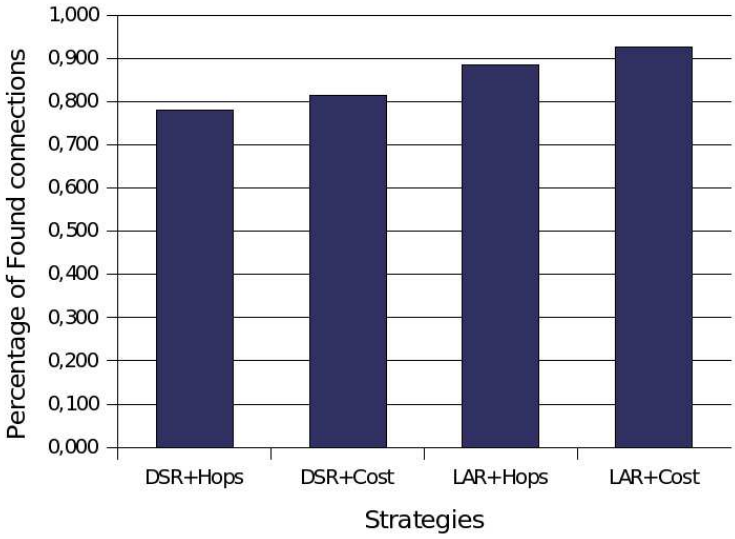

Figure 6.10: Percentage of Found connections for Test Case 1 with R_{TX_med}.

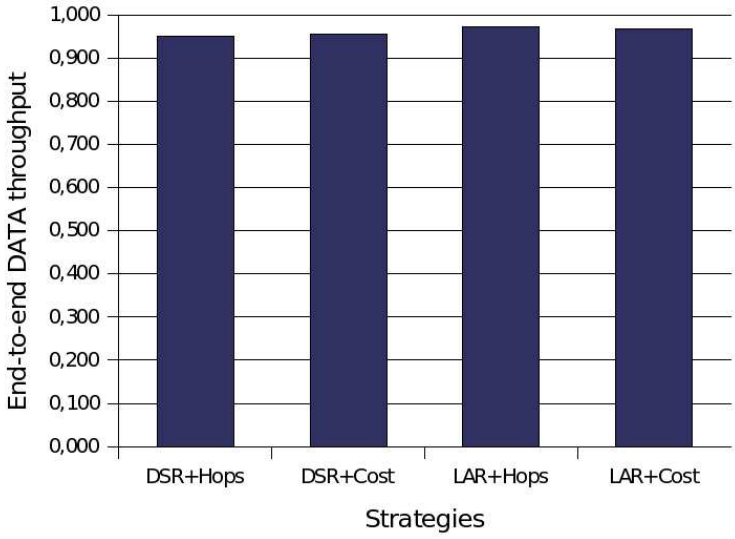

Figure 6.11: End-to-end DATA throughput for Test Case 1 with R_{TX_med}.

is around 10%, as it was in the case of transmission range set to R_{TX_med}. This is coherent with the fact that, given the number of terminals and the size adopted in the Test Cases, the average distance between two terminals is 40 m. As a consequence, the maximum advantage from the Cost function is obtained when the transmission range reaches 40 m, i.e. for R_{TX_med}. Further increases of the transmission range do not lead to any additional gain, since the energy saving obtained with the Cost function is not related to the maximum transmission distance reachable by a terminal, but to the actual distance between terminals.

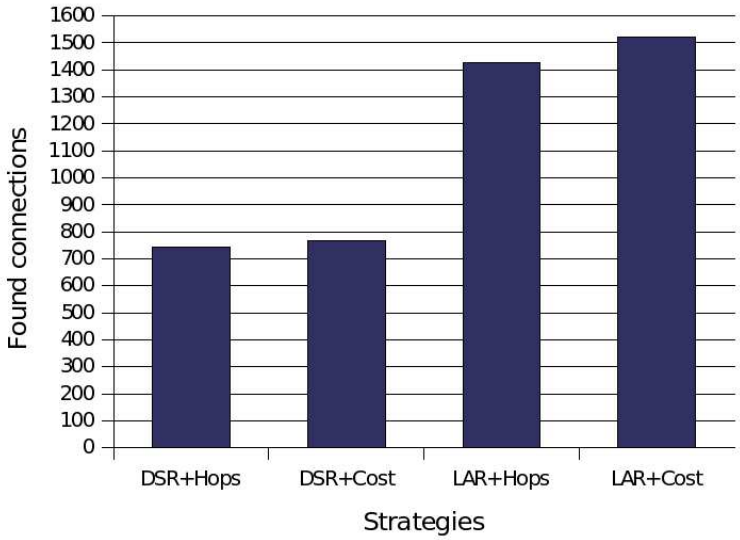

Figure 6.12: Found connections for Test Case 1 with R_{TX_high}.

6.3.2 Test case 2

This test case analyzed a scenario in which terminals were randomly deployed, and then moved following the *Inertia* mobility model, with mobility settings presented in Table 6.3.

Parameter	Value
V_{MAX}	6 m/s
ρ	0.5

Table 6.3: *Inertia* mobility settings for Test case 2.

The movement capability of terminals had a strong impact on network behavior, in particular in scenarios where network connectivity was more affected by mobility. This is the case when low transmission range R_{TX_low} is considered: the introduction of mobility led indeed to results which are completely different from those obtained in Test case 1 with still terminals.
Figure 6.13 shows the number of Found connections in this scenario.

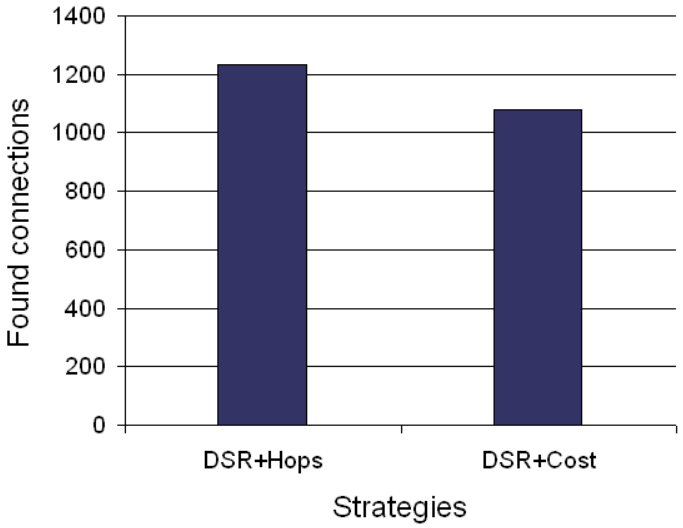

Figure 6.13: Found connections for Test Case 2 with R_{TX_low}.

Note that in conditions of low connectivity and mobility, the adoption of Cost instead of Hops does not bring any advantage. As already observed for Test Case 1, in fact, the low transmission range leads in most cases to an higher energy consumption when moving from a single hop to two or more hops. In presence of mobility, furthermore, the Cost-based strategies suffer from an additional increase of energy spent in signaling connection failures due to mobility by means of broadcast RRC packets. In fact, since the Cost function leads to a higher average number of hops, as shown in fig. 6.14, the paths selected with this metric are more subject to failures caused by terminal mobility.

If we consider a transmission range set to R_{TX_med}, we observe, on the contrary, that the adoption of the Cost function significantly increases the number of Found connections. In this conditions the best solution is thus the LAR+Cost strategy, as shown in fig. 6.15.

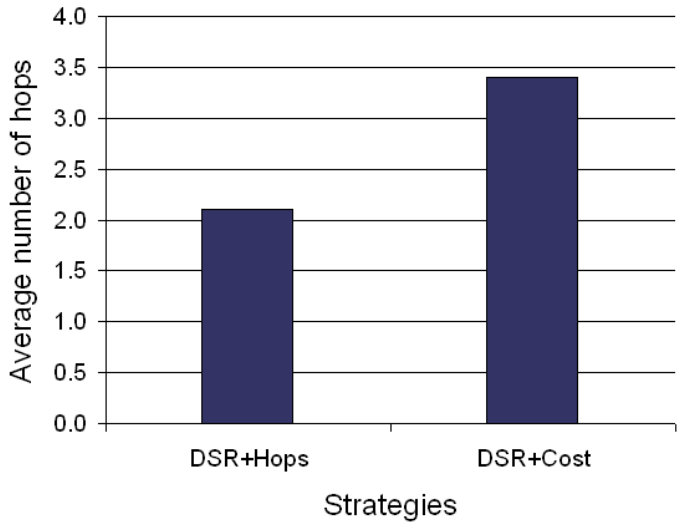

Figure 6.14: Average number of hops for Test Case 2 with R_{TX_low}.

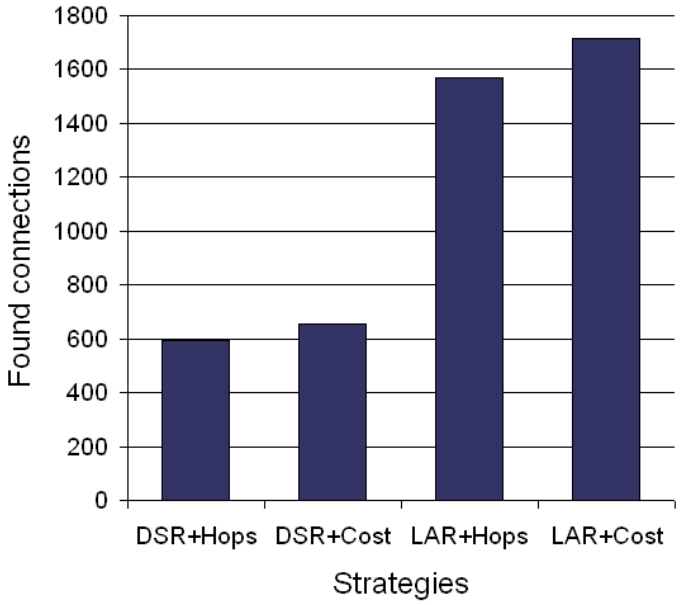

Figure 6.15: Found connections for Test Case 2 with R_{TX_med}.

This conclusion is also confirmed by the other long term parameter, i.e. the number of received DATA packets, shown in fig. 6.16.

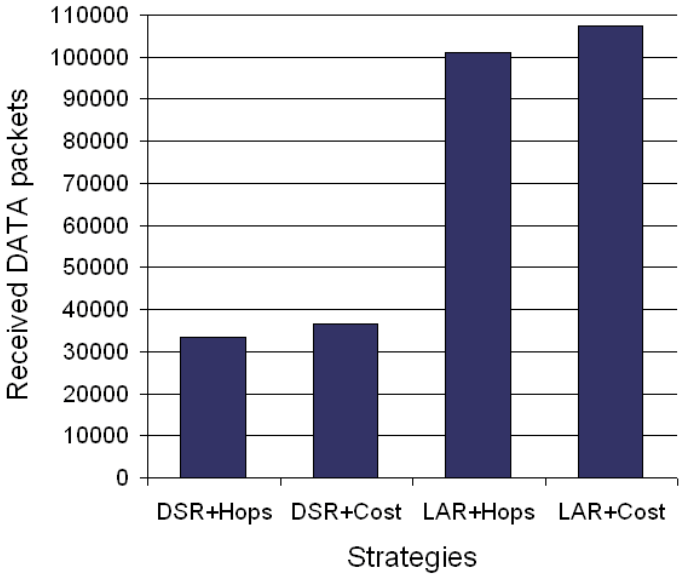

Figure 6.16: Received DATA packets for Test Case 2 with R_{TX_med}.

The short term parameters confirm that LAR+Cost is overall the best choice: the throughput, presented in fig. 6.17, is almost identical for both LAR-based strategies, and the same is true for the percentage of Found connections, in fig. 6.18.

The advantage obtained by the LAR-based strategies on the short term parameters is also influenced by the longer network lifetime. As an example, fig. 6.19 shows the evolution of the percentage of Found connections in time for a single simulation run. The plot shows that the difference on the average value (fig. 6.18) is due to the capability of LAR-based strategies of keeping the network in steady state for a longer period of time.

The case of high network connectivity confirms LAR-based strategies as the most efficient ones (see fig. 6.20), but in this case the performance gain achieved with the Cost function reduces to about 10%.

This can be explained by considering that, as already observed for Test Case 1, an increase of transmission range over the average distance between two terminals does not bring any additional advantage for Cost-based strate-

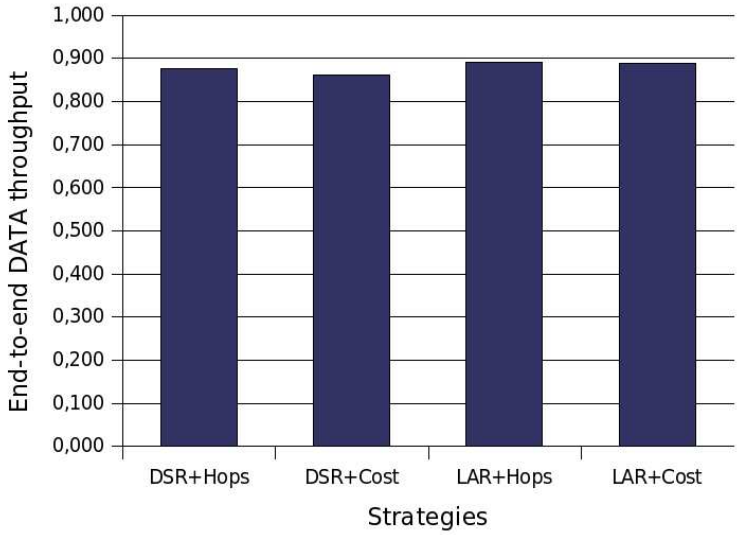

Figure 6.17: End-to-end DATA throughput for Test Case 2 with R_{TX_med}.

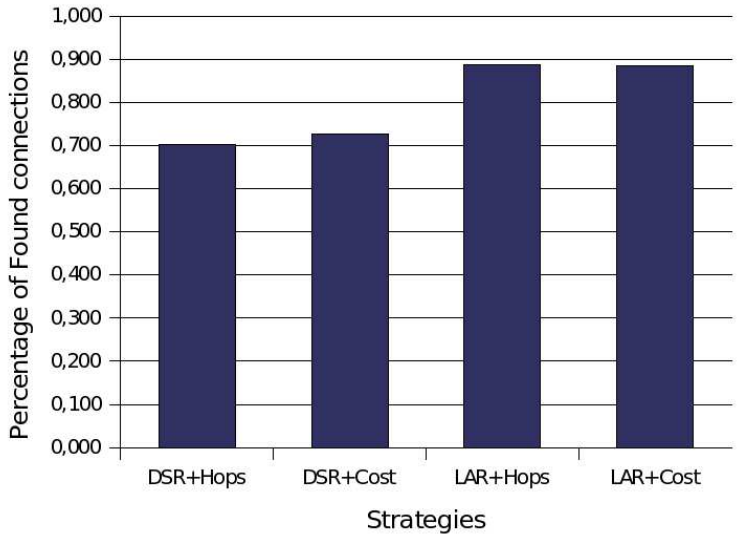

Figure 6.18: Percentage of Found connections for Test Case 2 with R_{TX_med}.

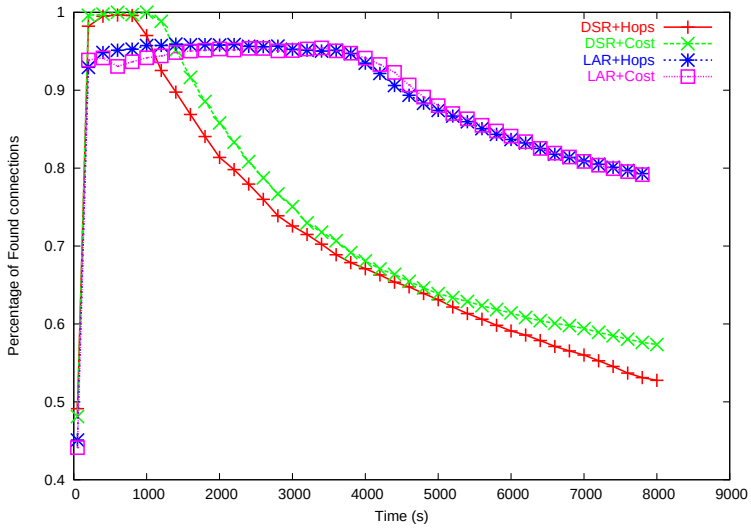

Figure 6.19: Percentage of Found connections for Test Case 2 with R_{TX_med} (single run).

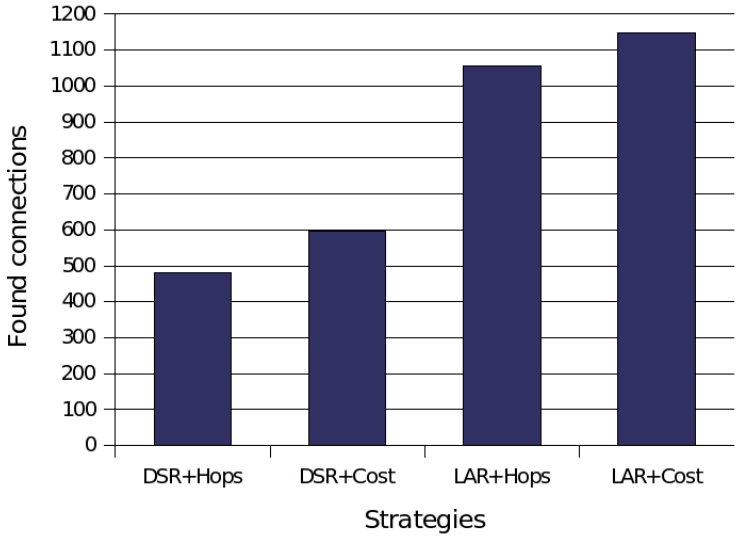

Figure 6.20: Found connections for Test Case 2 with R_{TX_high}.

Parameter	Value
V_{MAX}	6 m/s
Kerberos Range	$R_{TX} - V_{MAX}/3$ m
N_{neigh}	4

Table 6.4: *Kerberos* mobility settings for Test case 3.

gies. At the same time, the negative effect of connection failures due to mobility increases with transmission range, since RRC packets are transmitted at maximum power; this means that each RRC packet will consume higher energy, and that each packet will be forwarded for a higher number of times, thanks to the higher network connectivity. As a consequence, the efficiency of Cost-based strategies is reduced.

6.3.3 Test case 3

This test case analyzed a scenario in which terminals were randomly deployed, and then moved following the *Kerberos* mobility model, with mobility settings presented in Table 6.4.

The analysis of the performance of the four strategies in the case of the *Kerberos* mobility model is somewhat more difficult than in the other cases; the peculiar spatial distribution of terminals generated by this group mobility model must be taken into account for a correct evaluation of the simulation results, and the distinction between intra-group and inter-group network topology properties must be considered.
The global number of Found connections in the case of low transmission range is presented in fig. 6.21. The figure shows that, as much as in the case of Test Case 2, the presence of mobility allows for a higher network connectivity and thus leads to better results than in the case of still terminals. For both DSR-based strategies, however, performance is significantly worse than in the case of the *Inertia* mobility model.

This is due to the fact that the *Kerberos* model does not lead to an uniform distribution of nodes, but creates groups of nodes. Figure 6.23 shows in fact the average distance between two nodes in the network and the average distance between two nodes in the same group for both *Inertia* and *Kerberos*. As one could expect, the two distances are similar in the case of *Inertia*, since no special group behavior is defined in the model. Oppositely, in the case of *Kerberos* a large difference between the two values can be observed.

The lower performance observed for both strategies in Test Case 3 is due to the higher connectivity which characterizes the *Kerberos* model. Results in subsection 6.2.1 show in fact that this mobility model guarantees a higher average connectivity than the *Inertia* model; this translates in a higher number of

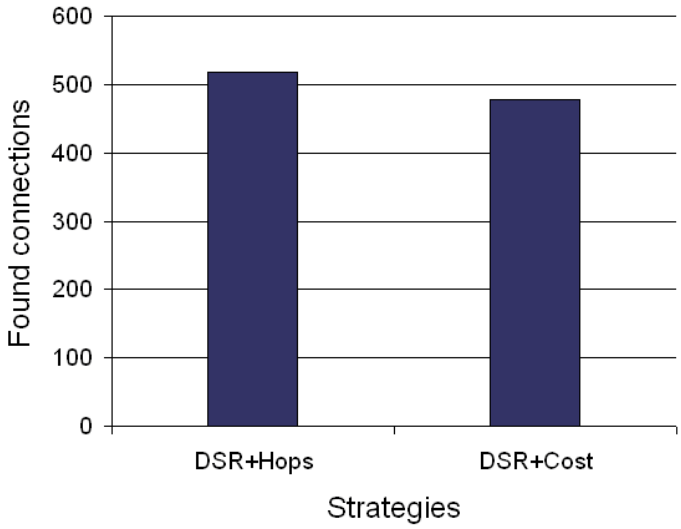

Figure 6.21: Found connections for Test Case 3 with R_{TX_low}.

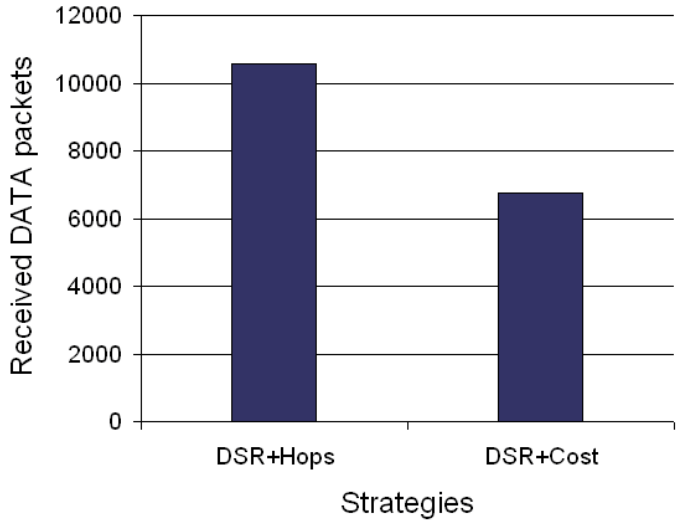

Figure 6.22: Received DATA packets for Test Case 3 with R_{TX_low}.

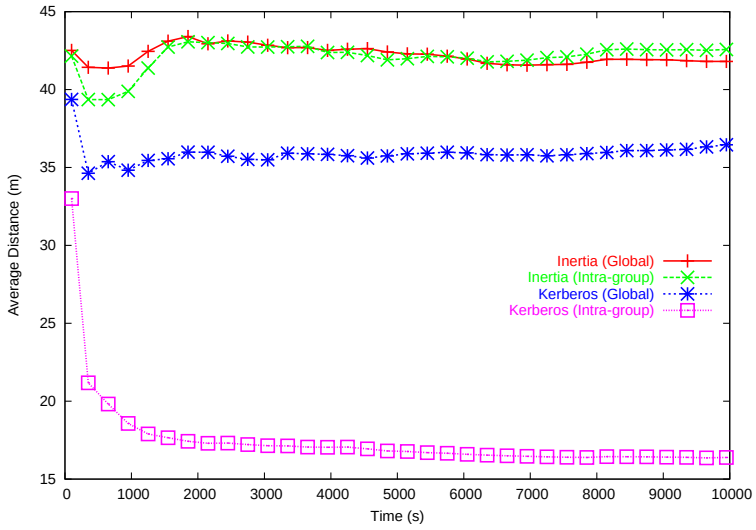

Figure 6.23: Average distances between two nodes in the network (Global) and between two nodes in the same group (Intra-group) for *Inertia* and *Kerberos*, the latter with *Kerberos Range* $= R_{TX\_low} - V_{MAX}/3$ m.

forwarded RRQ packets, and thus a shorter network lifetime. The higher connectivity, on the other hand, leads to a better performance on the short term, as presented in fig. 6.24, showing the percentage of Found connections for the first 5000 seconds of a simulation run for the same strategy (DSR+Hops) with the two mobility models.

Note that the Cost-based strategy is penalized by both the short transmission range and the negative effect of mobility on route duration, and achieves thus lower performance than the Hop-based strategy, similarly to what was observed in Test Cases 1 and 2. The number of Received DATA packets, presented in fig. 6.22, confirms furthermore the poorer performance of the DSR+Cost strategy.

The results with medium transmission range R_{TX_med}, presented in figs. 6.25 and 6.26, confirm that the adoption of the LAR strategies when the network is sufficiently connected increases the network lifetime. In this case too, however, the adoption of the Cost-based metric has a negative effect on network performance, so that the LAR+Hops strategy is still the overall best solution in this scenario. Note that this conclusion differs from what was obtained for the same transmission range in Test Cases 1 and 2. In particular, in Test Case 2 the negative effect of mobility was compensated by the energy saving achieved with a higher number of hops, so that for the transmission

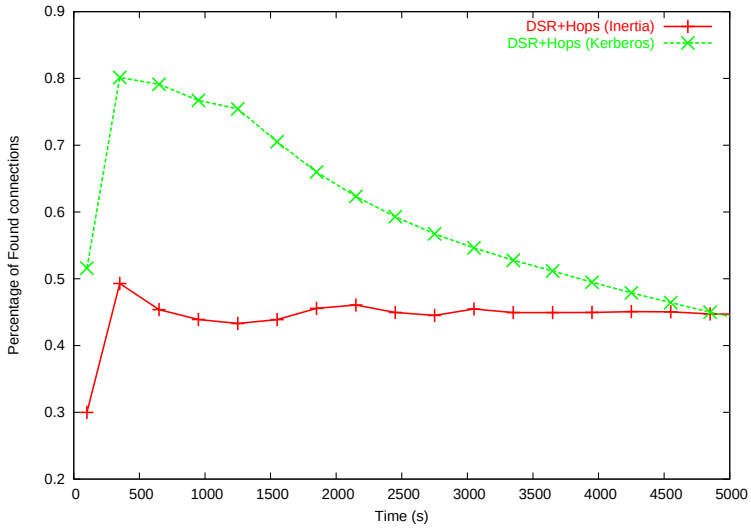

Figure 6.24: Percentage of Found connections in a single simulation run for the DSR+Hops strategy with *Inertia* and *Kerberos* mobility models.

range R_{TX_med}, the LAR+Cost led to the best performance (fig. 6.15). This is no longer true in Test Case 3, due the peculiar mobility pattern characterizing the *Kerberos* model. In subsection 6.2.1 it was showed in fact that the *Kerberos* mobility model is characterized by a shorter inter-group link duration if compared to *Inertia*: as a consequence, the adoption of inter-group links in a connection has a stronger negative effect on route stability. The Cost-based strategies are more sensible to this effect, since the selected routes are characterized by a high number of hops, and there is thus a significant probability for a route to include one or more inter-group links. The lower route stability is clearly indicated by the fact that, despite the highest number of Found connections, the LAR+Cost strategy is capable of delivering a lower number of DATA packets than the LAR+Hops strategy: this is due to the fact that routes selected with the LAR+Cost strategy have a shorter duration and thus the corresponding connections are able to deliver a lower number of packets before being broken for a lack of connectivity.

Results with high transmission range R_{TX_high} (figs. 6.27 and 6.28) further confirm that Cost-based strategies are heavily affected by the mobility pattern of the *Kerberos* model. Although in fact the higher network connectivity allows a higher average number of packets to be delivered in each connection for all the four strategies, yet a significant gap can be observed between Hop-based and Cost-based strategies.

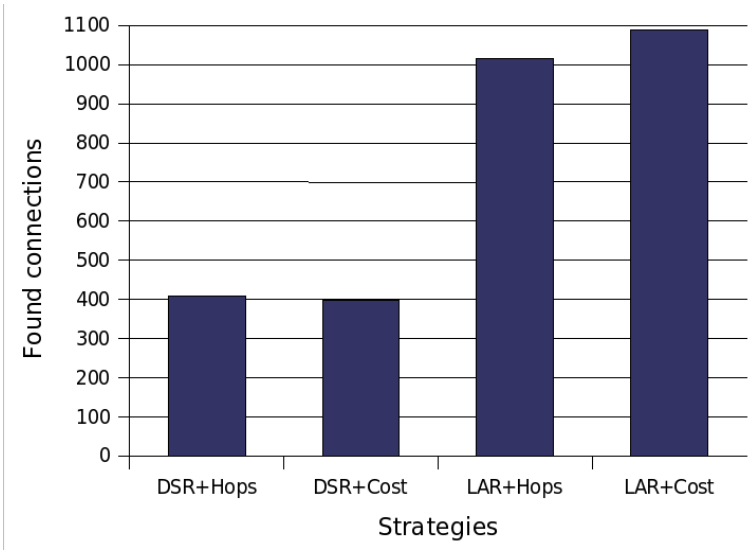

Figure 6.25: Found connections for Test Case 3 with R_{TX_med}.

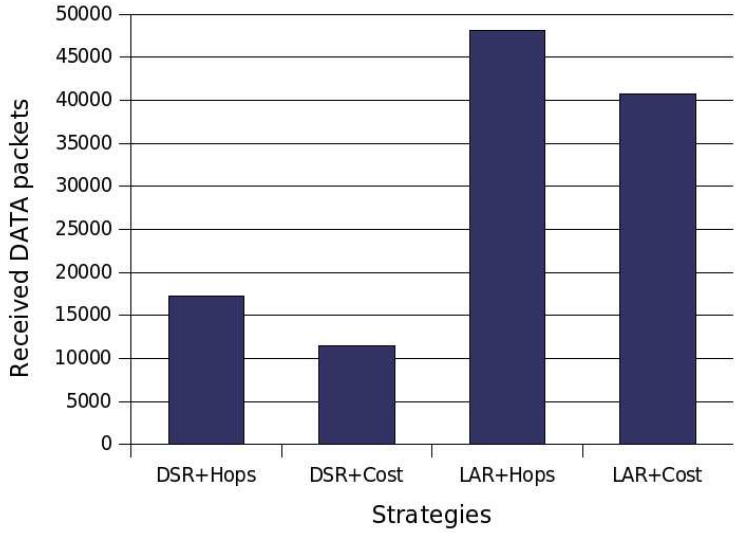

Figure 6.26: Received DATA packets for Test Case 3 with R_{TX_med}.

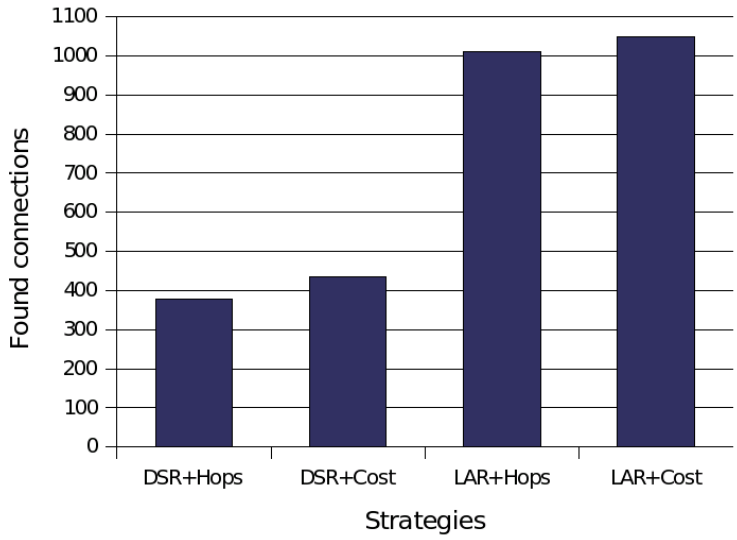

Figure 6.27: Found connections for Test Case 3 with R_{TX_high}.

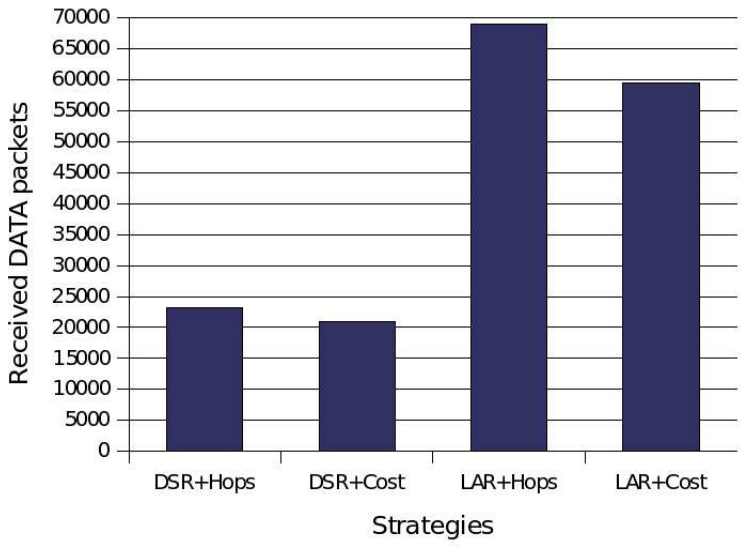

Figure 6.28: Received DATA packets for Test Case 3 with R_{TX_high}.

Quite interestingly, as the transmission range (and consequently the *Kerberos Range*) increases, the results in Test Case 3 approach those obtained

in Test Case 2; this is coherent with the fact that, as the *Kerberos Range* increases, there is a lower and lower probability for a node to remain isolated from its own group and thus to be forced to modify its mobility pattern. As a consequence, the *Kerberos* mobility model will more and more fall back to the standard *Inertia* mobility model.

6.4 Effect of packet decision model

In order to perform simulations presented in the previous section, the decision on the correct reception of a packet sent by a transmitter T to a receiver R was taken based on the following rules:

- packets received from a distance $d_{TR} \leq R_{TX}$ are subject to a decision taken by evaluating the average SNR per bit, taking into account both terminal noise and MUI noise generated by other packets in air.
  This allows to evaluate the bit error probability P_b. Starting from P_b, the packet error probability P_p is then evaluated based on the same hypothesis considered in section 2.1.1, and then the decision is taken by extracting a random value and comparing it with the P_p value.
- packets received from a distance $d_{TR} \geq R_{TX}$ are discarded a priori.

This approach was adopted for two reasons:

- Provide a clear definition of connected vs disconnected terminals
- Analyze network behavior independently from the effect of thermal noise: the R_{TX} is defined in fact as the distance at which the maximum transmit power allows for an average SNR per bit of X dB. The choice of X sufficiently high allows thus to guarantee that reception errors within the transmission range are almost always caused by interference. In the results presented in the previous section X was set to 20 dB, in order to guarantee such condition.

Note that the above approach did not affect the evaluation of interference noise, since the contribution of each packet in air to interference noise was evaluated up to the maximum distance between two terminals in the network, taking into account for each packet the distance between the transmitter and the victim receiver.

In a second set of simulations, we investigated the behavior of the above strategies when no packet is discarded a priori, in order to see how different route selections impact network performance when there is no clear differentiation between connected and disconnected terminals.

As an example, fig. 6.29 shows the behavior of the DSR+Hops and DSR+Cost strategies under this hypothesis with N=12 nodes with a $R_{TX} = 40$ m and the SNR X at this distance set to 20 dB. Figure 6.30 presents the results obtained in the same scenario with the assumptions made in the previous section.

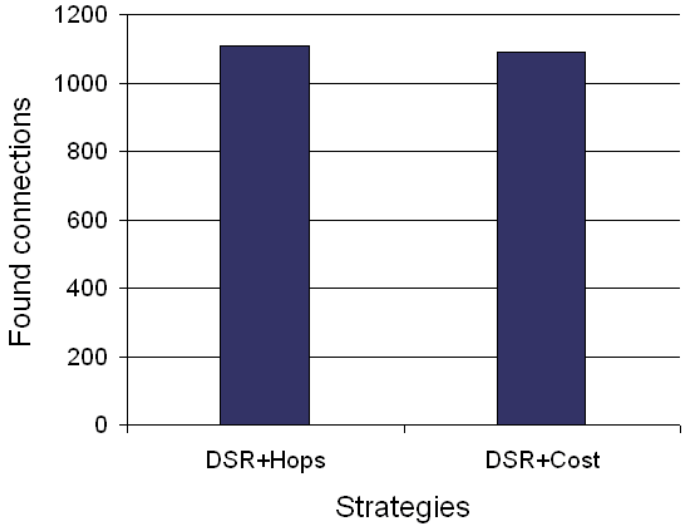

Figure 6.29: Number of Found Connections with no packets discarded a priori for the DSR+Hops and DSR+Cost strategies.

Interestingly, the possibility of having packets discarded for noise leads to a better performance of the Cost-based strategies, since the selection of a path formed by a higher number of hops guarantees in almost all cases that the target thermal SNR is met at the receiver. In the case of Hop-based strategies, oppositely, the choice of minimizing the number of hops may lead to the selection of routes that are subject to a higher packet error rate, due to the choice of one or more hops characterized by a low SNR. This effect compensates for the reduced efficiency of Cost-based strategies in conditions of low transmission range addressed in the previous section.
Interestingly, the introduction of this refined reception model also influences the fairness in power consumption throughout the network: the Cost-based and Hops-based strategies obtain in fact a similar number of overall Found connections, but the power consumption is clearly more uniform all over the network in the case of Cost-based strategy, as shown in fig. 6.31.

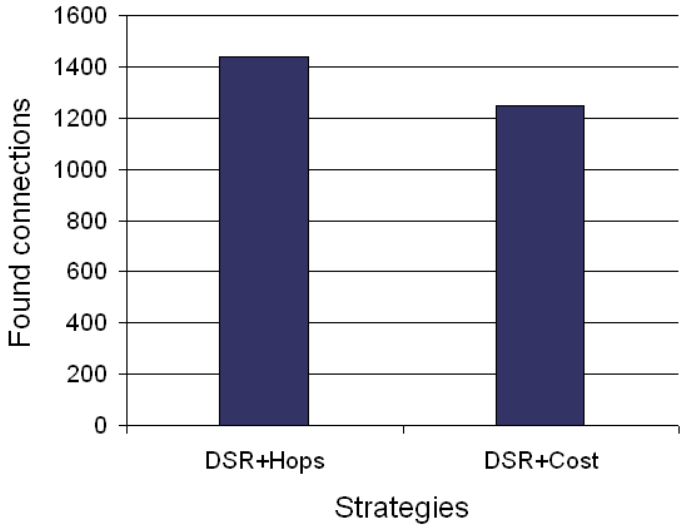

Figure 6.30: Number of Found Connections with the packet reception model of sec. 6.3 for the DSR+Hops and DSR+Cost strategies.

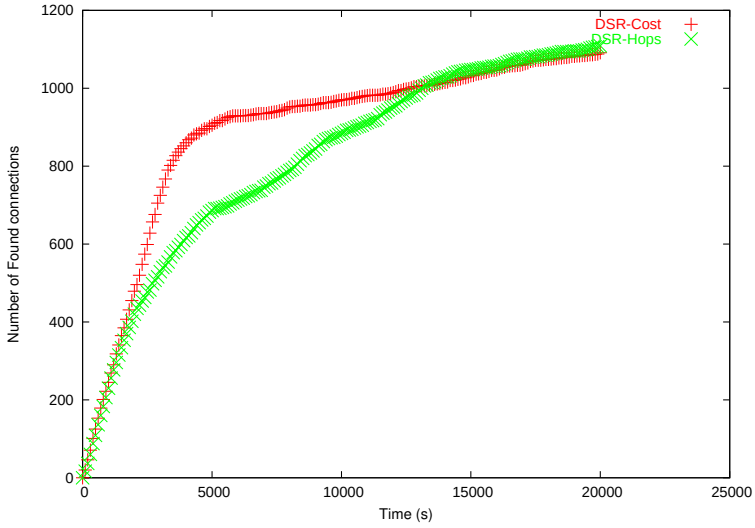

Figure 6.31: Number of Found Connections as a function of time with no packets discarded a priori, for the DSR+Hops and DSR+Cost strategies.

The higher fairness is also confirmed by the number of Found connections achieved before the first terminal runs out of energy, shown in fig. 6.32. This figure clearly shows that the adoption of the Cost-based function leads to a significantly higher network operation time.
As far as the effect of LAR is regarded, this second set of simulations confirms that the introduction of position information in routing leads to a significant increase of network lifetime, thanks to a reduced power consumption in the emission of control packets, as shown again in fig. 6.32, presenting the number of Found connections before the first terminal runs out of energy.

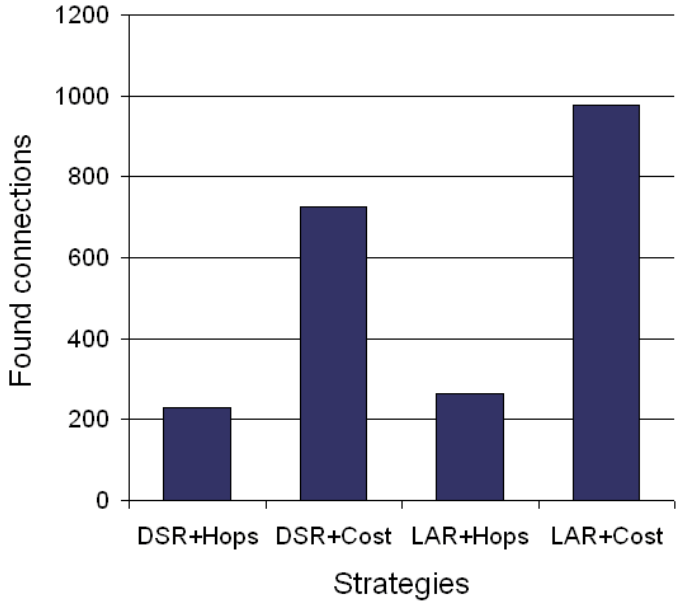

Figure 6.32: Number of Found Connections before the first terminal runs out of energy, with no packets discarded a priori.

Chapter 7

Conclusions

This work addressed the issue of power-aware, location-aided MAC and routing design for Ultra Wide Band networks. The theoretical analysis and the review of the existing solutions for several key aspects of Medium Access Control in wireless networks led to the conclusion that the specific features of UWB, identified in its high robustness to MUI and its accurate ranging capability, allow for innovative design of both MAC and routing modules.
The proposed solution is based on a Hybrid multi-channel MAC protocol enabled by the multiple access capability guaranteed by the use of Time Hopping codes; the proposed MAC also guarantees the capability of retrieving distance information without additional control overhead, by combining the ranging procedure with the link setup typical of a Hybrid multi-channel MAC protocol. Simulation results, coherently with the theoretical analysis available in the literature, confirmed that the proposed MAC is well suited for the case of low data rate UWB networks.
The distance information provided by the MAC is exploited in a twofold manner: first, it constitutes the basis for the definition of an UWB specific cost function, aiming at optimizing network behavior by including in the definition of the cost of a link several key aspects determining network performance, such as emitted synchronization overhead, emitted power, and Multi User Interference; second, it is the key input in the deployment of a distributed positioning system capable of providing position information in both indoor and outdoor environments. To this aim, the problem of distributed positioning in wireless networks was analyzed, and a protocol suitable for low data rate UWB networks was identified and implemented.
Finally, the position information made available by the positioning protocol is used at the routing layer for reducing the routing control overhead and thus improving the efficiency in the use of available power in each terminal. The location-based routing protocol was selected based on the analysis of both the

characteristics of existing location-based routing protocols and the specific requirements posed by an UWB-base distributed positioning system.
The effectiveness of the location-based, power-aware approach for MAC and routing in low bit rate UWB networks introduced above was then tested by means of simulations. The results of such analysis highlight that the exploitation of ranging information provided by the UWB physical layer may effectively extend network lifetime without significant effects on short term network performance.
The analysis also pointed out the direct relation between the scenario in which the network is deployed and the performance of the selected routing strategy. In particular, network connectivity, transmission range and mobility pattern of the nodes are all factors which can influence the performance of a routing strategy and, most important, may have a different effect on different strategies.
Results showed in fact that, depending on the scenario, the best network performance can be obtained by means of either a full exploitation of the positional information provided by UWB in both routing algorithm and metric (as in Test Case 1 and 2), or a partial one, limited to the adoption of the location information in the routing algorithm (as in Test Case 3).
The flexibility guaranteed by the cost function defined in section 5.2, however, enables a smooth transition from a fully power-aware metric to a traditional hop-based metric, and thus allows for a fine adaptation of the MAC and routing strategy to any network scenario, ranging between the two extreme cases of sparse networks of highly mobile nodes and dense networks of fixed nodes.

Bibliography

[1] D. B. Johnson and D. A. Maltz, *Dynamic Source Routing in Ad hoc Wireless Networks*. Kluwer Academic Publishers, 1996, ch. 5, pp. 153–181.

[2] M. G. Di Benedetto and G. Giancola, *Understanding Ultra Wide Band Radio Fundamentals*. Prentice Hall, 2004.

[3] "Revision of Part 15 of the Commission's rules Regarding Ultra-Wideband Transmission Systems: First report and order," Federal Communications Commission, Tech. Rep. 02-48, April 2002.

[4] S. B. Sorensen, "ETSI UWB Activities," in *ULTRA WIDE BAND (UWB) COLLOQUIUM*, July 2003, available for download at http://www.radio.gov.uk/topics/uwb/etsi-uwb_activities.pdf.

[5] ECC, "Draft ECC report on the protection requirements of radiocommunication systems below 10.6 ghz from generic uwb applications," CEPT, http://www.ero.dk/47BFEE2E-3CBD-4700-B086-1141C52628E0?frames=no\&, Tech. Rep. 64, November 2004, draft.

[6] M. G. Di Benedetto and B. R. Vojcic, "Ultra Wide Band (UWB) Wireless Communications: A Tutorial," *Journal of Communication and Networks, Special Issue on Ultra-Wideband Communications*, vol. 5, no. 4, pp. 290–302, December 2003.

[7] M. Z. Win and R. A. Scholtz, "Ultra-Wide Bandwidth Time-Hopping Spread-Spectrum Impulse Radio for Wireless Multiple-Access Communication," *IEEE Transactions on Communications*, vol. 48, no. 4, pp. 679–691, April 2000.

[8] R. Roberts, "XtremeSpectrum CFP Document," July 2003, available at http://grouper.ieee.org/groups/802/15/pub/2003/Jul03/03154r3P802-15_TG3%a-XtremeSpectrum-CFP-Documentation.pdf.

[9] "IEEE 802.15.3 MAC standard," Available at http://www.ieee.org/.

[10] J. C. Haartsen, "The Bluetooth Radio System," *IEEE Personal Communications*, vol. 7, no. 1, pp. 28–36, February 2000.

[11] "U.C.A.N. project official web page," http://www.ucan.biz/.

[12] F. Legrand, I. Bucaille, S. Hthuin, L. De Nardis, M. G. Di Benedetto, G. Spera, L. Blazevic, and P. Rouzet, "Description of positioning/routing algorithms and MAC protocol architecture," European Union project UCAN, Deliverable D42-2, July 2003.

[13] P. Kinney, "ZigBee Technology: Wireless Control that Simply Works," Available at http://www.zigbee.org/imwp/idms/popups/pop_download.asp?contentID=812, October 2003.

[14] "P.U.L.S.E.R.S. project official web page," http://www.pulsers.net/.

[15] M. L. Welborn, "System considerations for ultrawideband wireless networks," in *IEEE Radio and Wireless Conference*, August 2001, pp. 5–8.

[16] I. Guvenc and H. Arslan, "On the modulation options for UWB Systems," in *IEEE Military Communications Conference*, vol. 2, October 2003, pp. 892–897.

[17] X. Huang and Y. Li, "Generating Near-White Ultra-Wideband Signals with Period Extended PN Sequences," in *IEEE Vehicular Technology Conference*, May 2001, pp. 1184–1188.

[18] J. Foerster, "The performance of a Direct-Sequence Spread Ultra-Wideband system in the presence of Multipath, Narrowband Interference, and Multiuser Interference," in *IEEE Conference on Ultra Wideband Systems and Technologies*, May 2002, pp. 87–91.

[19] B. R. Vojcic and R. L. Pickholtz, "Direct-Sequence Code Division Multiple Access for Ultra-Wide Bandwidth Impulse Radio," in *IEEE Military Communications Conference*, October 2003, pp. 898–902.

[20] J. R. Foerster, V. Somayazulu, and S. Roy, "A Multi-Banded System Architecture for Ultra-Wideband Communications," in *IEEE Military Communications Conference*, October 2003, pp. 903–908.

[21] P.-C. Yeh, J. D. Choi, S. Zummo, M. D. Casciato, and W. E. Stark, "Performance Analysis of Coded Multi-Carrier Wideband Systems over Multipath Fading Channels," in *IEEE Military Communications Conference*, October 2003, pp. 909–914.

[22] A. Batra, "Multi-band OFDM Physical Layer Proposal for IEEE 802.15 Task Group 3a," September 2003, available at http://www.multibandofdm.org/papers/15-03-0268-01-003a-Multi-band-CFP-D%ocument.pdf.

[23] G. Giancola, L. De Nardis, and M. G. Di Benedetto, "Multi User Interference in Power-Unbalanced Ultra Wide Band systems: Analysis and Verification," in *IEEE Conference on Ultra Wideband Systems and Technologies*, November 2003, pp. 325–329.

[24] M. G. Di Benedetto, L. De Nardis, M. Junk, and G. Giancola, "$(UWB)^2$: Uncoordinated, Wireless, Baseborn medium access for UWB communication networks," 2005, to appear in Mobile Networks and Applications special issue on WLAN Optimization at the MAC and Network Levels.

[25] H. F. Harmuth, *Radiation of Nonsinusoidal Electromagnetic Waves*. New York: Academic Press, Inc., 1990.

[26] A. F. Kardo-Sysoev, "Generation and Radiation of UWB Signals," in *International Workshop on Ultra Wideband Systems*, June 2003.

[27] I. I. Immoreev and N. Sinyavin, "Features of ultra-wideband signals' radiation," in *IEEE Conference on Ultra Wideband Systems and Technologies*, May 2002, pp. 345–349.

[28] J. T. Conroy, J. L. LoCicero, and D. R. Ucci, "Communication techniques using monopulse waveforms," in *IEEE Military Communications Conference*, November 1999, pp. 1181–1185.

[29] M. Hamalainen, V. Hovinen, J. Iinatti, and M. Latva-aho, "In-band Interference Power Caused by Different Kinds of UWB Signals at UMTS/WCDMA Frequency Bands," in *IEEE Radio and Wireless Conference*, August 2001, pp. 97–100.

[30] M. Ghavami, L. B. Michael, S. Haruyama, and R. Kohno, "A Novel UWB Pulse Shape Modulation System," *Wireless Personal Communications*, vol. 23, no. 1, pp. 105–120, October 2002.

[31] B. Parr, B. Cho, K. Wallace, and Z. Ding, "A novel ultra-wideband pulse design algorithm," *IEEE Communications Letters*, vol. 7, no. 5, pp. 219–221, May 2003.

[32] L. Bin, E. Gunawan, and L. C. Look, "On the BER Performance of TH-PPM UWB Using Parr's Monocycle in the AWGN Channel," in *IEEE Conference on Ultra Wideband Systems and Technologies*, November 2003, pp. 403–407.

[33] H. Sheng, P. Orlik, A. M. Haimovich, L. J. Cimini Jr., and J. Zhang, "On the Spectral and Power Requirements for Ultra Wideband Transmission," in *IEEE International Conference on Communications*, vol. 1, May 2003, pp. 738–742.

[34] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.

[35] A. C. V. Gummalla and J. O. Limb, "Wireless Medium Access Control Protocols," *IEEE Communications Surveys*, vol. 3, no. 2, June 2000.

[36] H. Choi and N. Moayeri, "Evaluation procedure for 802.16 MAC Protocols," Available at http://grouper.ieee.org/groups/802/16/tg1/mac/pres/802161mp-00_16.pdf, 2000.

[37] K. Nahrstedt and R. Steinmetz, "Resource Management in Networked Multimedia Systems," *IEEE Computer*, vol. 28, no. 5, pp. 52–63, May 1995.

[38] X. Xiao and L. Ni, "Internet QoS: A Big Picture," *IEEE Network*, vol. 13, no. 2, pp. 8–18, March/April 1999.

[39] P. Karn, "MACA - A new Channel Access Protocol for Packet Radio," in *ARRL/CRRL Amateur Radio Ninth Computer Networking Conference*, 1990, pp. 134–140.

[40] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: A medium access protocol for wireless LANs," in *ACM Conference on Applications, Technologies, Architectures and Protocols for Computer Communication (SIGCOMM '94)*, 1994, pp. 212–225.

[41] F. Talucci and M. Gerla, "MACA-BI (MACA By Invitation): A wireless MAC protocol for high speed ad hoc networking," in *IEEE International Conference on Universal Personal Communications*, vol. 2, October 1997, pp. 913–917.

[42] C. L. Fullmer and J. J. Garcia-Luna-Aceves, "Floor Acquisition Multiple Access (FAMA) for Packet Radio Networks," in *Conference on Applications, Technologies, Architectures and Protocols for Computer Communication (SIGCOMM '95)*, 1995, pp. 262–273.

[43] B. P. Crow, I. Widjaja, J. G. Kim, and P. T. Sakai, "IEEE 802.11: Wireless Local Area Networks," *IEEE Communications Magazine*, vol. 35, no. 9, pp. 116–126, September 1997.

[44] F. A. Tobagi and L. Kleinrock, "Packet Switching in Radio Channels: Part II - The Hidden Terminal Problem in Carrier Sense Multiple Access and the Busy Tone Solution," *IEEE Transactions on Communications*, vol. COM–23, no. 12, pp. 1417–33, December 1975.

[45] J. Deng and Z. J. Haas, "Dual Busy Tone Multiple Access (DBTMA): A New Medium Access Control for Packet Radio Networks," in *IEEE International Conference on Universal Personal Communications*, vol. 2, October 1998, pp. 973–977.

[46] T. Makansi, "Transmitter-Oriented Code Assignment for Multihop Radio Networks," *IEEE Transactions on Computers*, vol. 35, no. 12, pp. 1379–1382, December 1987.

[47] M. Mouly and M. B. Pautet, *The GSM System for Mobile Communication.* Cell&Sys, 1992.

[48] J. Khun-Jush, G. Malmgren, P. Schramm, and J. Torsner, "Overview and performance of HIPERLAN type 2 - a standard for broadband wireless communications," in *IEEE Vehicular Technology Conference*, vol. 1, May 2000, pp. 112–117.

[49] N. Bambos, "Toward power-sensitive network architectures in wireless communications: Concepts, issues, and design aspects," *IEEE Personal Communications*, vol. 5, no. 1998, pp. 50–59, June 1998.

[50] S. Valaee and B. Li, "Distributed Call Admission control for Ad-hoc networks," in *IEEE Vehicular Technology Conference*, vol. 2, September 2002, pp. 1244–1248.

[51] Y. Yang and R. Kravets, "Contention-aware admission control for ad hoc networks," University of Illinois in Urbana-Champaign, Tech. Rep. 2003-2337, April 2003.

[52] F. Cuomo, C. Martello, A. Baiocchi, and F. Capriotti, "Radio Resource Sharing for Ad Hoc Networking with UWB," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1722–1732, December 2002.

[53] J. Nagle, "On Packet Switches with Infinite Storage," *IEEE Transactions on Communications*, vol. 35, no. 4, pp. 435–438, April 1987.

[54] A. Demers, S. Keshav, and S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm," in *ACM Conference on Applications, Technologies, Architectures and Protocols for Computer Communication (SIGCOMM '89)*, September 1989, pp. 1–12.

[55] S. Lu, V. Bharghavan, and R. Srikant, "Fair scheduling in wireless packet networks," *IEEE/ACM Transactions on Networking*, vol. 7, no. 4, pp. 473–489, August 1999.

[56] H. Aida, Y. Tamura, Y. Tobe, and H. Tokuda, "Wireless packet scheduling with signal-to-noise ratio monitoring," in *IEEE Conference on Local Computer Networks*, November 2000, pp. 32–41.

[57] N. Golmie, N. Chevrollier, and I. ElBakkouri, "Interference aware Bluetooth packet scheduling," in *IEEE Global Telecommunications Conference*, vol. 5, November 2001, pp. 2857–2863.

[58] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Transactions on Information Theory*, vol. 46, no. 2, pp. 388–404, March 2000.

[59] E. S. Sousa, "On channel power sensing in terrestrial spread spectrum packet radio networks," in *IEEE Conference on Computer Communications (INFOCOM)*, vol. 3, April 1989, pp. 1072–1077.

[60] R. E. Kahn, S. A. Gronemeyer, J. Burchfiel, and R. C. Kunzelman, "Advances in Packet Radio Technology," *Proceedings of the IEEE*, vol. 66, no. 11, pp. 1468–1496, November 1978.

[61] I. A. Getting, "The Global Positioning System," *IEEE Spectrum*, vol. 30, no. 12, pp. 36–38, 43–47, December 1993.

[62] P. Baldi, L. De Nardis, and M. G. Di Benedetto, "Modeling and Optimization of UWB Communication Networks through a flexible cost function," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1733–1744, December 2002.

[63] A. J. Goldsmith and S. B. Wicker, "Design Challenges for Energy-Constrained Ad-Hoc Wireless Networks," *IEEE Wireless Communications*, vol. 9, no. 4, pp. 8–27, August 2002.

[64] D. Raychaudhuri, "Performance Analysis of Random Access Packet-Switched Code Division Multiple Access Systems," *IEEE Transactions on Communications*, vol. COM-29, no. 6, pp. 895– 901, June 1981.

[65] S. Dastangoo, B. R. Vojcic, and J. N. Daigle, "Performance Analysis of Multi-Code Spread Slotted ALOHA (MCSSA) System," *IEEE Global Telecommunications Conference*, vol. 3, pp. 1839–1847, November 1998.

[66] E. S. Sousa and J. A. Silvester, "Spreading Code protocols for Distributed Spread-Spectrum Packet Radio Networks," *IEEE Transactions on Communications*, vol. COM–36, no. 3, pp. 272–281, March 1988.

[67] J. J. Garcia-Luna-Aceves and J. Raju, "Distributed Assignment of codes for multihop packet-radio networks," in *IEEE Military Communications Conference*, vol. 1, November 1997, pp. 450–454.

[68] M. S. Iacobucci and M. G. Di Benedetto, "Computer method for pseudo-random codes generation," National Italian patent RM2001A000592, 2002, under registration for international patent.

[69] A. Polydoros and J. Silvester, "Slotted Random Access Spread-Spectrum Networks: An Analytical Framework," *IEEE Journal on Selected Areas in Communications*, vol. 5, no. 6, pp. 989–1002, July 1987.

[70] J. Hallberg, M. Nilsson, and K. Synnes, "Positioning with Bluetooth," in *International Conference on Telecommunications*, vol. 2, February 2003, pp. 954–958.

[71] J. Hightower, C. Vakili, G. Borriello, and R. Want, "Design and Calibration of the SpotON Ad-Hoc Location Sensing System," August 2001, available at http://seattle.intel-research.net/people/jhightower/pubs/.

[72] J. G. Proakis, *Digital communications*, 3rd ed. McGraw-Hill International Editions, Inc., 1995.

[73] C. Drane, M. Macnaughtan, and C. Scott, "Positioning GSM telephones," *IEEE Communications Magazine*, vol. 36, no. 4, pp. 46–54, 59, April 1998.

[74] C. Savarese, "Robust Positioning Algorithms for Distributed Ad-Hoc Wireless Sensor Networks," Master's thesis, EECS Department, University of California at Berkeley, 2001.

[75] L. Doherty, L. El Ghaoui, and K. S. J. Pister, "Convex Position Estimation in Wireless Sensor Networks," in *IEEE Conference on Computer Communications (INFOCOM)*, vol. 3, April 2001, pp. 1655–1663.

[76] C. Savarese, J. Rabaey, and J. Beutel, "Location in distributed ad-hoc wireless sensor networks," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, 2001, pp. 2037–2040.

[77] D. Niculescu and B. Nath, "Ad hoc Positioning System (APS)," in *IEEE Global Telecommunications Conference*, vol. 5, November 2001, pp. 2926–2931.

[78] J. P. Hubaux, J. Y. Le Boudec, S. Giordano, M. Hamdi, L. Blazevic, L. Buttyan, and M. Vojnovic, "Towards mobile ad-hoc wans: Terminodes," in *IEEE Wireless Communications and Networking Conference*, vol. 3, September 2000, pp. 1052–1059.

[79] S. Capkun, M. Hamdi, and J. P. Hubaux, "GPS-free positioning in mobile Ad-Hoc networks," in *Hawaii International Conference On System Sciences*, 2001, pp. 3481–3490.

[80] R. A. Fleming and C. E. Kushner, "Spread Spectrum Localizers," U.S. Patent No. 6,002,708, 1997.

[81] H. Urkowitz, *Signal Theory and Random Processes.* Artech House, 1983.

[82] J. Y. Lee and R. A. Scholtz, "Ranging in a dense multipath environment using an UWB radio link," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1677–1683, December 2002.

[83] R. J. Fontana, E. Richley, and J. Barney, "Commercialization of an ultra wideband precision asset location system," in *IEEE Conference on Ultra Wideband Systems and Technologies*, November 2003, pp. 369–373.

[84] S. Singh and C. S. Raghavendra, "PAMAS: Power aware multiaccess protocol with signalling for ad hoc networks," *ACM SIGCOMM Computer Communication Review*, vol. 28, no. 3, pp. 5–26, July 1998.

[85] S. Banerjee and A. Misra, "Energy Efficient Reliable Communication for Multi-hop Wireless Networks," *Journal of Wireless Networks (WINET)*, 2003, available at http://www.cs.wisc.edu/~suman/pubs/winet03.pdf.

[86] Y. Xu, J. Heidemann, and D. Estrin, "Adaptive Energy-Conserving Routing for Multihop Ad Hoc Networks," USC/Information Science Institute, Tech. Rep. 2000-527, October 2000.

[87] V. Rodoplu and T. H. Meng, "Minimum Energy Mobile Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, pp. 1333–1344, August 1999.

[88] C. K. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE Communications Magazine*, vol. 39, no. 6, pp. 138–147, June 2001.

[89] M. G. Di Benedetto and P. Baldi, "A model for self-organizing large-scale wireless networks," in *International Workshop on 3G Infrastructure and Services*, 2001, pp. 210–213.

[90] L. De Nardis, P. Baldi, and M. G. Di Benedetto, "UWB ad-hoc networks," in *IEEE Conference on Ultra Wideband Systems and Technologies*, May 2002, pp. 219–224.

[91] B. Karp and H. T. Kung, "Gpsr: Greedy perimeter stateless routing for wireless networks," in *ACM International Conference on Mobile Computing and Networking (Mobicom)*, 2000, pp. 243–254.

[92] D. Kim, Y. Choi, and C. K. Toh, "Location-aware long lived-route selection in wireless ad hoc networks," in *IEEE Vehicular Technology Conference*, vol. 4, 2000, pp. 1914–1919.

[93] S. Basagni, I. Chlamtac, V. R. Syrotiuk, and B. A. Woodward, "A distance routing effect algorithm for mobility (DREAM)," in *ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom)*, 1998, pp. 25–30.

[94] Y. B. Ko and N. H. Vaidya, "Location aided routing in mobile ad hoc networks," in *ACM International Conference on Mobile Computing and Networking (Mobicom)*, 1998, pp. 66–75.

[95] ——, "Optimizations for location aided routing in mobile ad hoc networks," CS Dept, Texas A & M University, Tech. Rep. 98-023, 1998.

[96] Y.-C. Hu and D. Johnson, "Caching Strategies in On-Demand Routing Protocols for Wireless Ad Hoc Networks," in *ACM International Conference on Mobile Computing and Networking (MobiCom)*, August 2000, pp. 231–242.

[97] P.M. Melliar-Smith. E.M. Royer and L.E. Moser, "An Analysis of the Optimum Node Density for Ad hoc Mobile Networks," in *IEEE International Conference on Communications (ICC)*, June 2001, pp. 857–861.

[98] B. Liang and Z. J. Haas, "Predictive distance-based mobility management for multidimensional PCS networks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 718–732, June 2003.

[99] S. Basagni, I. Chlamtac, and V. R. Syrotiuk, "Dynamic source routing for ad hoc networks using the global positioning system," in *IEEE Wireless Communications and Networking Conference (WCNC)*, September 1999, pp. 301–305.

[100] L. De Nardis, M. G. Di Benedetto, and S. Falco, *UWB Communications Systems - A Comprehensive Overview.* Hindawi Publishing Corporation, 2005, ch. 4 - Higher layer issues, Ad-Hoc and Sensor Networks, pp. 1–59.

[101] M. Bergamo, "System design specification for mobile multimedia wireless network (MMWN)," DARPA - project DAAB07-95-C-D156, Tech. Rep., 1996.

[102] X. Hong, M. Gerla, G. Pei, and C. C. Chiang, "A Group Mobility Model for Ad Hoc Wireless Networks," in *ACM International Workshop on Modeling and Simulation of Wireless and Mobile Systems (MSWiM)*, 1999, pp. 53–60.

[103] K. H. Wang and B. Li, "Group Mobility and Partition Prediction in Wireless Ad-Hoc Networks," in *IEEE International Conference on Communications (ICC)*, vol. 2, April 2002, pp. 1017–1021.

[104] R. Min and A. Chandrakasan, "A framework for energy-scalable communication in high-density wireless networks," in *International Symposium on Low Power Electronics and Design (ISLPED)*, 2002, pp. 36–41.